Modeling Peptide-binding Interactions and Polymer-binding Interactions and their Role in Mass Spectrometry

Eric Martineau

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the Doctor of Philosophy degree in Chemistry

Supervisor: Professor Paul Michael Mayer

Department of Chemistry
Faculty of Science
University of Ottawa

À mes parents, Claudette et Jacques,
qui m'ont supporté dans mes projets
et qui ont toujours cru en moi.
Ils m'ont tout donné, et ce,
de toutes les façons possibles.

To my parents, Claudette and Jacques,
who supported me in my projects
and always believed in me.
They gave me everything
in any possible ways.

You can’t always get what you want
But if you try sometimes well you might find
You get what you need

(M. Jagger/K. Richards)
The Rolling Stones, 1969

Abstract

As a first project, collision-induced dissociation experiments were carried out using electrospray ionisation mass spectrometry on gas phase complexes involving different poly(methylmetacrylate) oligomers with three amino acids: glycine, leucine, and phenylalanine. After acquiring breakdown diagrams, RRKM modeling was used to fit the experimental data in order to obtain the 0 K activation energy (E_0) and the entropy of activation ($\Delta^{\ddagger}S$). These thermodynamic data were then used to understand the competing dissociation channels observed (except for gas phase complexes involving glycine that had only one dissociation channel). Molecular dynamics simulated annealing calculations were carried on the gas phase complexes to understand further the energetic and entropic effects involved as well as the 3D conformation of these complexes. Valuable insight information was found on the 3D conformations, on a qualitative level. Using rotational constants and vibrational harmonic frequencies, it was possible to evaluate the entropy variation between the experimentally observed competing channels. Reasonable agreement was found between the experimental and theoretical variations of entropies. Finally, the proton affinity of poly(methylmetacrylate) oligomers is being discussed. Even though no absolute values for the proton affinity were found, the experimental and computational results help to understand the variation that accompanies the oligomers length.

The second project presents the development an efficient and reproducible screening method for identifying low molecular weight compounds that bind to amyloid β peptides (Aβ) peptides using electrospray ionization mass spectrometry (ESI-MS). Low molecular weight (LMW) compounds

capable of interacting with soluble Aβ may be able to modulate/inhibit the Aβ aggregation process and serve as potential disease-modifying agents for Alzheimer's disease. The present approach was used to rank the binding affinity of a library of compounds to Aβ1-40 peptide. The results obtained show that low molecular weight compounds bind similarly to Aβ1-42, Aβ1-40, as well as Aβ1-28 peptides and they underline the critical role of Aβ peptide charge motif in binding at physiological pH. Finally, some elements of structure-activity relationship (SAR) involved in the binding affinity of homotaurine to soluble Aβ peptides are discussed. As a third project, the gas phase binding of small molecules to the Aβ1-40 peptide generated by electrospray ionization has been explored with collision-induced dissociation mass spectrometry and kinetic rate theory. This project presents a simple procedure used to theoretically model the experimental breakdown diagrams for the Aβ1-40 peptide complexed with a series of aminosulfonate small molecules, namely homotaurine, 3-cyclohexylamino-2-hydroxy-1-propanesulfonic acid (CAPSO), 3-(1,3,4,9-tetrahydro-2H-β-carbolin-2-yl) propane-1-sulfonic acid, 3-(1,3,4,9-tetrahydro-2H-β-carbolin-2-yl)butane-1-sulfonic acid, and 3-(cyclohexylamino) propane-1-sulfonic acid. An alternative method employing an extrapolation procedure for the microcanonical rate constant, k(E), is also discussed.

Résumé

En guise de premier projet, des complexes formés de différents oligomères de poly(méthylmetacrylate) avec trois acides aminés (glycine, leucine et phénylalanine) ont été dissociés par collisions en phase gazeuse à l'aide d'un spectromètre de masse à électronébulisation avec triple quadripôle et cellule de collisions. Les diagrammes de dissociation ont ensuite été modélisés à l'aide de la théorie RRKM afin de déterminer l'énergie d'activation à 0 K (E_0) et l'entropie d'activation ($\Delta^{\ddagger}S$). Ces données thermodynamiques ont par la suite été utiles pour comprendre les voies de dissociation en compétition (sauf pour les complexes en phase gazeuse impliquant la glycine qui n'avaient qu'une voie de dissociation). Des simulations de dynamique moléculaire à recuit simulé ont été effectuées sur les complexes en phase gazeuse afin de visualiser les conformations 3D ainsi que d'évaluer les effets énergétique et entropique des dissociations. D'importantes informations qualitatives ont été acquises sur les conformations 3D. En utilisant les constantes rotationnelles et les fréquences harmoniques de vibration, il a été possible d'évaluer la variation d'entropie entre les voies de dissociations compétitives observées expérimentalement. Une certaine correspondance a été observée entre les variations d'entropie expérimentale et théorique. Finalement, l'affinité protonique des oligomères de poly(méthylmetacrylate) est abordée. Bien qu'aucune valeur absolue d'affinité protonique n'ait été trouvée, les résultats expérimentaux et théoriques ont aidés à comprendre la variation reliée à la longueur de l'oligomère.

Comme deuxième projet, une méthode de criblage efficace et reproductible (par ESI-MS) est décrite dans le but d'identifier des molécules à faible masse moléculaire capables de s'associer à des peptides amyloïde β. Ces molécules de faible masse moléculaire, ayant la capacité de s'associer au peptide Aβ en phase aqueuse, ont le potentiel de moduler et/ou d'inhiber le processus d'agrégation et donc, d'agir comme médicament potential contre la maladie d'Alzheimer. L'approche présentée a été utilisée afin de classer une banque de molécules, selon leur affinité à s'associer au peptide Aβ1-40. Les résultats obtenus démontrent que ces molécules de faible masse moléculaire s'associent de façon similaire aux peptides Aβ1-42, Aβ1-40 et Aβ1-28 et soulignent le rôle essentiel de la séquence des charges des peptides au pH physiologique. Finalement, certaines relations entre la structure moléculaire et l'affinité d'association de l'homotaurine au peptide Aβ sont discutées. Comme troisième projet, des complexes formés de molécules de faible masse moléculaire avec le peptide Aβ1-40 ont été dissociés par collisions en phase gazeuse à l'aide d'un spectromètre de masse à électronébulisation "quadripôle/temps de vol" muni d'une cellule de collisions. L' analyse des résultats a été effecutée à l'aide de la théorie cinétique. Ce projet présente une procédure simplifiée pour modéliser théoriquement les diagrammes de dissociations pour les complexes formés du peptide Aβ1-40 et d'une série d'aminosulfonate de faible masse moléculaire: l'homotaurine, l'acide 3-cyclohexylamino-2-hydroxy-1-propanesulfonique, l'acide 3-(1,3,4,9-tetrahydro-2*H*-β-carbolin-2-yl)butane-1-sulfonique, l'acide 3-(1,3,4,9-tetrahydro-2*H*-β-carbolin-2-yl)butane-1-sulfonique et l'acide 3-(cyclohexylamino) propane-1-sulfonique. Une procédure alternative s'appuyant sur une méthode d'approximation pour la constante de vitesse microcanonique, k(E), est également abordée.

Acknowledgements

My supervisor, Professor Paul M. Mayer for accepting me in his group as a graduate student with a joint industrial-academic project. Through the years, Paul has been an amazing supervisor with his scientific ideas, his endless patience, and his support in many ways. I will always be grateful to him for his guidance during these Ph.D. years. When a scientist proposes a hypothesis, he/she often expects a probable result. Unfortunately, when the expected result does not happen, many would simply reject it pertaining that it did not work. A good scientist forces himself to think about it with a different perspective to find a logical explanation... and this important quality, I have learned it from Paul.

Dr. Ahmed M. Aman, who was first my colleague, and then became my group leader. He was the first one to hear my idea of doing a Ph.D. and instantly the first one to support me in my project. Coming from a theoretical background, he is the reason why I developed a great interest about mass spectrometry. I thank him for all the scientific discussions and for always allowing me to use the instruments afterhours to explore the various ideas that popped out of my mind.

Many executives who have granted me different permissions (access and publication rights) or scientific advises, came rain or sunshine at Neurochem Inc. These persons are: Xianqi Kong, Francine Gervais, Daniel Delorme, Shona McDiarmid, Judith Paquin, and Nigel Levins.

My colleagues at Neurochem Inc. who contributed to certain aspects of the article published in *Journal of the American Society for Mass Spectrometry*, presented in this thesis.

My colleagues from Paul's group who contributed to some of the projects or with whom I shared interesting discussions in different occasions: Justin B. Renaud, Anne-Marie Boulanger, Abdulrahman M. Alhazmi, Annick D. St-Amant, and Yawei Lin.

The thesis defence committee for taking the time to read my work and provide stimulating discussion and valuable comments during the defence. The committee was composed of: Professor Lars Konermann (from University of Western Ontario), Professor Jeffrey C. Smith (from Carleton University), Professors Alain St-Amant and Maxim V. Berezovski (both from University of Ottawa).

To everyone else who has been kind to me along the way without mentioning them.

Remerciements

Mon superviseur de thèse, le professeur Paul M. Mayer pour m'avoir accepté dans son groupe à titre d'étudiant gradué avec un projet conjoint industriel-académique. Au fil du temps, Paul fut un superviseur incroyable tant par ses idées scientifiques, que par sa patience sans borne et sans oublier support à différents niveaux. Je lui serai toujours reconnaissant pour sa bienveillance durant ces années de doctorat. Lorsqu'un(e) scientifique émet une hypothèse, il/elle s'attend souvent à un résultat probable. Malheureusement, lorsque ce résultat n'est pas observé, il est trop souvent rejeté selon la justification que l'expérience n'a pas fonctionné. Un bon scientifique va se repencher sur le problème et tenter d'y trouver une explication logique. Cette qualité importante, je l'ai apprise de Paul.

Le Dr. Ahmed M. Aman, qui fut d'abord un collègue de travail avant de devenir mon chef de groupe. Il fut le premier à entendre mon idée de me lancer dans un projet de doctorat et, instantanément, m'a supporté. Provenant d'une formation théorique durant ma maîtrise, Ahmed est la raison pour laquelle j'ai développé un grand intérêt pour la spectrométrie de masse. Je le remercie pour toutes les discussions scientifiques sans compter sa permission de me laisser utiliser les instruments du laboratoire après les heures de travail pour tenter plusieurs expériences qui étaient le fruit de mon imagination.

Plusieurs dirigeants m'ont donné diverses permissions (droits d'accès et de publication) ainsi que des conseils scientifiques, peu importe la situation de Neurochem Inc. Ces personnes sont: Xianqi Kong, Francine Gervais, Daniel Delorme, Shona McDiarmid, Judith Paquin, and Nigel Levins.

Mes collègues chez Neurochem Inc. qui ont contribué à certaines sections de l'article publié dans *Journal of the American Society for Mass Spectrometry*, présenté dans cette thèse.

Mes collègues du groupe de Paul qui ont contribué à certains projets ou encore avec qui j'ai partagé d'intéressantes discussions: Justin B. Renaud, Anne-Marie Boulanger, Abdulrahman M. Alhazmi, Annick D. St-Amant, and Yawei Lin.

Les membres du comité de défense de thèse, pour avoir pris le temps de lire mon travail, pour avoir généré des discussions stimulantes et avoir donné des commentaires pertinents pendant la défense. Le comité était formé du professeur Lars Konermann (de l'Université Western Ontario), du professeur Jeffrey C. Smith (de l'Université Carleton) et des professeurs Alain St-Amant et Maxim V. Berezovski (tous deux de l'Université d'Ottawa).

À tous ceux et celles dont la présence m'a été bénéfique, mais que je ne mentionne point.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ABNR	Adapted-Basis Newton-Raphson minimization
AD	Alzheimer's disease
Aβ	Amyloid Beta peptide
Amber	Assisted Model Building with Energy Refinement
AM1	Austin Method 1: a semi-empirical method
B3LYP	Becke 3 term, Lee, Yang, Par: a density functional method
BA	Binding affinity
BIRD	Blackbody Infrared Radiative Dissociation
N_A	Avogadro's number, equals to $6.022\ 14 \times 10^{23}\ mol^{-1}$
k_B	Boltzmann constant, equals to $1.380\ 66 \times 10^{-23}\ J\ K^{-1}$
k(T)	Canonical rate constant
E_{CM}	Centre-of-mass collision energy
CI	Chemical ionization
CE	Collision energy
CID	Collision-induced dissociation
CNDO	Complete Neglect of Differential Overlap
Da	Dalton
DNA	Deoxyribonucleic acid
P(E,T)	Distribution of internal energies at a particular temperature T
ρ(E)	Density of states at a particular energy E
EI	Electron impact

ESI	Electrospray ionization
E_0	0 K activation energy
$\Delta^{\ddagger}S$	Entropy of activation
FT-ICR	Fourier transform ion cyclotron resonance
R	Gas constant, equals to 8.314 51 J mol^{-1} K^{-1}
Gly	Glycine
GAG	Glycosaminoglycans
LDF	London Dispersion Forces
Leu	Leucine
LMW	Low molecular weight
MALDI	Matrix-Assisted Laser Desorption Ionization
m/z	Mass-to-charge ratio
MNDO	Moderate Neglect of Differential Overlap
MD	Molecular Dynamics
MD-SA	Molecular Dynamics Simulated Annealing
MM	Molecular Mechanics
MWD	Molecular weight distribution
mer	monomer unit
k(E)	Microcanonical rate constant
MS	Mass spectrometry
NMR	Nuclear magnetic resonance
ε_0	Vacuum permittivity, equals to $8.854\ 19 \times 10^{-12}$ C^2 J^{-1} m^{-1}
Phe	Phenylalanine

h	Planck constant, equals to $6.626\ 08 \times 10^{-34}$ J s
PES	Potential energy surface
PBA	poly(n-butylmetacrylate)
PMMA	poly(methylmetacrylate)
PVac	poly(vinylacetate)
PDB	Protein Data Bank
PG	Proteoglycan
PA	Proton Affinity
QET	Quasi-equilibrium theory
σ	Reaction degeneracy in RRKM theory or molecular symmetry number
RNA	Ribonucleic acid
RRK	Rice-Ramsperger-Kassel theory
RRKM	Rice-Ramsperger-Kassel-Marcus theory
SA	Simulated Annealing
c	speed of light, equals to $2.997\ 92 \times 10^{8}$ m s^{-1}
SD	Steepest Descent minimization
$N^{\ddagger}(E\text{-}E_0)$	Transition state sum of states above the 0 K activation energy (E_0)
T	Temperature
3D	Three-dimensional
TOF	Time-of-flight
TS	Transition state
vdW	van der Waals
VTST	Variational transition state theory

Chapter 1: Introduction

1.1 Objectives

The work presented in this thesis focuses on the intermolecular forces acting between chemical species in the gas phase.

Using CID mass spectrometry, RRKM modeling and molecular modeling, the first project investigates the dissociation of gas phase complexes involving different poly(methylmetacrylate) oligomers ($PMMA_n$) with three amino acids (AA): glycine, leucine, and phenylalanine. After acquiring experimental breakdown diagrams of diverse $PMMA_n$/AA systems, the 0 K activation energy, E_0, and the entropy of activation, $\Delta^{\ddagger}S$, will be calculated using RRKM theory. Finally, 3D conformations of the $PMMA_n$/AA systems obtained by molecular modeling will be presented to visualize and to understand the energetic and entropic trends found.

The second project describes an efficient and reproducible screening method for identifying low molecular weight compounds that bind to amyloid β peptides (Aβ) peptides using electrospray ionization mass spectrometry (ESI-MS). This screening method will help to demonstrate the importance of the peptides charge motifs on their binding affinity. Finally, some elements of structure-activity relationship (SAR) involved in the binding affinity of homotaurine to soluble Aβ peptides will be discussed.

The third project will discuss a procedure to theoretically model the experimental breakdown diagrams for the Aβ1-40 peptide complexed with a series of low molecular weight molecules (selected based on the knowledge acquired on the second project). An alternative procedure employing an extrapolation procedure for k(E) will also be discussed. This alternative procedure is actually an approximate RRKM method to treat large systems for which the sum-of-states is too high to be computationally calculated. The efficiency and robustness of this method will be tested on smaller systems, which have been previously treated with the original RRKM method in the first project.

1.2 One universe: four principal forces [1,2]

The universe as it is known today has evolved through the actions of four fundamental interactions in nature: the strong and the weak nuclear forces, electromagnetism, and the gravitational force. The strong nuclear force was the first one to appear after the Big Bang. This interaction is responsible for binding together quarks and gluons. This subsequently led to the formation of protons and neutrons, among other particles, to form the first atomic nucleus. This is the strongest of the four fundamental forces. First described by Enrico Fermi, the weak nuclear force is mainly responsible for the β decay, a radioactive decay in which a β particle (electron or positron) is emitted. This interaction is orders of magnitude weaker than the strong nuclear force. Electromagnetism is the force responsible for the interaction between two charged particles (or polar molecules) and is approximately 100 times weaker than the strong nuclear force. Historically, electricity and magnetism were considered as two distinct fields; however, the works of physicist James C. Maxwell allowed the reconciliation of both. The world of chemistry lies on this force: the interactions between protons and electrons in molecules as well as intermolecular between

chemical species are described by electromagnetism. Gravitation is responsible for the formation of massive objects such as galaxies, solar systems, stars, and planets. Sir Isaac Newton's law of universal gravitation states that any two particles attract one another through the product of their masses and inversely proportional to the square of the distance separating them. This force is the weakest of the four fundamental interactions and was the last to manifest itself as the universe cooled down. Altogether, the action of these forces is responsible for the different states of matter in this universe: solid, liquid, and gaseous. This thesis will mostly focus on one aspect of the electromagnetic force: the intermolecular forces between non-covalent molecular complexes in the gas phase.

1.3 Intermolecular forces [3, 4]

When two chemical species (neutral or charged) come to proximity of one another, intermolecular forces begin to arise between them. Intermolecular forces are responsible for the macroscopic physical characteristics of the solid state, the liquid state and the non-ideal behaviour of the gas state. In the solid state, the spatial arrangement of chemical species (atoms, ions and/or molecules) is very compact, and mostly, only vibrational motions occur. In general, the liquid state of a given substance is less dense compared to its solid state. This state of matter permits vibrations, rotations and translations between the particles. Despite the variety of motions, the chemical species remain in the liquid state due to the intermolecular forces acting between them. The gas state has a very low density, the greatest disorder and all three motions are present. According to the kinetic theory of gases, to be defined as ideal, a gas needs to meet the following requirements:

(i) The distance between the gas particles must be much greater than the size of the particles.

(ii) The particles are constantly moving but they do not collide together (or very little).

(iii) The gas particles do not interact together (no repulsion, no attraction).

(iv) The mean kinetic energy of an ensemble of gas particles is directly proportional to its temperature, as shown by the Maxwell-Boltzmann distribution.

A non-ideal gas is an ensemble of particles in the gas phase that does not meet at least one of the preceding criteria. As mentioned in the previous section, since the present research is focussing on interacting chemical species in the gas phase, therefore, the non-ideal gas will be considered.

Although intermolecular forces are predominantly attractive on long distances between particles, they become repulsive at very short distance. The main characteristics of those forces are presented in this section. The concept of an electric dipole is central to define the polarity of molecules and intermolecular forces. An electric dipole, μ, is a separation of charges (+q, -q) over a distance, ℓ. The magnitude of this vector is given by equation 1.1. Units of the dipole are Coulomb-meter (C m) or simply Debye (D).

$$\mu = q.\ell \tag{1.1}$$

A first category of forces is known under the term van der Waals (vdW) forces. These forces are: the London Dispersion Forces (or instantaneous dipole-induced dipole), the Keesom interaction (or dipole-dipole), and the Debye interaction (or dipole-induced dipole). Although not considered

as a vdW force, the hydrogen bond (H-bond) can be considered as an extreme, atom-specific case of the Keesom interaction. The four intermolecular forces mentioned occur between neutral chemical species. There is also a second category of intermolecular forces, stronger than the vdW forces and the H-bond. These are the ion-dipole interaction (which occur between a neutral species and an ion), and the ion-ion (or ionic) interaction.

1.3.1 London Dispersion Forces [5]

This intermolecular force arises when two non-polar molecules (molecules without any dipole) come close to one another (Figure 1.1). Non-polar molecules can be considered as having a symmetrical distribution of electrons in their molecular orbitals. Due to the electrons movement inside the molecule, the distribution of charges may lead to an instantaneous separation of them, which is the origin of a dipole. Once this instantaneous dipole exists, it can induce a dipole on another unpolarized molecule within a short distance. The interaction between both dipoles is known as a dispersion force. In the early 1930's, German-American scientist Fritz London, first described a theoretical model for this dispersion force. It is often referred nowadays as the London Dispersion Forces (LDF). Although originally its explanation was based on quantum mechanics, the following approximate empirical equation 1.2 is used to quantify the behaviour of this interaction, u_{disp}. [4]

$$u_{disp}(r) = -\frac{3}{2}\left(\frac{I_1 I_2}{I_1 + I_2}\right)\frac{\alpha_1 \alpha_2}{(4\pi\varepsilon_0)^2}\frac{1}{r^6} \qquad (1.2)$$

where: I_j is the ionisation energy of chemical species j, α_j is the polarizability of molecule j, ε_0 is the vacuum permittivity, and r is the distance between both chemical species.

The LDF depend on the first ionization energy of both molecules involved. It is proportional to the product of the polarizability of both molecules (the extent of which electronic wavefunctions can distort). Its attractive term varies as $1/r^6$ and u_{disp} increases with the molecule or the atom size (because of the larger polarizability). Although LDF are the weakest vdW interaction, if it did not exist, no non-polar molecules could exist in the solid or the liquid state.

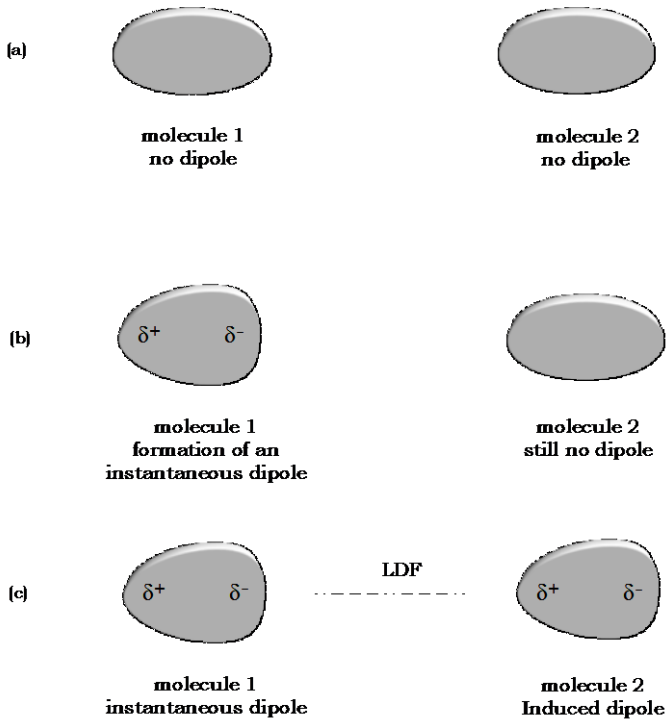


Figure 1.1: (a) Two non-polar molecules come to a certain range of one another. (b) Due to a symmetry break in the orbitals, an instantaneous dipole is created on molecule 1. (c) The instantaneous dipole of molecule 1 induces a dipole on molecule 2. Figure adapted from reference [6].

1.3.2 Keesom interactions

Mathematically described by Dutch physicist Willem Hendrik Keesom in 1921, a dipole-dipole interaction is an attractive force that occurs between two polar molecules when they align their permanent dipoles in a "head-to-tail" or an antiparallel orientation (Figure 1.2). The Keesom interaction, $u_{d,d}$ is described mathematically by equation 1.3 [4].

$$u_{d,d}(r) = -\frac{2\mu_1^2\mu_2^2}{(4\pi\varepsilon_0)^2(3k_BT)}\frac{1}{r^6} \tag{1.3}$$

where: μ_j is the dipole moment of molecule j, ε_0 is the vacuum permittivity, k_B is Boltzmann constant, T is the temperature, and r is the distance between both chemical species.

In fact, a more realistic portrait of the Keesom interaction considers the different possible orientations of the dipoles as a Boltzmann average. The dipole-dipole interaction is directly proportional to the product of the permanent dipoles of both molecules involved and inversely proportional to the temperature. At high temperature, both molecules have more rotational energy, which makes it harder to align their dipole and therefore decrease the strength of the interaction between both molecules. Also, as for the LDF, $u_{d,d}$ varies as $1/r^6$.

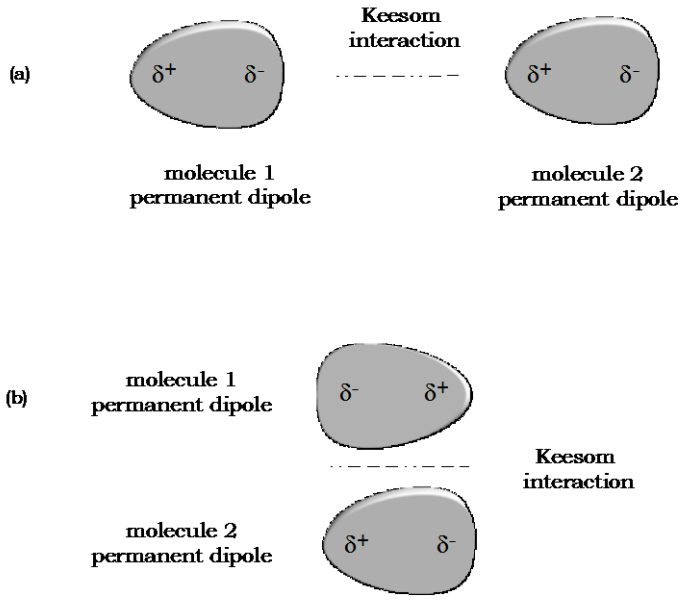


Figure 1.2: Two possible orientations for observing a Keesom interaction between polar molecules. Configuration (a) is referred as a parallel or "head-to-tail" orientation whereas configuration (b) is referred as an antiparallel orientation. Figure adapted from reference [6].

1.3.3 Debye interactions

The last vdW force occurs when a polar molecule approaches a non-polar molecule. This attractive force arises as the electric field, E, of a polar molecule induces a dipole on a non-polar molecule. A mathematical form of the Debye interaction, u_{ind} is described in equation 1.4 [4].

$$u_{ind}(r) = -\frac{\mu_1^2 \alpha_2}{(4\pi\varepsilon_0)^2 r^6} - \frac{\mu_{ind\,2}^2 \alpha_1}{(4\pi\varepsilon_0)^2 r^6} \tag{1.4}$$

$$\text{in which: } \mu_{ind,2} = \alpha E \tag{1.5}$$

where: $\mu_{ind,2}$ represents the induced dipole from the non-polar molecule, α represents the polarizability, E is the electric field generated by the molecule having the permanent dipole, ε_0 is vacuum permittivity, and r is the distance between both chemical species.

As opposed to the Keesom interactions, Debye interactions are independent of temperature. The induced dipole of the second molecule is directly oriented ("head-to-tail") to the permanent dipole of the polar molecule. This interaction is of intermediate strength between those of Keesom and LDF.

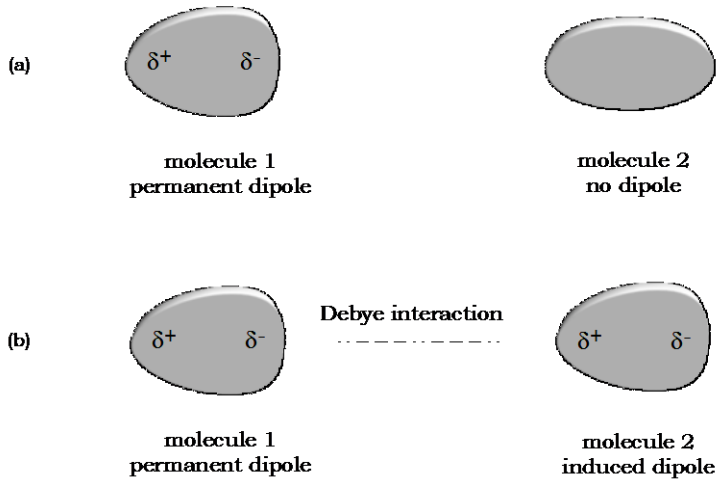


Figure 1.3: When a molecule with a permanent dipole approaches a non-polar molecule (a), it induces a dipole on the second molecule, which leads to the Debye interaction (b). Figure adapted from reference [6].

1.3.4 Hydrogen bonding

Hydrogen bonding (H-bond) can be viewed as a highly polarized Keesom interaction due to the high electronegativity of certain atoms involved. Typically, the orbital 1s of the hydrogen atom is covalently bonded to one of the three following atoms N, O, F, or Cl (the H-bond donor) and then interact non-covalently with one lone pair of electrons from a neighbouring close-shell atom N, O, F, or Cl (the H-bond acceptor) (Figure 1.4). The H-bond length and strength mainly depend on the H-bond donor covalent bond length, the H-bond angle, the temperature, and the dielectric of the media. H-bonds play an important role in the functional 3D conformation of DNA, RNA, proteins,

peptides as well as some of the non-covalent interactions between biomolecules and smaller substrates in living organism.

X-H --- :X (where : « X » = N, O, Cl, or F)

Figure 1.4: General representation of an H-bond.

1.3.5 Ion-dipole interactions

When atoms gain or lose electrons, they become ions. Electrolytes (acids, bases, and soluble ionic salts) dissociate in water into their respective cation and anion. Their solubility is explained by this behaviour since it leads to an attraction between the charge of this ion (q) and the dipole (μ) of water (Figure 1.5). This interaction is mathematically represented by equation 1.6. [3]

$$u_{i,d}(r) = -\frac{q_1 \mu_2}{4\pi\varepsilon_0} \frac{1}{r^2} \qquad (1.6)$$

where: ε_0 is the vacuum permittivity and r is the distance between both chemical species.

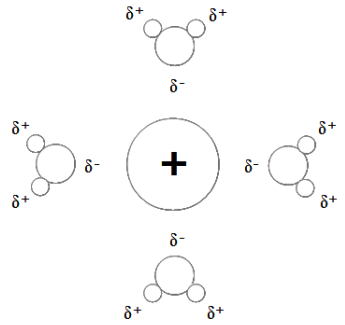


Figure 1.5: A possible representation of an ion-dipole interaction between four polar molecules orienting the partially negative atom towards the positive charge of the cation. Such surrounding is referred as a solvation cage.

1.3.6 Ion-ion (ionic) interactions

The last type of intermolecular interaction arises when two charged chemical species come close to one another. Described by French physicist Charles Augustin de Coulomb, this is the strongest non-bonded interaction possible. Mathematically it corresponds to equation 1.7.

$$u_{ionic}(r) = -\frac{q_1 q_2}{4\pi\varepsilon_0 r} \tag{1.7}$$

where: q_1 and q_2 are the charges of both ions interacting, ε_0 is the vacuum permittivity, and r is the distance between both ions.

1.4 Methods to study intermolecular forces

Scientific knowledge on intermolecular forces helps to understand several phenomena in the universe. In 1873, Johannes van der Waals, demonstrated that the very existence of the condensed phases of matter (solid and liquid) was a direct consequence of attractive intermolecular forces [7]. His work also states that the reason why the liquid phase has such a low compressibility is partly due to the short-range repulsions that exist between the particles. The H-bond network of water molecules is giving its unique property of having its highest density at 4°C. The capillary action is a consequence of the balance between the cohesion of the molecules in the liquid phase (surface tension) and the adhesive forces between the liquid phase and the container (solid phase). These forces are best known for their effect between distinct molecules; however, they also exist inside them, between atoms not directly bonded. As an example, specific low energy conformation(s) of a molecule can lead to and intramolecular H-bond, such as the case of salicylic acid shown in Figure 1.6. The presence of this H-bond has an impact on the 3D structure and the interactions this molecule will have with its environment.

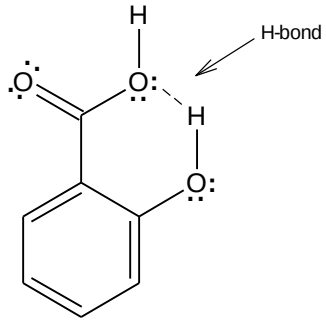


Figure 1.6: An intramolecular H-bond between the hydrogen of the alcohol function and an oxygen of the carboxylic function.

These are only a few examples of the ubiquitous presence of intermolecular forces. The next section will briefly present the mathematical origin of these forces and some of the methods that were developed and used to study them.

1.4.1 Mathematical origins found in the Schrödinger's equation [8]

As mentioned previously, the different intermolecular forces can be represented by empirical equations, which demonstrate their effect and dependence over a range of distances. When averaged over the vibrational motion of two systems (chemical species) interacting together, these forces (especially the vdW) can be represented by a Lennard-Jones potential (Figure 1.7).

Figure 1.7: A generic Lennard-Jones potential of two systems interacting. It shows the dependence of the potential energy, U(r) over the distance, r.

At an infinite distance, both systems do not interact together. The total energy, E at this point can be considered as the sum of energies of both independent systems, A and B.

$$E_{tot}(\infty) = E_A + E_B \tag{1.8}$$

When both systems approach one another, they begin to attract mutually through the existing intermolecular forces. The potential energy of the interacting system, U(r), then becomes:

$$U(r) = E_{tot}(r) - E_{tot}(\infty) \tag{1.9}$$

In other words, the potential energy corresponds to the force, F, required to bring both systems together:

$$U(r) = \int_r^{\infty} F(r)dr \tag{1.10}$$

where:
$$F(r) = -\frac{\partial U(r)}{\partial r} \tag{1.11}$$

The lowest energy on the potential energy well depth corresponds to the equilibrium distance, r_{min}, between both systems. When the distance between both systems tends to become shorter than the equilibrium distance, repulsions arise sharply. Typically, the Lennard-Jones potential shows an attractive term, $U^A(r)$, at distances $r > r_{min}$ and a repulsive term, $U^R(r)$, at distances $r < r_{min}$.

$$U(r) = U^A(r) + U^R(r) \tag{1.12}$$

Equation 1.12 is often represented by the generic 6-12 potential used in force fields.

$$U(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{1.13}$$

where: ε represents the potential energy well depth and is equal to ΔE (r_{min}), and σ represents the distance at which $U(r) = 0$.

These empirical equations take their origin from the Schrödinger's equation, shown here in its time-independent form.

$$\left\{-\frac{\hbar^2}{2\mu}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + U(r)\right\}\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{1.14}$$

where: $\hbar$ is Planck constant divided by 2π, μ is the reduced mass for a 2-body system (1.15) moving through space, Ψ is the wavefunction which represents the particle's motion and E is the total energy of the system.

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{1.15}$$

Altogether, the terms in between brackets that represent the kinetic and potential energy, is called a Hamiltonian, which operates on the wavefunction, Ψ (the eigenfunction) to return the product of the energy (the eigenvalue) and the wavefunction. It can be re-written in a reduced form 1.16, where $\mathcal{H}$ represents the Hamiltonian operator.

$$\mathcal{H}\,\Psi = E\,\Psi \tag{1.16}$$

Considering the binary system composed of A and B to explain the mutual interaction, equation 1.16 can be expressed as:

$$\mathcal{H}_{AB}(r_A, r_B; r)\,\Psi = E_{AB}(r)\,\Psi \tag{1.17}$$

where: the total Hamiltonian operator is given by:

$$\mathcal{H}_{AB} = \mathcal{H}_A + \mathcal{H}_B + \lambda U_{AB}(r) \tag{1.18}$$

For which the wavefunction at short range is a series of coefficients derived from the perturbation theory.

$$\Psi = \lambda^{(0)}\Psi^{(0)} + \lambda^{(1)}\Psi^{(1)} + \lambda^{(2)}\Psi^{(2)} + \ldots \tag{1.19}$$

where: (0) represents the ground state; (1), a first-order perturbation; and (2), a second-order perturbation. As the system separates, λ decreases down to 0 at $r = \infty$.

The overall interaction energy corresponding to the total intermolecular forces is thus given by:

$$\Delta E = \underbrace{\left\langle \Psi_A^{(0)}\Psi_B^{(0)} \middle| \mathcal{H}_{AB} \middle| \Psi_A^{(0)}\Psi_B^{(0)} \right\rangle}_{\text{Coulombic interaction (1st order)}} - \underbrace{\sum_k \frac{\left| \left\langle \Psi_A^{(0)}\Psi_B^{(0)} \middle| \mathcal{H}_{AB} \middle| \Psi_A^{(k)}\Psi_B^{(0)} \right\rangle \right|^2}{E_A^{(k)} - E_A^{(0)}}}_{\text{Polarization of A (2nd order)}} \tag{1.20}$$

$$- \underbrace{\sum_l \frac{\left| \left\langle \Psi_A^{(0)}\Psi_B^{(0)} \middle| \mathcal{H}_{AB} \middle| \Psi_A^{(0)}\Psi_B^{(l)} \right\rangle \right|^2}{E_B^{(l)} - E_B^{(0)}}}_{\text{Polarization of B (2nd order)}} - \underbrace{\sum_k \sum_l \frac{\left| \left\langle \Psi_A^{(0)}\Psi_B^{(0)} \middle| \mathcal{H}_{AB} \middle| \Psi_A^{(k)}\Psi_B^{(l)} \right\rangle \right|^2}{\left(E_A^{(k)} + E_B^{(l)}\right) - \left(E_A^{(0)} + E_B^{(0)}\right)}}_{\text{Dispersion}} + \text{ higher order terms}$$

This formulation is known under the long-range Rayleigh-Schrödinger perturbation theory [8]. One inconvenience with this theory lies in its low efficiency at short distances because there is no term for the electron exchange and the infinitesimal overlap. It remains somehow efficient to, at least grasp, the concept of intermolecular forces acting between two systems using a quantum formulation. This theory has its advantages and drawbacks and there are other methods to account for intermolecular forces, however they are beyond the scope of this thesis.

1.4.2 Computational methods [9]

When computers began to be decently powerful, they became extremely helpful to study intermolecular forces between chemical entities. Computational methods are classified in three categories: quantum mechanics, semi-empirical and molecular mechanics methods.

Quantum mechanics methods find their origin in the Schrödinger's equation, and generally deal with all of the electrons in the system. Exact solutions of this equation exist only for very few systems such as the particle in a box, the harmonic oscillator, or the hydrogen atom. There are only approximate solutions for polyelectronic species because of the antisymmetry principle and the wavefunction that can adopt different functional forms (for which one is not necessarily more "correct" than the others). Several approximate methods are well established to deal with these difficulties such as the Hartree-Fock (HF) approximation and the density functional theory (DFT) [10]. Among all the computational methods, quantum mechanics methods generally provide the best accuracy in terms of molecular geometry and energy, over semi-empirical and molecular mechanics methods. However, these methods are only applicable to small systems because they are very expensive in calculation time.

Semi-empirical methods are also based on the Schrödinger equation, but they take into account several approximations, empirical data, and electron correlation effects. They represent an alternative for larger systems where a full quantum treatment is too computationally expensive. Initial methods such as CNDO were parameterized to fit the quantum mechanics results but had limited success in reproducing experimental results. Michael Dewar later introduced AM1 [11], a semi-empirical method that was parameterized to reproduce experimental molecular geometries, ionization energies, and heats of formation. Although faster than the quantum mechanics methods, these semi-empirical methods generally lead to less accurate results. Therefore, they are a compromise between quality and computational costs.

Molecular mechanics methods (MM) are different as they treat the systems as "balls and springs" for respectively atoms and bonds. Such methods are efficient for large systems such as floppy organic molecules, peptides, and proteins, to name only a few. Such methods require a force field (which is an empirical equation taking into consideration different possible "ball and spring" motions). If the system needs to be studied in a time-dependent fashion (molecular dynamics), the force field is coupled to a time-step integrator that solves the Newtonian equations of motion. If the system needs to be simply geometry optimized, the force field is coupled to a minimization algorithm. These methods helped to understand the intermolecular force involved in docking and protein folding, to mention only those. One project of this thesis involves a molecular mechanics method, more specifically molecular dynamics (MD). The method will be presented in more details in Chapter 2.

1.4.3 Experimental methods

X-ray crystallography may be the experimental method that provided the best insights into molecular structures and intermolecular forces. In this experimental technique, an X-ray beam is projected through a crystal structure, which makes it diffract [3]. The spacing, d, between the crystal lattice layers is determined using Bragg Law (1.21).

$$n\lambda = 2d \sin \theta \tag{1.21}$$

where: n is an integer and corresponds to the order of reflection, λ is the X-ray wavelength, θ is the angle between the incident beam of X-rays and the lattice layer.

Using this experimental technique, it is possible to generate a 3D structure of the crystal. One great leap forward in biological chemistry was made by Nobel Prize winner Dorothy Crowfoot Hodgkin who initially solved the 3D structure of cholesterol, vitamin B12, penicillin, and insulin [12]. Over decades, 3D structures of larger biomolecules began to be elucidated. According to the Protein Data Bank (PDB) [13], there are more than 87000 3D structures of biological macromolecules known as of today. More than 76000 were elucidated by X-ray crystallography. Nowadays, not only these 3D structures are available to the scientific community, but also, many show a small molecule docked into different cavities of receptors. Many parameters for computational methods were derived from experimental observations on these structures.

Nuclear Magnetic Resonance (NMR) is another experimental technique by which it is possible to elucidate structures and observe the manifestation of intermolecular forces [14]. Basic one-dimensional ^{1}H NMR is a powerful tool to determine the connectivity of the atoms inside a molecule as well as certain conformational effects. According to the PDB [13], over 9000 structures of biological macromolecules were elucidated by protein NMR, a technique pioneered by Nobel Prize winner Kurt Wüthrich [15, 16]. The basis of ^{1}H NMR lies on the fact that protons have a spin and hence, a nuclear magnetic moment. In presence of a magnetic field, this nuclear magnetic moment can take two orientations (parallel and antiparallel) of different energies. Upon absorption of a specific electromagnetic radiation (typically a radio frequency pulse), a resonance between these two energy levels occurs. The results are then processed into a spectrum of the absorption verses the NMR frequencies. The chemical shifts resulting (the position of the peaks on the NMR spectrum) are influenced by neighbouring bonded and non-bonded atoms. NMR is also possible with other nuclei that have a non-zero spin (e.g. ^{13}C, ^{15}N). To study proteins, more

complex NMR techniques are being employed such as multidimensional NMR, in which there are at least two radio frequency pulses. In such experiments, information on through-bond and through-space interactions is being collected. Typically, a series of multidimensional NMR spectra are taken to acquire different information (atom positions, connectivity, and distances) on isotope-labelled structures. If a sufficient amount of chemical shifts has been assigned to atoms, they are computed to generate a narrow ensemble of representative folded conformations in solution. Once conformations are obtained, they can be validated against statistical models to reduce the risk of error. Drawbacks from this technique are mainly the costs involved.

Mass spectrometry (MS) is a widely used technique to identify ions of different mass and charge in a sample. Another topic of interest with MS is to study intermolecular attractions between chemical species. In a basic MS experience, a sample is first vaporized and ionized before reaching an electromagnetic sector that will separate the ions based on their mass-to-charge ratio. Initial techniques employed harsh ionizations (such as electron impact (EI) and chemical ionization (CI)) that were fragmenting molecules. New softer ionization methods do not initially fragment the ions in the sample and allow the observation of intact charged species and non-covalent bonded complexes. Mass spectrometers cost often less than NMR; data acquisition is faster, and generally easier to interpret for large systems. One disadvantage of MS techniques is that they provide very little or no information about the 3D structure itself. At the very best, ion-mobility experiments can provide the collision cross section of chemical species [17], and Hydrogen-Deuterium exchange [18] as well as covalent labeling [19] can provide information on the solvent accessible areas of a protein (as an example). Therefore, experimental binding data obtained by MS can be

complemented by computational simulations to give an insight into the molecular structures of the chemical system.

All research projects of this thesis used MS to study the formation or the dissociation of bound complexes in the gas phase. More specifically, three types of MS (and related instruments) will be presented in Chapter 2: Electrospray ionization mass spectrometry (ESI-MS), coupled with a time-of-flight module (ESI-QTOF-MS), and Collision induced dissociation (CID) experiments using a triple-quadrupole mass spectrometer and the ESI-QTOF-MS mass spectrometer.

1.5 Gas phase interactions between polymers and amino acids

1.5.1 Generalities on polymers [20]

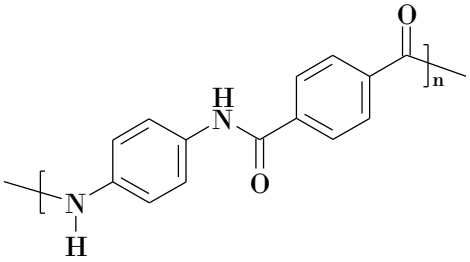
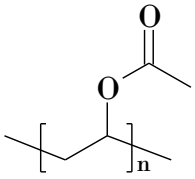
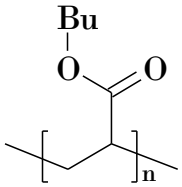
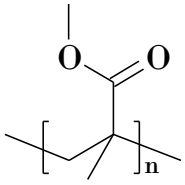
A polymer is a macromolecule consisting of small structural repeating units called monomers. Contrarily to proteins and peptides where the units are the amino acids, a polymer generally consists of a constant series of unique structural units (homopolymer) or sometimes a few different structural units (copolymer). Polymers are obtained by synthesis where the monomer units are being covalently linked together. This chemical process, called polymerization, can lead to different chain lengths. A polymer sample can be characterized by its polydiversity, D, which is the ratio of the average molecular weight by weight, $<M_w>$, over the average molecular weight by number, $<M_n>$.

$$D = \frac{\langle M_w \rangle}{\langle M_n \rangle} = \left(\frac{\sum m_i^2 N_i}{\sum m_i N_i} \right) \Big/ \left(\frac{\sum m_i N_i}{\sum N_i} \right) \tag{1.22}$$

where: m_i is the mass (or molecular weight) of chain "i" and N_i is the number of chains "i".

The polydiversity is an essential characteristic of a polymer sample since it provides an indication its molecular weight distribution (MWD). Narrow distributions are characterized by a value of D close to 1, whereas broad distributions have values of 2 or greater. Nowadays, polymers have many uses. Table 1.1 shows some examples of polymers, their monomer units, and their purpose. For each use, there is an ideal polydispersity providing the required qualities of the polymer sample. Information on the MWD of polymer can be gathered by analytical techniques such as, osmometry and light scattering. The polydispersity of a polymer sample can be optimized by polymer fractionation, which consists in removing the smaller or the longer oligomers.

Table 1.1: Some examples of polymers and their purposes.

Name	Monomeric unit	MW (Da)	Purposes
Kevlar™		238	in ballistic and body armour
PVac poly(vinylacetate)		86	in wood glue and adhesive for porous materials
PBA poly(butylacrylate)		128	useful for its transparency and its breaking resistance
PMMA poly(methylmetacrylate)		100	in the fabrication of Plexiglass™

For more than two decades now, mass spectrometry has become an efficient analytical tool to study polymers in general, their MWD, and their fragmentation patterns. In the early days of MS, only harsh ionization methods such as electron impact (EI) and chemical ionization (CI) were available. As mentioned before, these ionization methods were efficient to fragment small molecules. Based on the fragments observed, it was possible to propose a mechanism that would

lead to the parent chemical species. However, these ionization methods were not suitable for large organic molecules, polymers, and proteins because they would generate too many fragments. Nowadays, softer ionization methods are available for such large molecules: one of them is the matrix-assisted laser desorption ionization (MALDI) [21, 22]. In this two-step process, a matrix containing the analyte first absorbs the UV from a laser beam. This causes a desorption on the matrix surface layer and thus the analyte, which is then ionized before being sprayed into the remaining part of the mass spectrometer. The other popular technique is the electrospray ionization (ESI) [23-25], in which a liquid sample is first being desolvated before its ionization. A more detailed description of the ionization mechanism will be presented in Chapter 2. As a concrete example of the efficiency of ESI-MS, Figure 1.8 shows a mass spectrum acquired from another project in our group in which they were studying the protonation of PMMA with triglycine in an acidic environment [26]. In a typical mass spectrum, the y-axis is the intensity of the peaks, which, represents the abundance of the chemical species, and the x-axis represents the mass-to-charge ratio (m/z) of the chemical species. Using this experimental technique and supported by computational techniques, one project of this thesis will investigate the energetics involved in the competing dissociation channels of protonated PMMA/amino acid complexes, (PMMA)(AA)(H^+), in the gas phase.

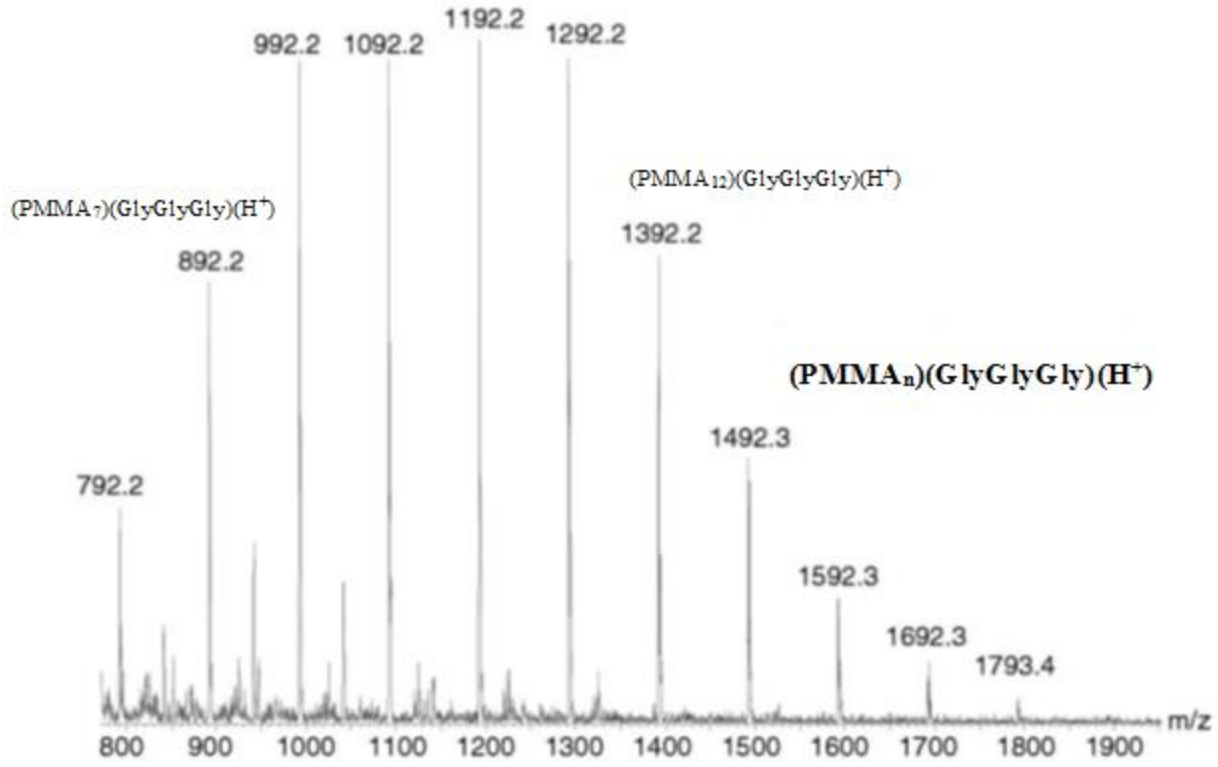


Figure 1.8: PMMA protonated with triglycine in an acidic environment $(PMMA_n)(GlyGlyGly)(H^+)$. The distribution of oligomers in the sample is clear as the major peaks all have a difference of mass of 100 Da, which correspond to an MMA monomer unit. Thus, in this sample, there was a distribution of PMMA oligomers (from 6 to 16 monomer units) bound to triglycine. Complexes with 7 $(PMMA_7)(GlyGlyGly)(H^+)$ and 12 $(PMMA_{12})(GlyGlyGly)(H^+)$ monomer units are noted on this spectrum. Figure reproduced with permission from reference [26].

1.5.2 Competing dissociation pathways of $(PMMA_n)(AA)(H^+)$ complexes

Gas phase collisions with neutral inert molecules are an efficient way to increase the internal energy of an ion in the field free region of a mass spectrometer. McLafferty *et al.* [27] first demonstrated the usefulness of collision-induced dissociation (CID) experiments in ion structure determination. Over the years, CID mass spectrometry has become useful to determine the sequence and the structure of a polymer. However, compared to biological macromolecules, polymers do not ionize easily in solution due to their generally low proton affinity. Therefore,

metal cations often help to overcome this limitation. Polymers such as PMMA and poly(*n*-butylmethacrylate) (PBMA) have been studied using diverse mass spectrometry techniques [28-32]. It was found that the resulting ion m/z depends on which oligomer fragment retains the metal cation. Our group previously studied the fragmentation of PVAc ionized with Li^+ generated by CID mass spectrometry [33]. Using MD, the sequential loss of acetate side-chains were determined to be following a stepwise mechanism of increasing energy. In a different project, the fragmentation of poly(dimethylsiloxane) oligomers with metal ions (Li^+, Na^+, K^+, Ag^+) and organic ions (NH_4^+, $CH_3CH_2NH_3^+$) was investigated [34]. It was found that the CID fragmentation of the oligomer ions strongly depends of the metal ion while the organic cations transfer their proton to the oligomer. Subsequently, amino acid or small peptides were used to ionize oligomers of PMMA and PBA [26]. It was found that both PMMA and PBA protonated oligomers undergo a unique fragmentation pattern by eliminating neutral molecules from the monomer side chains until all that is left is the backbone hydrocarbon. Ionization with a metal cation is thus a rather harsh ionization method leading to fragmentation products. The scope of this project is to investigate the effect of softer ionizers on the dissociation pathways of a polymer. For this purpose, non-covalent complexes formed of different PMMA oligomers and three ionizers of different proton affinity (PA) (glycine: 887 kJ mol^{-1}, leucine: 915 kJ mol^{-1}, phenylalanine: 993 kJ mol^{-1}) [35] were formed and investigated in the gas phase using CID mass spectrometry with the ESI-triple quadrupole-MS instrument. RRKM modeling (which will be explained in sections 1.7 and 1.8) will help to understand the energy and entropy effects during the dissociations. Molecular Dynamics Simulated Annealing (MD-SA) simulations (which will be explained in Chapter 2) will help to visualise the 3D conformations of the (PMMA)(AA)(H^+) complexes and support the energy and entropy effects obtained by RRKM modeling.

1.6 Gas phase interactions between amyloidogenic peptides and LMW molecules

1.6.1 The role of the intermolecular forces in proteins and peptides [36]

Proteins and peptides are biopolymers composed of a sequence of residues (amino acids) linked together by amide bonds. Amino acids are composed of a backbone having an amine and a carboxylic function. The 20 amino acids found in proteins, all have this common backbone (Figure 1.9) except for the proline, which has a cycle.

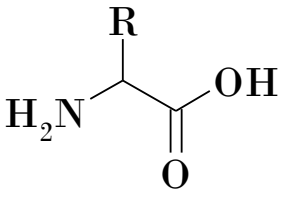


Figure 1.9: Common backbone of all amino acids where R is the side-chain.

What differentiate the amino acids are the side-chains, R, which give each of them unique physical and chemical properties. The carbon bearing the side-chain is chiral, except for the amino acid glycine for which the side-chain is a hydrogen. Both L and D enantiomers are possible. Although the L form is more biologically relevant and abundant in nature, the D form is sometimes observed like in the peptidoglycan cell walls of bacteria [37] or as a neurotransmitter in the brain [38]. As shown in Table 1.2, amino acids can be classified in three categories: hydrophobic (having non-polar side-chains), hydrophilic (having polar side chains) and charged under physiological pH (having ionic side chains). The physical nature of those amino acids will directly influence the type of intermolecular forces that will act between them in a protein or a peptide. The sequence of residues (the primary structure) may lead to a number of secondary structures, such as α-helices, parallel and anti-parallel β-sheets to mention only the most frequently encountered. These

secondary structures are linked together by loops, which have no defined conformation. They are sometimes referred as random coils.

Table 1.2: Chemical structures, names, 3-letter and 1-letter abbreviations, as well as molecular weight of the 20 common amino acids.

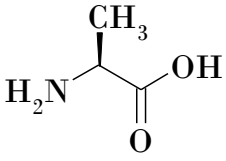
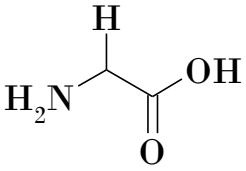
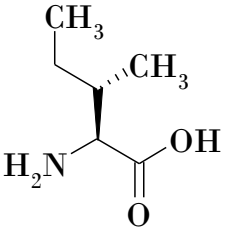
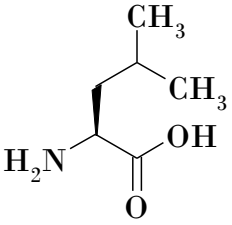
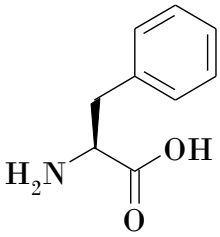
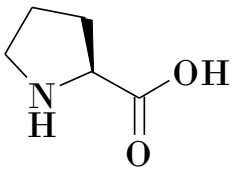
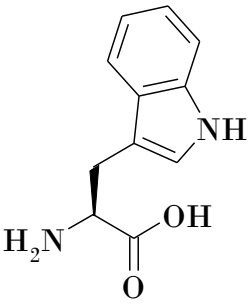
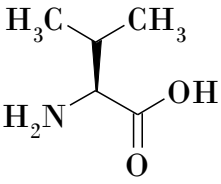
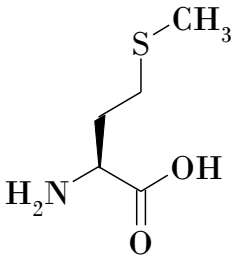
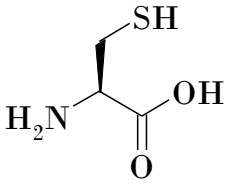
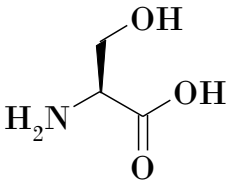
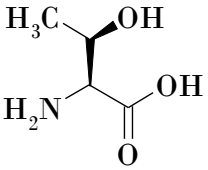
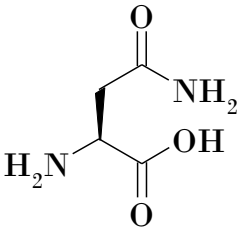
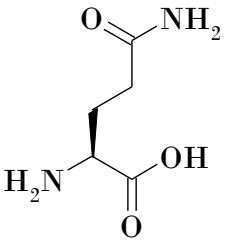
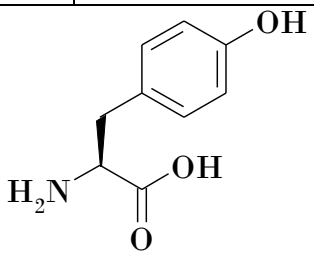
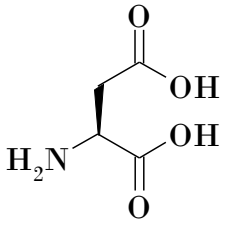
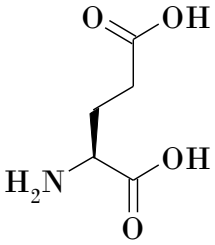
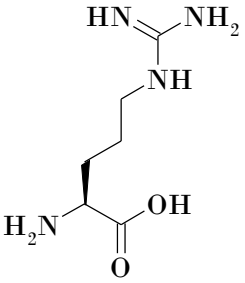
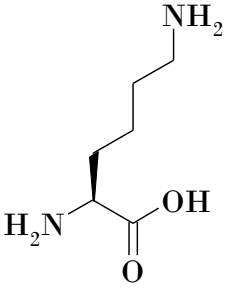
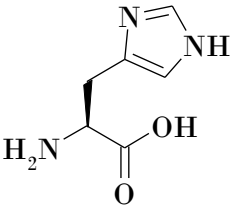
Hydrophobic amino acids			
			
Alanine (Ala, A) [MW: 89 Da]	Glycine (Gly, G) [MW: 75 Da]	Isoleucine (Ile, I) [MW: 131 Da]	Leucine (Leu, L) [MW: 131 Da]
			
Phenylalanine (Phe, F) [MW: 165 Da]	Proline (Pro, P) [MW: 115 Da]	Tryptophan (Trp, W) [MW: 204 Da]	Valine (Val, V) [MW: 117 Da]

Table 1.2: continued

Hydrophilic amino acids			
			
Methionine (Met, M) [MW: 149 Da]	Cysteine (Cys, C) [MW: 121 Da]	Serine (Ser, S) [MW: 105 Da]	Threonine (Thr, T) [MW: 119 Da]
			
Asparagine (Asn, N) [MW: 132 Da]	Glutamine (Gln, Q) [MW: 146 Da]	Tyrosine (Tyr, Y) [MW: 181 Da]	
Charged amino acids under physiological pH			
				
Aspartic acid (Asp, D) [MW: 133 Da]	Glutamic acid (Glu, E) [MW: 147 Da]	Arginine (Arg, R) [MW: 174 Da]	Lysine (Lys, K) [MW: 146 Da]	Histidine (His, H) [MW: 155 Da]

Altogether, the primary and secondary structures form the tertiary structure (the 3D conformation) which in turn is responsible of the biological role of the protein. The 3D structures are a direct consequence of the intermolecular forces acting between the amino acids and the protein environment as well as the intramolecular attractions and repulsions. As an example, a globular protein in an aqueous environment will generally have its hydrophilic amino acids on its "globe-like spherical" surface while the hydrophobic amino acids are interacting together inside the core. Membrane proteins, like potassium and sodium channels, have their hydrophobic amino acids interacting with the phospholipid bilayer inside the membrane (by London Dispersion Forces) while the interior of the channel will contain mostly hydrophilic and charged amino acids that interact with the transiting cations (by ion-dipole and ion-ion attractions and/or repulsions). Fibrous proteins and peptides tend to form biological tissues and filaments. Under certain conditions, those filaments can accumulate to form insoluble aggregates because of the interactions between the charged and hydrophilic amino acids. The peptides treated in this research are of the latter nature.

1.6.2 The role of intermolecular forces in medicinal chemistry [39]

Research in fields such as biology, biochemistry, pharmacology, and medicine has shown direct links between certain diseases and given receptors. Receptors are generally biomolecules such as proteins that are located on the surface of a cell. Their role is to receive a chemical signal from an external source in its environment. When the signal received is not proper (not enough or too much), the biochemical pathway related to this receptor may not be functioning properly. One objective of medicinal chemistry is to design and synthesize drug molecules that can simulate these

signals. A classical approach in medicinal chemistry is known as the "lock and key model". As the name implies, a key with the proper structure will unlock a specific door lock such as a proper drug molecule with the required 3D conformation and pharmacophore groups will dock inside a specific cavity in a receptor. Once inside the receptor, the drug molecule is, in most cases, non-covalently bound to the amino acids inside the cavity through intermolecular forces. The complex formed can induce an effect on the structure of the receptor and help re-establishing its biological function. Concisely, agonists are drug molecules that will trigger an action on the cell upon binding to the receptors, whereas antagonists will occupy the cavity without inducing any action upon binding. In the present research, the interaction between the drug-like molecules and the peptide is rather a non-classical approach because the peptide itself, under physiological conditions, has no defined 3D conformation and hence, no cavity. The binding occurring in this case will serve to maintain a fibrous peptide in its aqueous environment to prevent its aggregation, eventually.

1.6.3 The potential role of Aβ peptides in Alzheimer's disease

Alzheimer's disease (AD) is an incurable, progressive neurodegenerative condition leading to a gradual decline in cognitive function. Intracellular accumulation of neurofibrillary tangles consisting of tau protein and extracellular deposition of senile plaques composed of amyloid β peptides (Aβ) represent histopathological hallmarks of AD [40-42]. Aβ is considered to play a key role in AD pathogenesis [43] and the amyloid cascade hypothesis [44] stipulates that this results from the sequential β- and γ-secretase cleavage of the amyloid precursor protein (APP) [45,46]. Several Aβ variants, including Aβ1-40 and Aβ1-42, produced by this amyloidogenic pathway are considered toxic under certain conditions. Non-pathologic low concentrations of Aβ are found in

biological fluids throughout life. Abnormal metabolism such as reduced clearance or increased production leads to elevated Aβ concentration and aggregation [47]. At physiological pH, soluble Aβ mainly adopts a random coil conformation [48] where peptides may interact to form intermolecular β-sheets leading to the production of intermediate species such as oligomers, fibrils, and eventually plaques [49, 50]. Whether or not the production of oligomers and fibrils occurs through a common mechanism is still debated [51]. Certain forms of oligomeric Aβ are believed to be toxic to neurons [52] while it has been shown that Aβ fibrils are harmful to cell membranes and elicit inflammatory responses from glial cells [53].

Proteoglycans (PGs) and glycosaminoglycans (GAGs) found in amyloid deposits contribute to Aβ fibril formation by promoting the transition from a random coil to a β-sheet conformation [54, 55]. It was hypothesized that by mimicking some GAGs structural properties, low molecular weight (LMW) compounds, could interfere with the interaction of Aβ with heparin sulphate proteoglycans of the basement membrane during the course of amyloid deposition [56]. One promising therapeutic perspective consists in identifying LMW compounds able to slow down or even stop these events in amyloidogenic pathway [57]. It was postulated that LMW compounds must first bind or interact with Aβ to interfere in the aggregation and fibril formation. Efforts were invested in developing assays such as Thioflavin T fluorescence (ThT) [58] and circular dichroism (CD) [59] to evaluate the interaction/binding of Aβ with LWM compounds. Very limited success has been achieved primarily due to variability in different lots of synthetic and recombinant Aβ, both within and among laboratories. Modifications in the Aβ synthetic and purification protocols may affect its conformational and fibrillogenic properties [60, 61]. The presence of impurities may alter Aβ kinetics of aggregation [62] and different storage conditions may result in Aβ chemical

modifications [63]. Although serious efforts have been invested in addressing these problems [64, 65], a reliable assay was highly desirable at this point.

ESI-MS is an established technique for detecting non-covalent protein/LMW molecules complexes [66-68]. Studies of such complexes between specific LMW compounds and Aβ have been reported previously by other groups [69-72] and at Neurochem Inc. [73-75]. One debate in using ESI-MS is whether the detected gas phase complex maintains a structural resemblance to the solution-phase complex. Although ESI-MS does not give insight into the 3D structure of the complex, reasonable agreements have been found between gas phase and solution-phase binding affinities [67]. However, in source collisions can lead to the disruption of weakly bound complexes. On the other hand, droplet cooling due to the evaporation of solvent can lead to an apparent enhanced binding of the gas phase complex and the formation of higher order complexes, which were not pre-existing in the solution phase [76]. Nanoelectrospray MS has been shown to represent better the solution phase complex because smaller droplets undergo less desolvation [77, 78]. For these reasons, it is very difficult to determine absolute solution-phase binding thermodynamics from the mass spectrum. However, if care is taken to maintain constant solution properties (pH, salt content, and low analyte concentrations to prevent aggregation) as well as constant ESI source parameters, the observed trends in binding between analytes and the target peptide observed in ESI mass spectra will be consistent with those in solution [79]. The scope of this study is focused on the application of an efficient and reproducible moderate throughput ESI-MS screening method for the ranking of the binding affinity (BA) of a library of LMW compounds to Aβ1-42, Aβ1-40, and Aβ1-28. The role of side-chain charge is investigated by substitution of charged amino acids by neutral amino acids on the Aβ1-28 sequence. Recent literature reported

the effect of pyro-glutamate-modified N-terminal truncated Aβ (Aβ3pE-40) to influence the oligomerization process [80] and to induce learning impairment as well as neuronal apoptosis [81]. Due to the increasing interest, we also investigated the BA of LMW compounds to this peptide. Finally, there is a discussion on the peptide-binding characteristics of homotaurine, a natural product [82, 83] for the potential treatment of AD [84].

1.7 RRK and RRKM theories for unimolecular reactions [85, 86]

Statistical theories, such as the Rice-Ramsperger-Kassel-Marcus quasi-equilibrium theory (RRKM/QET), were developed to understand and quantify the energetics of unimolecular reactions, among other, produced by CID mass spectrometry experiments. The current section traces an overview of them.

A unimolecular reaction (1.23) is defined as any system evolving in time as a result of prior excitation and follows an exponential decay (1.24).

$$AB \rightarrow \text{products} \tag{1.23}$$

$$-\frac{d[AB]}{dt} = k[AB] \tag{1.24}$$

The fact that the dissociation rate constant was dependent on the internal energy of the molecule was first recognized by Rice and Ramsperger [87, 88] and later on by Kassel [89]. Generally, statistical theories of unimolecular reactions begin by calculating the microcanonical rate constant, k(E).

The canonical rate constant, k(T), is connected to the microcanonical rate constant by averaging k(E) over the energy distribution:

$$k(T) = \int_{E_0}^{\infty} P(E,T)k(E)dE \tag{1.25}$$

where: E_0 is the 0 K activation energy and P(E,T) is the distribution of internal energies at a given temperature, T.

Lindemann was the first to propose a unimolecular reaction mechanism by which reacting molecule, AB, in a thermal system is energized by collisions with another molecule, M, present in the reaction vessel. The proposed mechanism is expressed as:

$$AB + M \xrightarrow{k_1} AB^* + M \tag{1.26}$$

$$AB^* + M \xrightarrow{k_{-1}} AB + M \tag{1.27}$$

$$AB^* \xrightarrow{k_2} A + B \tag{1.28}$$

If AB*, the sufficiently excited molecule, is produced in a steady-state fashion, the overall rate of product formation can be expressed as:

$$\text{rate} = k_2[AB^*] = \frac{k_1 k_2[A][M]}{k_{-1}[M]+k_2} = k_{uni}[AB] \tag{1.29}$$

In equation 1.29, the unimolecular rate constant, k_{uni}, is dependent on the collision activation rate constant, k_1, the deactivation rate constant, k_{-1}, and the unimolecular rate constant, k_2.

Rice, Ramsperger and Kassel developed a conceptual framework to understand unimolecular reactions. A system is composed of an ensemble of s identical harmonic oscillators, one of which has an energy E_0, the required energy for the dissociation. This critical oscillator is in fact the reaction coordinate. A fundamental assumption in this theory is that the energy flows statistically among all the oscillators. The dissociation rate is proportional to the probability of the critical oscillator having an energy equal to or greater than $E_0 = mh\upsilon$ in a system where the total energy is given by $E = nh\upsilon$ (h is the Planck constant and υ is the oscillator vibrational frequency). The probability of the reacting molecule being in a dissociative state is expressed as:

$$\text{Probability} = \frac{(n-m+s-1)!n!}{(n-m)!(n+s-1)!} \tag{1.30}$$

where: n is the total number of vibrational quanta in the system and m is the number of vibrational quanta in the critical oscillator.

The expression can be converted into a rate constant expression if it is multiplied by the transition state frequency. This leads to the RRK rate constant expression:

$$k(E) = \upsilon\left(\frac{n-m}{n}\right)^{s-1} = \upsilon\left(\frac{E-E_0}{E}\right) \tag{1.31}$$

This expression shows that k(E) has two important features: first it is increased with excess energy ($E-E_0$) and second it is strongly dependent on the number of critical oscillators. Unfortunately, the expression is incapable of giving a correct rate constant even within an order of magnitude. There are a number of reasons to justify this flaw. One reason is the assumption that the ensemble of classical oscillators (n-m) >> s is generally inappropriate for most classical systems. Another

reason point lies in the fact that the RRK theory neglects the zero point energy. Two quantum theories solved the problems of the RRK theory by treating in details the vibrational and rotational degrees of freedom adequately. These two theories are known as RRKM and quasi-equilibrium theory (QET). The RRKM/QET expression is:

$$k(E) = \frac{\sigma N^{\ddagger}(E-E_0)}{h\rho(E)} \tag{1.32}$$

where: σ is the reaction degeneracy, $N^{\ddagger}(E\text{-}E_0)$ is the transition state sum of states from 0 to $(E\text{-}E_0)$, and $\rho(E)$ is the parent ion density of states at an energy, E.

The reaction degeneracy depends on the symmetry number of the reactant, σ_r, and the transition state, $\sigma^{\ddagger}$. Both symmetry numbers are related in the expression for the reaction degeneracy as:

$$\sigma = \frac{\sigma_r}{\sigma^{\ddagger}} \tag{1.33}$$

For the sum and density of states, often only the vibrational degrees of freedom are considered (rotational degrees of freedom are ignored). For a molecule (or an ion) having s vibrational degrees of freedom at an internal energy of E, the sum of states represents all the possibilities of distributing the energy among the s oscillators such that the total energy will be equal to or less than E. Unlike the RRK theory, each oscillator has its own vibrational frequency and thus its unique energy content. N(E) is the total number of states and is unitless.

The density of states represents the number of vibrational configurations with an energy content between E and E + δE. As a comparison: the sum of states represents the number of possibilities for the reactant ion to pass through the transition state (at an energy E-E_0) while the density of states at an energy, E, represents the number of ways to get lost in the molecular ion phase space. In the RRKM expression, the sum of states is the numerator and the density of states at the denominator because at the transition state, a fraction of the ion's total energy has been channelled into the reaction coordinate (critical oscillator). The remaining energy, (E-E_0), is then distributed among the other s-1 oscillators. The reaction could also pass "over" the transition state at an energy ε above E_0, leaving the remaining internal energy (E-E_0-ε) partitioned among the rest of the vibrational modes of the ion (Figure 1.10).

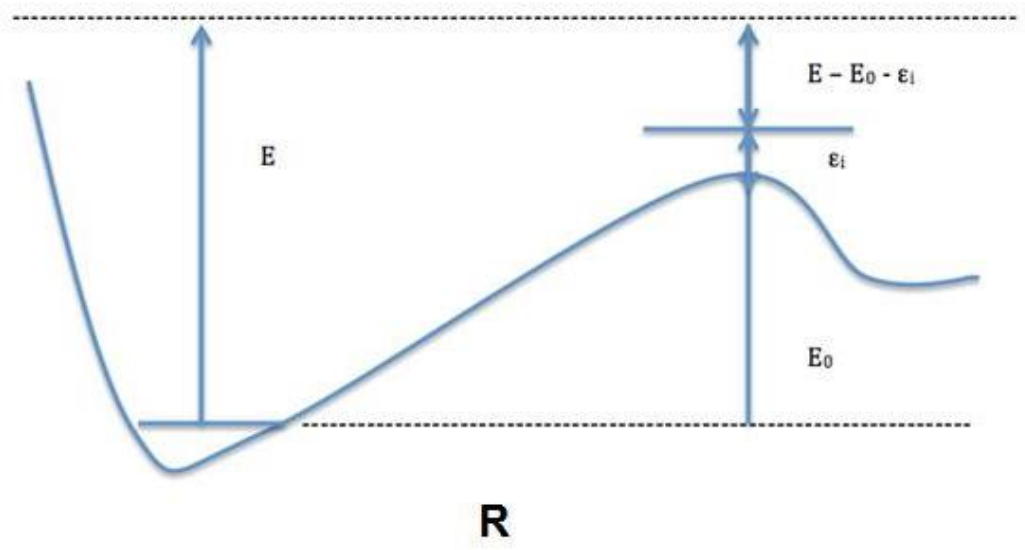


Figure 1.10: Dissociation reaction where the reaction coordinate passes over an energy barrier. E is the total ion energy, E_0 is the activation energy (required to pass the barrier), ε_i is the reaction coordinate translation energy. Figure adapted from reference [86].

A transition state is a short-lived species, which makes it hard to characterize when using statistical theories. In addition, the frequencies related to the transition state are difficult to choose. In several cases, it is hard to identify a single structure for the transition state because it varies with the ion's internal energy. Wigner hypothesized that a transition state structure corresponded to the configuration at which a reaction trajectory passed only once [90]. In this assumption lies the basis of the transition state theory. In the application of many statistical theories, the structure and especially the vibrational frequencies of the transition state must be known in addition of those of the reactants. Vibrational frequencies for species in their ground state are nowadays not very difficult to obtain, at least theoretically. However, for the transition states, it is not always an easy task. It was postulated that the vibrational frequencies of the transition state could be estimated approximately the same as those of the reactant ion. The transition states frequencies can also be approximated by scaling the lowest vibrational modes of the neutral reactant molecule.

1.7.1 Variational transition state theory (VTST)

Another way of defining a transition state is the point at which the phase space corresponds to a minimum on the dividing configuration space between the reactant and the product. For reactions, which have a significant energy barrier, the transition state is defined as the saddle point of highest energy on the minimum energy path between the reactant and the product. Reactions, like bond cleavage or intermolecular dissociation do not have a maximum between the reactant and the product. It does become tedious to be able to evaluate the entropy of activation, $\Delta^{\ddagger}S$; in statistical theory terms, the transition state is referred as "generalized" and does not necessarily pass through a saddle point.

Since ionic dissociation have a negligible reverse activation energy, Lifshitz suggested that ionic systems are well represented by VTST [91]. In this theory, the rate constant, k(E,J) is a function of the reaction coordinate, R. An advantage of the VTST is that, on the PES, it provides the upper limit for k(E,J). Therefore, the optimal choice of transition state in the dividing surface between the reactant and the product corresponds to the smallest rate constant. The point on the PES where:

$$\frac{\partial \mathrm{N}^{\ddagger}(\mathrm{E,J,R})}{\partial \mathrm{R}} = 0 \tag{1.34}$$

The rate constant, can be determined from the minimum sum of states, $N^{\ddagger}$, using the RRKM/QET expression, expressed in this case as:

$$\mathrm{k(E,J,R^{\ddagger})} = \frac{\mathrm{N}^{\ddagger}(\mathrm{E}-\mathrm{V(r)})}{\mathrm{h}\rho(\mathrm{E,J})} \tag{1.35}$$

where: $R^{\ddagger}$ is the separation (or bond distance in cleavage reactions) at the transition state and V(r) represents the potential energy curve.

This point location is also known as the "entropic bottleneck" as shown on Figure 1.11. During the dissociation, there is an increase in V(r), which reduces the available energy (E-V(r)) and in turn, reduces the density of states. On the other hand, the decrease in vibrational frequencies will increase the sum of states. The two opposing forces point towards the minimum in the sum of states at a separation $R^{\ddagger}$ (which corresponds to the transition state).

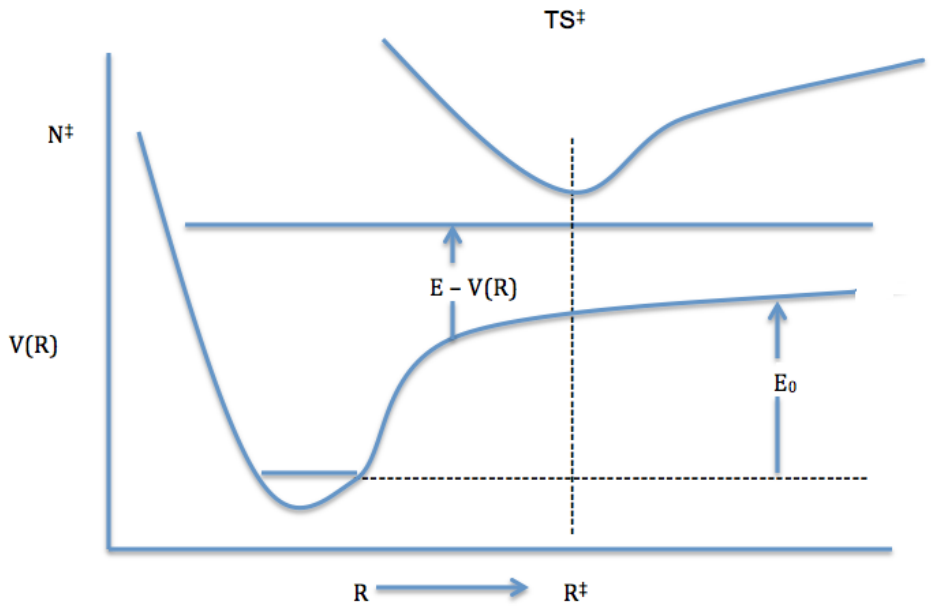


Figure 1.11: Unimolecular reaction for which there is no real energy barrier. The activation energy, E_0, corresponds to the energy required for a complex to be dissociated. The position of the transition state is given at the reaction coordinate $R^{\ddagger}$. Figure adapted from reference [85].

1.7.2 Entropy of activation ($\Delta^{\ddagger}S$)

In the canonical formalism, the rate constant, k(T), is expressed in terms of entropy of activation, $\Delta^{\ddagger}S$ and activation enthalpy $\Delta^{\ddagger}H_0$:

$$k(T) = \frac{k_B T}{h} e^{\Delta^{\ddagger}S/R} e^{\Delta H_0^{\ddagger}/RT} \tag{1.36}$$

where: $\Delta^{\ddagger}S$ is the difference in entropy between the reactant and the transition state.

It gives an idea about the nature of the reaction itself. If $\Delta^{\ddagger}S$ has a positive value, the transition state is less ordered than the reactant and referred as “loose”. However, if $\Delta^{\ddagger}S$ has a negative value,

the transition state is more ordered than the reactant ion and is referred as "tight". Generally, "loose" transition states are more associated with bond cleavage reactions while "tight" transition states are associated with rearrangement reactions. If $\Delta^{\ddagger}S$ is null, it means that both the reactant and the transition state have the same vibrational and rotational properties. The entropy of activation is given by the following formula:

$$\Delta^{\ddagger} S = k_B \ln \frac{\prod Q_i^{\ddagger}}{\prod Q_i} + \frac{E_{int}^{\ddagger} - E_{int}}{T} \tag{1.37}$$

where: Q and $Q^{\ddagger}$ correspond to the rotational (q_{rot}) (1.38) and vibrational (q_{vib}) (1.39) partition functions of the parent ion and the transition state respectively. T is the temperature and k_B is Boltzmann constant.

$$q_{rot} = \frac{1}{\sigma}\left(\frac{k_B T}{hc}\right)^{3/2}\left(\frac{\pi}{V}\right)^{1/2} \tag{1.38}$$

where: σ is the symmetry number of the molecule which is the number of indistinguishable orientations of the complex, c is the speed of light, and V is the molecular volume.

$$q_{vib} = \sum \frac{1}{1 - e^{-h\upsilon_j / k_B T}} \tag{1.39}$$

where: υ_j represent the vibrational frequencies

During the fitting process of experimentally determined rate constants, both E_0 and $\Delta^{\ddagger}S$ have an effect on the k(E). In a breakdown diagram, the magnitude of k(E) is mostly related to E_0 while the slope of k(E) is mostly related to $\Delta^{\ddagger}S$. The transition state frequencies can be determined by ab initio calculation or by MD normal mode calculations. If not possible, the transition state frequency is estimated by taking the reactant frequencies less one corresponding to the reaction coordinate. The transition state can be tightened or loosened by scaling the six lowest vibrational frequencies by a common factor. It this factor is less than 1, it will increase $\Delta^{\ddagger}S$ whereas if it is greater than 1, it will decrease $\Delta^{\ddagger}S$).

1.8 Experimental approaches to evaluating E_0 and $\Delta^{\ddagger}S$ for larger systems

The binding of a substrate to a protein is fundamental to all life and is a property that has been extensively explored in biological media. Recently, there has been interest in the use of mass spectrometry to determine the binding of protein–substrate complexes in solution. Given that mass spectrometry is a gas phase technique, this endeavor has been undertaken with two broad approaches. The first attempts to transfer the protein–substrate solution into the gas phase with either electrospray ionization (ESI) or matrix-assisted laser desorption ionization (MALDI) in such a way that the ratio of bound to unbound protein is unchanged, thereby directly providing the binding ratio and subsequent thermodynamics of binding in solution [92-94]. The other attempts to directly measure the gas phase binding of the complex and compare the result with that derived from studies of the binding in solution (see below).

Much has been written of the difficulties involved with the former approach due to the many changes that occur to the solution during the process of transferring the ions into the gas phase, most notably with electrospray ionization. These include salt and pH effects when the droplets desolvate, cooling upon desolvation and non-specific electrostatic interactions which can dominate in the gas phase [92-103]. That being said, on a case-by-case basis, there is evidence that the procedure can be successful [104-107]. For the latter approach, there are few reliable methods for the determination of precise gas phase binding energies. Probably the best is the measurement of the dissociation of a thermalized protein–substrate complex as a function of temperature with the BIRD technique [108-119]. The resulting Arrhenius plot directly yields the enthalpy and entropy of activation in the dissociation, the former of which typically reflects the gas phase thermodynamic binding energy of the complex. There has also been some success employing thermal decomposition with quadrupole ion traps [120-122]. While BIRD is an excellent approach to this problem, the apparatus used (usually Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry with customized ICR cell heating) is not routinely available. What is fairly easy to obtain is the collision induced dissociation mass spectrum of a protein–substrate complex as a function of collision energy. These CID breakdown curves inherently contain information on the gas phase binding of the complex. The usual problem is extracting this information. Many qualitative and quantitative approaches have been described depending on the nature of the experiment [123]. The most successful have been for slow-heating methods such as CID performed in ion traps, [119, 124-130] CID and surface-induced dissociation in FT-ICR [131-137] and single-collision threshold measurements [138-140]. Recently, Knyazev and Stein have reported a combined Monte Carlo/RRKM/classical trajectory model for the CID of a few small

ions [141, 142]. While extremely successful, the methods employed may not be suitable to large systems such as for peptides.

This project describes a very simple model for extracting information from the CID breakdown curves for complexes between the Aβ1-40 peptide and small molecule substrates with RRKM theory [85, 86] to obtain their gas phase binding energies. As will be seen, even though the model is simple, since the systems under investigation are similar, the information obtained is likely reliable in a relative sense. Due to the sheer size of this system, we also developed an alternative approach to the rigorous sums- and densities-of-states calculations required in RRKM theory based on a Hill equation extrapolation of k(E). The performance of this approximate, but faster, approach to k(E) is tested for several polymer-based systems as well as for the Aβ1-40 peptide.

Chapter 2: Experimental and computational techniques

2.1 Introduction

The binding of amino acids to PMMA and LMW molecules to Aβ peptides can be observed in the gas phase through the use of ESI-MS. The dissociation pathways of (PMMA)(AA)(H^+) complexes have been produced with an ESI-triple quadrupole-MS. The binding affinity of LMW molecules on the Aβ peptides have been determined using a simple ESI-MS. For the dissociation pathways of the Aβ1-40/LMW molecules, an electrospray ionization Quadrupole/Time-of-flight (ESI-QTOF-MS) was employed. Since the mass spectrometers available that were used do not provide any information on the 3D conformations of the complexes formed, molecular modeling was also employed. Both molecular dynamics and semi-empirical calculations were performed. This chapter will present in details, the instruments and the computational techniques behind each approach.

2.2 The ESI-MS instrument [143]

Mass spectrometry is an analytical technique in which ions are generated and then detected by their mass-to-charge ratio (m/z). A mass spectrometer is generally composed of three main parts: the ionisation source, the ion analyzer, and the detector. In the early days, harsh ionization such as electron impact (EI) or chemical ionization (CI) were used. In EI, the sample to analyze was subjected to an electron beam, which would fragment following different patterns depending on the chemical functions present in the molecules. In CI, the sample was subjected to a collision beam of inert molecules, which would create a different fragmentation pattern. These ionization

sources were referred as harsh because they would dissociate the core of the molecules into several fragments. For small molecules, certain mechanisms have been proposed in order to "reconstruct" the original molecule from its observed fragments. Far larger molecules, such as polymer and proteins, these types of ionisation techniques are useless for it is impossible to find back the sequence. Mass spectrometry gained considerable popularity in the proteomic field when electrospray ionisation (ESI) was invented by Nobel Prize winner John Bennett Fenn. ESI is a soft ionization technique because it generates multiply charged ions, which allow to accurately calculating the mass of large molecules [24-26]. Here is a description of the main components of ESI-MS (Waters ZQ4000) used to perform the experiments.

2.2.1 The ESI source

The liquid sample (aqueous or not) containing the analyte of interest is infused (by syringe) or injected from a liquid chromatograph (LC) into the ion source inlet. Once inside the capillary needle, a high pressure is applied which nebulizes the sample into a mist and a surrounding electrode applies a voltage of approximately 4 kV to charge the droplets and form ions. Once ejected from the inlet, the mist enters a desolvating capillary where heat and a flow of drying gas (such as N_2) are applied. As shown on Figure 2.1, when the droplets desolvate, they become smaller and their electric field density increases.

The like charges begin to repel one another until they break apart into smaller droplets at the Rayleigh limit [144]:

$$Q_R = 8\pi(\varepsilon_0 \gamma R^3)^{1/2} \quad (2.1)$$

where: Q_R is the droplet charge, γ is the surface tension of the solvent and R is the droplet radius.

The fission of droplets into smaller ones of either positive of negative charges is known as the coulombic explosion. The desolvated molecular ions then reach the entrance of the ion analyzer while the neutral solvent droplets are being pumped out of the source. Figure 2.2 shows the principal components of a typical ESI source.

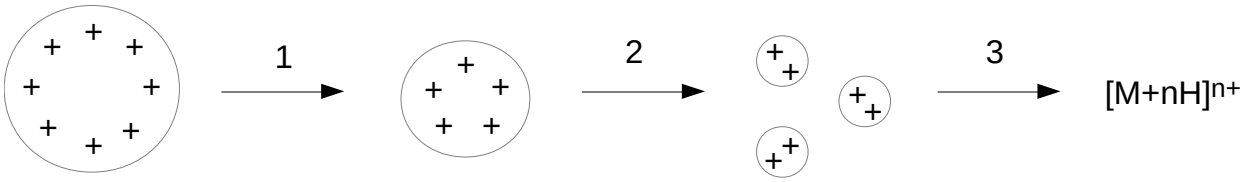


Figure 2.1: After their formation, some droplets may have an excess of positive (or negative charges). (1) As the solvent evaporates from the droplet, it becomes smaller and charges come closer to one another. (2) This causes the droplets to break into smaller ones, and as the solvent and neutral species evaporate, the multiply charged ions are produced. (3) Figure adapted from reference [20].

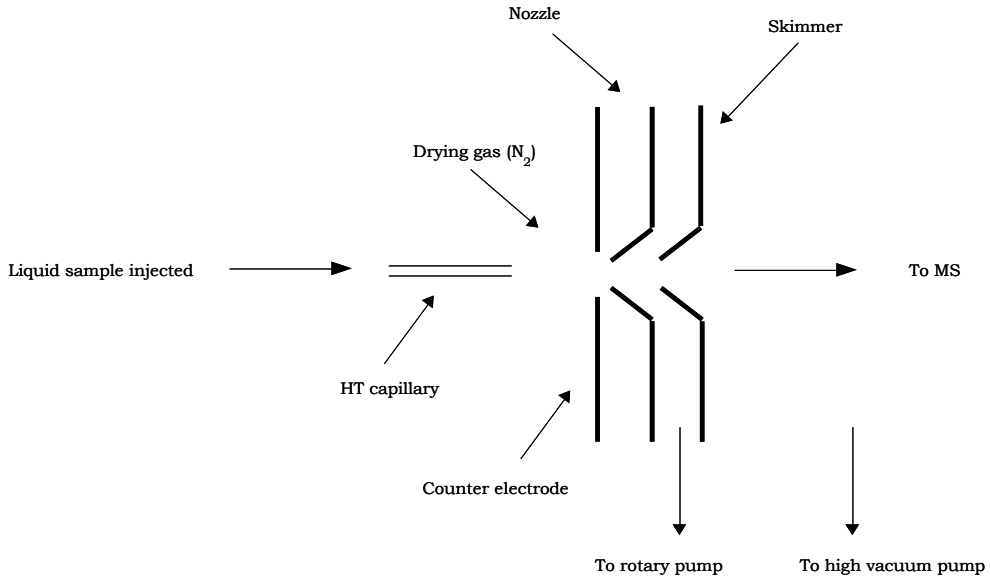


Figure 2.2: An ESI source. The sample injected passes through a high electric discharge capillary before being subjected to a drying gas to help its evaporation, and form a spray. The droplets produced then continue towards a counter electrode, which pushes the ions and neutral species through a nozzle and a skimmer. The conical shaped nozzle and skimmer refine the spray and provide an electrical potential, to help for the extraction of the sample ions. While neutral species are being extracted by the pumps, the ions follow their momentum into the MS analyzer. Figure adapted from reference [143].

2.2.2 The ion analyzer

It consists generally of a quadrupole (or sometimes a hexapole), i.e. four cylindrical rods (length: 50-250 mm; diameter: 6-15 mm) that are placed at 90° from one another (each rod is at the apex of a "square"). Figure 2.3 shows an ion analyzer having a quadrupole. In order to "fly" inside the quadrupole, the beam of ions may pass through accelerating lenses (with a potential of approximately 5 V). While one opposed pair of rods has a potential of $+[U + V\cos(\omega t)]$ applied, the other pair has a potential of $-[U + V\cos(\omega t)]$ applied. U is a fixed potential while $V\cos(\omega t)$ is

a radio frequency (RF) of amplitude, V, and frequency, ω. The expression for the electric potential, F, is given by:

$$F = \left[\frac{(x^2 - y^2)}{r^2}\right](U + V\ \cos(\omega t)\ ) \tag{2.2}$$

where: x and y are the distance between rods in the x and y axes respectively while r represents the radius of the space in the middle of the quadrupole.

A given pair of rods cannot keep the same potential, as ions having an opposite charge would crash on them. The alternating potentials are switching from pair to pair over time as the cos(ωt) function. The middle of the four rods has a null resultant electrical field, because when x = y, then F = 0. The ions then oscillate among the quadrupoles in a motion, around the center field free region, that will depend on their mass, U, V, and ω. The proper combination of these parameters will allow only the ions having a stable motion to reach the detector while all other ions will have greater amplitude (unstable motion), and hit one of the rods. Typical values for ω are in the 1-2 MHz range, U revolves about 1000 V and the maximum RF voltage (V) about 6000 V [143]. The proposed values can be calculated by solving the Mathieu equations applied to the motions of ions. Solutions to this equation yield two factors: a, which is related to the potential, U (2.3) and q, which is related to the amplitude, V (2.4).

$$a = \frac{8zU}{m(\omega r)^2} \tag{2.3}$$

$$q = \frac{4zV}{m(\omega r)^2} \tag{2.4}$$

Combining these two equations define how to control the stable motion of the ions.

$$a/q = 2U/V \tag{2.5}$$

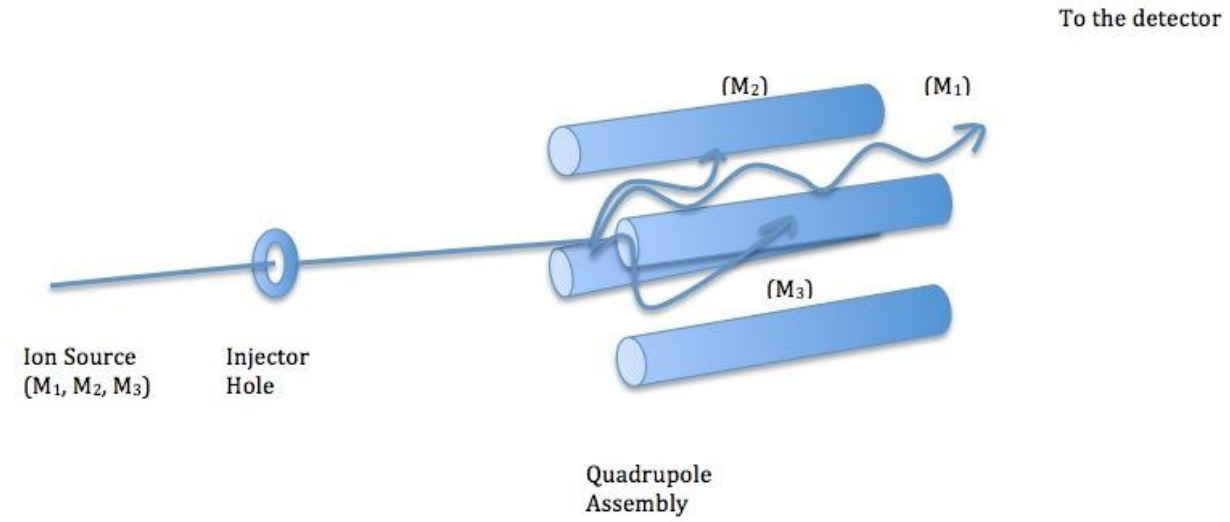


Figure 2.3: The ion beam shown here contains ions of molecular weight M_1, M_2 and M_3. Depending on the parameters selected, some ion(s) (here, M_1) will pass through the middle of the four rods towards the detector while some others (M_2 and M_3) will be deflected to the rods (and never be detected). Figure adapted from reference [143].

For a defined stable motion, a high value of the ratio 2U/V results in few ions being transmitted (narrow range of m/z) but a maximal resolution to the detector. On the other hand, a low value of a/q results in a greater number of ions being transmitted (wide range of m/z), but a decreased resolution. In general, instruments nowadays continuously vary U and V during a run (keeping the ratio 2U/V constant). This allows a wide m/z range of ions to transfer successively through the quadrupole and obtain a relatively high resolution.

2.2.3 Point ion detector

Leaving the analyzer, the ions will hit a detector. MS having a magnetic sector sometimes have a point array collector, because they receive ions of different m/z simultaneously. Quadrupole instruments often use point ion detector as they receive ions successively. A common type of point ion detector is the electron multiplier (Figure 2.4). The principle of the electron multiplier is based on the fact that when an ion hits a metal surface with great velocity, it ejects secondary electrons from the metal. Such a detector is composed of a series of dynodes (metal plates). Ions coming from the analyzer are directed towards the first dynode, which in return will eject a few electrons. These few electrons are then accelerated through an electric potential towards a second dynode and so on. Commercial electron multipliers have 10 to 12 dynodes. Therefore, for a 10 dynodes detector, if one considers that each electron strike ejects 2 "new" electrons, a beam of approximately 2^{10} electrons produce a detectable electric current. The current intensity detected for each m/z value will provide a signal that will be processed to give the usual mass spectra (abundance vs. m/z). The resolution of the mass spectrum peaks is given by equation 2.6:

$$\text{Resolution} = {}^{m}/_{\Delta m} \tag{2.6}$$

where: m is the mass of a singly charged ion peak on the spectrum and Δm is the width of the peak at a specified height.

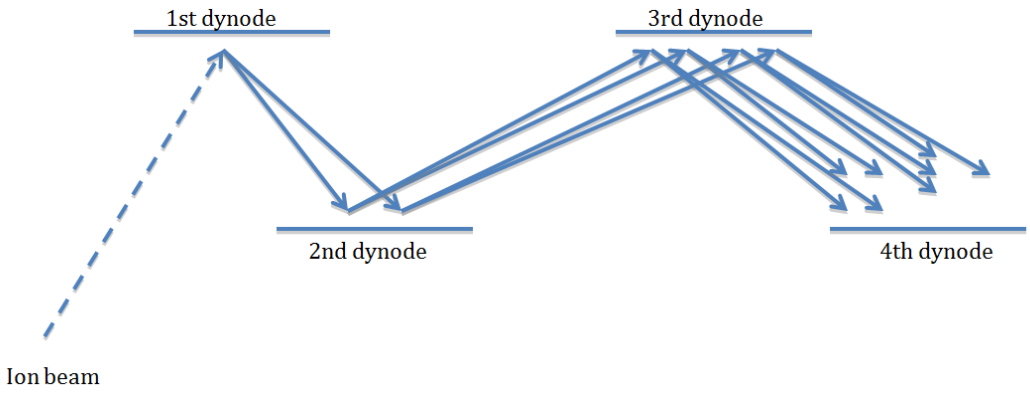


Figure 2.4: An electron multiplier. Each time a dynode is struck; electrons are ejected and directed towards the next dynode. After a series of dynodes, the electrons are multiplied to generate a detectable current.

2.3 The ESI-triple quadrupole-MS instrument

The basics of this instrument resemble the ESI-MS described previously, but the sequence of events is not the same. The main difference lies in its three sequential quadrupoles that are being used in CID experiments (Figure 2.5). The ion source employed is still an ESI. The first quadrupole (Q1) detects the ions produced and is used to select the precursor (parent) ion that will be transmitted to the second quadrupole (Q2). The second quadrupole is located inside the collision cell where a certain pressure of an inert gas (typically argon) is applied. The pressure applied is controlled by the user. The number of fragments produced by precursor ion is directly proportional to the pressure of the inert gas. Only ionic fragments will be transmitted and then analyzed by the last quadrupole (Q3). Finally, a detector similar to the electron multiplier will capture the current of all the transmitted ions.

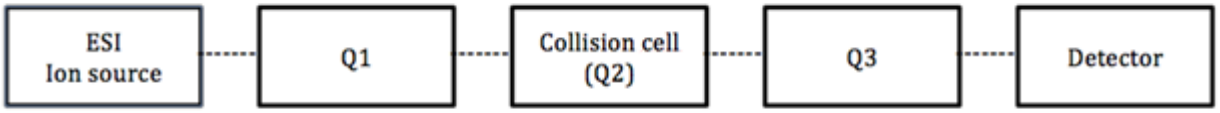


Figure 2.5: Sequence of events in the ESI-triple quadrupole-MS instrument.

2.4 The ESI-QTOF-MS instrument

This instrument essentially functions as the ESI-triple quadrupole-MS described previously. The main difference is the time-of-flight (TOF) module that replaces the third quadrupole of the ESI-triple quadrupole-MS. The ESI-QTOF-MS instrument (Figure 2.6) is very popular in laboratories dealing with proteomics because there is no theoretical upper mass limitation. Although in practice there is one, it is still higher than most other MS. This is the main reason why this instrument was employed for the dissociation of Aβ1-40/LMW molecules. In the ESI-QTOF-MS instrument mentioned, the ions are being generated in an ESI ion source before passing through a first quadrupole analyzer and the collision cell. After exiting the collision cell, the ions enter the orthogonal acceleration (oa) TOF module, as shown in Figure 2.6.

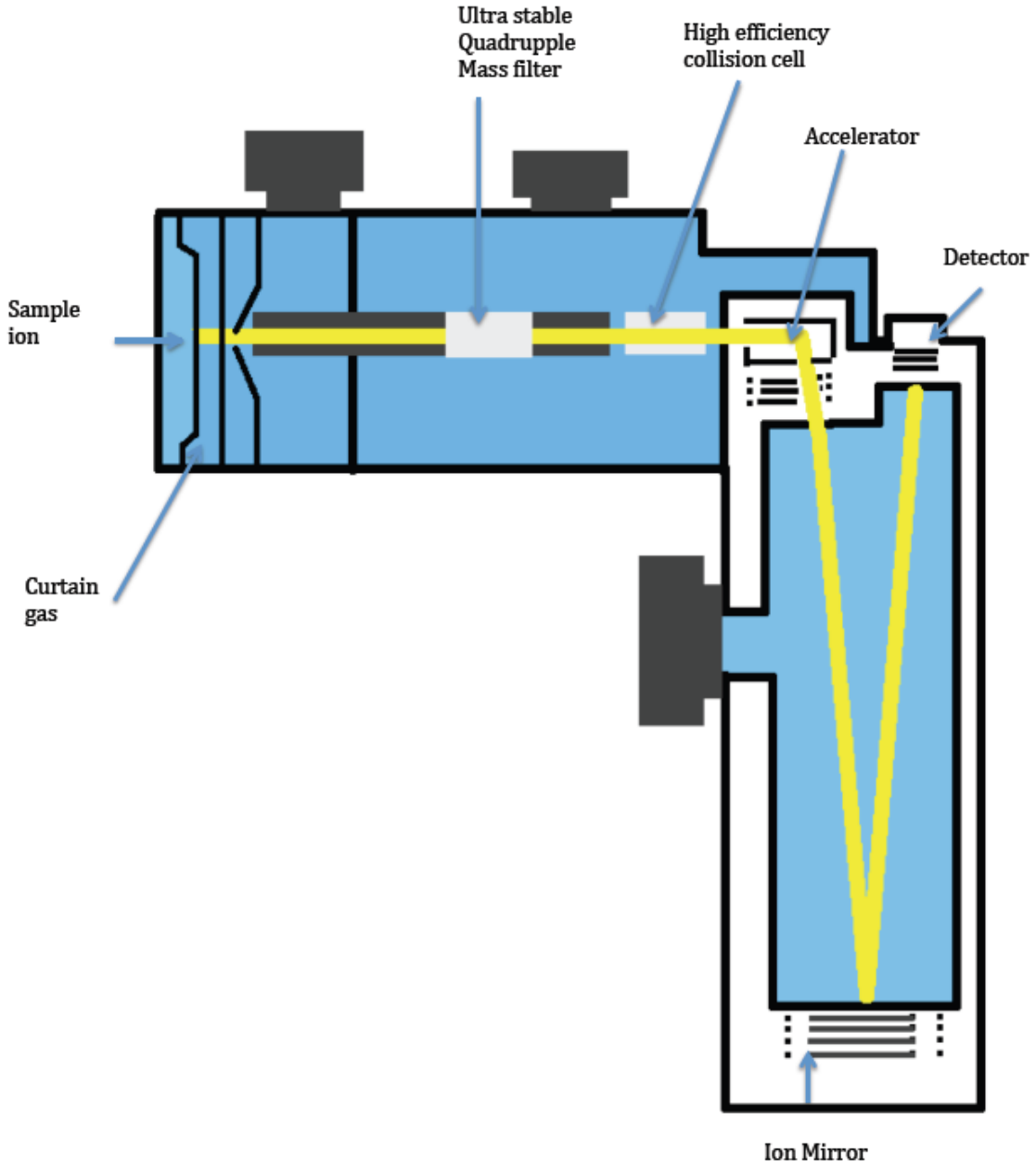


Figure 2.6: Sequence of events in the TOF module of the instrument. A continuous ion beam from the quadrupole enters the accelerator and is being pulsed with an orthogonal field gradient (right angle). This orthogonal acceleration (oa) forces the ions to travel in a modified trajectory inside a drift tube before reaching an ion mirror that reverses the direction of ions and correct small energy differences. The ion beam is then directed towards the detector (typically a multichannel plate ion collector) which records the precise arrival time of each ion.

The ion optics principle of the TOF is expressed as the drift time of an ion passing through a tube, which depends on its mass and its charge. At the entrance of the TOF module, the ions are pulsed: they leave the source at the same time before being accelerated through an electric field, E. The kinetic energy, K, of the ions is related to the electric field via equation 2.7.

$$K = \frac{mv^2}{2} = z \cdot e \cdot E \tag{2.7}$$

where: m and v are the mass and the final velocity (after acceleration) of the ion respectively, z is number of charges on the ion, e is the charge of the electron.

The following rearrangement of equation puts in evidence the relationship between the velocity of the ion and the other parameters.

$$v = \sqrt{\frac{2z \cdot e \cdot E}{m}} \tag{2.8}$$

For a drift tube of distance, d, in the QTOF module, it becomes possible to determine the time required, t, for the ion to reach the detector.

$$t = {}^{d}/_{v} = {}^{d}\Big/\sqrt{\frac{2z \cdot e \cdot E}{m}} \tag{2.9}$$

By rearranging equation 2.9 to separate the parameters that are actually constant, equation 2.10 is obtained:

$$t = \left(\sqrt{m/z}\right) \cdot \left(\frac{d}{\sqrt{2e \cdot E}}\right) = \sqrt{m/z} \times \text{constant} \tag{2.10}$$

It is possible to see from equation 2.10 that the drift time of an ion is proportional to the square root of its mass-to-charge ratio. As an example, an ion of m/z 400 will arrive 4 times slower at the detector as compared to an ion of m/z 25.

2.5 Molecular Mechanics / Molecular Dynamics [8]

Quantum mechanics (QM) calculations help providing very accurate geometry and energy for molecules. However, the computational cost of these methods is rising significantly with the size of the system to study. Therefore, for moderate to large systems, such as those studied in this thesis, an alternative is required: molecular mechanics / molecular dynamics (MM/MD) methods. In such computational methods, atoms in molecules are defined by as "balls" (atoms with partial charges) linked by "springs" (chemical bonds) and not as nuclei surrounded by electrons in orbitals. Because of the Born-Oppenheimer approximation (electrons are moving much faster than nuclei and therefore adjust their positions instantly), there are no more electronic degrees of freedom to consider, which helps saving calculation time. MM/MD works on the concept that the chemical properties of a chemical group are transferrable. As an example, a carboxylic group will remain approximately the same from one molecule to another, if the molecules are relatively similar.

2.5.1 The General Amber force field

To define the movements between atoms, force fields are used. Force fields are a parameterized sum of mathematical functions to evaluate the potential energy of molecular conformations. The parameterization and form of such force fields are based to reproduce experimental data or QM

calculations. Before using a force field, the user must specify the types of atoms involved in the system in terms of hybridization and connectivity. A force field is functional and developed as a whole; in other words, one cannot take a term from one force field and transfer it into another.

The available empirical force field employed for all systems studied here is the general Amber force field [145] which mathematical form is:

$$E_{pair} = \sum_{bonds} k_r(r - r_{eq})^2 + \sum_{angles} k_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{\upsilon_n}{2} \times [1 + \cos(n\omega - \gamma)] + \sum_{i<j} \frac{q_i q_j}{\varepsilon R_{ij}} + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right] \quad (2.11)$$

where: the first term represents the bond stretching, the second term is the angle strain, the third term is the torsion energy, the fourth term relates to the non-bonded 1-4 electrostatic interactions, and the last term treats the non-bonded 1-4 van der Waals interactions.

Bond stretching is best represented by a Morse potential. However, since this mathematical form is computationally expensive, a simple harmonic potential is used where r is the bond length, r_{eq} is the bond length at equilibrium, and k_r is a parameterized force constant similar to Hook constant. This mathematical form is valid if the bond stretching is considered harmonic near the reference bond length. The angle bending term is also approximated by a harmonic function in which θ is the angle between three atoms linked by two bonds, θ_{eq} is the angle at equilibrium, and k_θ is a parameterized constant.

As opposed to stretching and angle bending terms, the torsional term has multiple energy minima separated by finite energy barriers. In this simple version, the υ_n is the "barrier height", n is the multiplicity of the torsion, ω is the torsion's value, and γ is a phase factor to correct the position of the energy minima/maxima. Some more "focussed" force fields that are parameterized for small molecules such as MM2 (for hydrocarbons) [146] or MM3 (for conjugated alkenes) [147, 148], also have a term for improper torsions. Nevertheless, in the current force field, it would add a severe computational cost to gain insight into conformational effects, which can be neglected in the systems studied here. The general Amber force field does not have any cross terms (to account for coupled movements), which saves on computational costs.

When elements with different electronegativity are bound, it affects the charge distribution. This behaviour is represented by the 1-4 non-bonded electrostatic interactions that obey Coulomb's law where q_i and q_j are the partial charges of non-bonded atoms, ε_0 is the vacuum permittivity, and r_{ij} is the distance separating the pair of atoms. The "1-4" refers to the position in the connectivity of atoms: a given pair of non-bonded atoms separated by two atoms (positions 2 and 3) in between them. This choice of 1-4 was made to differentiate the non-bonded terms from the bonded terms.

The partial charges need to be assigned to each atom before using MM. Several methods have been established to calculate these partial atomic charges. One classic approach is to use QM to calculate a molecular electrostatic potential (MESP), $\phi(r)$, i.e. the force acting on a unit positive charge. It is mathematically the sum of repulsions with the nuclei, $\phi_{nucl}(r)$, and the attractions to the electronic charge distribution ($\rho(r)$), $\phi_{elec}(r)$. The MESP is usually expressed as:

$$\phi(r) = \phi_{nucl}(r) + \phi_{rep}(r) = \underbrace{\sum_{A}^{nuclei} \frac{Z_A}{|r-R_A|}}_{attraction} - \underbrace{\int \frac{\rho(r')}{|r-r'|} dr'}_{repulsion} \quad (2.12)$$

where: Z_A is the charge of the nuclei, R_A is the position of nuclei, and r' is a point in space in the electronic charge distribution.

Although the concept is robust, for large systems, it requires a long time to calculate since QM is necessary. For all MM/MD calculations, the Austin Method 1 – Bond Charge Corrections (AM1-BCC) method was used [149, 150]. This method calculates the partial charges with the semi-empirical method AM1 and then applies a correction to match the quality of charges calculated at the HF/6-31G* ESP-derived charges.

The last term is the vdW interactions and is represented by a classical Lennard-Jones (or 6-12) potential. At very short distances, the repulsive nuclear term (A_{ij}/r^{12}) dominates and its numerical value increases sharply. At longer distances, the second term (B_{ij}/r^6) represents more adequately the London Dispersion Forces (LDF) as they die off as $1/r^6$. In fact, even if this last term is known for representing vdW interactions, it represents more the LDF while the Keesom interactions are being included in the electrostatic term due the presence of dipoles.

2.5.2 Molecular Mechanics minimization methods

Finding the global minimum energy and its corresponding 3D conformation of a large system can be challenging and it has been known for over 40 years by scientists interested in protein folding

[151]. For a system containing N atoms, if the translations and rotations are ignored, its potential energy surface (PES) has 3N-6 degrees of freedom. A typical multidimensional PES has one global minimum, a certain number of local minima, saddle points and transition states. Locating these points on the PES requires evaluating all first derivatives of the energy with respect to all internal degrees of freedom when it is equal to zero ($\frac{\partial E}{\partial x_i} = 0$). To further determine if the extremum corresponds to a minimun or a maximum on the PES, the second derivative is evaluated as well. If $\frac{\partial^2 E}{\partial {x_i}^2} < 0$ the point corresponds to a maximum, however, if $\frac{\partial^2 E}{\partial {x_i}^2} > 0$, the point corresponds to a minimum. The first derivative of the energy with respect to each internal degree of freedom is a vector called a gradient. There are different strategies to optimize a conformation and minimize its related energy. Two of these minimizers were used to optimize the 3D conformation of the system under study: steepest descent (SD) and adapted-basis Newton-Raphson (ABNR). Steepest descent is an algorithm that brings the displacement downhill towards the nearest minimum of energy on the PES by minimizing the gradient's value. In this first-order minimization method, the 3N Cartesian coordinates are represented by a 3N-dimensional unit vector, s_k:

$$s_k = -\frac{g_k}{|g_k|} \tag{2.13}$$

where: g is the gradient vector.

This algorithm is slower, but generally used as a first minimization method because it ensures to find a minimum on the PES. The Newton-Raphson minimization method is reliable when the PES is more or less harmonic. If the second derivatives are not too computationally demanding, they are calculated in a matrix called Hessian. When the Hessian matrix too is long to calculate,

algorithms are used to generate an approximate Hessian matrix (like ABNR can do in large systems). Mathematically, the minimization is calculated using the following expression:

$$x_{k+1} = x_k - g_k H_k^{-1} \quad (2.14)$$

where: H_k^{-1} is the inverse Hessian matrix and g_k the gradient. For each iteration, k, the new positions, x_{k+1}, are obtained based on the current positions, x_k.

If the conformation is far from a minimum, a method like steepest descent is preferable. However, to refine a conformation already pre-minimized by steepest descent, ABNR is more efficient. Most of the time, minimizations cannot be done in one single step, but rather through an iterative process. Therefore, a convergence criterion is required. One strategy used as a convergence criterion is to calculate the gradient of energy with respect to the coordinates and calculate a root mean square (RMS) deviation between the actual gradient value and the preceding one.

$$RMS = \sqrt{\frac{g^T g}{3N}} \quad (2.15)$$

where: N is the number of degrees of freedom and g^T is the transposed gradient.

2.5.3 Molecular Dynamics Simulated Annealing (MD-SA)

Once the initial system was charge equilibrated and minimized, the PES of the system can be explored using molecular dynamics. MD is a deterministic method in which successive configurations of a system are generated over time by integrating Newton's laws of motion. In a finite difference method, for each time step, δt, particles in the system are moved, changing the forces exerted inside this system. Time steps should normally respect the molecular stretching motions of 0.5 ps to a maximum of 2 ps. From the coordinates, the forces acting on the particles and their velocities, the displacements of the particles are calculated and this process repeats itself for another time step. Again, there are a variety of finite difference methods such as the popular Verlet [152] or the leap-frog [153] algorithm. However, for the current polymer complex systems, we are more interested in sampling the PES to reach the global minimum than obtaining a realistic MD simulation over time. Thus, the self-guided Langevin dynamics [154] was used to enhance conformational sampling of the PES. The Langevin equation of motion for a particle i, is written as:

$$m_i \frac{d^2 x_i(t)}{dt^2} = F_i\{x_i(t)\} - \gamma_i \frac{dx_i(t)}{dt} m_i + R_i(t) \qquad (2.16)$$

The variable m_i is the mass of particle i. F_i is a guided force calculated from a potential of mean force during the simulation to accelerate the particles motions to enhance the conformational sampling. The factor γ_i is the collision frequency: a high value provides a high temperature control while a low value allows the temperature to fluctuate more. Finally, R_i is a random fluctuation uncorrelated with the particles velocities.

The first stage of an MD simulation is the heating phase in which the temperature of the system is slowly increased and transferred into the particles kinetic energy. Then, an equilibration stage is necessary to allow the kinetic energy to flow inside the overall system. Then, the MD simulation can start. To help the conformation sampling of the PES, the MD is coupled a simulated annealing (SA) algorithm which consists in a series of controlled heating and cooling phases. It takes its name from metallurgy, where annealing is a treatment to relieve strain from a metal. This computational method helps in finding an optimal value (hopefully the best) for problems that have multiple solutions [155]. The mathematical relationship between the temperature and the kinetic energy of the system is given by the relation:

$$\sum_{i=1}^{N} \frac{|p_i|^2}{2m_i} = \frac{k_B T}{2}(3N - N_c) \qquad (2.17)$$

where: p_i is the momentum of one of the N particles and N_c is the number of constraints applied to the system (e.g. fixed distance constraint).

By heating the system to higher temperatures, it provides sufficient kinetic energy for the molecules to escape local potential energy minimum they can be located in. The cooling phase allows the release of kinetic energy and finds a new minimum on the PES. By repeating this annealing schedule several times, the PES can be sampled reasonably. By the end of this iterative process, a certain number of low energy conformations are found.

2.5.4 Normal mode analysis

If the system is too large and in case ab initio methods cannot converge, the normal mode analysis can provide rotational constants and harmonic vibrational frequencies. Although not as accurate as frequencies provided by ab initio methods, it is essential to locate minima and saddle points on the PES. In addition, it will provide the necessary vibrational frequencies for the application of RRKM theory to complexes dissociation. The Nmode module [156] of the Amber 9 suite of programs [157] allows the calculation of the vibrational Langevin modes (which are similar to the normal modes) [158]. The vibrational frequencies, υ, are found using the following equation:

$$\upsilon_i = \sqrt{\frac{\lambda_i}{2\pi}} \qquad (2.18)$$

where: λ_i are the eigenvalues of the mass inverted Hessian matrix.

Normal mode analysis works well for stiff molecules but not as well for floppy molecules since the thermal energy can flow inside the molecule causing conformational changes. Therefore, each time a normal mode analysis is done, it is at least required to ensure that the conformation is in a well defined energy minimum found by the MD-SA simulation.

2.6 The semi-empirical method AM1 [9, 11]

In computational chemistry, when a chemical system is difficult to converge into a local or global minimum energy on the PES, a trick of the trade is to use a "cheap" minimization method first. Once the proper minimum is obtained, if the system needs a subsequent optimization, the scientist

can use a more accurate method. In case of large or very floppy systems, QM methods are not really efficient because they take a long time to converge. However, semi-empirical methods represent a suitable alternative between QM and MM. One of the most popular semi-empirical methods is the Austin Method 1 (AM1).

Semi-empirical methods trace their origin in the Hartree-Fock (HF) approximation to treat the Schrödinger's equation. The HF method is based on the variational theorem, which states that even though there is no "correct" solution to the Schrödinger's equation, if the energy of a system is obtained from an approximation of the wavefunction, it will always be higher than the global minimum energy. Therefore, the best wavefunction will provide the "closest energy" to the true global minimum. In semi-empirical methods, the HF method uses several approximations and parameters are imported from experimental data. An older method, MNDO (Modified Neglect of Differential Overlap), uses an approximation of the differential diatomic overlap integral. AM1 is based on this method. It uses a modified version of the core-core term by reducing the repulsion at short interatomic distances with the use of off-center spherical Gaussian functions. The energy for two atoms, E_{AB} is given in the AM1 formalism by equation 2.19.

$$E_{AB} = E_{MNDO} + \frac{Z_A Z_B}{R_{AB}}$$

$$\times \left\{ \sum_i K_{Ai} \cdot \exp[-L_{Ai}(R_{AB} - M_{Ai})^2] + \sum_j K_{Bj} \cdot \exp[-L_{Bj}\left(R_{AB} - M_{Bj}\right)^2] \right\} \quad (2.19)$$

where: E_{MNDO} is the energy calculated by the MNDO method, Z is the nuclear charge of an atom, R is the interatomic distance, L is the width of the Gaussian function, K and M are parameters related to individual atoms.

Chapter 3: Experimental and computational protocols

3.1 Experimental protocols

This section describes the experimental protocols developed for the different projects. The list of chemicals and mass spectrometers employed are also provided.

3.1.1 Dissociation of $(PMMA_n)(AA)(H^+)$ complexes

Acetonitrile was purchased from Fisher Scientific (Pittsburgh, PA, USA). PMMA (Mp = 1450, Mw/Mn = 1.07, batch # 20223-9) and PMMA (Mp = 2680, Mw/Mn = 1.09, batch 20227-6) were purchased from Polymer Laboratories (Amherst, MA, USA). Glycine, DL-leucine, and DL-phenylalanine were purchased from Sigma-Aldrich (Milwaukee, WI, USA).

PMMA was dissolved in acetonitrile to prepare 1 mg mL^{-1} solution. Each amino acid was dissolved in deionized water to prepare 1 mg mL^{-1} solution as well. Equal parts of the PMMA and given amino acid solutions were mixed together to form 1:1 polymer/amino acid solution. To ensure the complexes formation, 50 μL of 5 % formic acid were added to the solutions.

CID MS experiments were performed using a Micromass Quattro LC triple quadrupole mass spectrometer (Waters Micromass, Manchester, England) equipped with a Z-spray ion source. The polymer/amino acid solutions were infused through a 50 μL loop using a syringe with a mobile phase of 8:2 acetonitrile/water at a rate of 40 μL min^{-1}. Since a polymer solution contains oligomers

of various lengths, a series of PMMA/amino acid complexes should be observed. First, a complete mass spectrum (m/z 50 to 1800) was acquired to ensure that the complexes were formed and detected by the instrument. The m/z ratio peak of a given complex was then selected to be transmitted through the collision cell to observe its daughter ions. In order to obtain a breakdown diagram for such a system, the collision cell energy was increased by increment of 5 eV until the parent ion was completely dissociated. Table 3.1 reports the polymer/amino acid complexes that were subjected to this protocol.

Table 3.1: $(PMMA_n)(AA)(H^+)$ complexes dissociated in acidic conditions.

PMMA oligomer	Glycine	Leucine	Phenylalanine
7mer	$(PMMA_7)(Gly)(H^+)$	$(PMMA_7)(Leu)(H^+)$	$(PMMA_7)(Phe)(H^+)$
10mer	$(PMMA_{10})(Gly)(H^+)$	$(PMMA_{10})(Leu)(H^+)$	$(PMMA_{10})(Phe)(H^+)$
13mer	$(PMMA_{13})(Gly)(H^+)$	$(PMMA_{13})(Leu)(H^+)$	$(PMMA_{13})(Phe)(H^+)$

When operating the mass spectrometer to acquire full spectra, the collision cell energy was set to 0 eV while the entrance and exit of the collision cell were set to 35 V. When MS/MS spectra were acquired, the entrance and exit of the collision cell were set to 0 V, since the collision cell energy was taking non-zero values. All other parameters were held constant and as reported in Table 3.2.

Table 3.2: Parameter settings of the Micromass Quattro LC triple quadrupole mass spectrometer for the detection and dissociation of $(PMMA_n)(AA)(H^+)$ complexes.

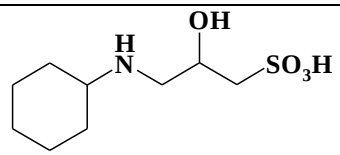
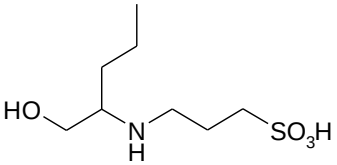
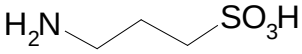
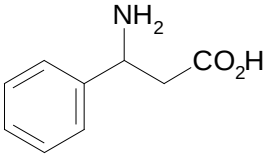
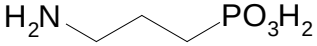
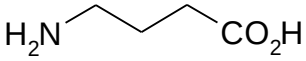
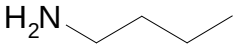
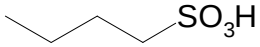
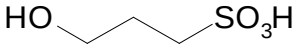
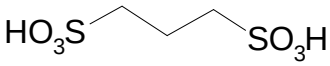
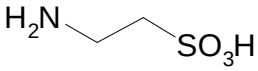
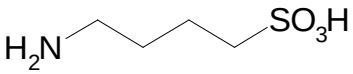
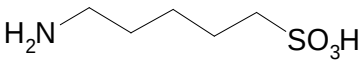
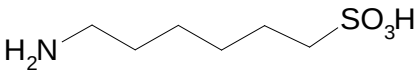
Parameter	Setting
Electrospray ionization source	
Ionization mode	ES+
Capillary (kV)	4.00
HV lens (kV)	0.23
Cone (V)	40
Skimmer lens (V)	10
Skimmer (V)	1.0
RF lens (V)	0.4
Source temperature (°C)	80
Drying gas (N_2) flow (L h^{-1})	85
MS 1 and MS 2	
Ion energy (V)	2.0
Low mass resolution	13.0
High mass resolution	13.0
Multiplier [MS 2 only] (V)	657
Collision cell	
Collision gas	Argon
Gas cell average pressure† (mbar)	1.0×10^{-4} to 1.0×10^{-3}
Detector	
Cycle time (s)	3.20
Scan duration (s)	2.96
Interscan delay (s)	0.24
Retention window (min)	3.00
m/z†	50 to 1800

†depending on the experiment

3.1.2 Binding affinity determination of Aβ peptide/LMW molecule complexes

Aβ1-28, Aβ1-40, Aβ1-42, and Aβ3pE-40 peptides were purchased from rPeptide (Bogart, GA, USA) and three additional peptides were synthesized at Neurochem Inc. (Figure 3.1). In the sequences shown in Figure 3.1, histidines are not identified as protonated amino acids. This is supported by the fact that both histidines 13 and 14 are also known as the Aβ zinc binding site. At pH below 6, the binding of the zinc cation is weaker due to the protonated histidines, whereas at physiological pH, the Aβ has more propensity to bind the zinc cation [159]. LMW compounds were purchased from commercial sources or synthesized at Neurochem Inc. (Table 3.3). Stock solution (50 μM) of each peptide was prepared in 0.1 % trifluoroacetic acid (TFA, in deionized H_2O), which contributes to minimize peptide self-assembly by maintaining the N-terminal part as an α-helix [160, 161]. Each stock solution was then sonicated for 5 minutes. To determine the binding affinity (BA) of LMW compounds, a stock solution of Aβ peptide was mixed with a compound solution in deionized H_2O to obtain a test solution with final concentrations of 30 μM and 150 μM for the peptide and the compound, respectively. The pH of the test solution was adjusted to 7.4 ± 0.2 using 0.1 % NaOH and a Beckman Φ35 pH meter equipped with a Corning Semi-Micro Combination pH electrode. ESI-MS measurements were performed on a Waters ZQ 4000 mass spectrometer equipped with a Waters 2795 sample manager (Waters Micromass, Manchester, England). MassLynx 4.1 was employed for data processing and analysis. The sample solutions were injected in the ESI-MS using a mobile phase composed of 95 % H_2O and 5 % ethanol at a flow rate of 150 μL min^{-1}. All other parameter settings were held constant throughout the experiments as reported in Table 3.4.

Table 3.3: Compounds tested for their binding affinity to the different Aβ peptides.

ID	Structure[a]	MW (Da)	Name(s) [CAS]
1		237.3	3-Cyclohexylamino-2-hydroxy-1-propanesulfonic acid; CAPSO [73463-39-5]
2		225.3	3-(1-Hydroxy-2-pentyl)amino-1-propanesulfonic acid [720699-42-3]
3		139.2	3-Amino-1-propanesulfonic acid; homotaurine; Vivimind™ [3687-18-1]
4		165.2	3-Amino-3-phenylpropanoic acid [103-01-5]
5		139.1	3-Aminopropyl-1-phosphonic acid [13138-33-5]
6		103.1	4-Aminobutyric acid; GABA [56-12-2]
7		73.1	n-Butylamine [109-73-9]
8		138.2	1-Butanesulfonic acid [2386-47-2]
9		140.2	3-Hydroxy-1-propanesulfonic acid [15909-83-8]
10		204.2	1,3-propanedisulfonic acid; Eprodisate; Kiacta™ [13881-91-9]
11		125.2	2-aminoethanesulfonic acid; Taurine [2386-47-2]
12		153.2	4-amino-1-butanesulfonic acid [14064-34-7]
13		167.2	5-amino-1-pentanesulfonic acid [37043-68-8]
14		181.2	6-amino-1-hexanesulfonic acid [72372-71-5]

[a] Only non-salt forms are given. Some compounds were sodium salt forms.

amino acid position

1 10 20 28 40

- - + - - + - - +

H_2N-**DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA**-OH

Aβ1-42 (4514 Da)

- - + - - + - - +

H_2N-**DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV**-OH

Aβ1-40 (4330 Da)

- - + - - + - - +

H_2N-**DAEFRHDSGYEVHHQKLVFFAEDVGSNK**-OH

Aβ1-28 (3263 Da)

- - -

H_2N-**GASFGHASGYEVHHQELVFFAESVGSNS**-NH_2

all (-) peptide (2950 Da)

+ + +

H_2N-**GASFRHASGYSVHHQKLVFFASAVGSNK**-NH_2

all (+) peptide (2989 Da)

+ - - + - - +

pyro-**FRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV**-OH

Aβ3pE-40 (4126 Da)

Figure 3.1: Peptides used in the study. The numbers on top represent the amino acid number in the sequence. The "+" and "-" signs represent the negative or positive charge respectively in its side-chain at physiological pH. Each peptide sequence is followed by its name and molecular weight in parentheses.

Table 3.4: Parameter settings of the Waters ZQ 4000 mass spectrometer for determining the binding affinity of Aβ peptides/LMW molecules.

Parameters	Setting
Electrospray ionization source	
Ionization mode	ES+
Capillary (kV)	3.20
Cone (V)	18
Extractor (V)	1.0
RF lens (V)	0.3
Source temperature (°C)	90
Desolvation temperature (°C)	150
Cone gas (N_2) flow (L h^{-1})	50
Desolvation gas (N_2) flow (L h^{-1})	300
MS	
Ion energy (V)	0.5
Low mass resolution	12.0
High mass resolution	12.0
Multiplier (V)	680
Detector	
Cycle time (s)	1.0
Scan duration (s)	0.9
Interscan delay (s)	0.1
Number of scans	300
m/z range	100 to 2000

3.1.3 Dissociation of Aβ peptide/LMW molecule complexes

The Aβ1-40/LMW molecule solutions were prepared using the same protocol as described in the previous section. If a given complex peak was less than 10 % in intensity, the concentration of LMW molecule was increased up to a maximum of 600 μM. ESI-MS/MS measurements were performed on a Micromass QTOF micro mass spectrometer (Waters Micromass, Manchester, England). First, a complete mass spectrum (m/z 100 to 2000) was acquired to ensure that the complexes were formed and detected by the instrument. The m/z ratio peak of a given complex was then selected to be transmitted to the collision cell to observe its daughter ions. In order to obtain a breakdown diagram for such a system, the lab-frame collision energy was increased by increments of 2 eV until the parent ion was completely dissociated. The complex solutions were infused in the mass spectrometer using a syringe pump at a rate of 10 μL min^{-1}. In this mass spectrometer, the collision energy could not be turned off because the ions would not fly across. The Aβ1-40/LMW molecule complexes under study are reported in Table 3.5. For the development of the RRKM method, the LMW molecules that were dissociated from the Aβ1-40 peptide are listed in Table 3.6. When acquiring the data for the breakdown diagrams, the collision cell energy was increased while other parameters were held constant. Table 3.7 reports the parameters settings for these experiments on the ESI-QTOF-MS.

Table 3.5: Aβ1-40/LMW molecule complexes dissociated by CID MS.

m/z region	Aβ1-40		
	1 LMW molecule (1M)	2 LMW molecules (2M)	3 LMW molecules (3M)
5+	$[A\beta1\text{-}40\cdot1M\cdot xNa]^{5+}$	$[A\beta1\text{-}40\cdot2M\cdot xNa]^{5+}$	$[A\beta1\text{-}40\cdot3M\cdot xNa]^{5+}$
4+	$[A\beta1\text{-}40\cdot1M\cdot xNa]^{4+}$	$[A\beta1\text{-}40\cdot2M\cdot xNa]^{4+}$	$[A\beta1\text{-}40\cdot3M\cdot xNa]^{4+}$

Note: the "x" represents the number of sodium cations corresponding to the highest intensity complex peak.

Table 3.6: LMW molecules used to form complexes with the Aβ1-40 peptide and develop the approximate RRKM method.

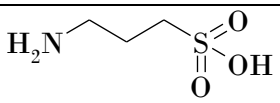
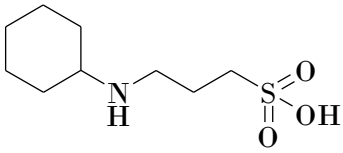
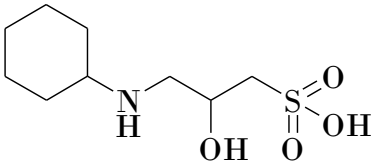
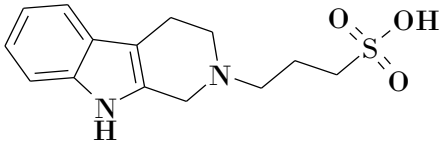
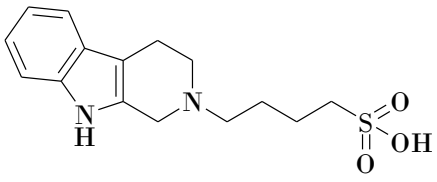
ID	Molecules	Structures
M1	Homotaurine	
M2	3-(cyclohexylamino) propane-1-sulfonic acid	
M3	Cyclohexylamino-2-hydroxy-1-propanesulfonic acid (CAPSO)	
M4	3-(1,3,4,9-tetrahydro-2*H*-β-carbolin-2-yl) propane-1-sulfonic acid	
M5	3-(1,3,4,9-tetrahydro-2*H*-β-carbolin-2-yl) butane-1-sulfonic acid	

Table 3.7: Parameter settings of the Micromass QTOF micro mass spectrometer for the detection and dissociation of the Aβ1-40/LMW molecule complexes.

Parameter	Setting
ESI-QTOF-MS	
Ionization mode	ES+
Capillary (kV)	2.80
Sample cone (V)	15.0
Extraction cone (V)	2.0
Desolvation temperature (°C)	120.0
Source temperature (°C)	90.0
Ion energy (V)	2.0
Low Mass Resolution	14.0
High Mass Resolution	14.0
RF Lens 1	0.9
Pre/post filter	5.0
RF Lens 2	9.0
Aperture (V)	7.0
Plate one	3.0
Entrance	105.0
Gas cell RF	180.0
Plate two	-3.0
Pusher cycle time	47.0 (auto)
Pusher Frequency	21276.60
MCP detector (V)	2350.0
Acceleration (V)	200.0
Tube Lens (V)	90.0
TOF flight tube (V)	5630.0
Reflectron (V)	1780.0
Pusher offset	0.6
Pusher	830.0

Table 3.7: continued

Puller	635.0
TDC start (mV)	800.0
TDC stop (mV)	61.0
Centroid threshold	1.0
Minimum points	1.0
Resolution	5000.0
Lteff	1079.81
veff	5630.00
Detector	
Cycle time (s)	1.0
Scan duration (s)	0.9
Interscan delay (s)	0.1
Run time (min)	1.0
m/z range	50 to 2000

3.2 Computational Procedures

This section describes the computational procedures developed for the different projects. First, the MD-SA procedure to model the $(PMMA_n)(AA)(H^+)$ complexes is described, followed by the AM1 optimization of these complexes. Finally, the procedure to model the breakdown diagrams using the RRKM theory is presented.

3.2.1 MD-SA procedure and geometry optimization

The investigation of low energy conformations of PMMA oligomers and $(PMMA_n)(AA)(H^+)$ complexes was done using a MD-SA script programmed inside the Amber 9 modeling

environment [157]. Since the amino acids are small molecules, they were minimized using standard procedure of combined SD and ABNR with automated convergence energy criteria of in "mild" SA conditions. The MD-SA described here was inspired and adapted from previous publications [162-164]. The $(PMMA_n)(AA)(H^+)$ systems investigated were formed with an oligomer of PMMA (4mer, 7mer, or 13mer) and an amino acid (glycine, leucine, or phenylalanine). As shown in Table 3.8, simulations were performed on systems with different protonation sites since each complex formed has, experimentally, two dissociation pathway possibilities: the amino acid keeping the proton, leaving a neutral oligomer (3.1) or the oligomer keeping the proton, leaving a neutral amino acid (3.2).

$$(PMMA_n)(AA)(H^+) \rightarrow PMMA_n + (AAH^+) \quad (3.1)$$

$$(PMMA_n)(AA)(H^+) \rightarrow (PMMA_nH^+) + AA \quad (3.2)$$

Modeling these systems will provide insights on the effect of: (i) having the same amino acid on different PMMA oligomers, (ii) having a different amino acid on the same PMMA oligomer, and (iii) the different protonation sites for $(PMMA_n)(AA)(H^+)$ complexes.

Table 3.8: PMMA oligomers, amino acids, and complexes investigated computationally.

Molecular species	Charge distributions[†]
$PMMA_4$	$PMMA_4$, $PMMA_4H_1^+$, $PMMA_4H_2^+$, $PMMA_4H_3^+$, $PMMA_4H_4^+$
$PMMA_7$	$PMMA_7$, $PMMA_7H_1^+$, $PMMA_7H_2^+$, $PMMA_7H_3^+$, $PMMA_7H_4^+$, $PMMA_7H_5^+$, $PMMA_7H_6^+$, $PMMA_7H_7^+$
$PMMA_{13}$	$PMMA_{13}$, $PMMA_{13}H_7^+$, $PMMA_7H_{13}^+$
Glycine	Gly, ($GlyH^+$)
Leucine	Leu, ($LeuH^+$)
Phenylalanine	Phe, ($PheH^+$)
$(PMMA_4)(Gly)(H^+)$	$(PMMA_4)(GlyH^+)$, $(Gly)(PMMA_4H_1^+)$, $(Gly)(PMMA_4H_2^+)$, $(Gly)(PMMA_4H_3^+)$, $(Gly)(PMMA_4H_4^+)$
$(PMMA_4)(Leu)(H^+)$	$(PMMA_4)(LeuH^+)$, $(Leu)(PMMA_4H_1^+)$, $(Leu)(PMMA_4H_2^+)$, $(Leu)(PMMA_4H_3^+)$, $(Leu)(PMMA_4H_4^+)$
$(PMMA_4)(Phe)(H^+)$	$(PMMA_4)(PheH^+)$, $(Phe)(PMMA_4H_1^+)$, $(Phe)(PMMA_4H_2^+)$, $(Phe)(PMMA_4H_3^+)$, $(Phe)(PMMA_4H_4^+)$
$(PMMA_7)(Gly)(H^+)$	$(PMMA_7)(GlyH^+)$, $(Gly)(PMMA_7H_1^+)$, $(Gly)(PMMA_7H_2^+)$, $(Gly)(PMMA_7H_3^+)$, $(Gly)(PMMA_7H_4^+)$, $(Gly)(PMMA_7H_5^+)$, $(Gly)(PMMA_7H_6^+)$, $(Gly)(PMMA_7H_7^+)$
$(PMMA_7)(Leu)(H^+)$	$(PMMA_7)(LeuH^+)$, $(Leu)(PMMA_7H_1^+)$, $(Leu)(PMMA_7H_2^+)$, $(Leu)(PMMA_7H_3^+)$, $(Leu)(PMMA_7H_4^+)$, $(Leu)(PMMA_7H_5^+)$, $(Leu)(PMMA_7H_6^+)$, $(Leu)(PMMA_7H_7^+)$
$(PMMA_7)(Phe)(H^+)$	$(PMMA_7)(PheH^+)$, $(Phe)(PMMA_7H_1^+)$, $(Phe)(PMMA_7H_2^+)$, $(Phe)(PMMA_7H_3^+)$, $(Phe)(PMMA_7H_4^+)$, $(Phe)(PMMA_7H_5^+)$, $(Phe)(PMMA_7H_6^+)$, $(Phe)(PMMA_7H_7^+)$
$(PMMA_{13})(Gly)(H^+)$	$(PMMA_{13})(GlyH^+)$, $(Gly)(PMMA_{13}H_7^+)$, $(Gly)(PMMA_{13}H_{13}^+)$
$(PMMA_{13})(Leu)(H^+)$	$(PMMA_{13})(LeuH^+)$, $(Leu)(PMMA_{13}H_7^+)$, $(Leu)(PMMA_{13}H_{13}^+)$
$(PMMA_{13})(Phe)(H^+)$	$(PMMA_{13})(PheH^+)$, $(Phe)(PMMA_{13}H_7^+)$, $(Phe)(PMMA_{13}H_{13}^+)$

[†] The subscript on H for the oligomer indicates the position of the protonated carbonyl function. On an amino acid, the protonation was on the amine.

For each system, the charges were assigned using the AM1-BCC model [149, 150]. Each MD-SA simulation was initiated by a pre-minimization phase of 1500 steps of steepest descent followed by 1500 steps of adapted-basis Newton-Raphson (ABNR). The system was subsequently heated from 0 to 300 K at a rate of 200 K ps^{-1} and then equilibrated for 2 ps at 300 K using 1 fs time-step. For complexes where the protonation site was on the amino acid, a distance constraint was gradually applied between the centers of both chemical species whereas for complexes where the protonation site was on the oligomer, two distance constraints were gradually applied in order to maintain the entities within a 10 Å radius of each other. These constraints were used to prevent both entities of the complex to drift away from each other when subjected to the mid-cycle annealing schedule in the following step of the simulation. For systems composed of a single molecule, no constraint was applied. Although inspired from previous publications, the simulated annealing schedule had to be adapted to the present systems. Three aspects were taken into consideration: (i) the damping temperature, (ii) the mid-cycle temperature, and (iii) the total number of simulated annealing cycles.

For the damping temperature optimization, each PMMA/amino acid complex was subjected to two different cooling schedules: (1) 100 K ps^{-1} and (2) 10 K ps^{-1}. The longer oligomers were subjected to the slower damping conditions since they need more time to sample properly the PES. The kinetic energy provided to the system is directly proportional to the temperature, thus, longer oligomers may require higher mid-cycle annealing temperatures to overcome local minima on the potential energy surface. Three different higher mid-cycle temperatures were considered: 700 K, 900 K, and 1100 K. Once a consensus was reached for given systems on both the damping schedule

and the mid-cycle temperature, the whole simulations were set for 1500 or 2000 simulated annealing cycles.

In all cases, the annealing schedule was first begun by a drastic heating phase at a rate of 200 K ps^{-1} using time steps of 0.5 fs. After spending 2 ps at the mid-cycle temperature, the system was cooled down to 300 K respecting the damping schedule mnetioned above. The lowest energy conformation was selected at the end of the annealing and cooled down to 0 K. A post-minimization phase of 1500 steps of SD followed by 1500 steps of ABNR was applied to the 0 K conformation, which was subsequently saved. As each annealing cycle was completed, a low energy conformation was saved. Once the whole simulation was terminated after the total number of SA cycles, a 50 ps standard MD simulation at 300 K and without distance constraint was performed. The parameters settings for this step are the same (except for the number of steps) as the equilibration phase. This step was added to the procedure to ensure that there were no anomalies with the complex formed. For each system, the lowest energy conformer generated during the whole simulation was then selected for studying the dissociation pathways. Tables 3.9 to 3.12 report the parameter settings for each step of the MD-SA cycle.

Table 3.9: Parameter settings for the heating phase (0 K to 300 K).

Parameter	Setting
Number of MD steps [nstlim]	2000
Time step [dt] (ps)	0.001
NTF	1 (complete interaction is calculated)
NTT	3 (using Langevin dynamics)
Collision frequency [gamma_ln] (ps^{-1})	1.0 (looser temperature control)
Initial temperature [tempi] (K)	0.0
Final temperature [temp0] (K)	300.0
Nonbonded cutoff (Å)	None

Table 3.10: Parameter settings for the equilibration phase at 300 K.

Parameter	Setting
Number of MD steps [nstlim]	3000
Time step [dt] (ps)	0.001
NTF	1 (complete interaction is calculated)
NTT	3 (using Langevin dynamics)
Collision frequency [gamma_ln] (ps^{-1})	5.0 (tighter temperature control)
Initial temperature [tempi] (K)	300.0
Final temperature [temp0] (K)	300.0
Nonbonded cutoff (Å)	None

Table 3.11: Parameter settings for the simulated annealing phase.[†]

Parameter	Setting
Number of MD steps [nstlim]	18000
Time step [dt] (ps)	0.0005
NTF	1 (complete interaction is calculated)
NTT	3 (using Langevin dynamics)
Collision frequency [gamma_ln] (ps^{-1})	1.0 (tighter temperature control)
Initial temperature [tempi] (K)	300.0
Mid-cycle temperature (K)	900.0*
Final temperature [temp0] (K)	300.0
Heating rate (K ps^{-1})	200
Number of MD steps at mid-cycle temperature	3000
Cooling rate (K ps^{-1})	100*
Number of MD equilibration steps in between cooling phase steps	2000
Nonbonded cutoff (Å)	None

[†]Example provided with a damping schedule of 100 K·ps^{-1} and a mid-cycle temperature of 900 K. The parameters indicated with an (*) were modified for different SA schedule.

Table 3.12: Parameter settings for the cooling phase (300 K to 0 K).

Parameter	**Setting**
Number of MD steps [nstlim]	2000
Time step [dt] (ps)	0.001
NTF	1 (complete interaction is calculated)
NTT	3 (using Langevin dynamics)
Collision frequency [gamma_ln] (ps^{-1})	1.0 (looser temperature control)
Initial temperature [tempi] (K)	300.0
Final temperature [temp0] (K)	0.0
Nonbonded cutoff (Å)	None

The lowest energy conformations of amino acids and 4mer systems generated by MD-SA simulations were further geometry optimized, using AM1 [11]. Semi-empirical geometry optimizations were performed using the Gaussian 03 suite of programs [165]. Calculation of the vibrational harmonic frequencies for each system was obtained from the AM1 geometry optimizations (for amino acids and 4mer systems) or from Nmode inside the Amber 9 suite of programs (for 7mer and 13mer systems)

3.2.2 RRKM modeling of breakdown diagrams

As mentioned previously, the vibrational harmonic frequencies of the longer oligomer systems were obtained using Nmode. Since, no modeling information was available on the Aβ1-40/LMW molecules complexes, another strategy was applied. To apply the RRKM theory, the vibrational harmonic frequencies of the Aβ1-40 peptide were taken from the method of Moon *et al.* [166]. In their paper, they found that as long as $\Delta^{\ddagger}S$ remains the same, RRKM results for E_0 are quite insensitive to the specific set of frequencies determined at the B3LYP/6-31G** level of theory. This behaviour is well known for small molecules. In the present study, since the small molecules are very similar in structure, differences in $\Delta^{\ddagger}S$ should be small as well. For LMW molecules, geometry optimizations followed by frequency calculations were performed at the AM1 level of theory using the Gaussian 03 suite of programs. As recommended by Scott and Radom [167], since vibrational harmonic frequencies were acquired using two different levels of theory, those of the peptide were scaled by a factor of 0.9614, whereas those of the LMW molecules were scaled by a factor of 0.9532. Frequencies of both the peptide and the LMW molecules were combined together for each complex undergoing gas phase dissociation. Missing in the current model is any

discussion of the actual conformation of the peptide/LMW molecule complexes and the effect that it would have on the vibrational frequencies. Since each of the five small molecules tested are chemically similar, there should not be a significant conformation difference between the five complexes. In addition, since the transition state vibrational frequencies are being adjusted to fit the experimental data, the actual choice of vibrational frequencies should not be of great importance. The fact that the approximate Hill equation, extrapolations for k(E) outline later appear to work, suggests this to be true. In the end, since it is the relative binding energies of these species that should be the most reliable piece of information, the current choice of vibrational frequencies should be sufficient.

The simple procedure used to model the experimental breakdown curves is now described. The experimental observations are dependent on (i) the time-scale of the dissociation events in the collision cell, (ii) the complex internal energy distribution after collision, which in turn will depend on the centre-of-mass collision energy and the gas density in the cell, and (iii) the microcanonical rate constant, k(E), for the dissociation of the complex. Since the gas density is held constant, and the complexes are expected to have similar collision cross-sections, this parameter was left out of the modeling. Thus, what follows is explicitly for the gas pressure described in the experimental section. Changing gas pressure will affect the value of the α parameter described below. In addition, ion scattering has been found by Knyazev and Stein [141, 142] to make a significant contribution to the parent vs. product breakdown curves in CID. They observed from their simulations that product ions are scattered more than parent ions due to kinetic energy release contributing to their velocities. In our study, the kinetic energy released into the large product ion is only ~5 % of the total (since on average the product ion is ~95 % of the parent ion mass), which

should mean scattering losses are similar for both product and parent. This would not be the case if we were fitting the small sodiated substrate product ion present in most systems. Thus, we have chosen to ignore scattering as a significant contributor to our breakdown curves.

The time scale of events in the collision cell is determined by the length of the collision cell (10 cm) and the laboratory frame translational energy of the ion. This latter quantity is a function of the voltage change leading into the collision cell (E_{LAB}) and the charge on the ion. As an example, for a charge state of z, the translational energy (T) is given by:

$$T = z \cdot E_{LAB} \tag{3.3}$$

The time-scale of the observations, t, is then solved with $T = \frac{1}{2} m_{ion} v^2$, where m_{ion} is the molar mass of the complex and v is the ion's velocity derived from T. The post-collision internal energy distribution is a function of the centre-of-mass collision energy as opposed to the laboratory frame collision energy, which involves the usual conversion:

$$E_{CM} = E_{lab} \left(\frac{m_{Ar}}{m_{Ar} + m_{ion}} \right) \tag{3.4}$$

where: m_{Ar} is the mass of the argon collision gas used in all of the experiments.

As the E_{CM} changes, so will the resulting internal energy distribution. There have been a number of studies that have explored this relationship under multiple collision conditions in a quadrupole collision cell [168-171]. For simplicity, it was chosen to reflect the post-collision internal energy distribution of the complex ions with an effective temperature, T_{eff}, through the equation:

$$T_{eff} = 300 + \alpha E_{CM} \quad (3.5)$$

where: 300 reflects a pre-collision internal temperature of 300 K and α (which has units K eV^{-1}) is an adjustable parameter reflecting the effect of a given collision energy will have upon the effective internal energy distribution.

While there is no evidence for a linear scaling of the effective temperature with E_{CM}, due to (a) the small absolute range of E_{CM} exhibited in the breakdown diagrams, (b) the fact that all of the systems studied dissociated over the same range of E_{CM} and (c) the fact that all of the species will be of similar degrees of freedom, this assumption should not affect the relative binding energies obtained from the modeling. The resulting internal energy distribution $P(E,E_{CM})$ would be calculated according to the equation:

$$P(E, E_{CM}) = \frac{\rho(E) e^{-E/RT_{eff}}}{Q(E_{CM})} \quad (3.6)$$

where: $\rho(E)$ is the vibrational density of states of the ion (in our case calculated with the direct-count algorithm of Beyer and Swinehart [172]), and $Q(E_{CM})$ the total vibrational partition function at the temperature assigned to a particular E_{CM}.

Knyazev and Stein [140, 141] have recently demonstrated that post-collision ion internal energy distributions can be non-Boltzmann in character. In their reported case for the small

benzylammonium ion, the P(E) of the ions at E_{lab} = 10 eV (E_{CM} = 2.7 eV) is definitely not Boltzmann, and appears more like a Boltzmann distribution moved up along the internal energy axis (Figure 6 in ref. [141]). The discrepancy between the modeled distribution and a thermal distribution increased with increasing collision energy. This would clearly draw into doubt our use of a thermal distribution to approximate the P(E) of the peptide/substrate complexes. However, in the present study, due to the size of the peptide E_{CM} is much lower (< 1.0 eV). If agreement between the actual P(E) and a thermal P(E) decreases as E_{CM} increases, it is arguable that in the present system with its very low E_{CM}, using a thermal distribution to approximate P(E) will be better than it was for the small ions studied by Knyazev and Stein.

At ion temperatures above 300 K, the population distribution $P(E,E_{CM})$ can have significant values up to an internal energy of 35 eV. Although the polymers systems deal tractable values of ρ(E), the Aβ1-40/LMW molecule complexes investigated have over 1800 vibrational modes and at internal energies greater than ~9 eV, ρ(E) takes on values $>10^{308}$, the double precision floating point limit for 64-bit compilers. Therefore, the exact RRKM solution can only be run on 128-bit processors. The microcanonical rate constant k(E) was calculated according to the usual RRKM expression [85, 86]:

$$k(E) = \frac{\sigma N^{\ddagger}(E-E_0)}{h\rho(E)} \tag{3.7}$$

Vibrational densities and sums of states were obtained with the direct count algorithm. Rotations were omitted. The transition state vibrational frequencies were estimated by taking those of the complex, scaling the lowest six by a common factor (thereby adjusting $\Delta^{\ddagger}S$) and one mode was

removed to reflect the reaction coordinate. The fraction of ions dissociating, F, as a function of internal energy E is given by :

$$F(E) = 1 - e^{-k(E)t} \tag{3.8}$$

where: k(E) is the microcanonical rate constant and t is the maximum time of the observation.

The total observation is then given by multiplying F(E) and $P(E,E_{CM})$ and integrating over internal energy. For an ion undergoing a single dissociation channel, this provides the theoretical breakdown diagram. There are three adjustable parameters in the above model, E_0, $\Delta^{\ddagger}S$ (by way of the transition state scaling factor) and, α. Fitting of the experimental breakdown curves was accomplished by obtaining theoretical curves for a systematic range of each variable. For each theoretical curve, the absolute deviation with the experimental data was calculated at each E_{CM} value and a score assigned. To find the best fit, approximately three to four thousand theoretical curves were calculated and compared to the experimental data. The three adjustable parameters were fine-tuned in order to minimize the score for the fit. The error bars on the derived values for E_0, $\Delta^{\ddagger}S$ and α reflect the resulting possible best fits to the data only. They obviously do not reflect the inherent assumptions built into the model, which will ultimately be responsible for the accuracy of the values.

Chapter 4: Energy and entropy effects in the dissociation of $(PMMA_n)(AA)(H^+)$ complexes

4.1 Introduction

The scope of this chapter is to study the competing dissociation pathways of different PMMA oligomers complexed with three amino acids having different proton affinity (PA) (glycine: 887 kJ mol^{-1}, leucine: 915 kJ mol^{-1}, phenylalanine: 993 kJ mol^{-1}). In the $PMMA_n$/AA complexes studied here, two dissociation channels were competing:

$$PMMA_nH^+ + AA \xleftarrow{k_{PMMA_nH^+}} (PMMA_n)(AA)(H^+) \xrightarrow{k_{AAH^+}} PMMA_n + AAH^+ \qquad (4.1)$$

One possible dissociation channel is an oligomer neutral loss and protonated amino acid while the other dissociation channel is a protonated oligomer and an amino acid neutral loss.

4.2 The effect of the amino acid PA on the dissociation of $(PMMA_n)(AA)(H^+)$

Since it was hypothesized that the PA of an amino acid (AA) may play a role in the dissociation of $(PMMA)(AA)(H^+)$ complexes, CID mass spectrometry experiments were performed on the complexes formed with three PMMA oligomers and each of the three amino acids.

Figure 4.1 shows three CID mass spectra for a $(PMMA_{10})(Gly)(H^+)$ complex at 25 and 35 eV respectively. At 25 eV, there is only a peak at m/z 1078 due to the presence of the parent ion. There

was not enough energy for any fragmentation to occur. There is no peak showing the presence of the charged glycine or oligomer. When the voltage was increased to 35 eV, a certain fragmentation of this complex appeared at m/z 971, indicating the oligomer $(PMMA_{10})(H^+)$ had lost a neutral methanol fragment. The fragmentation mechanism for this methanol loss is explained in reference [26]. At threshold, for the $(PMMA_{10})(Gly)(H^+)$ complex, it seems that the dissociation channel preferred has a protonated oligomer and a neutral amino acid loss.

$$(PMMA_{10})(Gly)(H^+) \rightarrow [\ (PMMA_{10}H^+)\ - CH_3OH\ ] + Gly \qquad (4.2)$$

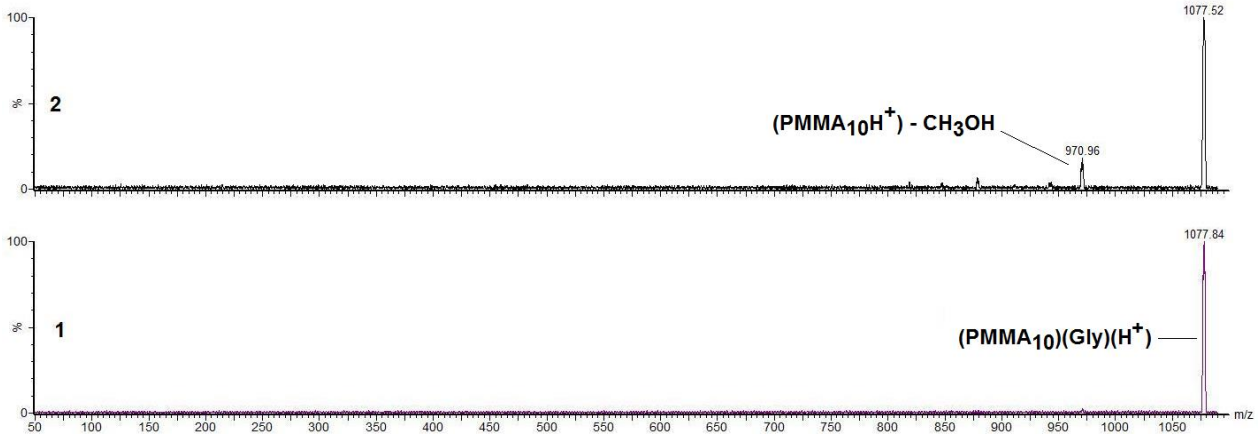


Figure 4.1: CID mass spectra of a $(PMMA_{10})(Gly)(H^+)$ complex at collision energy 25 eV (1) and 35 eV (2). The CID mass spectrum (2) explicitly shows the initial appearance of protonated oligomer less a methanol loss.

Figure 4.2 shows three CID mass spectra for a $(PMMA_{10})(Leu)(H^+)$ complex at respectively 20 and 35 eV. At 20 eV, there is only a peak at m/z 1134 due to the presence of the parent ion. There is no peak showing the presence of the charged leucine or oligomer. When the voltage was increased to 35 eV, the fragment that appeared at m/z 131 indicates the presence of the protonated

leucine. At threshold, for the $(PMMA_{10})(Leu)(H)^+$ system, it seems that the dissociation channel preferred has a protonated amino acid and a neutral oligomer loss (although at higher CE, the second channel begin to appear).

$$(PMMA_{10})(Leu)(H^+) \rightarrow \quad PMMA_{10} + LeuH^+ \qquad (4.3)$$

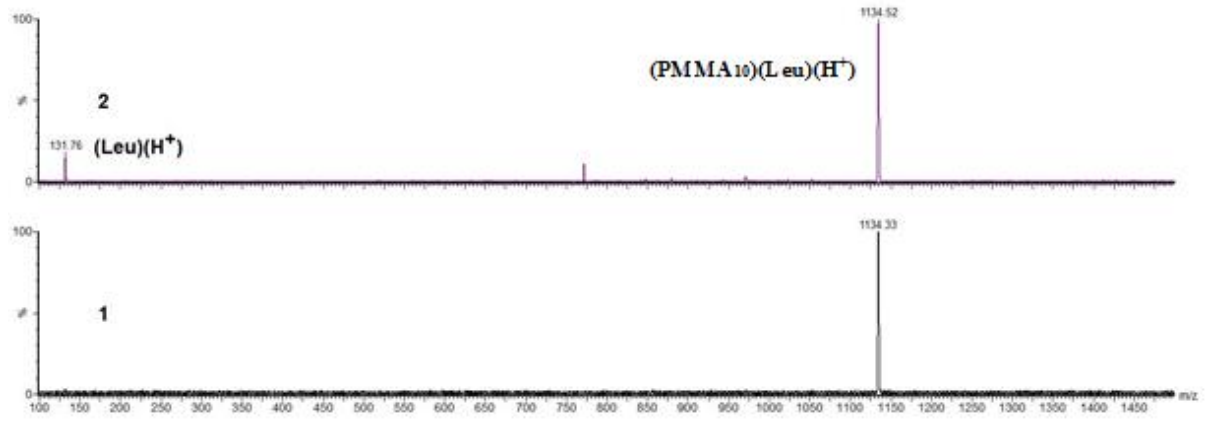


Figure 4.2: CID mass spectra of a $(PMMA_{10})(Leu)(H^+)$ complex at collision energy 20 eV (1) and 35 eV (2). The CID mass spectrum (2) explicitly shows the initial appearance of protonated amino acid.

Figure 4.3 shows three CID mass spectra for a $(PMMA_{10})(Phe)(H^+)$ complex at 15 and 25 eV respectively. At 15 eV, there is only a peak at m/z 1168 due to the presence of the parent ion. There is no peak showing the presence of the charged phenylalanine or oligomer. When the voltage was increased to 25 eV, the fragment that appeared at m/z 166 indicates the presence of the protonated phenylalanine. At threshold, for the $(PMMA_{10})(Phe)(H^+)$ complex, it seems that the dissociation channel preferred has a protonated amino acid and a neutral oligomer loss (although at higher CE, the second channel begin to appear).

$$(PMMA_{10})(Phe)(H^+) \rightarrow \quad PMMA_{10} + PheH^+ \qquad (4.4)$$

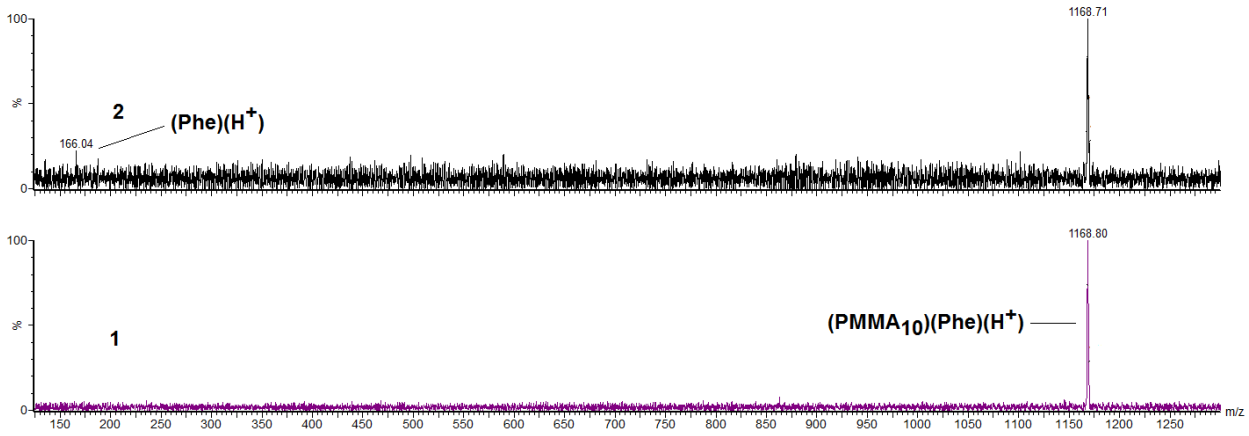


Figure 4.3: CID mass spectra of a $(PMMA_{10})(Phe)(H^+)$ complex at collision energy 15 eV (1) and 25 eV (2). The CID mass spectrum (2) explicitly shows the initial appearance of protonated amino acid.

In summary, the protonated amino acid dissociation channel is preferred when a leucine or a phenylalanine is involved, whereas, the protonated oligomer dissociation channel is preferred when a glycine is involved. These results are coherent with the PA of the three amino acids stated initially: amino acids having a higher PA dissociate with the proton. One question at this point is: does the PA of the amino acid always dictate the preferred dissociation channel? To try answering this question, a series of amino acids were selected in addition to the three original (arginine, lysine, tryptophan, and alanine) and a series of PMMA oligomers of different length were selected (3mer, 4mer, 7mer, 13mer, and 17mer). Solutions of each amino acid with each oligomer were prepared in a 1:1 ratio and were subjected to CID mass spectrometry at threshold collision energy. Table 4.1 shows the results obtained in our group [173].

Table 4.1: Fragment ion observed at threshold in CID mass spectra of $(PMMA_n)(AA)(H^+)$ complexes (A = AAH^+, P = $PMMA_nH^+$).

	Observed Threshold Fragment Ion†						
oligomer	Arg (1051)	Lys (996)	Trp (949)	Phe (923)	Leu (915)	Ala (902)	Gly (887)
3mer	A	A	A	A	A/P	P	P
4mer	A	A	A	A	A/P	P	P
7mer	A	A	A	A	A/P	P	P
13mer	A	A	A/P	P	P	P	P
17mer	A	A	P	P	P	P	P

†Amino acid PA in parentheses (kJ mol^{-1}) [35].

From the results obtained, arginine and lysine, the two amino acids having the highest PA, always dissociate protonated, whereas alanine and glycine, the two amino acids having the lowest PA, never dissociate protonated. For the phenylalanine, when the oligomer is shorter (3mer, 4mer, 7mer) the amino acid dissociates protonated, but when the oligomer is longer (13mer, 17mer), it is the oligomer that dissociates protonated. There seems to be a competition between the PA of the amino acid and the length of the oligomer. For the tryptophan and the leucine, there are cases where both dissociation channels occur simultaneously. Again, for these two amino acids, the longer oligomers seem to be able to dissociate with the proton. For the leucine, the dissociation channel of the protonated amino acid is never favoured, as it is always competing with the protonated oligomer for the shorter ones (3mer, 4mer, 7mer).

4.3 RRKM modeling of $(PMMA_n)(AA)(H^+)$ complexes

To investigate further these tendencies, the three initial amino acids (glycine, leucine, and phenylalanine) were selected with three oligomers (7mer, 10mer, and 13mer) to form the different gas phase complexes. CID experiments were performed on each of them to obtain breakdown diagrams for all nine systems. RRKM modeling was performed afterwards to determine E_0 and $\Delta^{\ddagger}S$ of each dissociation channel.

4.3.1 Dissociation of the $(PMMA_7)(AA)(H^+)$ complexes

The experimental breakdown curves and their corresponding theoretical fits for the $(PMMA_7)(AA)(H^+)$ complexes dissociation are shown in Figure 4.4. The corresponding E_0 and $\Delta^{\ddagger}S$ are reported in Table 4.2. The dissociation of the $(PMMA_7)(AA)(H^+)$ complex (Figure 4.4a) led to a single dissociation channel corresponding to the appearance of the protonated oligomer (4.5). The glycine dissociated as a neutral loss and did not appear on the mass spectrum. As shown in Table 4.1, this behaviour was somehow expected since the glycine is the amino acid with the lowest proton affinity. The RRKM fitting of the breakdown curve yields an $E_0 = 0.83 \pm 0.08$ eV and a $\Delta^{\ddagger}S = -14 \pm 2$ J mol^{-1} K^{-1}.

$$(PMMA_7)(Gly)(H^+) \xrightarrow{k_{PMMA_7H^+}} PMMA_7H^+ + Gly \qquad (4.5)$$

The dissociation of the $(PMMA_7)(Leu)(H^+)$ (Figure 4.4b) led to two competitive dissociation channels. One in which the oligomer is protonated and the amino acid is a neutral loss ($E_0 = 0.83 \pm 0.08$ eV, $\Delta^{\ddagger}S = -14 \pm 2$ J mol^{-1} K^{-1}) and the second in which the amino acid is protonated and the

oligomer is neutral (E_0 = 0.91 ± 0.08 eV, $\Delta^{\ddagger}S$ = -3 ± 2 J mol^{-1} K^{-1}). The higher proton affinity of the leucine compared to the glycine allowed a second dissociation channel to occur.

$$PMMA_7H^+ + Leu \xleftarrow{k_{PMMA_7H^+}} (PMMA_7)(Leu)(H^+) \xrightarrow{k_{LeuH^+}} PMMA_7 + LeuH^+ \qquad (4.6)$$

Similarly, the dissociation of the $(PMMA_7)(Phe)(H^+)$ (Figure 4.4c) also led to two competitive dissociation channels. The fitting of the channel leading the protonated oligomer yields an E_0 = 0.81 ± 0.08 eV and a $\Delta^{\ddagger}S$ = -17 ± 2 J mol^{-1} K^{-1} whereas the channel leading to a protonated phenylalanine was found to have an E_0 = 0.90 ± 0.08 eV and a $\Delta^{\ddagger}S$ = -5 ± 2 J mol^{-1} K^{-1}.

$$PMMA_7H^+ + Phe \xleftarrow{k_{PMMA_7H^+}} (PMMA_7)(Phe)(H^+) \xrightarrow{k_{PheH^+}} PMMA_7 + PheH^+ \qquad (4.7)$$

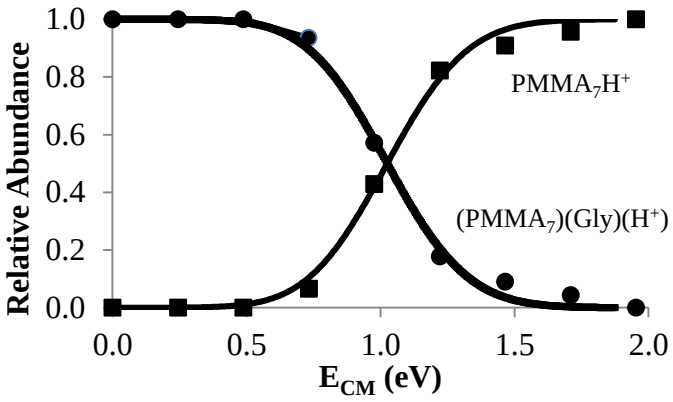
Reasonable error limits were determined based on a range of visually acceptable fits (Table 4.2). These error limits relate to the fitting of the experimental data and the model used in the fitting but not the assumption used in the RRKM theory itself. The scaling factor α was kept at 200 for all complexes, which is consistent with reported values for similar systems [174]. It is not possible to assess that these thermodynamic parameters are absolute since the interplay of E_0 and $\Delta^{\ddagger}S$ are complementary on the ion k(E) during the fitting process. However, when taken together, E_0 for all three $PMMA_7H^+$ channels range from 0.81 to 0.86 eV whereas the two AAH^+ channels are slightly higher at 0.90 and 0.91 eV. This suggests that the dissociation channels leading to $PMMA_7H^+$ are more energetically favourable than those leading to the $PMMA_7$ as a neutral loss. However, it is the opposite in terms of the entropy of activation. Upon dissociation, the AAH^+

channels have a much higher entropy (-3 to -4 J mol^{-1} K^{-1}) than the channels in which the AA is a neutral loss (-14 to -17 J mol^{-1} K^{-1}). Therefore, the AAH^+ dissociation channel is entropically more favourable over the AA dissociation channel. In addition to the fact that leucine and phenylalanine have higher PA than glycine, these greater values in $\Delta^{\ddagger}S$ also contribute to the dissociation channels involving the protonated amino acids.

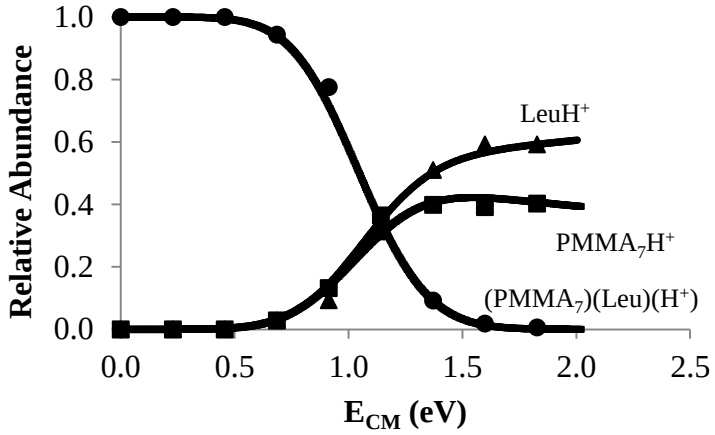
Table 4.2: E_0 and $\Delta^{\ddagger}S$ for the $(PMMA_7)(AA)(H^+)$ complexes dissociations.

Complex dissociated	Fragment ion dissociation channel	E_0 (± 0.08 eV)	$\Delta^{\ddagger}S$ (± 2 J mol^{-1} K^{-1})	α
$(PMMA_7)(Gly)(H^+)$	$PMMA_7H^+$	0.83	-14	200
$(PMMA_7)(Leu)(H^+)$	$LeuH^+$	0.91	-3	200
	$PMMA_7H^+$	0.86	-14	200
$(PMMA_7)(Phe)(H^+)$	$PheH^+$	0.90	-5	200
	$PMMA_7H^+$	0.81	-17	200

(a)



(b)



(c)

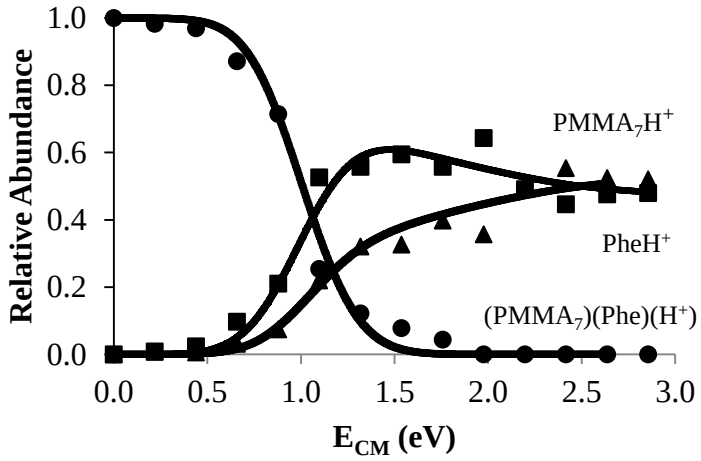


Figure 4.4: (a) Experimental breakdown curves (markers) and theoretical fits (lines) of the $(PMMA_7)(Gly)(H^+)$ complex. (b) Experimental breakdown curves (markers) and theoretical fits (lines) of the $(PMMA_7)(Leu)(H^+)$ complex. (c) Experimental breakdown curves (markers) and theoretical fits (lines) of the $(PMMA_7)(Phe)(H^+)$ complex.

4.3.2 Dissociation of the $(PMMA_{10})(AA)(H^+)$ complexes

The experimental breakdown curves and their corresponding theoretical fits for the $(PMMA_{10})(AA)(H^+)$ complexes dissociation are shown in Figure 4.5. The corresponding E_0 and $\Delta^{\ddagger}S$ are reported in Table 4.3. The dissociation of the $(PMMA_{10})(Gly)(H^+)$ (Figure 2a) led to a single dissociation channel in which the glycine is the neutral loss and the polymer is protonated. The RRKM fitting of the breakdown curve yields an E_0 = 0.96 ± 0.08 eV and a $\Delta^{\ddagger}S$ = -11 ± 2 J mol^{-1} K^{-1}. As opposed to the $(PMMA_7)(Gly)(H^+)$ dissociation, the longer oligomer did increase the activation energy by 0.13 eV. This was accompanied by a small increase in the entropy of activation of 3 ± 2 J mol^{-1} K^{-1}.

$$(PMMA_{10})(Gly)(H^+) \xrightarrow{k_{PMMA_{10}H^+}} PMMA_{10}H^+ + Gly \qquad (4.8)$$

The dissociation of the $(PMMA_{10})(Leu)(H^+)$ (Figure 2b) led to two competitive dissociation channels. One in which the polymer is protonated and the amino acid is a neutral loss (E_0 = 0.95 ± 0.08 eV, $\Delta^{\ddagger}S$ = -9 ± 2 J mol^{-1} K^{-1}) and the second in which the amino acid is protonated and the polymer is a neutral loss (E_0 = 0.92 ± 0.08 eV, $\Delta^{\ddagger}S$ = -6 ± 2 J mol^{-1} K^{-1}).

$$PMMA_{10}H^+ + Leu \xleftarrow{k_{PMMA_{10}H^+}} (PMMA_{10})(Leu)(H^+) \xrightarrow{k_{LeuH^+}} PMMA_{10} + LeuH^+ \qquad (4.9)$$

As with the $PMMA_7$ olifomer, the dissociation of the $(PMMA_{10})(Phe)(H^+)$ (Figure 2c) also led to two competitive dissociation channels. The RRKM fitting of the channel leading the protonated

polymer yields an E_0 = 0.85 eV and a $\Delta^{\ddagger}S$ = -5 ± 2 J mol^{-1} K^{-1} whereas the channel leading to a protonated phenylalanine was found to have an E_0 = 0.83 ± 0.08 eV and a $\Delta^{\ddagger}S$ = 2 ± 2 J mol^{-1} K^{-1}.

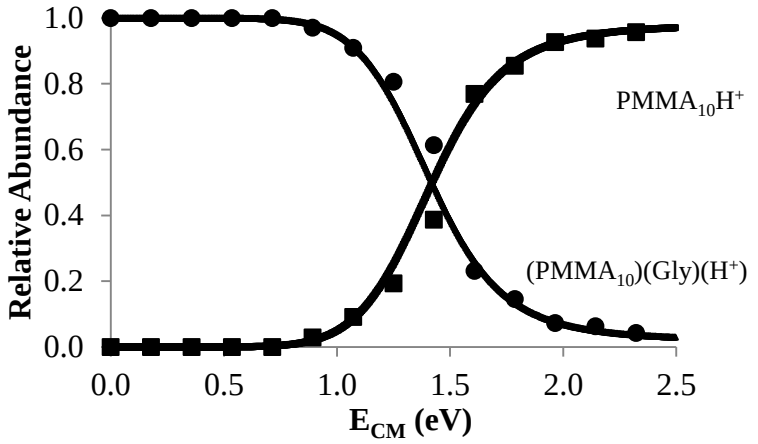
$$PMMA_{10}H^+ + Phe \xleftarrow{k_{PMMA_{10}H^+}} (PMMA_{10})(Phe)(H^+) \xrightarrow{k_{PheH^+}} PMMA_{10} + PheH^+ \qquad (4.10)$$

As for the $(PMMA_7)(AA)(H^+)$ RRKM modeling, the scaling factor α was set at 200. As a general tendency, it is observed that the $PMMA_{10}H^+$ channels have higher E_0 and $\Delta^{\ddagger}S$ than the corresponding $PMMA_7H^+$ channels. As opposed to the $(PMMA_7)(Phe)(H^+)$ dissociation, the $PMMA_{10}H^+$ channel is energetically more favourable over the two other 10mer channels. As opposed to the $LeuH^+$ channel coming from the $(PMMA_7)(Leu)(H^+)$ dissociation, the $LeuH^+$ channel of the $(PMMA_{10})(Leu)(H^+)$ has a lower $\Delta^{\ddagger}S$. In addition, the channel leading to the dissociation of the $PheH^+$ fragment even has a positive $\Delta^{\ddagger}S$.

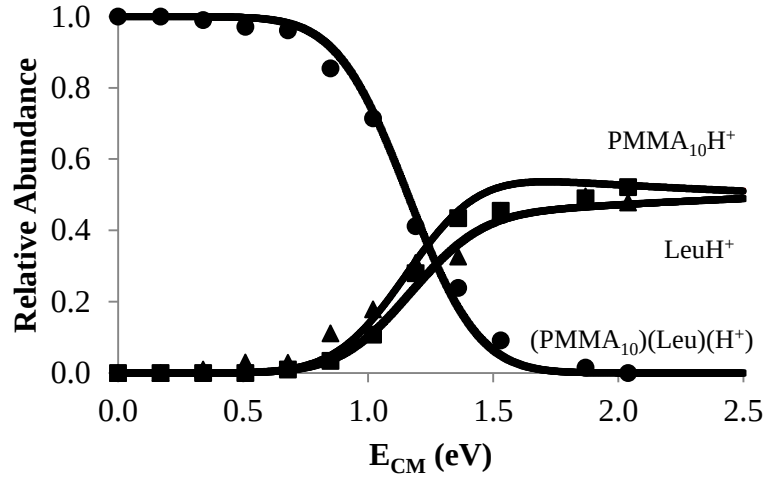
Table 4.3: E_0 and $\Delta^{\ddagger}S$ for the $(PMMA_{10})(AA)(H^+)$ complexes dissociations.

Complex dissociated	Fragment ion dissociation channel	E_0 (± 0.08 eV)	$\Delta^{\ddagger}S$ (± 2 J mol^{-1} K^{-1})	α
$(PMMA_{10})(Gly)(H^+)$	$PMMA_{10}H^+$	0.96	-11	200
$(PMMA_{10})(Leu)(H^+)$	$LeuH^+$	0.92	-6	200
	$PMMA_{10}H^+$	0.95	-9	200
$(PMMA_{10})(Phe)(H^+)$	$PheH^+$	0.83	2	200
	$PMMA_{10}H^+$	0.85	-5	200

(a)



(b)



(c)

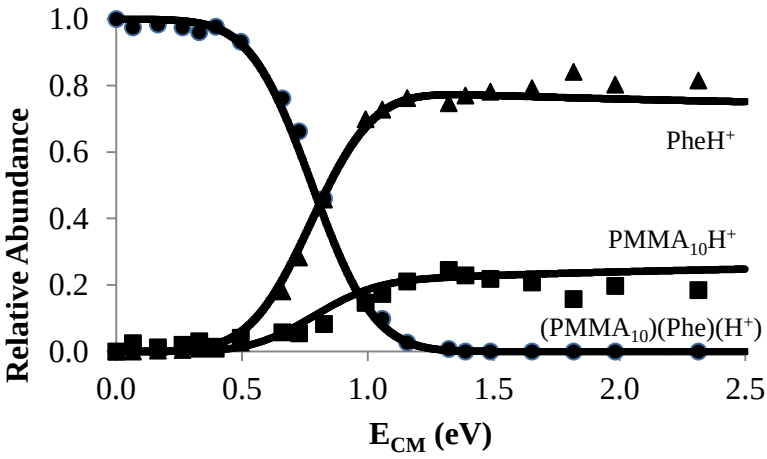


Figure 4.5: (a) Experimental breakdown curves (markers) and theoretical fits (lines) of the $(PMMA_{10})(Gly)(H^+)$ complex. (b) Experimental breakdown curves (markers) and theoretical fits (lines) of the $(PMMA_{10})(Leu)(H^+)$ complex. (c) Experimental breakdown curves (markers) and theoretical fits (lines) of the $(PMMA_{10})(Phe)(H^+)$ complex.

4.3.3 Dissociation of the $(PMMA_{13})(AA)(H^+)$ complexes

The experimental breakdown curves and their corresponding theoretical fits for the $(PMMA_{13})(AA)(H^+)$ complexes dissociation are shown in Figure 4.6. The corresponding E_0 and $\Delta^{\ddagger}S$ are reported in Table 4.4. The dissociation of the $(PMMA_{13})(Gly)(H^+)$ (Figure 4.6a) led to a single dissociation channel in which the glycine is the neutral loss and the polymer is protonated. The RRKM fitting of the breakdown curve yields an E_0 = 1.30 ± 0.08 eV and a $\Delta^{\ddagger}S$ = 5 ± 2 J mol^{-1} K^{-1}. As compared to both $(PMMA_7)(Gly)(H)^+$ and $(PMMA_{10})(Gly)(H^+)$ dissociations, the longer oligomer did significantly increase the activation energy and, to a lesser extent, the entropy of activation.

$$(PMMA_{13})(Gly)(H^+) \xrightarrow{k_{PMMA_{13}H^+}} PMMA_{13}H^+ + Gly \qquad (4.11)$$

The dissociation of the $(PMMA_{13})(Leu)(H^+)$ (Figure 4.6b) led to two competing dissociation channels: one in which the polymer is protonated and the amino acid is a neutral loss (E_0 = 1.29 ± 0.08 eV, $\Delta^{\ddagger}S$ = 4 ± 2 J mol^{-1} K^{-1}) and the second in which the amino acid is protonated and the polymer is neutral (E_0 = 1.27 ± 0.08 eV, $\Delta^{\ddagger}S$ = 3 ± 2 J mol^{-1} K^{-1}). Here again, as compared to both $(PMMA_7)(Leu)(H^+)$ and $(PMMA_{10})(Leu)(H^+)$ complexes dissociations, the longer oligomer did significantly increase the activation energy and, to a lesser extent, the entropy of activation.

$$PMMA_{13}H^+ + Leu \xleftarrow{k_{PMMA_{13}H^+}} (PMMA_{13})(Leu)(H^+) \xrightarrow{k_{LeuH^+}} PMMA_{13} + LeuH^+ \qquad (4.12)$$

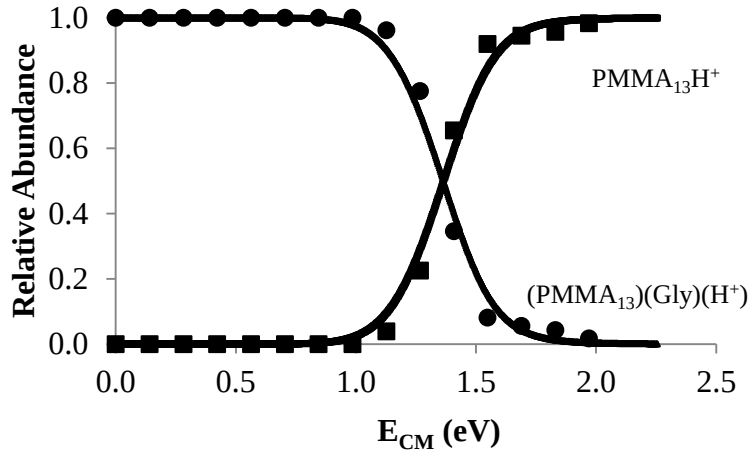
The dissociation of the $(PMMA_{13})(Phe)(H^+)$ (Figure 4.6c) also led to two competing dissociation channels. The RRKM fitting of the channel leading to the protonated polymer yields an E_0 = 1.12 ± 0.08 eV and a $\Delta^{\ddagger}S$ = 6 ± 2 J mol^{-1} K^{-1} whereas the channel leading to a protonated phenylalanine was found to have an E_0 = 1.24 ± 0.08 eV and a $\Delta^{\ddagger}S$ = 8 ± 2 J mol^{-1} K^{-1}. As a general tendency, it is observed that the $PMMA_{13}H^+$ channels have higher E_0 and $\Delta^{\ddagger}S$ as compared with the $PMMA_7H^+$ and $PMMA_{10}H^+$ channels. It is found that the $(PMMA_{13})(Phe)(H^+)$ dissociation led to channels energetically more favourable over the two other 13mer channels. In addition, the $(PMMA_{13})(Phe)(H^+)$ dissociation channels are entropically more favourable over the corresponding channels coming from the $(PMMA_{13})(Leu)(H^+)$.

$$PMMA_{13}H^+ + Phe \xleftarrow{k_{PMMA_{13}H^+}} (PMMA_{13})(Phe)(H^+) \xrightarrow{k_{PheH^+}} PMMA_{13} + PheH^+ \qquad (4.13)$$

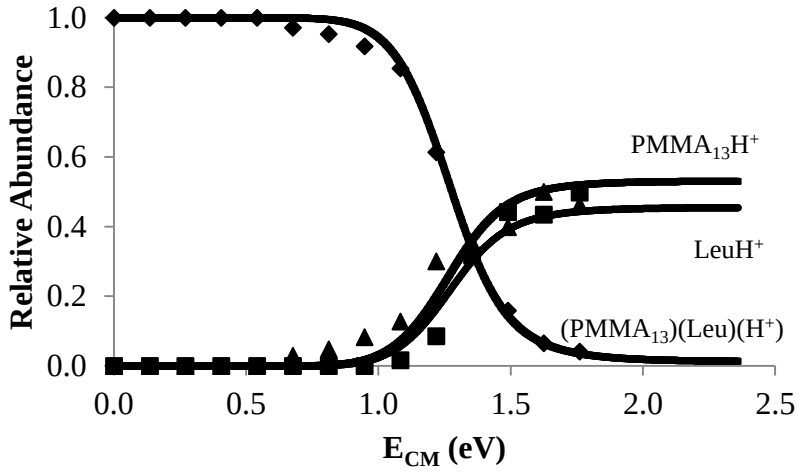
Table 4.4: E_0 and $\Delta^{\ddagger}S$ for the $(PMMA_{13})(AA)(H)^+$ complexes dissociations.

Complex dissociated	Fragment ion dissociation channel	E_0 (± 0.08 eV)	$\Delta^{\ddagger}S$ (± 2 J mol^{-1} K^{-1})	α
$(PMMA_{13})(Gly)(H^+)$	$PMMA_{13}H^+$	1.30	5	300
$(PMMA_{13})(Leu)(H^+)$	$LeuH^+$	1.27	3	300
	$PMMA_{13}H^+$	1.29	4	300
$(PMMA_{13})(Phe)(H^+)$	$PheH^+$	1.24	8	300
	$PMMA_{13}H^+$	1.12	6	300

(a)



(b)



(c)

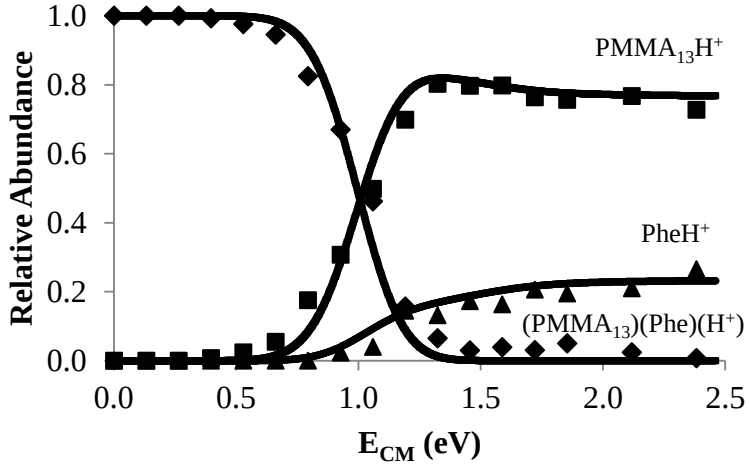


Figure 4.6: (a) Experimental breakdown curves (markers) and theoretical fits (lines) of the $(PMMA_{13})(Gly)(H^+)$ complex. (b) Experimental breakdown curves (markers) and theoretical fits (lines) of the $(PMMA_{13})(Leu)(H^+)$ complex. (c) Experimental breakdown curves (markers) and theoretical fits (lines) of the $(PMMA_{13})(Phe)(H^+)$ complex.

For the RRKM fitting of these ($PMMA_{13}$)(AA)(H^+) complexes, the scaling factor α was set to 300, which is higher than in the other systems involving $PMMA_7$ and $PMMA_{10}$ for which $\alpha = 200$. This is slightly higher value compared to the polymer/substrate systems studied in other projects [174]. The scaling factor had to be increased due the steep dissociation breakdown curves obtained experimentally (Figure 4.6). As mentioned before, MS is a powerful technique to study non-covalent complexes in the gas phase, however, no information is provided on the bound 3D conformation given the available instruments. In the next section, the MD-SA results and 3D conformations of a series of modeled complexes are presented.

4.4 Molecular modeling of the gas phase complexes

The molecular dynamics simulated annealing (MD-SA) conditions were presented in section 3.2.1. In MD, the objective is to find the global minimum of a system, which is the lowest energy conformation possible. In practice, this gets harder as the number of degrees of freedom increase for a given system. Therefore, for complex systems, such as polymers, finding the global minimum can be a tricky task, however, if it is not found, it is certainly hoped to find a representative local minimum that is sufficiently close in energy.

4.4.1 MD-SA parameters optimization

A certain number of parameters had to be optimized in order to ensure a proper sampling of the potential energy surface (PES). The first parameter that required controlling was the mid-cycle temperature at each annealing cycle. In MD, a temperature increase provides energy, which is distributed among the various degrees of freedom of the system. The larger the system, the more

energy is required. The type of interactions existing inside the system is also important, as it takes more energy to break ion-dipole over vdW interactions. A mid-cycle temperature that is too low will not provide enough energy to a system that would be stuck on a slightly deeper local minimum. On the contrary, a mid-cycle temperature that is too high will bring the system in "hectic" energy jumps as well as increased calculation times. Based on previous protocols proposed [162-164], mid-cycle temperatures of 700 K, 900 K, and 1100 K were used for 4mer, 7mer, and 13mer systems respectively. These temperatures proved to be efficient as no significant lower energy minima were found when test runs were performed using higher mid-cycle temperatures.

The second parameter to be controlled was the damping temperature (or the cooling schedule). A drastic cooling schedule is not efficient as it traps the system inside the nearest local minimum available on the PES. A slower cooling schedule allows a more proper sampling of the different available local minima on the PES. Again, if the cooling schedule is too slow, it will increase significantly the calculation time. Two cooling schedules were explored: temperature decrease at 100 K ps^{-1} or 10 K ps^{-1}. For all three systems, the slower damping temperature of 10 K ps^{-1} proved to be more efficient in finding lower energy conformations.

The last parameter was the number of annealing cycles that was chosen to be 1500 or 2000. For the 4mer systems, 1500 were far enough as several low energy conformers were found before the first 1000 annealing cycles. As shown in Table 4.5, in 8 calculations out of 20, the lowest energy conformers were found, after the first 1000 steps. However, the conformations found there were < 1 kcal mol^{-1} compared the lowest energy conformer found in the first 1000 annealing cycles.

Similar results were observed for the 7mer systems as 13 calculations out of 32 had their lowest energy conformation between 1000 and 1500 annealing cycles (Table 4.9). For the 13mer systems, 2000 annealing cycles were preferable, as 8 calculations out of 12 had their lowest energy conformation after the first 1000 annealing cycles (Table 4.10).

For each $PMMA_n$ complex systems, comparing the energy difference between the initial conformation (E_0) and the lowest energy conformation (E_f), shows that the control of these three parameters allowed finding reasonable low energy conformations. There were energy difference (E_f-E_0) averages of -26 kcal mol^{-1}, -38 kcal mol^{-1}, and -59 kcal mol^{-1} for the 4mer, 7mer, and 13mer systems respectively.

4.4.2 Modeling the $PMMA_4$ systems: conformational features and energetics

According to the experimental results, two different dissociation channels were competing (except for the $(PMMA_4)(Gly)(H^+)$ complex which had only one). Therefore, for each system, it was of interest to study different aspects: (i) the protonation site, (ii) the initial complexes, and (iii) the dissociation products. For the protonation site, MD-SA simulations were performed on $PMMA_4H^+$ oligomers and complexes bearing the proton at each carbonyl position. As shown in Table 4.5, the three complexes having the proton located on the amino acid have the lowest energy. This gives an insight information when the complex is bound in the gas phase; the proton is on the amino acid, regardless of its proton affinity. Comparing the same relative energies, it seems that if there should be a proton transfer between the amino acid and the oligomer, the energetically favourable protonation sites would be the carbonyl functions located in the middle of the polymer sequence.

Table 4.5: Numerical results obtained for $(PMMA_4)(AA)(H^+)$ complexes and $PMMA_4$ oligomers by MD-SA.

Complex	Generalized Amber force field					AM1		
	iteration[a]	E_0[b] (kcal mol^{-1})	E_f[c] (kcal mol^{-1})	Relative Energy[d] (kcal mol^{-1})	Relative Energy (kJ mol^{-1})	SPE[e] (Hartree)	SPE (kJ mol^{-1})	Relative Energy[f] (kJ mol^{-1})
$(PMMA_4)(GlyH^+)$	274	20.6	17.2	0.0	0.0	-0.5056	-1327.5	0.0
$(Gly)(PMMA_4H_1^+)$	843	63.7	49.6	32.3	135.2	-0.4601	-1208.0	119.5
$(Gly)(PMMA_4H_2^+)$	1480	56.1	41.5	24.2	101.3	-0.4590	-1205.2	122.4
$(Gly)(PMMA_4H_3^+)$	1287	56.0	40.8	23.5	98.5	-0.4607	-1209.6	117.9
$(Gly)(PMMA_4H_4^+)$	1091	64.2	48.8	31.5	132.0	-0.4634	-1216.6	111.0
$(PMMA_4)(LeuH^+)$	1114	54.5	16.2	0.0	0.0	-0.5379	-1412.2	0.0
$(Leu)(PMMA_4H_1^+)$	420	76.5	52.9	36.7	153.7	-0.4711	-1236.9	175.3
$(Leu)(PMMA_4H_2^+)$	491	73.7	46.2	30.0	125.5	-0.4709	-1236.3	176.0
$(Leu)(PMMA_4H_3^+)$	1054	73.4	45.1	29.0	121.3	-0.4734	-1242.8	169.4
$(Leu)(PMMA_4H_4^+)$	1198	85.0	50.9	34.8	145.5	-0.4859	-1275.6	136.6
$(PMMA_4)(PheH^+)$	231	70.9	17.4	0.0	0.0	-0.4692	-1232.0	0.0
$(Phe)(PMMA_4H_1^+)$	1373	87.8	52.8	35.4	148.1	-0.4042	-1061.3	170.7
$(Phe)(PMMA_4H_2^+)$	1138	75.1	44.3	26.9	112.4	-0.3853	-1011.5	220.4
$(Phe)(PMMA_4H_3^+)$	526	75.1	45.7	28.3	118.5	-0.4015	-1054.0	177.9
$(Phe)(PMMA_4H_4^+)$	922	90.6	52.2	34.8	145.5	-0.4034	-1059.0	172.9
$PMMA_4$	304	50.8	38.1	0.0	0.0	-0.5633	-1478.9	0.0
$PMMA_4H_1^+$	48	87.0	73.6	35.5	148.6	-0.2910	-763.9	715.0
$PMMA_4H_2^+$	505	75.7	64.3	26.2	109.5	-0.2852	-748.9	730.1
$PMMA_4H_3^+$	733	75.5	62.8	24.7	103.4	-0.2993	-785.7	693.2
$PMMA_4H_4^+$	364	89.3	72.6	34.6	144.6	-0.2978	-781.8	697.1

[a] SA iteration at which the lowest energy conformation was obtained.
[b] E_0 is the energy of the initial complex conformation before the SA.
[c] E_f is the energy of the lowest energy conformation obtained.
[d] Relative energy from the force field to the corresponding $(PMMA_4)(AAH^+)$ (normalized to 0.0).
[e] AM1 Single point energy (SPE) of the lowest energy conformation obtained during the SA.
[f] Relative energy from AM1 SPE to the corresponding $(PMMA_4)(AAH^+)$ (normalized to 0.0).

Table 4.6 shows the results for the MD-SA of leucine and phenylalanine, while glycine was simply minimized using standard SD and ABNR methods. The MD-SA applied there was having "mild" annealing conditions, as these chemical species are fairly small. The amino acid modeling was performed to study the dissociated species coming out of each channel later on.

Table 4.6: Numerical results obtained for the amino acids MD-SA.[a]

amino acid	Generalized Amber force field					AM1		
	iteration	E_0 (kcal mol^{-1})	E_f (kcal mol^{-1})	Relative Energy (kcal mol^{-1})	Relative Energy (kJ mol^{-1})	SPE (Hartree)	SPE (kJ mol^{-1})	Relative Energy (kJ mol^{-1})
Gly	---	---	-9.7	0.0	0.0	-0.1491	-391.5	0.0
$GlyH^+$	---	---	3.2	12.8	53.8	0.1149	301.6	693.1
Leu	87	-1.3	-2.9	0.0	0.0	-0.1754	-460.4	0.0
$LeuH^+$	1	16.1	13.2	16.1	67.5	0.0728	191.1	651.4
Phe	88	1.5	-1.3	0.0	0.0	-0.1000	-262.5	0.0
$PheH^+$	3	18.0	12.8	14.1	59.0	0.1358	356.6	619.1

[a] except for Gly and $GlyH^+$ which were minimized with SD and ABNR only.

The AM1 SPE scoring of the MD-SA conformations obtained confirms that the lowest energy complexes are those for which the proton sits on the amino acid. However, there is no clear indication as to where would be the protonation site on the polymer if there should be a proton transfer. Table 4.7 shows the results of an AM1 geometry optimization on the lowest energy conformation for each system found by MD-SA. For the $(PMMA_4)(Gly)(H^+)$ and $(PMMA_4)(Phe)(H^+)$ systems, the AM1 optimized structures now show that the preferred protonation sites on the oligomers would be the carbonyl functions in the middle of the polymer sequence. However, there seems to be disagreement with the $(PMMA_4)(Leu)(H^+)$ systems. This discrepancy may be due to the difference in parameterization between both methods. In other words, a local minimum found with the Amber force field may correspond to a saddle point using AM1.

Table 4.7: Numerical results obtained for the $PMMA_4$ and the $PMMA_4H_n^+$ after AM1 geometry optimization.

complex	iteration	Optimized[a] (Hartree)	Optimized (kJ mol^{-1})	ZPE[b] (Hartree)	Optimized corrected[c] (kJ mol^{-1})	Relative Energy[d] (kJ mol^{-1})
$(PMMA_4)(GlyH^+)$	274	-0.5789	-1519.9	0.6337	66.1	0.0
$(Gly)(PMMA_4H_1^+)$	843	-0.5398	-1417.2	0.6298	158.9	92.8
$(Gly)(PMMA_4H_2^+)$	1480	-0.5494	-1442.3	0.6304	135.3	69.2
$(Gly)(PMMA_4H_3^+)$	1287	-0.5482	-1439.4	0.6313	140.6	74.5
$(Gly)(PMMA_4H_4^+)$	1091	-0.5403	-1418.6	0.6294	156.5	90.5
$(PMMA_4)(LeuH^+)$	1114	-0.6113	-1605.0	0.7476	266.0	0.0
$(Leu)(PMMA_4H_1^+)$	420	-0.5628	-1477.7	0.7432	382.3	116.3
$(Leu)(PMMA_4H_2^+)$	491	-0.5699	-1496.2	0.7441	365.9	99.9
$(Leu)(PMMA_4H_3^+)$	1054	-0.5710	-1499.2	0.7447	364.3	98.3
$(Leu)(PMMA_4H_4^+)$	1198	-0.5800	-1522.8	0.7452	342.1	76.1
$(PMMA_4)(PheH^+)$	231	-0.5404	-1418.8	0.7456	447.1	0.0
$(Phe)(PMMA_4H_1^+)$	1373	-0.4953	-1300.5	0.7434	559.9	112.8
$(Phe)(PMMA_4H_2^+)$	1138	-0.5065	-1329.7	0.7432	530.1	83.1
$(Phe)(PMMA_4H_3^+)$	526	-0.5029	-1320.3	0.7444	542.5	95.5
$(Phe)(PMMA_4H_4^+)$	922	-0.4967	-1304.1	0.7432	555.8	108.7
$PMMA_4$	304	-0.6128	-1608.9	0.5371	-264.7	0.0
$PMMA_4H_1^+$	48	-0.3609	-947.4	0.5483	424.8	689.6
$PMMA_4H_2^+$	505	-0.3678	-965.6	0.5490	408.3	673.0
$PMMA_4H_3^+$	733	-0.3706	-973.0	0.5493	401.6	666.3
$PMMA_4H_4^+$	364	-0.3718	-976.1	0.5491	398.0	662.7

[a] Energy calculated after the AM1 geometry optimization of the lowest energy conformation obtained by MD-SA at the iteration specified.

[b] AM1 zero point energy (ZPE)

[c] AM1 energy after correction with the ZPE.

[d] Relative AM1 energy (optimized and corrected) to the corresponding $(PMMA_4)(AAH^+)$.

Table 4.8: Numerical results obtained for the amino acids after AM1 geometry optimization.

amino acid	iteration	Optimized (Hartree)	Optimized (kJ mol^{-1})	ZPE (Hartree)	Optimized corrected (kJ mol^{-1})	Relative Energy (kJ mol^{-1})
Gly	---	0.1020	267.8	0.0932	501.1	0.0
$GlyH^+$	---	-0.1619	-425.2	0.0805	-223.7	724.8
Leu	87	0.0571	150.0	0.2074	669.2	0.0
$LeuH^+$	1	-0.1943	-510.0	0.1951	-21.8	691.0
Phe	88	0.1229	322.8	0.2059	838.2	0.0
$PheH^+$	3	-0.1161	-304.9	0.1934	179.2	659.0

Since the 3D conformations were primarily obtained using MD-SA and have provided slightly more correlated results, the theoretical energetics based on MD-SA were compared for each channel in an energy diagram (Figure 4.7).

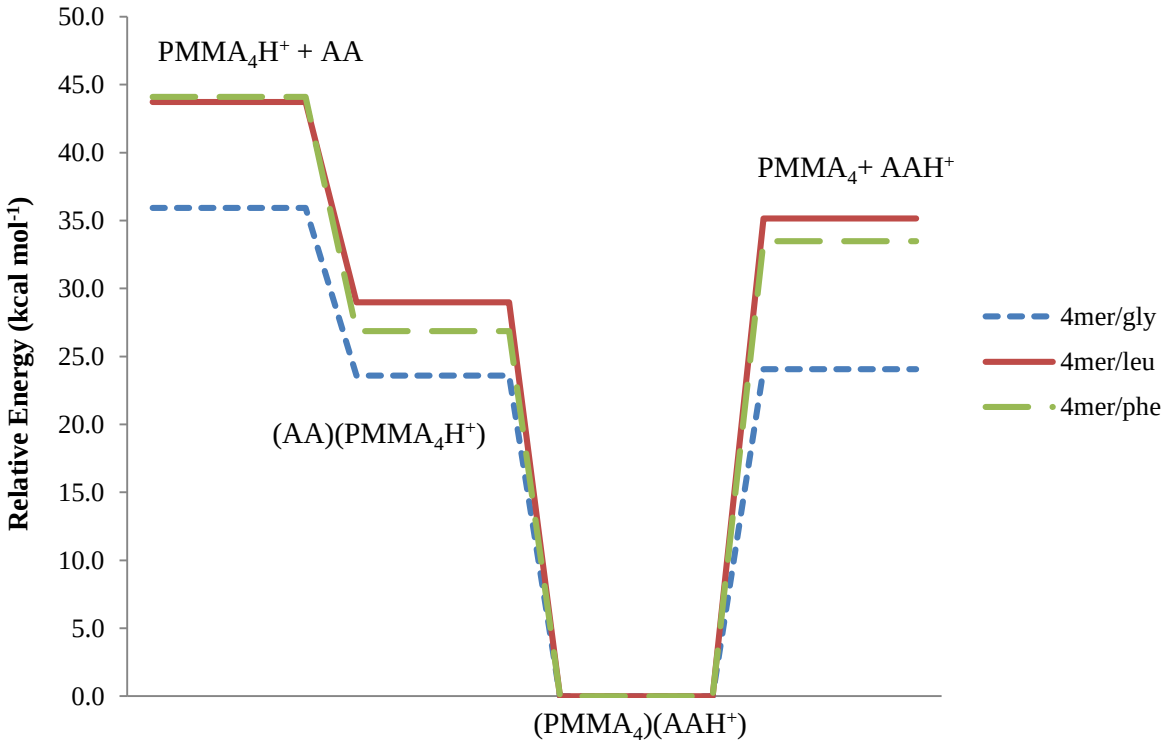
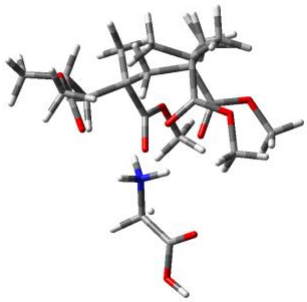


Figure 4.7: Modeling of the lowest energy dissociation channels for the $PMMA_4$ systems.

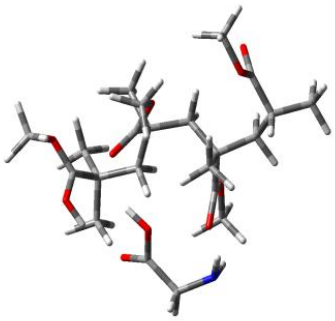
All channels were normalized to the energy of the initial gas phase complex $(PMMA_4)(AAH^+)$, which corresponds to the lowest energy conformation in this system. Even if there was only one channel for the dissociation of the $(PMMA_4)(Gly)(H^+)$ complex, it was modeled to see if some useful information could be highlighted from there. Even though, the objective here is not to propose a dissociation mechanism, it is interesting to see that for the $(PMMA_4H^+ + AA)$ channel, there could be a proton transfer from the amino acid to the oligomer before complete dissociation.

Although the general energetic trend of competing dissociating channels is represented here (as previously discussed), certain aspects do not correlate with the experimental observations. The theoretical energetics do not match the experimental ones. Experimentally, the ($PMMA_4H^+$ + AA) dissociating channel is preferred and is energetically favourable for the systems with leucine and phenylalanine. This is not the case here, as the modeled ($PMMA_4H^+$ + AA) channels seem to require higher energy for complete dissociation when compared to their respective ($PMMA_4$ + AAH^+) channel. Unless there would be an unknown transition state for the proton transfer between the amino acid and the oligomer, the modeling cannot entirely explain this experimental result. In addition, when looking at the ($PMMA_4$)($GlyH^+$) dissociation channel, it seems that it would be very easy to dissociate through the protonated glycine channel, but this is not observed experimentally. The fact that the glycine side-chain is smaller than the two other amino acids, it cannot do additional interactions (such as LDF) with the oligomer to stabilize the overall complex. Once the proton is transferred on the oligomer, it is shared within the oligomer and leaving the neutral glycine because of the insufficient interactions remaining.

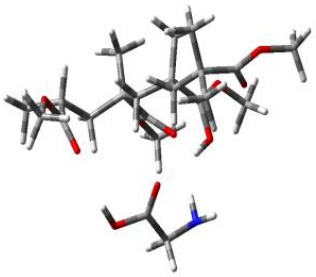
In terms of conformations, Figure 4.8 shows an example of all $PMMA_4$ complexes with the different possibilities for the protonation site. It can be seen that when the amino acid bears the proton, there is a local coordination of oligomer carbonyl functions around the charge (Figure 4a, f and k). As can be seen, in cases where the oligomer is protonated, there is a tendency for the leucine and the phenylalanine to orientate their side-chain closer to the oligomer to increase their interaction together in the complex. Such behaviour is not possible with the glycine. Qualitatively, all complexes have a similar molecular volume and therefore a similar cross-section. This is in agreement with the use of a common α value for the RRKM modeling of the dissociating channels.



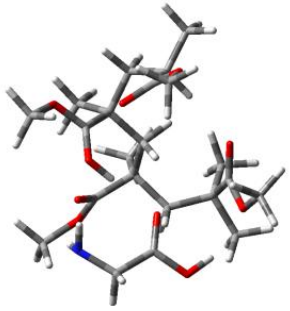
(a) $(PMMA_4)(GlyH^+)$



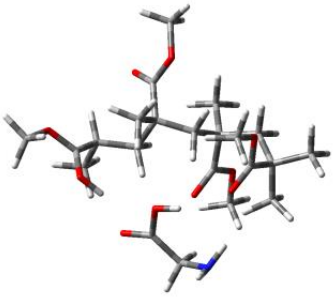
(b) $(Gly)(PMMA_4H_1^+)$



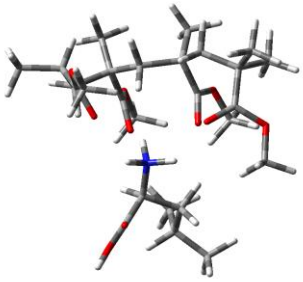
(c) $(Gly)(PMMA_4H_2^+)$



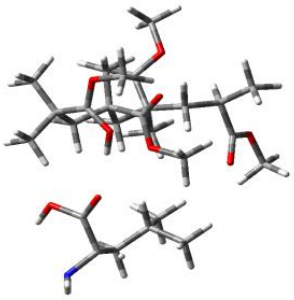
(d) $(Gly)(PMMA_4H_3^+)$



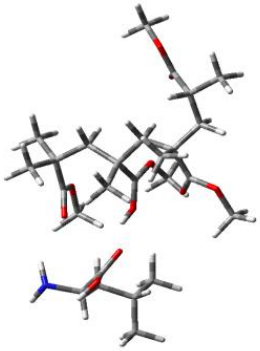
(e) $(Gly)(PMMA_4H_4^+)$



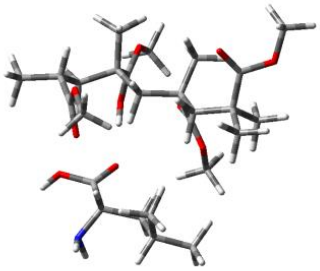
(f) $(PMMA_4)(LeuH^+)$



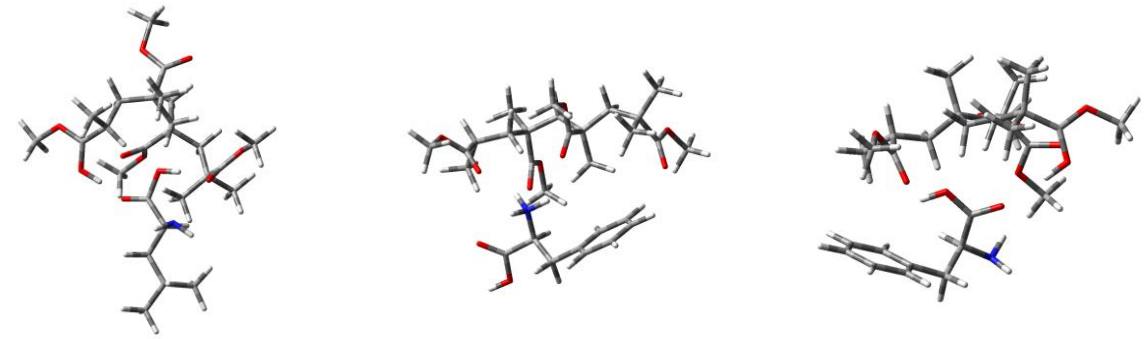
(g) $(Leu)(PMMA_4H_1^+)$



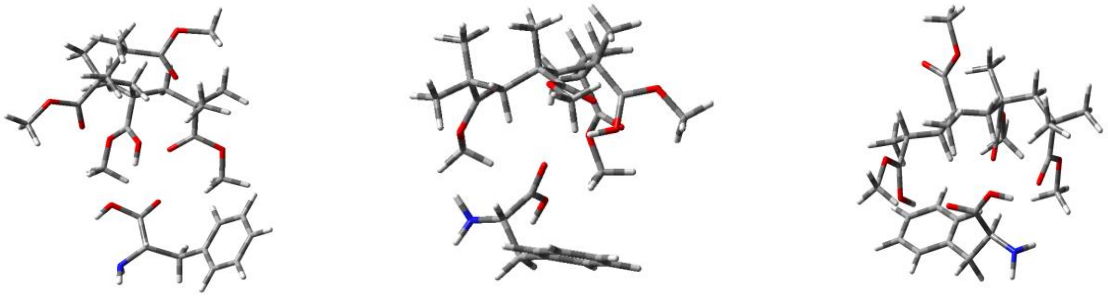
(h) $(Leu)(PMMA_4H_2^+)$



(i) $(Leu)(PMMA_4H_3^+)$



(j) (Leu)($PMMA_4H_4^+$) (k) ($PMMA_4$)($PheH^+$) (l) (Phe)($PMMA_4H_1^+$)



(m) (Phe)($PMMA_4H_2^+$) (n) (Phe)($PMMA_4H_3^+$) (o) (Phe)($PMMA_4H_4^+$)

Figure 4.8: Examples of low energy 3D conformations obtained by MD-SA simulations for the $PMMA_4$ oligomers bound to glycine (a)-(e), leucine (f)-(j), and phenylalanine (k)-(o).

4.4.3 Modeling the $PMMA_7$ systems: conformational features and energetics

According to the experimental results, two different dissociation channels were competing (except for the ($PMMA_7$)(Gly)(H^+) complex which had only one). For the protonation site, MD-SA simulations were performed on $PMMA_7H^+$ oligomers and complexes bearing the proton at each carbonyl position. As shown in Table 4.9, the three complexes having the proton located on the amino acid have the lowest energy. This gives an insight on the fact that when the complex is

bound in the gas phase, the proton is on the amino acid, regardless of its proton affinity. Comparing the same relative energies, it seems that if there should be a proton transfer between the amino acid and the oligomer, the energetically favourable protonation sites would be the carbonyl functions located in the middle of the polymer sequence. Taken together, these observations are very similar to those of the $PMMA_4$ systems. The AM1 SPE scoring also behaved as in the case of the $PMMA_4$ systems: it confirms the energetic stability of the complex where the amino acid is protonated; however, it shows no trend as to find a preferable protonation carbonyl chain on the oligomer. The MD-SA results are shown in the energy diagram of the competing dissociation channels (Figure 4.9).

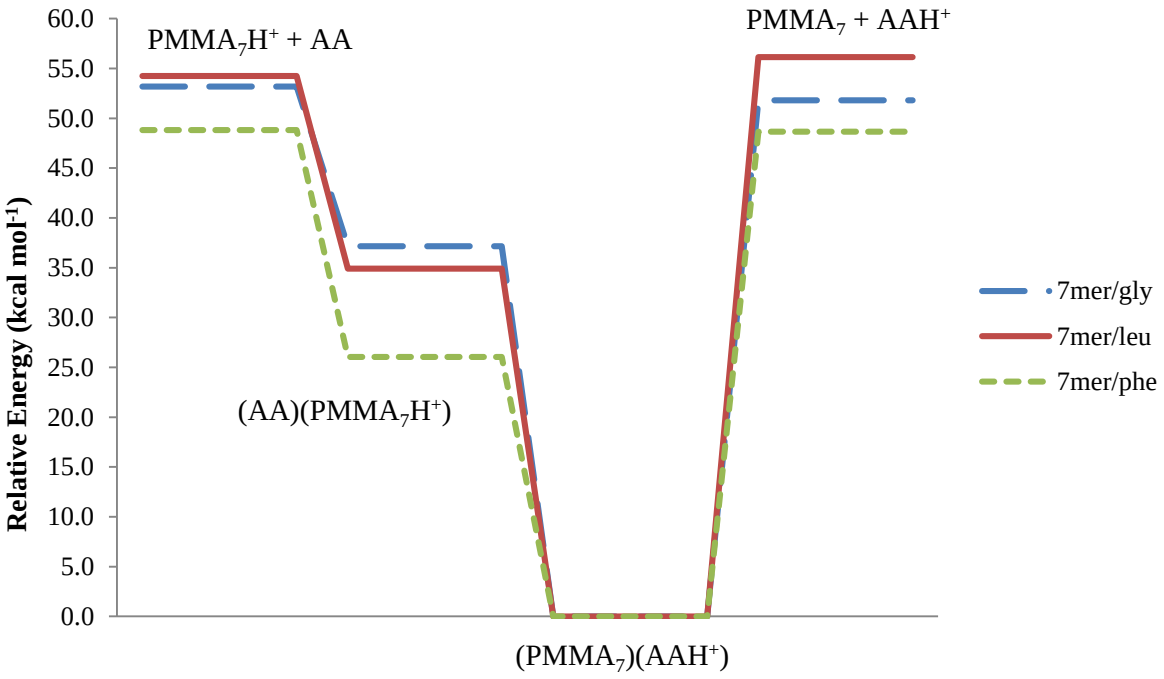
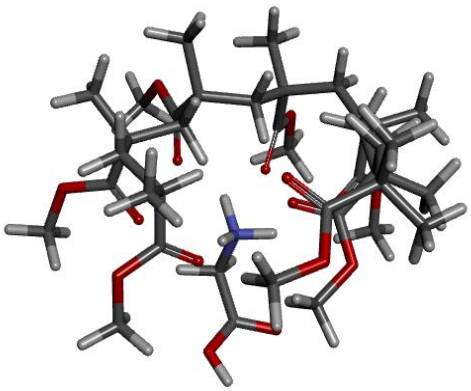


Figure 4.9: Modeling of the lowest energy dissociation channels for the $PMMA_7$ systems.

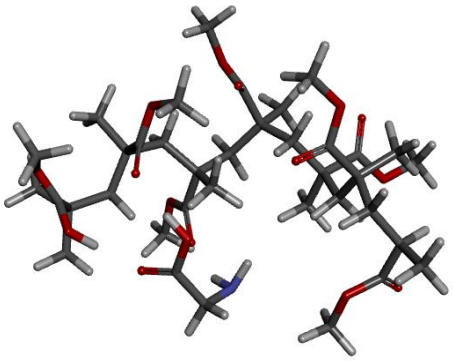
All channels were normalized to the energy of the initial gas phase complex ($PMMA_7$)(AAH^+), which corresponds to the lowest energy conformation in this system. Even if there was only one channel for the dissociation of the ($PMMA_7$)(Gly)(H^+) complex, it was modeled to see if some useful information could be highlighted from there. At first, this modeling diagram respects the general energetics tendencies of the competing dissociation channels. However, there seems to be a better qualitative agreement with the experimental results. Both competing dissociation channels are rather close to one another in terms of energy. In the case of the ($PMMA_7$)($GlyH^+$) complex, the increased flexibility of the oligomer allows to coordinate more carbonyl groups around the protonated amine (Figure 4.10a). However, for (Gly)($PMMA_7H^+$) complex, the oligomer is not as tighly wrapped around the glycine (Figure 10b). For the ($PMMA_7$)($LeuH^+$) complex, the flexible side-chain of leucine can also generate additional interactions with the oligomer in addition to the quasi-cyclic conformation adopted by the oligomer around the charge of the amino acid (Figure 4.10c). For the (Leu)($PMMA_7H^+$) complex, the majority of the carbonyl side chains of the oligomer do not even interact with the leucine (Figure 4.10d). Inside the ($PMMA_7$)($PheH^+$) complex, the phenylalanine has a bigger rigid side-chain that restricts its conformational flexibility; the interactions with the oligomer are then less optimal since it cannot wrap efficiently the amino acid (Figure 4.10e), and it becomes even worse when the oligomer is protonated (Figure 4.10f). It is interesting to see that when the oligomer bears the proton, one of the reasons why the complex may be less stable is due to H-bond competing between the oxygen of the amino acid and the oxygen of another proximal carbonyl group. Even if the energetic tendency is in better agreement with the experimental values, there is still no apparent reason that would explain why the ($PMMA_7$)(Gly)(H^+) would not dissociate through a second channel ($PMMA_7$ + $GlyH^+$).

Table 4.9: Numerical results obtained for the $(PMMA_7)(AA)(H^+)$ complexes, $PMMA_7$ and the $PMMA_7H_n^+$ by MD-SA.

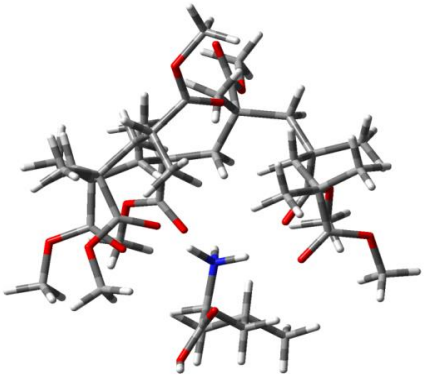
Complex	Generalized Amber force field					AM1		
	iteration	E_0 (kcal mol^{-1})	E_f (kcal mol^{-1})	Relative Energy (kcal mol^{-1})	Relative Energy (kJ mol^{-1})	SPE (Hartree)	SPE (kJ mol^{-1})	Relative Energy (kJ mol^{-1})
$(PMMA_7)(GlyH^+)$	484	101.1	36.8	0.0	0.0	-0.8687	-2280.6	0.0
$(Gly)(PMMA_7H_1^+)$	1300	115.7	89.4	52.6	220.1	-0.8211	-2155.8	124.8
$(Gly)(PMMA_7H_2^+)$	1021	102.7	81.6	44.8	187.5	-0.8329	-2186.8	93.8
$(Gly)(PMMA_7H_3^+)$	229	97.7	73.9	37.2	155.5	-0.8057	-2115.3	165.3
$(Gly)(PMMA_7H_4^+)$	1118	100.1	75.9	39.1	163.7	-0.8535	-2240.8	39.8
$(Gly)(PMMA_7H_5^+)$	138	106.2	79.3	42.5	177.7	-0.8152	-2140.3	140.3
$(Gly)(PMMA_7H_6^+)$	1130	103.0	81.1	44.4	185.6	-0.8140	-2137.0	143.6
$(Gly)(PMMA_7H_7^+)$	536	124.5	94.7	58.0	242.6	-0.8146	-2138.5	142.0
$(PMMA_7)(LeuH^+)$	625	113.4	42.5	0.0	0.0	-0.9403	-2468.6	0.0
$(Leu)(PMMA_7H_1^+)$	336	129.5	88.5	46.1	192.8	-0.8343	-2190.5	278.1
$(Leu)(PMMA_7H_2^+)$	1160	117.3	80.0	37.5	157.0	-0.8445	-2217.2	251.4
$(Leu)(PMMA_7H_3^+)$	1388	111.6	77.4	35.0	146.3	-0.8606	-2259.4	209.2
$(Leu)(PMMA_7H_4^+)$	1456	117.2	79.3	36.9	154.2	-0.8675	-2277.4	191.2
$(Leu)(PMMA_7H_5^+)$	1010	119.4	77.3	34.9	146.0	-0.8577	-2251.8	216.8
$(Leu)(PMMA_7H_6^+)$	726	120.2	82.5	40.1	167.8	-0.8371	-2197.8	270.8
$(Leu)(PMMA_7H_7^+)$	616	135.1	97.3	54.8	229.4	-0.8342	-2190.2	278.4
$(PMMA_7)(PheH^+)$	1429	123.0	49.5	0.0	0.0	-0.8337	-2188.8	0.0
$(Phe)(PMMA_7H_1^+)$	303	130.7	88.2	38.6	161.7	-0.8019	-2105.4	83.4
$(Phe)(PMMA_7H_2^+)$	1313	117.4	80.4	30.9	129.4	-0.7836	-2057.3	131.5
$(Phe)(PMMA_7H_3^+)$	837	114.0	75.5	26.0	108.9	-0.7950	-2087.1	101.6
$(Phe)(PMMA_7H_4^+)$	399	119.3	81.7	32.2	134.7	-0.7969	-2092.2	96.6
$(Phe)(PMMA_7H_5^+)$	671	122.3	81.9	32.4	135.5	-0.7530	-1976.8	211.9
$(Phe)(PMMA_7H_6^+)$	802	124.2	88.1	38.6	161.6	-0.7822	-2053.7	135.1
$(Phe)(PMMA_7H_7^+)$	807	137.0	97.2	47.6	199.4	-0.7714	-2025.3	163.5
$PMMA_7$	674	99.9	85.4	0.0	0.0	-0.9255	-2429.8	0.0
$PMMA_7H_1^+$	173	131.4	113.9	28.6	119.5	-0.6504	-1707.5	722.2
$PMMA_7H_2^+$	458	120.1	101.7	16.3	68.4	-0.6653	-1746.6	683.2
$PMMA_7H_3^+$	193	114.7	99.6	14.2	59.6	-0.6543	-1717.8	712.0
$PMMA_7H_4^+$	1349	118.9	104.7	19.3	80.7	-0.6518	-1711.3	718.4
$PMMA_7H_5^+$	1443	121.2	103.1	17.7	74.2	-0.6465	-1697.4	732.4
$PMMA_7H_6^+$	1447	122.2	110.7	25.3	105.8	-0.6329	-1661.7	768.1
$PMMA_7H_7^+$	214	136.2	123.1	37.8	158.0	-0.6478	-1700.8	728.9



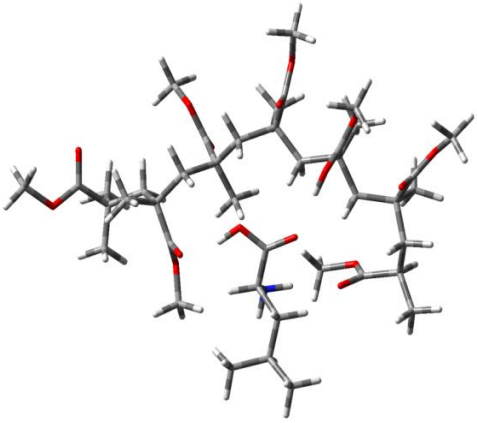
(a) ($PMMA_7$)($GlyH^+$)



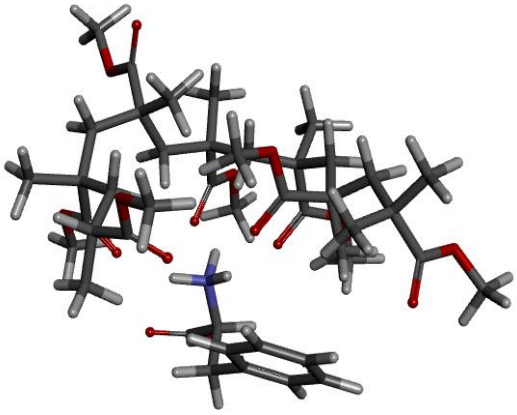
(b) (Gly)($PMMA_7H^+$)



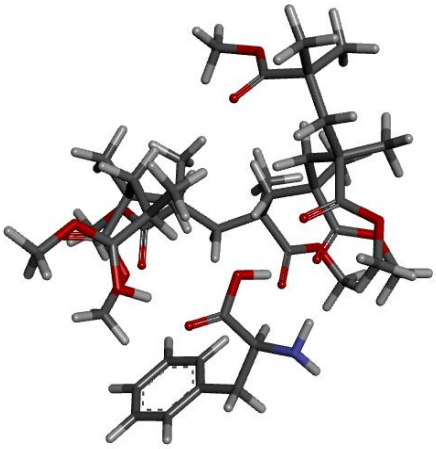
(c) ($PMMA_7$)($LeuH^+$)



(d) (Leu)($PMMA_7H^+$)



(e) ($PMMA_7$)($PheH^+$)



(f) (Phe)($PMMA_7H^+$)

Figure 4.10: Examples of low energy 3D conformations obtained by MD-SA simulations for the $PMMA_7$ oligomers bound to glycine (a)-(b), leucine (c)-(d), and phenylalanine (e)-(f).

4.4.4 Modeling the $PMMA_{13}$ systems: conformational features and energetics

According to the experimental results, two different dissociation channels were competing (except for the $(PMMA_{13})(Gly)(H^+)$ complex which had only one). For the protonation site, MD-SA simulations were performed on $PMMA_{13}H^+$ oligomers and complexes bearing the proton at two carbonyl positions only. Since there was a trend found that the carbonyl in the middle of the sequence have greater chances of bearing the proton, there was no need to verify every single possibility. As shown in Table 4.10, the three complexes having the proton located on the amino acid have the lowest energy. Again, and according to the calculations, when the complex is bound in the gas phase, the proton is on the amino acid, regardless of its proton affinity. Comparing the same relative energies, it seems that if there should be a proton transfer between the amino acid and the oligomer, the energetically favourable protonation sites would be the carbonyl functions located in the middle of the polymer sequence. Of interest here, the energy difference between a proton bound in the middle vs. bound at the extreme end, is less than in the other systems involving shorter oligomers. The $PMMA_{13}$ is long enough so that no matter where the proton is bound, the remaining part of the oligomer will interact with the amino acid or with itself. The AM1 SPE scoring did not bring any new insight again in this case.

The MD-SA results are shown in the energy diagram of the competing dissociation channels (Figure 4.11). Here, the three ($PMMA_{13}H^+$ + AA) dissociation channels are lower in energy than their corresponding ($PMMA_{13}$ + AAH^+) channels. However, experimentally, the $(PMMA_{13})(Leu)(H^+)$ system has its values of E_0 interchanged between the competing channels; the difference in activation energy is small though (ΔE_0 = 0.02 eV), which is within the experimental error. From the energy diagrams, it is also possible to see that the dissociation

channels are all having higher energy barriers, which is in agreement with the RRKM results found. As shown in Figure 4.12, longer oligomers can more easily coordinate around the protonated amino acid. If the amino acid is protonated, the complex formed with the oligomer seems tighter (Figure 4.12 (a), (c), and (e)). However, if the proton is located on the oligomer, the complexes seems looser and adopting a slightly more elongated form (Figure 4.12 (b), (d), and (f)). When comparing the complexes in which the amino acid is protonated vs. the complexes in which the oligomer is protonated, they seem to have different cross-sections. Gas phase complexes, in which the glycine is involved, have a better coordination of the carbonyl around the amino acid, due in part to the small side-chain. However, complexes in which a leucine and a phenylalanine are involved have more steric hindrance due to the bigger side-chains. Although both leucine and phenylalanine side-chains remain outside the globe-like complex, they have the ability to interact better with the remaining part of the oligomer.

Table 4.10: Numerical results obtained for the $(PMMA_{13})(AA)(H^+)$ complexes, $PMMA_{13}$, $PMMA_{13}H_7^+$, and $PMMA_{13}H_{13}^+$ by MD-SA.

Complex	Amber force field					AM1		
	iteration	E_0 (kcal mol^{-1})	E_f (kcal mol^{-1})	Relative Energy (kcal mol^{-1})	Relative Energy (kJ mol^{-1})	SPE (Hartree)	SPE (kJ mol^{-1})	Relative Energy (kJ mol^{-1})
$(PMMA_{13})(GlyH^+)$	1113	177.2	89.6	0.0	0.0	-1.6317	-4283.8	0.0
$(Gly)(PMMA_{13}H_7^+)$	1741	191.4	139.7	50.1	209.7	-1.5781	-4143.3	140.5
$(Gly)(PMMA_{13}H_{13}^+)$	1281	199.0	142.0	52.4	219.3	-1.5726	-4128.6	155.2
$(PMMA_{13})(LeuH^+)$	292	187.7	105.9	0.0	0.0	-1.6619	-4363.1	0.0
$(Leu)(PMMA_{13}H_7^+)$	1515	199.5	153.8	47.8	200.1	-1.5828	-4155.5	207.6
$(Leu)(PMMA_{13}H_{13}^+)$	1781	207.1	160.5	54.6	228.2	-1.5836	-4157.7	205.5
$(PMMA_{13})(PheH^+)$	447	192.8	111.2	0.0	0.0	-1.5738	-4131.8	0.0
$(Phe)(PMMA_{13}H_7^+)$	932	202.9	152.3	41.1	171.9	-1.5169	-3982.5	149.3
$(Phe)(PMMA_{13}H_{13}^+)$	1288	210.0	157.0	45.8	191.6	-1.5133	-3973.0	158.8
$PMMA_{13}$	1351	174.8	162.1	0.0	0.0	-1.6643	-4369.4	0.0
$PMMA_{13}H_7^+$	360	201.5	164.9	2.8	11.8	-1.4154	-3715.9	653.5
$PMMA_{13}H_{13}^+$	1511	208.5	168.7	6.6	27.7	-1.4137	-3711.5	657.9

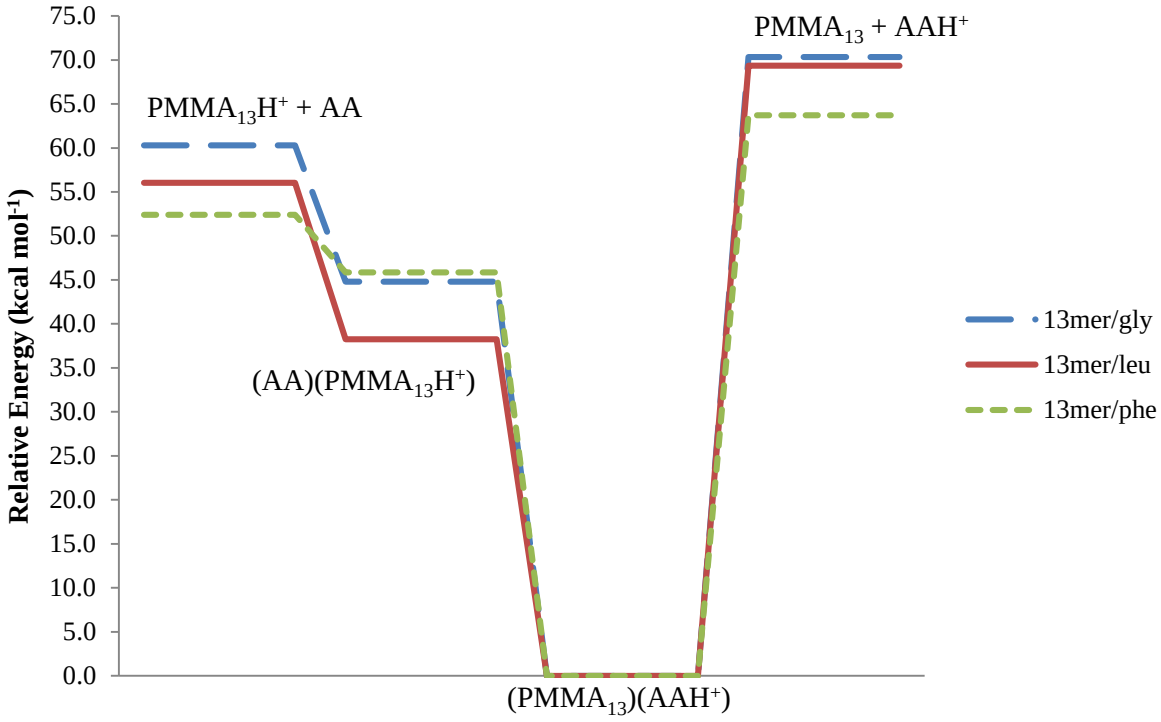
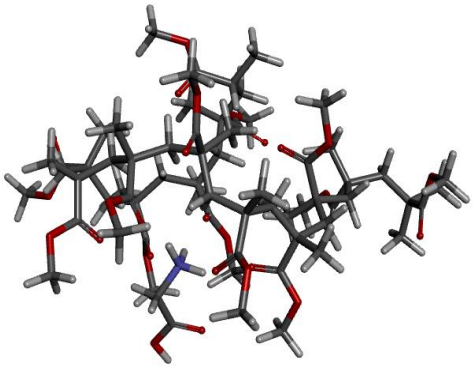
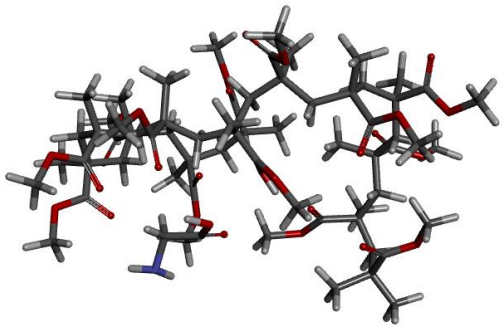


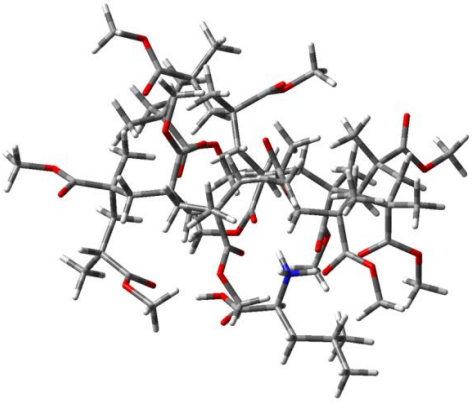
Figure 4.11: Modeling of the lowest energy dissociation channels for the $PMMA_{13}$ systems.



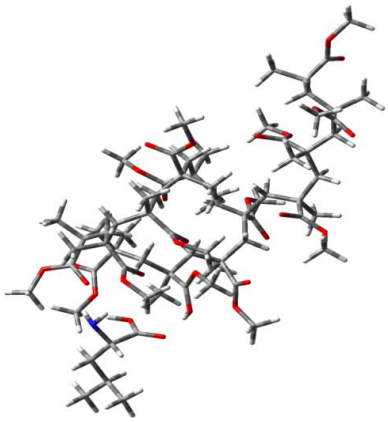
(a) ($PMMA_{13}$)($GlyH^+$)



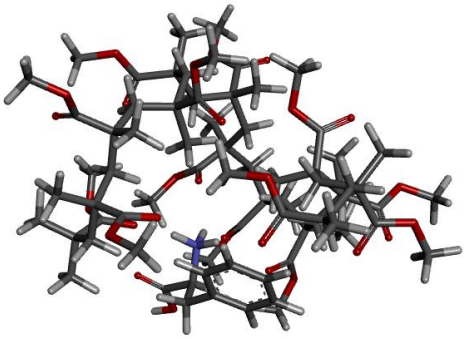
(b) (Gly)($PMMA_{13}H^+$)



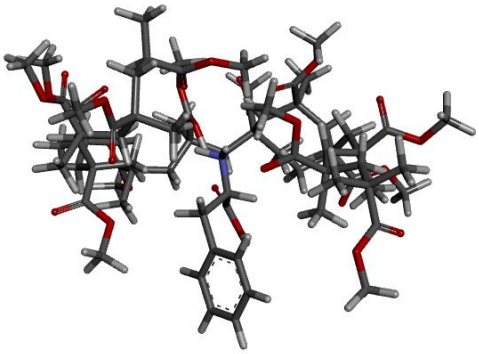
(c) ($PMMA_{13}$)($LeuH^+$)



(d) (Leu)($PMMA_{13}H^+$)



(e) ($PMMA_{13}$)($PheH^+$)



(f) (Phe)($PMMA_{13}H^+$)

Figure 4.12: Examples of low energy 3D conformations obtained by MD-SA simulations for the $PMMA_{13}$ oligomers bound to glycine (a)-(b), leucine (c)-(d), and phenylalanine (e)-(f).

4.4.5 An insight on the 0 K activation energy difference (ΔE_0)

Although only qualitative information could be obtained from the modeling results presented, no rigorous quantitative information could be extracted. This is partially due the affordable theoretical model used (MD) that is not precise enough to obtain accurate results. One semi-quantitative trend of interest was that no matter the $(PMMA_n)(AA)(H^+)$ modeled, the lowest energy conformation was always the bound $(PMMA_n)(AAH^+)$ complex. All dissociation pathways (Figures 4.7, 4.9 and, 4.11) were normalized to this lowest energy conformation. Rearranging these dissociation pathways by normalizing them to their $(PMMA_nH^+ + AA)$ dissociation channel, it was possible to gain insight into the difference in activation energy, ΔE_0, between both competing channels. These re-normalized dissociation pathways are shown in Figures 4.13 to 4.15 for the $PMMA_4$, $PMMA_7$, and $PMMA_{13}$ systems respectively.

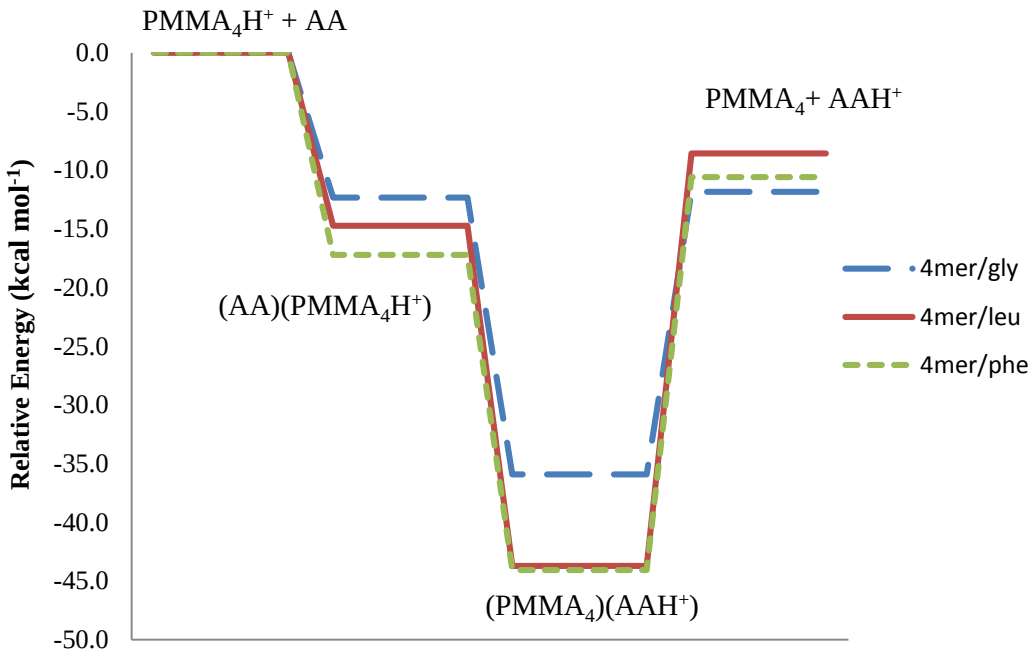


Figure 4.13: Modeling of the lowest energy dissociation channels for the $PMMA_4$ systems, normalized to the $(PMMA_4H^+ + AA)$ dissociation channel.

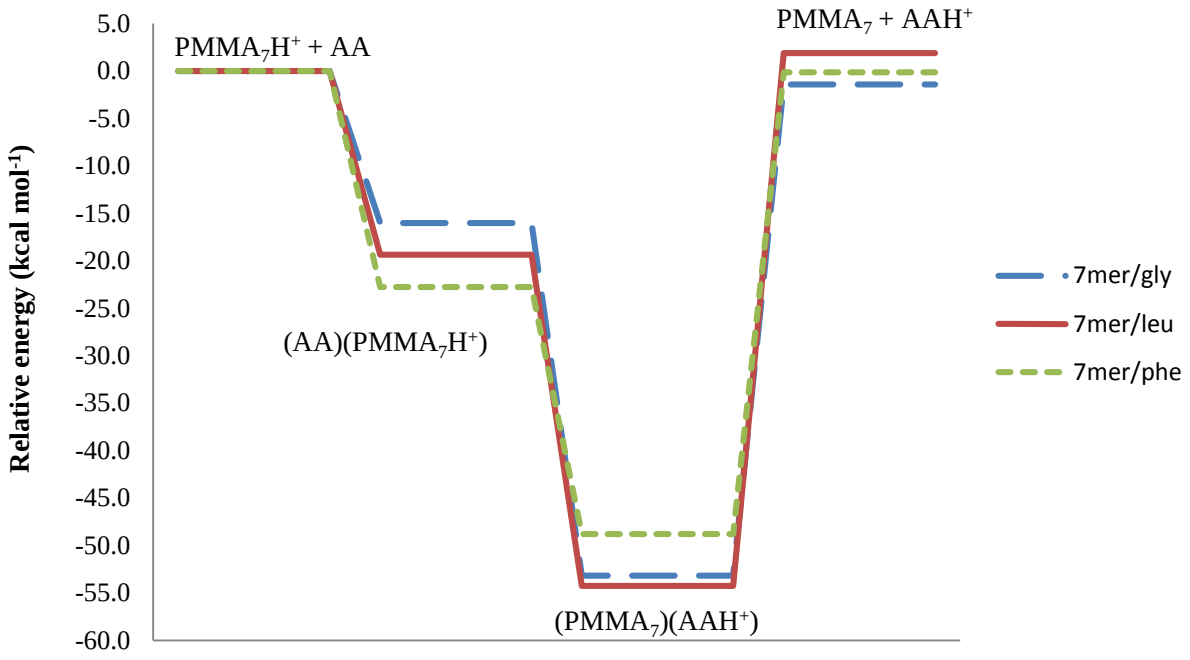


Figure 4.14: Modeling of the lowest energy dissociation channels for the $PMMA_7$ systems, normalized to the ($PMMA_7H^+$ + AA) dissociation channel.

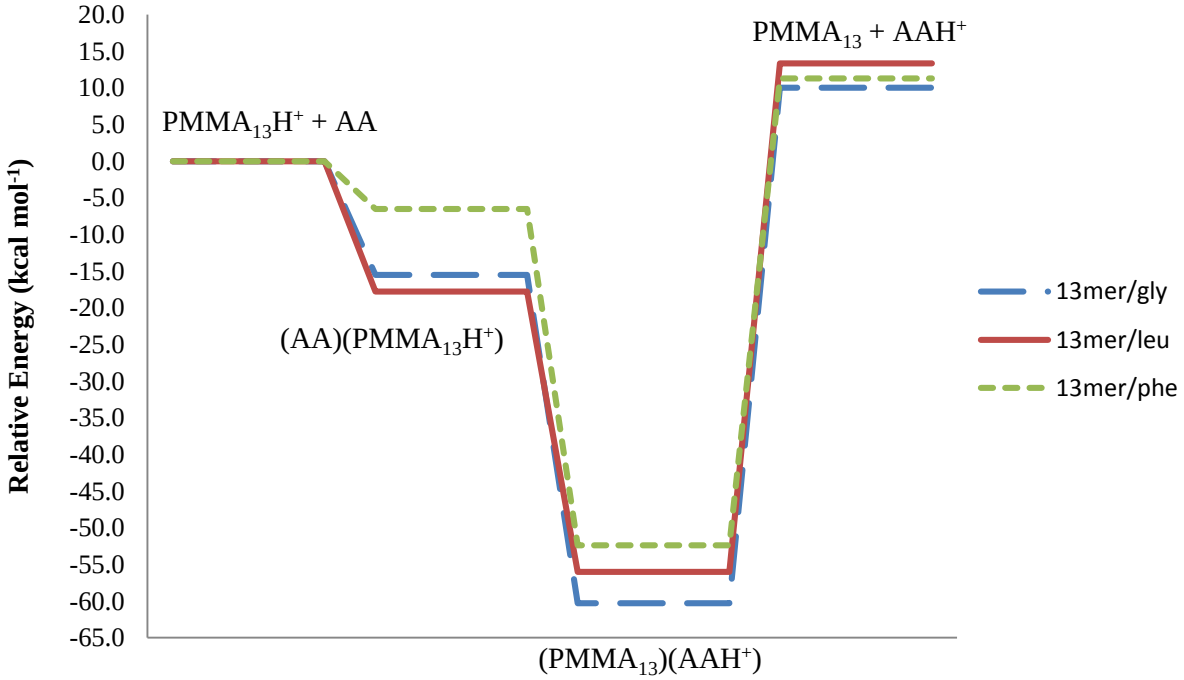


Figure 4.15: Modeling of the lowest energy dissociation channels for the $PMMA_{13}$ systems, normalized to the ($PMMA_{13}H^+$ + AA) dissociation channel.

From these three figures, it can be seen that the energy levels of the ($PMMA_n + AAH^+$) channels are very similar for all amino acid with a given oligomer system. Another trend found here is that these ($PMMA_n + AAH^+$) channels energy levels increase from the $PMMA_4$ to the PMM_{13} systems as compared to their respective competing ($PMMA_nH^+ + AA$) channels. It becomes interesting to calculate the difference in activation energy, ΔE_0, between two related competing channels.

$$\Delta E_0 = E_0\,(PMMA_nH^+ + AA) - E_0\,(PMMA + AAH^+) \qquad (4.14)$$

Based on the modeling results, it was hypothesized that there could be a proton transfer between the amino acid to one carbonyl group of the oligomer, before completely dissociating. Hence, there would be a competition between the energy level of the $(AA)(PMMA_nH^+)$ complex and the products of the ($PMMA_n + AAH^+$) dissociation channel.

$$\Delta E_0' = E_0\,(AA)(PMMA_nH^+) - E_0\,(PMMA + AAH^+) \qquad (4.15)$$

These theoretical ΔE_0 and $\Delta E_0'$ values as well the experimental ΔE_0 (converted in kcal mol^{-1}) are shown in Table 4.11 for all oligomer complexes involving leucine and phenylalanine. The values of ΔE_0 were not calculated for the oligomer systems involving glycine since there was only one dissociation pathway observed experimentally.

Table 4.11: Experimental and theoretical differences in activation energy for competing dissociation channels.

System	Experimental ΔE_0 (kcal mol^{-1})	Theoretical ΔE_0 (kcal mol^{-1})	Theoretical ΔE_0' (kcal mol^{-1})
$PMMA_4$/Leu	---	+8.6	-6.1
$PMMA_4$/Phe	---	+10.6	-8.2
$PMMA_7$/Leu	-1.2	-1.9	-21.2
$PMMA_7$/Phe	-2.1	+0.1	-22.7
$PMMA_{10}$/Leu	+0.7	---	---
$PMMA_{10}$/Phe	+0.5	---	---
$PMMA_{13}$/Leu	+0.5	-13.3	-25.6
$PMMA_{13}$/Phe	-2.8	-11.3	-17.8

As expected, no correlation is obtained between experimental and theoretical ΔE_0. Experimentally, all ΔE_0 are included between -2.8 and 0.7, whereas the theoretical values diverge significantly. Even if for the $PMMA_7$ systems, both experimental and theoretical values are in the same range, a real correlation cannot be concluded since all other values are considerably different. However, there is an interesting trend that appears with the theoretical ΔE_0': the $PMMA_4$ systems have relatively low differences compared to the $PMMA_7$ and $PMMA_{13}$ systems. This may support the results reported in Table 4.1: since the energy level for the proton transfer is higher (compared to the lowest energy conformation), it may partially explain why leucine and phenylalanine can dissociate protonated when the oligomer is shorter. For the $PMMA_7$ systems, there were no changes at threshold dissociation energy even if the (AA)(PMM_7H^+) energy level was more accessible. For the $PMMA_{13}$ systems, the dissociation pathway preferred is the one of the protonated oligomer at threshold energy. Even though, these are not definitive conclusions, it might be worthy to keep in mind these insights gained from the modeling.

4.4.6 Modeling the ($PMMA_n$)(AA)(H^+): entropy effects

So far, in section 4.4, much attention has been brought in order to find reasonable agreements with experimental results; however, in addition to the difficulties encountered, energy is not the only factor to consider. Entropy may potentially play an important role in the dissociation of the ($PMMA_n$)(AA)(H^+) systems. Experimentally, entropies of activation have been calculated using RRKM modeling. In this section, the entropy is calculated from the results obtained by AM1 and MD-SA simulations. According to statistical thermodynamics, it is possible to calculate the translational entropy (S_{trans}), the rotational entropy (S_{rot}), and the vibrational entropy (S_{vib}) for non-linear molecules using equations 4.16, 4.17, and 4.18 respectively.

$$S_{trans} = \frac{5}{2} N_A k_B + N_A k_B \ln\left(\left(\frac{2\pi M k_B T}{N_A h^2}\right)^{3/2}\left(\frac{k_B T}{P_0}\right)\right) \tag{4.16}$$

$$S_{rot} = \frac{3}{2} N_B k_B + N_A k_B \ln\left(\frac{\pi^{1/2}}{\sigma}\left(\frac{8\pi^2 k_B T}{h^2}\right)^{3/2}\left(I_A I_B I_C\right)^{1/2}\right) \tag{4.17}$$

$$s_{vib} = N_A k_B \sum_{j=1}^{3n-6}\left(\frac{(h\nu_j/k_B T)e^{h\nu_j/k_B T}}{1-e^{h\nu_j/k_B T}} - \ln\left(1-e^{h\nu_j/k_B T}\right)\right) \tag{4.18}$$

where: N_A, k_B, and h are Avogadro's number, Boltzmann constant, and Planck constant, respectively; M is the molar mass; P_0 is 101 325 Pa; σ is the symmetry number; I_A, I_B, and I_C are the three principal moments of inertia; υ_j is the vibration frequency associated with the jth normal mode. All calculations were made at T = 300 K.

The moments of inertia in equation 4.17 are calculated based on the rotational constants given by Nmode or AM1 (through Gaussian). The equation to transform the rotational constants, $B_{A,B,C}$, into moments of inertia is:

$$I = \frac{h}{8\pi^2 B} \tag{4.19}$$

Vibrational frequencies for the amino acids and $PMMA_4$ systems have been calculated using AM1 whereas those for the $PMMA_7$ and $PMMA_{13}$ systems have been calculated using Nmode. To make sure, there are no scaling factors to apply between both sets of vibrational frequencies, a quick verification was done. The sum of all vibrational frequencies has been calculated for a $PMMA_4$ and a $PMMA_7$ system using AM1 and Nmode. In both cases, there was a difference < 1 % between the sum of frequencies for each system between the values obtained by AM1 and Nmode. Therefore, no scaling factor needed to be applied, since differences in entropy will be considered.

Results from entropy calculations are shown in Table 4.12. The translational and rotational entropies are alike in many cases. S_{trans} depends on the molar mass and the temperature (which was held constant). Therefore, when two systems or chemical species have a molar mass of ± 1 proton, it does not change the result significantly, since these are large systems with high molecular weight (>700 Da). S_{rot} depends on the temperature and the moments of inertia. Again, when two systems have only one proton difference, this does not change significantly the overall shape of the system. Therefore, when two systems have approximately the same shape, their S_{rot} should be rather similar. The only contribution that changes significantly is S_{vib}, because vibrational frequencies can be significantly different between two systems.

Table 4.12: Translational, rotational, and vibrational (in J mol^{-1} K^{-1}) of selected species chemical species involved in the dissociation channels.[a]

Chemical species	S_{trans}	S_{rot}	S_{vib}[b]
Gly	163	107	51
$GlyH^+$	163	107	60
Leu	170	107	147
$LeuH^+$	170	123	152
Phe	172	124	153
$PheH^+$	173	129	156
$PMMA_4$	184	151	602
$PMMA_4H^+$	185	148	543
$(PMMA_4)(GlyH^+)$	186	153	710
$(Gly)(PMMA_4H^+)$	186	154	723
$(PMMA_4)(LeuH^+)$	187	156	830
$(Leu)(PMMA_4H^+)$	187	157	838
$(PMMA_4)(PheH^+)$	188	158	831
$(Phe)(PMMA_4H^+)$	188	160	871
$PMMA_7$	191	166	993
$PMMA_7H^+$	191	166	992
$(PMMA_7)(GlyH^+)$	192	168	1154
$(Gly)(PMMA_7H^+)$	192	168	1127
$(PMMA_7)(LeuH^+)$	193	170	1202
$(Leu)(PMMA_7H^+)$	193	169	1251
$(PMMA_7)(PheH^+)$	193	169	1300
$(Phe)(PMMA_7H^+)$	193	172	1248
$PMMA_{13}$	198	183	2040
$PMMA_{13}H^+$	198	184	2050
$(PMMA_{13})(GlyH^+)$	199	184	2199
$(Gly)(PMMA_{13}H^+)$	199	184	2240
$(PMMA_{13})(LeuH^+)$	199	184	2294
$(Leu)(PMMA_{13}H^+)$	199	184	2303
$(PMMA_{13})(PheH^+)$	199	184	2317
$(Phe)(PMMA_{13}H^+)$	199	185	2333

[a] when several structures were available for a given chemical species (e.g. $PMMA_nH^+$), vibrational harmonic frequency were calculated the lowest energy conformation.

[b] Vibrational harmonic frequencies were calculated with AM1 for $PMMA_4$ systems and amino acids, and with Nmode for $PMMA_7$ and $PMMA_{13}$ systems.

To make entropy results as significant as possible, the total difference of entropy, ΔS_{tot}, for each species on the dissociation channels was calculated using the data of Table 4.12. For a common basis of comparison, the entropy of the $(PMMA_n)(AAH^+)$ was set to 0 and all other values were reported relative to this one.

$$\Delta S_{tot} = \Delta S_{trans} + \Delta S_{rot} + \Delta S_{vib} \quad (4.20)$$

Table 4.13 shows the results for the difference in entropy between both channels for the $PMMA_4$ systems. As explained above, it is possible to see that the translational and rotational contributions cancel out (or affect very little). The total variation of entropy is mostly dictated by the ΔS_{vib}. For all three $PMMA_4$/AA systems, the ($PMMA_4$ + AAH^+) dissociation channel is entropically favourable over the ($PMMA_4H^+$ + AA) dissociation channel. Even for the non-existing channel ($PMMA_4$ + $GlyH^+$), the entropy would have been in the same order as well. This would maybe suggest that the PA of glycine is so low that it would take a much more significant entropic effect upon dissociation of the complex.

Table 4.14 shows the results for the between both channels for the $PMMA_7$ systems. Translational and rotational contributions to the entropy cancel out once again. For all three $PMMA_7$/AA systems, the ($PMMA_7$ + AAH^+) dissociation channel is entropically favourable over the ($PMMA_7H^+$ + AA) dissociation channel. However, the difference in entropy between both dissociation channels is smaller.

Table 4.13: Difference in translational, rotational, vibrational, and total entropy (in J mol^{-1} K^{-1}) of selected chemical species involved in the $(PMMA_4)(AA)(H^+)$ dissociation channels.[a]

species	ΔS_{trans}	ΔS_{rot}	ΔS_{vib}	ΔS_{tot}
$Gly + PMMA_4H^+$	160	102	-116	146
$(Gly)(PMMA_4H^+)$	0	1	13	14
$(PMMA_4)(GlyH^+)$	0	0	0	0
$PMMA_4 + GlyH^+$	161	105	-48	218
$Leu + PMMA_4H^+$	166	116	-140	142
$(Leu)(PMMA_4H^+)$	0	1	8	8
$(PMMA_4)(LeuH^+)$	0	0	0	0
$PMMA_4 + LeuH^+$	166	118	-76	208
$Phe + PMMA_4H^+$	168	119	-136	152
$(Phe)(PMMA_4H^+)$	0	2	40	42
$(PMMA_4)(PheH^+)$	0	0	0	0
$PMMA_4 + PheH^+$	168	122	-73	217

Table 4.14: Difference in translational, rotational, vibrational, and total entropy (in J mol^{-1} K^{-1}) of selected chemical species involved in the $(PMMA_7)(AA)(H^+)$ dissociation channels.[a]

species	ΔS_{trans}	ΔS_{rot}	ΔS_{vib}	ΔS_{tot}
$Gly + PMMA_7H^+$	161	104	-111	155
$(Gly)(PMMA_7H^+)$	1	1	25	27
$(PMMA_7)(GlyH^+)$	0	0	0	0
$PMMA_7 + GlyH^+$	163	105	-101	167
$Leu + PMMA_7H^+$	167	119	-62	224
$(Leu)(PMMA_7H^+)$	0	1	49	50
$(PMMA_7)(LeuH^+)$	0	0	0	0
$PMMA_7 + LeuH^+$	168	119	-57	230
$Phe + PMMA_7H^+$	170	126	-156	140
$(Phe)(PMMA_7H^+)$	0	2	53	55
$(PMMA_7)(PheH^+)$	0	0	0	0
$PMMA_7 + PheH^+$	170	126	-151	145

Finally, Table 4.15 shows the results for the difference in entropy between both channels for the $PMMA_{13}$ systems. As explained above, it is possible to see that the translational and rotational

contributions to the total entropy, cancel out. Both dissociation channels have approximately the same variation in total entropy and in the cases of systems involving leucine and phenylalanine, the protonated oligomer channel is slightly more entropically favourable.

Table 4.15: Difference in translational, rotational, vibrational, and total entropy (in J mol^{-1} K^{-1}) of selected chemical species involved in the $(PMMA_{13})(AA)(H^+)$ dissociation channels.[a]

species	ΔS_{trans}	ΔS_{rot}	ΔS_{vib}	ΔS_{tot}
$Gly + PMMA_{13}H^+$	162	105	-99	168
$(Gly)(PMMA_{13}H^+)$	0	0	41	40
$(PMMA_{13})(GlyH^+)$	0	0	0	0
$PMMA_{13} + GlyH^+$	162	106	-99	169
$Leu + PMMA_{13}H^+$	168	122	-97	193
$(Leu)(PMMA_{13}H^+)$	0	0	9	9
$(PMMA_{13})(LeuH^+)$	0	0	0	0
$PMMA_{13} + LeuH^+$	169	122	-102	189
$Phe + PMMA_{13}H^+$	171	127	-114	184
$(Phe)(PMMA_{13}H^+)$	0	1	17	17
$(PMMA_{13})(PheH^+)$	0	0	0	0
$PMMA_{13} + PheH^+$	171	128	-120	179

With these results, it becomes interesting to compare the differences in total entropy $\Delta(\Delta S_{tot})$ between respective competing channels.

$$\Delta(\Delta S_{tot}) = \Delta S_{tot}\ (PMMA_nH^+ + AA) - \Delta S_{tot}\ (PMMA + AAH^+) \qquad (4.21)$$

Table 4.16 compares the experimental results obtained by RRKM modeling on the difference of entropy of activation, $\Delta(\Delta^{\ddagger}S)$, with the theoretical results obtained from the modeling, $\Delta(\Delta S_{tot})$.

Table 4.16: Experimental and theoretical differences in entropy for respective competing dissociation channels.

System	Experimental $\Delta(\Delta^{\ddagger}S)$ (J mol^{-1} K^{-1})	Theoretical $\Delta(\Delta S_{tot})$ (J mol^{-1} K^{-1})
$PMMA_4$/Leu	---	-66
$PMMA_4$/Phe	---	-65
$PMMA_7$/Leu	-11	-6
$PMMA_7$/Phe	-12	-5
$PMMA_{10}$/Leu	-3	---
$PMMA_{10}$/Phe	-7	---
$PMMA_{13}$/Leu	+1	+4
$PMMA_{13}$/Phe	-2	+5

At first, it is very interesting to see that for systems where both experimental and theoretical results are available, $\Delta(\Delta^{\ddagger}S)$ resemble to the $\Delta(\Delta S_{tot})$. In terms of entropy, the modeling results seem more representative of reality. It is important to note here that the entropy of activation, $\Delta^{\ddagger}S$, represents the change in entropy between the initial complex to the transition state, whereas the difference in total entropy, $\Delta S_{tot,}$ represents the variation of entropy between the initial complex and the dissociated products. The reason why these values have a certain similarity together implies that the conformation of the transition state might be close to the conformation of the dissociated products. An interesting trend is observed for both experimental and theoretical values: as the length of the oligomer increases, the variation of entropy increases as well. In other words, longer oligomers favour the protonated dissociation channels. This tendency explains the observations in Table 4.1: in addition to the PA of the oligomer, the length of the oligomer provides a favourable entropy change to help the ($PMMA_nH^+$ + AA) dissociation channel. Although the $\Delta(\Delta S_{tot})$ for the shorter oligomers may not be quantitative (or representative), these oligomers have such a strong entropy penalty, that it helps the ($PMMA_n$ + AAH^+) dissociation channel to occur. At this point, it

seems that both energetic and entropic effects are required to explain the behaviour of these competing dissociation channels.

4.5 The proton affinity of an oligomer

The proton affinity (PA) is a gas phase thermochemical quantity defined as the negative of the enthalpy change for the gas phase transfer reaction between a proton and another chemical species [175]. Several methods such as equilibrium measurements, bracketing experiments and the kinetic method [176] are already used to accurately measure the PA of small molecules [177]. Nevertheless, what is happening for larger oligomers with the same functional group such as oligomers of PMMA? As presented before, there are two competing unimolecular dissociation channels. It is possible to write their respective rate constant as:

$$k_{PMMA_nH^+} = \frac{k_B T}{h} \frac{Q^{\ddagger}_{PMMA_nH^+}}{Q_{complex}} e^{-E_{0,\ PMMA_nH^+}/RT} \qquad (4.22)$$

and

$$k_{AAH^+} = \frac{k_B T}{h} \frac{Q^{\ddagger}_{AAH^+}}{Q_{complex}} e^{-E_{0,\ AAH^+}/RT} \qquad (4.23)$$

where: k_B, h, R, and T are Boltzmann constant, Planck constant, the gas constant, and the temperature respectively. E_0 is the 0 K activation energy and Q is the total partition function.

In the experiment, we are probing the relative rate constants for these two processes:

$$\frac{k_{PMMA_nH^+}}{k_{AAH^+}} = \frac{Q^{\ddagger}_{PMMA_nH^+}}{Q_{complex}} \frac{Q_{complex}}{Q^{\ddagger}_{AAH^+}} e^{-\Delta E_0/RT} \tag{4.24}$$

which simplifies to:

$$\frac{k_{PMMA_nH^+}}{k_{AAH^+}} = \frac{Q^{\ddagger}_{PMMA_nH^+}}{Q^{\ddagger}_{AAH^+}} e^{-\Delta E_0/RT} \tag{4.25}$$

The relative rate constants can be approximated by the relative peak intensities in the CID mass spectrum, R. At threshold, if we observe $PMMA_nH^+$ and not AAH^+ (like for the $(PMMA_4)(Gly)(H^+)$ complexes), we assume a constant peak ratio, R, of approximately 10 to 1.

$$R = \frac{Q^{\ddagger}_{PMMA_nH^+}}{Q^{\ddagger}_{AAH^+}} e^{-\Delta E_0/RT} \tag{4.26}$$

ΔE_0 in this case simply becomes the difference in PA of the oligomer and the amino acid:

$$R = \frac{Q^{\ddagger}_{PMMAnH^+}}{Q^{\ddagger}_{AAH^+}} e^{-PA/RT} \tag{4.27}$$

Each partition function in the above expression can be written out fully:

$$R = \frac{g_{PMMA_nH^+}}{g_{AAH^+}} \frac{q^{\ddagger}_{trans,\ PMMA_nH^+}}{q^{\ddagger}_{trans,\ AAH^+}} \frac{q^{\ddagger}_{rot,\ PMMA_nH^+}}{q^{\ddagger}_{rot,\ AAH^+}} \frac{q^{\ddagger}_{vib,\ PMMA_nH^+}}{q^{\ddagger}_{vib,\ AAH^+}} e^{-\Delta PA/RT} \tag{4.28}$$

where: g is the degeneracy for both competing dissociation pathways.

The translational partition function is classically represented as:

$$q_{trans} = \frac{V(2\pi m k_B T)^{3/2}}{h} \tag{4.29}$$

where: V is the volume of the complex, m is the mass of the complex, k_B is Boltzmann constant, T is the temperature, h is the Planck constant.

Previously, it was observed that there were some differences in the conformational shapes of (Leu)($PMMA_{13}H^+$) and ($PMMA_{13}$)($LeuH^+$). However, these conformational shapes were representative of one complex and not an ensemble average. In addition to the fact that the modeling provided qualitative results, it can be reasonably assumed that both molecular volumes will approximately be the same. The mass and the temperature also remain the same no matter the dissociation channel; therefore, the translational partition functions of both competing channels cancel in the equation (4.28).

The rotational partition function for a non-linear entity is classically described as:

$$q_{rot} = \frac{1}{\sigma}\left(\frac{kT}{hc}\right)^{3/2}\left(\frac{\pi}{V}\right)^{1/2} \tag{4.30}$$

where: σ is the symmetry number of the molecule which is the number of indistinguishable orientations of the complex and c is the speed of light.

For a small molecule, such as ammonia, protonation changes significantly its symmetry number. As an example, the protonation of NH_3 to NH_4^+ change its symmetry point group from C_{3V} to T_d. However, since the polymer systems treated here are fairly large and asymmetrical, their symmetry number do not change. As for the transitional partition function, the volume and the temperature do not change, therefore the rotational partition functions of both competing channels cancel in the equation (4.28).

The vibrational partition function is usually represented as:

$$q_{vib} = \sum \frac{1}{1-e^{-h\upsilon_j/k_BT}} \tag{4.31}$$

where: υ_j represent the vibrational frequencies.

Without explicit information about the transition states and their vibrational frequencies, it is a difficult to derive a quantitative value for the PA of a given oligomer. However, it is possible to derive the expected difference in PA for the two oligomeric chains. As an example, the expressions for a 7mer and the 13 mer are:

$$R = \frac{g_{PMMA_7H^+}}{g_{AAH^+}} \frac{q^{\ddagger}_{vib,\ PMMA_7H^+}}{q^{\ddagger}_{vib,\ AAH^+}} e^{-\Delta PA_{7-AA}/RT} \tag{4.32}$$

and:

$$R = \frac{g_{PMMA_{13}H^+}}{g_{AAH^+}} \frac{q^{\ddagger}_{vib,\ PMMA_{13}H^+}}{q^{\ddagger}_{vib,\ AAH^+}} e^{-\Delta PA_{13-AA}/RT} \tag{4.33}$$

The degeneracy on the transition state where the proton is bound on the amino acid can be considered as g = 1, since this is the only acceptable location on the amino acid. For the oligomer, the degeneracy is a more difficult problem to address since the proton can theoretically be bound to any carbonyl group. This is in addition to the fact that the protonated carbonyl can share the binding with another carbonyl group or two (as it was sometimes observed in the low energy conformations of $PMMA_{13}H^+$ generated by MD-SA). The bidentate or tridentate binding can occur with adjacent or non-adjacent side-chains. This combinatorial problem can be addressed using binomial coefficients.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{4.34}$$

where: $k \in \mathbb{N}$ and $k \leq n$

For the $PMMA_7$ oligomers, there are seven carbonyl groups (n = 7) and bidentate binding only (k = 1). Thus, the degeneracy for the $PMMA_7H^+$ oligomer can be evaluated by:

$$g_{PMMA_7H^+} = n \cdot \binom{n-1}{1} = 7 \cdot \binom{6}{1} = 42 \tag{4.35}$$

For the $PMMA_{13}H^+$ oligomer, with n = 13, and k = 2 (bidentate and tridentate binding possibilities), the degeneracy is evaluated by:

$$g_{PMMA_{13}H^+} = n \cdot \binom{n-1}{1} + n \cdot \binom{n-1}{2} = 13 \cdot \binom{12}{1} + 13 \cdot \binom{12}{2} = 1014 \tag{4.36}$$

Taking into consideration the degeneracies proposed, equations 4.32 and 4.33 can be rewritten as:

$$R = 42\,\frac{q^{\ddagger}_{vib,\ PMMA_7H^+}}{q^{\ddagger}_{vib,\ AAH^+}}\,e^{-\Delta PA_{7-AA}/RT} \tag{4.37}$$

and:

$$R = 1014\,\frac{q^{\ddagger}_{vib,\ PMMA_{13}H^+}}{q^{\ddagger}_{vib,\ AAH^+}}\,e^{-\Delta PA_{13-AA}/RT} \tag{4.38}$$

Dividing equation 4.37 by equation 4.38 gives:

$$24.1 = \frac{q^{\ddagger}_{vib,\ PMMA_7H^+}}{q^{\ddagger}_{vib,\,AAH^+,\ 7mer}}\,\frac{q^{\ddagger}_{vib,\ AAH^+,\ 13mer}}{q^{\ddagger}_{vib,\ PMMA_{13}H^+}}\,e^{-\Delta PA_{7-13}/RT} \tag{4.39}$$

In section 4.4.6, it was hypothesized that the transition states for the species in equation 4.39 would probably resemble those of the products. The four vibrational partition functions above were calculated using the vibrational frequencies of the products obtained for each dissociation channel. The vibrational partition functions for the 13mer channels are the same as the 7mer channels except for additional vibrational frequencies in the products. These additional vibrational frequencies come from the extra oligomer chain not involved in the transition state binding region, and will be the same for the two channels. Since they appear in the partition function as a product, they cancel out, which means that:

$$\frac{q^{\ddagger}_{vib,\ PMMA_7H^+}}{q^{\ddagger}_{vib,\ AAH^+,\ 7mer}} = \frac{q^{\ddagger}_{vib,\ PMMA_{13}H^+}}{q^{\ddagger}_{vib,\,AAH^+,\ 13mer}} \tag{4.40}$$

And equation 4.40 simplifies into:

$$24.1 = e^{-\Delta PA_{7-13}/RT} \quad (4.41)$$

Yielding:

$$\Delta PA_{7\text{-}13} = -RT \ln 24.1 \quad (4.42)$$

At the temperature at which the experiments are performed (300 K), this result suggests that there would only be a difference of 7.9 kJ·mol^{-1} between the proton affinity of a 7mer and a 13mer. This value seems rather strange at first. From Table 4.1, it was observed, at threshold dissociation energy, that there was a competition between both dissociating channels for complexes involving leucine with $PMMA_3$, $PMMA_4$, and $PMMA_7$. It is possible to argue that the PA of leucine is approximately equal to the PA of the oligomer, which is 915 kJ mol^{-1}. For the three same PMMA oligomers, complexes with phenylalanine all dissociate through the protonated amino acid channel whereas complexes with alanine all dissociate through the protonated oligomer channel. Taken together, it is possible to estimate that the PA of the three PMMA oligomers is greater than 902 kJ mol^{-1} but less than 923 kJ mol^{-1}. Based on these data, the PA of the $PMMA_3$, $PMMA_4$, $PMMA_7$ can be estimated (with standard deviation) at 915 ± 11 kJ mol^{-1}. Using a derivation similar to the one obtained for the $\Delta PA_{7\text{-}13}$, but for the $\Delta PA_{4\text{-}7}$, it is possible to re-write equation 4.42, as:

$$\Delta PA_{4\text{-}7} = -RT \ln 3.5 \quad (4.43)$$

where: the degeneracy for $PMMA_7H^+$ and for $PMMA_4H^+$ equal 42 and 12 respectively.

At the temperature at which the experiments are performed (300 K), this result suggests that there would only be a difference of 3.1 kJ·mol^{-1} between the proton affinity of a 4mer and a 7mer. In that respect, the ΔPA_{4-7} seems rather a small difference but is included within the experimental error.

Still in Table 4.1, the $(PMMA_{13})(Trp)(H)^+$ complex dissociates through two competing channels, thus, it can be argued that the PA of the $PMMA_{13}$ is approximately the same as the PA of tryptophan (949 kJ mol^{-1}). As previously, it is possible to estimate that the $PMMA_{13}$ PA will be greater than 923 kJ mol^{-1} but less than 996 kJ mol^{-1}. Based on these data, the PA of $PMMA_{13}$ can be estimated (with standard deviation) at 949 ± 37 kJ mol^{-1}. However, this value is less reliable than the first one since the standard deviation is rather large. Between 949 kJ mol^{-1} and 915 kJ mol^{-1}, the difference is 34 kJ mol^{-1}, which is not even close to the ΔPA_{7-13} = 7.9 kJ mol^{-1} calculated.

Taken together, these results suggest that the variation of PA between PMMA oligomers may not be linear on the whole series studied. This experimental observation is not unique. Methyl acetate, which has the same functional ester group as PMMA, has a PA of 822 kJ mol^{-1} [178]. Ethyl acetate, and iso-Propyl acetate have experimental PA of 836 and 837 kJ mol^{-1} respectively [178]. These PA values are too low compared to the estimated 915 ± 11 kJ mol^{-1} for the $PMMA_3$. Of course, one major difference between the previous esters and PMMA, is the number of carbonyl groups on the same molecule. This presence of several functional groups on the same molecule has an effect on the PA. As an example with alcohols: ethanol and butanol have a PA of 776 and 788 kJ mol^{-1} respectively whereas 1,2-ethanediol and 1,4-butanediol have a PA of 816 and 916 kJ mol^{-1}

respectively [178]. Considering these experimental data, not only the presence of a second functional group increase significantly the PA, but the distance separating the functional groups also has a significant effect on PA. Therefore, it is not surprising to see that a molecule having several ester functional groups (PMMA), has a greater PA. Therefore, what would partially explain the fact that shorter oligomers ($PMMA_3$, $PMMA_4$, and PMM_7) have very little difference in their PA vs. the longer oligomers ($PMMA_{13}$ and $PMMA_{17}$) probably lies in a change from bidentate to higher -dentate binding.

In the previous section, it was shown, from the modeling results, that the longer oligomers had an entropic effect in favour of the protonated oligomer dissociation pathways. However, even when considering a switch of bidentate to tridentate binding with the longer oligomers, the computational results did not reflect this change. We have seen from Tables 4.13 to 4.15 that translational and rotational contributions cancel out and the overall entropy difference depend strongly on the vibrational contribution. The vibrational frequencies have been leading us to results that do not really represent reality of the complexes studied. Therefore, to obtain meaningful ΔPA, it may imply that equation 4.39 does not simplify this way. A next step to solve this problem would be finding an efficient modeling method of higher level of theory, such as a QM/MM to obtain more realistic 3D conformations, relative energies, and vibrational frequencies. The difference in transition state moment of inertia and vibrational frequencies may not cancel out after all. In addition, to obtain more significant experimental results, dissociation of $(PMMA_n)(AA)(H^+)$ complexes involving all oligomer lengths and all 20 amino acids should be done. This should lead to more reliable values of PA for the different PMMA oligomers.

4.6 Summary

In this project, the dissociation of gas phase $(PMMA_n)(AA)(H^+)$ complexes were investigated using CID MS, RRKM modeling and MD-SA simulations. At threshold energy, it was found that two competing dissociation channels were possible: one in which the oligomer is dissociating protonated leaving a neutral amino acid ($PMMA_nH^+ + AA$), and the second in which the amino acid is dissociating protonated leaving a neutral oligomer ($AAH^+ + PMMA_n$). Two major factors influence the appearance of the dissociation channels. On one hand, there is the proton affinity of the amino acid vs. the proton affinity of the oligomer. On the other hand, the length of the oligomer itself has an influence on the dissociation channel taken. It was found that for the $(PMMA_n)(Gly)(H^+)$, there is only one dissociation channel: ($PMMA_nH^+ + AA$) whereas for the other complexes (with phenylalanine and leucine), two dissociation channels were competing depending on the length of the oligomer selected. Complexes formed with $PMMA_{13}$ always dissociated through the protonated oligomer channel, regardless of the three amino acids selected. Complete CID MS breakdown diagrams were acquired for complexes formed of each of the three amino acids (glycine, leucine, and phenylalanine) and each of the three PMMA oligomer (7mer, 10mer, and 13 mer). RRKM modeling of the $(PMMA_7)(AA)(H^+)$ breakdown curves showed that ($PMMA_nH^+ + AA$) dissociation channels had lower E_0 but higher $\Delta^{\ddagger}S$ whereas the ($PMMA_nH^+ + AA$) dissociation channels (for leucine and phenylalanine) had higher E_0 but smaller $\Delta^{\ddagger}S$. RRKM modeling of the $(PMMA_{10})(AA)(H^+)$ breakdown curves showed that ($PMMA_nH^+ + AA$) dissociation channels had higher E_0 but lower $\Delta^{\ddagger}S$ whereas the ($PMMA_nH^+ + AA$) dissociation channels (for leucine and phenylalanine) had lower E_0 but higher $\Delta^{\ddagger}S$. RRKM modeling of the $(PMMA_{10})(AA)(H^+)$ breakdown curves showed mix results as to the relative E_0 and $\Delta^{\ddagger}S$, however, the α value for these complexes was 300 vs. 200 for the systems involving 7mer and 10mer. As

shown in our group previously [174], the α value is directly proportional to the cross-section of the complex and inversely proportional to the vibrational heat capacity of the parent ion complex. Further ion-mobility experiments will be required to understand this sudden increase in α value.

In order to obtain information on the 3D conformations, as well as their energetic and entropic effect, MD-SA simulations were carried on the 4mer, 7mer and 13mer systems. Three parameters of the SA protocol were first optimized. First, it was found that for the 4mer and 7mer systems, 1500 SA cycles were sufficient vs. 2000 cycles for the 13mer systems. Second, the mid-cycle temperature had to be increased depending on the length of the oligomer: 700 K for the 4mer systems, 900 K for the 7mer systems, and 1100 K for the 13mer systems. Finally, a slower damping temperature schedule of 10 K ps^{-1} was preferred to allow better conformational sampling. The modeling of the gas phase complexes provided useful insights on the 3D conformations. For each system, it was found that $(PMMA_n)(AAH^+)$ were always having lower energies compared to their respective $(AA)(PMMA_nH^+)$ complexes. In the majority of cases where the oligomer was protonated, it was found that it was more energetically stable for the protonation site to be in the middle of the oligomer. Unfortunately, the level of theory used was not sufficient to obtain a correlation between the experimental and the theoretical results. Solely based on the modeling results, it is not possible to fully explain the competition between both dissociation channels. The only qualitative trend that could be observed is that there seems to exist a competition between the energy levels of $(AA)(PMMA_nH^+)$ and $(PMMA_n + AAH^+)$: as the oligomer length increases, it becomes energetically favourable to dissociate through the protonated oligomer channel.

Using the vibrational frequencies and the rotational constants from the MD-SA or AM1 calculations, it was possible to evaluate the ΔS_{tot} for each dissociation channel. The translational and rotational contributions to ΔS_{tot} canceled out between both dissociation channels. The main contribution to the ΔS_{tot} came from the vibrational contribution. In order to compare the theoretical entropy results obtained, the $\Delta(\Delta S_{tot})$ were calculated for the 7mer and 13mer systems, and reasonable agreement was obtained when compared the $\Delta^{\ddagger}S$ calculated by RRKM theory. These similarities suggest that even if the 3D conformations are not completely accurate, the transition states of these dissociation channels are probably close the structures of the products. Finally, as a follow up on another project initiated in our group, the question was raised as to what is the PA of a PMMA oligomer? Using the results shown in Table 4.1, it was possible to provide an estimate that the PA of the 3mer, 4mer and 7mer was approximately 915 ± 11 kJ mol^{-1}, whereas theoretically, a ΔPA_{4-7} of 3.1 kJ mol^{-1} was calculated. This calculated difference in PA is relatively small, but is included inside the experimental error. For the 13mer, the experimental PA was estimated to be 949 ± 37 kJ mol^{-1} for a ΔPA_{7-13} of 34 kJ mol^{-1}. This difference is greater than the theoretical ΔPA_{7-13} of 7.9 kJ mol^{-1} (even if this result was considering a change from bidentate binding for the 7mer to tridentate binding for the 13mer). In order to obtain better experimental $PMMA_n$ PA values, ion-mobility experiments will have to be conducted on several $(PMMA_n)(AA)(H^+)$ complexes. On the computational level, a higher level of theory will have to be used to obtain refined 3D conformations and hence, better energetics, rotational constants, and vibrational frequencies.

Chapter 5: Investigation of the non-covalent interactions between anti-amyloid agents and Amyloid β peptides by ESI-MS [179]

5.1 Introduction

It was hypothesized that anti-amyloid agents (LMW aminosulfonate molecules) could alter the aggregation process of Aβ peptides if they first have the ability to bind to them. To do so, a reliable ESI-MS moderate throughput assay was required to determine the binding affinity of LMW molecules to different Aβ peptides. In this section, we discuss the unique charge motif and common features shared by the Aβ peptides, the binding affinity (BA) as a method to classify the binding strength, and well as the LMW molecules chemical requirements to bind. All results were obtained using the ESI-MS moderate throughput assay.

5.2 Discussion

Aβ1-40 and Aβ1-42 peptides share an identical amino acid sequence except for two residues, Ile and Ala, added to the C-terminal end of the latter (Figure 3.1). Aβ1-28 peptide is a C-terminal-truncated version of the Aβ1-40/42 peptides containing only the first 28 AA residues from the N-terminal. The C-terminal part of the Aβ1-40/42 peptides (AA 29-40/42) is the hydrophobic part. All charged or polar side-chains are located in the first 28 residues of the N-terminal. The charge distribution in the peptides has a unique pattern: a positive charge following two negative charges at pH 7.4 ± 0.2 (assuming that histidines remain uncharged under such conditions). The charge motif is repeated three times from residue 1 (D) to residue 28 (K). Under experimental conditions at physiological pH, the mass spectra of the three peptides (Aβ1-42, Aβ1-40, and Aβ1-28) show

typical peaks corresponding to species having 4 and 5 positive charges. In certain cases a 3 positive charge region is visible as well. Each charged species appears as a peak cluster due to the various number of sodium ions associated to the peptide. Consequently, when a LMW compound is bound to a peptide, the non-covalent complex also appears as a peak cluster at the corresponding m/z ratio, due to the association of the complex with various numbers of sodium ions. To evaluate the BA of a compound to the Aβ peptide, the following equation 5.1 was used:

$$\mathrm{BA}(\%) = \frac{\sum \text{complex peaks height}}{\sum(\text{peptide peaks height} + \text{complex peaks height})} \times 100 \qquad (5.1)$$

In equation 5.1 the sums are over all ionic species present in the mass spectrum for the complex or peptide (all peaks have the same widths). For purposes of reproducibility, the average binding (mean ± S.D.) was calculated from six measurements. For a result to be accepted, an S.D. ≤ 10% was required. Other groups have developed methods to calculate an association constant (K_A) [78, 97, 102, 180]. Although more accurate, the calculation of K_A can be time-consuming for a first-line screening method. It is worth mentioning that our assay uses much less compound and peptide to observe complex formation than in a method previously described [69]. Previously we have shown [75] that, the maximum BA between a given compound and the Aβ1-40 peptide rises from pH 1 to approximately pH 6 and then decreases with further increase of pH. Hence, the maximum BA is close to the physiological pH. In this assay, the Aβ1-40 peptide was selected over the Aβ1-42 and Aβ1-28 peptides for two reasons: first, the Aβ1-40 peptide is less prone to aggregation and has a better water solubility than the Aβ1-42 peptide [50] and second, it contains a C-terminal hydrophobic region (absent in the Aβ1-28 peptide) which makes it more physiologically relevant.

To further justify the selection of the Aβ1-40 peptide for this assay, it was required to verify if this peptide could differentiate the binding strength of various compounds and if the BA ranking order remained consistent whichever peptide (Aβ1-28, Aβ1-40, or Aβ1-42) was used. Table 5.1 summarizes the binding results between representative compounds and the three Aβ peptides.

Table 5.1: Binding affinity for compounds 1-4 with Aβ1-42, Aβ1-40, and Aβ1-28.

Compound	BA (%)		
	Aβ1-42	Aβ1-40	Aβ1-28
1	54 ± 3	49 ± 1	49 ± 1
2	38 ± 2	43 ± 2	41 ± 4
3	28 ± 3	33 ± 1	33 ± 3
4	14 ± 1	13 ± 1	9 ± 1

For a visual examination of binding complexes, Figure 5.1 compares the binding profiles of the four selected compounds to the Aβ1-42 peptide. The m/z region from 900 to 1125 contains the peak-clusters of the peptide and the complexes bearing 5 positive charges (+5-region). The m/z region from 1125 to 1400 contains the species having 4 positive charges (+4-region). The m/z region corresponding to species having 3 positive charges (+3-region) is not shown on the spectra. On all mass spectra, the peak-clusters centered approximately at m/z 925, 1151, and 1526 (not shown) are the free peptide having +5, +4, and +3 charges, respectively. In each region, to the right side of the free peptide peak cluster are the 1:1 compound-peptide complex (first peak cluster), the 2:1 compound-peptide complex (second peak cluster), and the 3:1 compound-peptide complex (third peak cluster). This observation indicates, at least in some cases, that more than one molecule of a compound can bind to one peptide ion. Employing the intensities of each peak in equation 5.1 allows the calculation of BA. Compounds in Table 5.1, represent four different categories (strong,

BA > 45%; moderate, 35 ≤ BA< 45%; weak, 20 ≤ BA < 35%, and non-binding, BA < 20%). One aspect of ESI that hinders the quantitative assignment of the BA in solution is the tendency to form non-covalent complexes in the gas phase during the electrospray process. The compounds in Table 5.1 are all structurally similar and thus differences between solution and gas phase behaviour should also be conserved across the compound library, making comparisons meaningful.

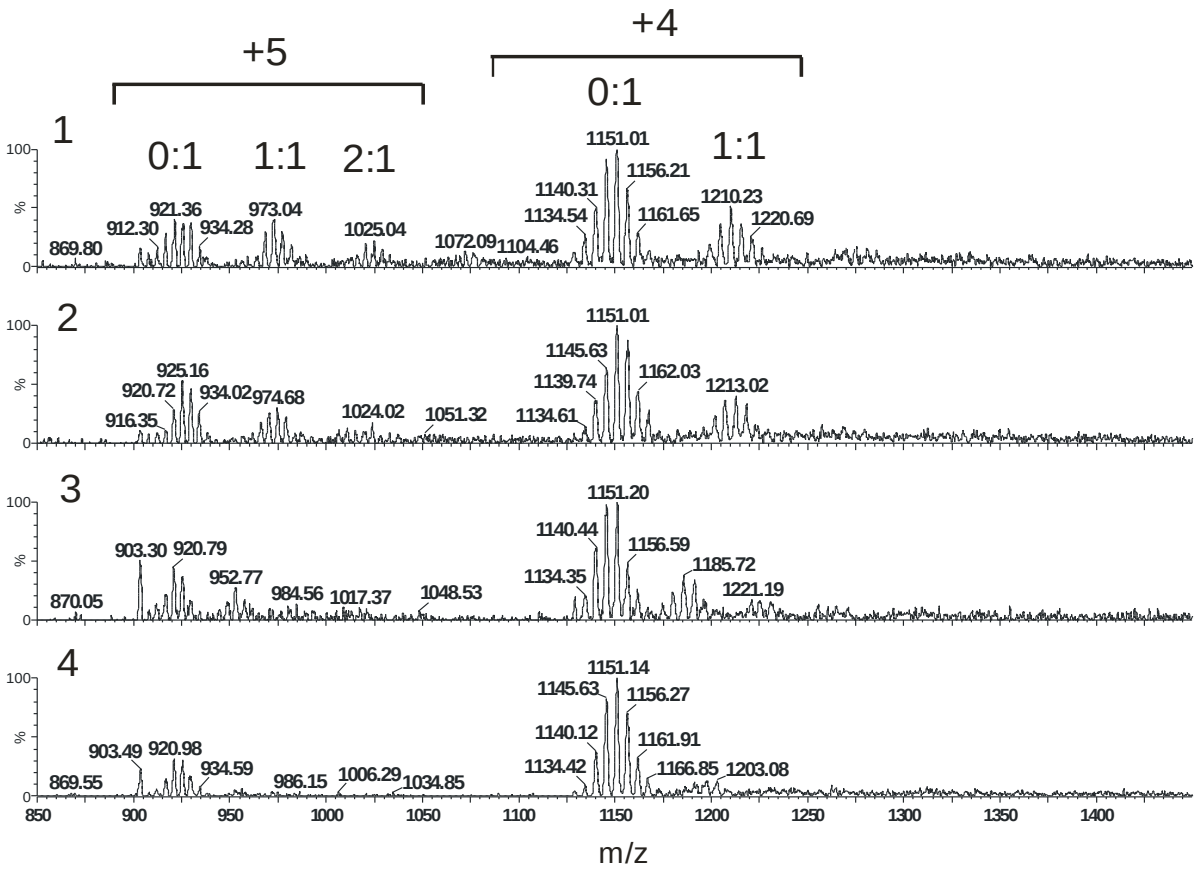


Figure 5.1: Binding profiles of compounds 1-4 to Aβ1-42. +5 and +4 charge regions denoted, along with peak clusters due to the protonated peptide (0:1), 1:1 compound: peptide complex and 2:1 compound: peptide complex.

As shown in Table 5.1, the BA ranking order remains the same for each given compound to all three peptides. Furthermore, the "absolute binding affinity" (as measured in binding percentage)

for a specific compound is similar when tested against all three peptides. These observations indicate that the C-terminal hydrophobic part of the Aβ peptide does not play a significant role in binding these LMW compounds, suggesting therefore that the N-terminal part is mostly responsible for the non-covalent binding. The non-covalent complex formation occurs through dipole– dipole or charge–charge interactions [97]. In this assay, we have chosen to use Aβ1-40 as the peptide for screening purposes since it is the least expensive.

During method development, another important consideration was given to ensure that these low structural complexity compounds were binding specifically to Aβ peptides and not to any peptide in general. At physiological pH, the N-terminal sequence of the Aβ peptide possesses a specific charge motif, which may play an important role in the peptide binding to LMW compounds. To test this hypothesis, two peptides (all (-), all (+), each having 28 AA residues; see Figure 3.1) were synthesized and studied for their binding behaviour. At physiological pH the "all (-) peptide" has three negatively charged residues in the sequence and no positive charges (except for the N-terminal residue) and the "all (+) peptide" has three positively charged residues. Table 5.2 compares the BA of compounds 1-4 with these two peptides and Aβ1-28. When comparing the BA of compounds 1-4 with all three peptides having 28-AA residues, the most striking observation is that none of them bind significantly to the all (+) peptide (Figure 5.2). One important feature of the mass spectra is the absence of sodium cluster peaks for this system because there are no side-chains charged negatively under experimental conditions. In the all (+) peptide, all these AA residues have been substituted by neutral ones; thus, there may be one (or a very few) sodium cations that are bound to the species present on these mass spectra.

Table 5.2: Binding affinity for compounds 1-4 with peptides of different charge motif.

Compound	BA (%)		
	Aβ1-28 (native)	all (-)	all (+)
1	49 ± 1	35 ± 2	30 ± 2
2	41 ± 4	33 ± 3	24 ± 1
3	33 ± 3	16 ± 2	0
4	9 ± 1	6 ± 1	0

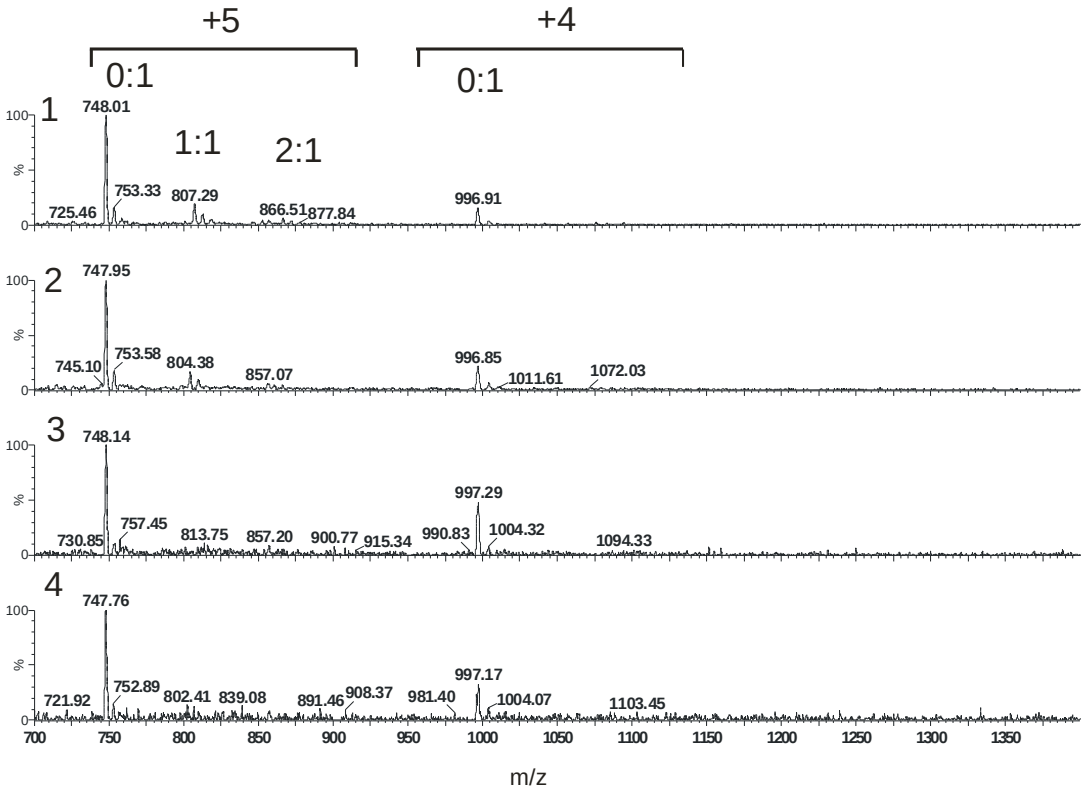


Figure 5.2: Binding profiles of compounds 1-4 to the all (+) peptide. +5 and +4 charge regions denoted, along with peak clusters due to the protonated peptide (0:1), 1:1 compound: peptide complex and 2:1 compound: peptide complex.

Figure 5.3 shows the binding profile of the four compounds to the all (-) peptide. The peak-cluster centered at m/z 755 represents the free peptide having four positive charges, whereas the one at m/z 999 is the +3 charged species of the peptide. In this case, the BA was decreased by at least

25% with all four compounds compared with the native peptide. Interestingly, the decrease of the BA is not proportional to the lower number of negative charges in the peptide. When comparing observations made with the all (+) peptide, it may be concluded that binding of these compounds is dominantly mediated by the negatively-charged side chains in the Aβ peptide.

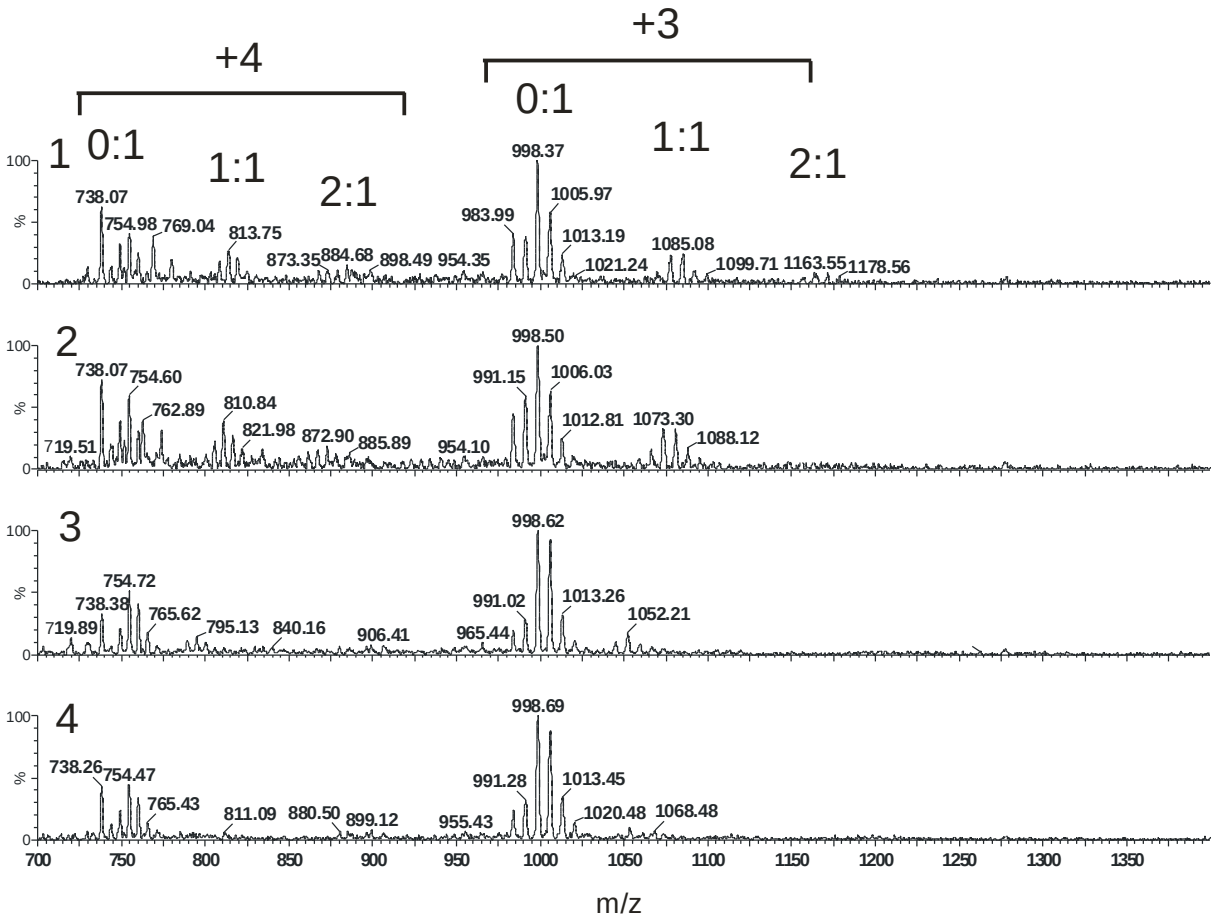


Figure 5.3: Binding profile of the four compounds to the all (-) peptide. The peak-cluster centered at m/z 755 represents the free peptide having 4 positive charges, whereas the one at m/z 999 is the +3 charged species of the peptide.

The BA of the Aβ1-40 peptide with compounds **1-4** was compared to their binding with the Aβ3pE-40 peptide under exactly same conditions. Aß3pE-40 was shown to be very amyloidogenic and can potentially initiate the amyloid aggregation process. Due to the low solubility of Aβ3pE-40 in aqueous media, experiments for both peptides were performed in a mixture of water (80%) and ethanol (20%). The data are summarized in Table 5.3 and the mass spectra are given in Figure 5.4. As shown in Table 5.3, the BA with the Aβ1-40 peptide increases by 5 to 10% for compounds **1-4** when the solvent contains 20% ethanol. Their BA with the Aβ3pE-40 peptide was approximately 5 to 10% lower compared to Aβ1-40, but the ranking order stayed the same. The lower BA with the Aβ3pE-40 peptide might be due either to removal of two negative charges from the N-terminal or to disappearance of a charge motif after the N-terminal truncation of Aβ1-40 peptide.

Table 5.3: Binding affinity for compounds 1-4 with Aβ3pE-40 and Aβ1-40 in 20% ethanol.

Compound	BA in 20% ethanol (%)	
	Aβ1-40	Aβ3pE-40
1	57 ± 1	49 ± 1
2	52 ± 1	39 ± 1
3	37 ± 1	30 ± 2
4	22 ± 1	15 ± 1

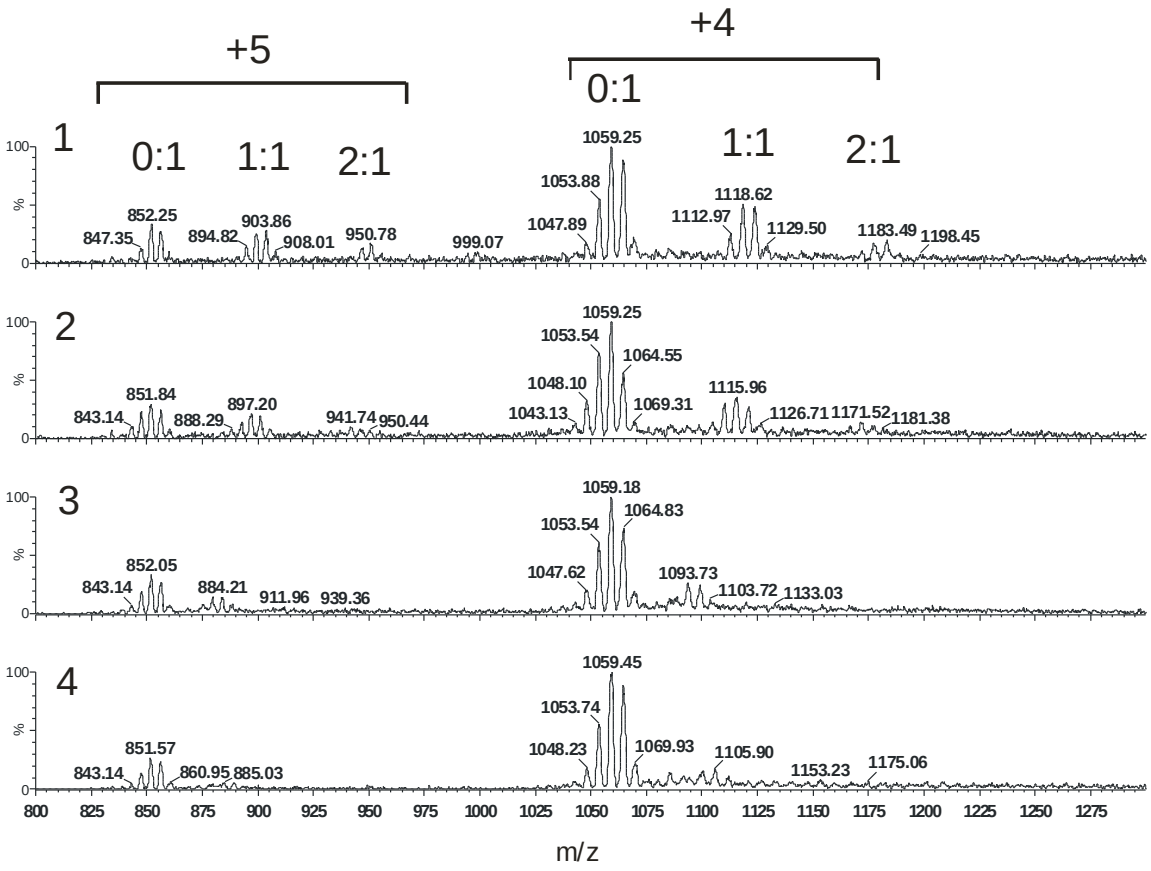


Figure 5.4: Binding profiles of compounds 1-4 to the Aβ3pE-40. +5 and +4 charge regions denoted, along with peak clusters due to the protonated peptide (0:1), 1:1 compound: peptide complex and 2:1 compound: peptide complex.

Effort was also directed to the structure-activity relationship (SAR) for binding of the Aβ1-40 peptide with LMW compounds. Homotaurine (compound 3, Vivimind^{TM}) aims at targeting soluble Aβ peptide and was used as a reference point for comparing the effect of each component in the structure: the amino group, the sulphonate group, and the alkyl spacer in between. Table 5.4 reports the effect of different substitutions on the sulfonic group. None of the substitutions shown in Table 5.4 is advantageous to maintain BA. Only the phosphonate substituent retained significant binding. It is important to mention here that GABA, a bioisostere of homotaurine and an important

neurotransmitter present in the human brain, does not bind at all to soluble Aβ1-40 peptide. Table 5.5 reports the effect of amine substitution of compound 3. Again in this case, none of the reported substitutions was advantageous to maintain BA. Only methyl substitution was able to preserve some binding. Eprodisate (compound 10), an investigational clinical drug candidate targeting serum amyloid A [181], is selective for this peptide and does not bind to the Aβ1-40 peptide. Finally, the length of the alkyl spacer was investigated using five different chain spacers tested against the Aβ1-40 peptide (Table 5.6). The alkyl spacer does not play a significant role in maintaining BA to the Aβ1-40 peptide, although it seemed that the homotaurine three-carbon spacer between amino and sulfonate groups was optimal. The lack of sensitivity of BA to the length of the carbon spacer might be due to the flexibility of the linear structure in such simple compounds. Altogether, this information on the binding components of homotaurine show that even though it is not a strong binder to Aβ peptides, it contains key elements to be considered as an anti-amyloid agent. The mechanism of action of homotaurine for the treatment of AD has been extensively studied, *in vitro* and *in vivo*, and results will be reported separately.

Table 5.4: Effect of replacing the sulfonic group on the binding affinity with Aβ1-40.

Compound	H_2N-$(CH_2)_3$-R	BA to Aβ1-40 (%)
3	-SO_3H	33 ± 1
5	-PO_3H_2	24 ± 1
6	-CO_2H	0
7	-CH_3	0

Table 5.5: Effect of replacing the amino group on the binding affinity with Aβ1-40.

Compound	R-$(CH_2)_3$-SO_3H	BA to Aβ1-40 (%)
3	-NH_2	33 ± 1
8	-CH_3	26
9	-OH	0
10	-SO_3H	0

Table 5.6: Effect of changing the length of the methylene group spacer chain on the binding affinity with Aβ1-40.

Compound	H_2N-$(CH_2)_n$-SO_3H	BA to Aβ1-40 (%)
11	n = 2	29 ± 2
3	3	33 ± 1
12	4	30 ± 3
13	5	29 ± 1
14	6	27 ± 2

5.3 Summary

An efficient and reproducible ESI-MS binding assay was developed for semi-quantitative evaluation of the binding affinity of LMW compounds to Aβ peptides. The assay proved to be consistent on the binding affinity (BA) ranking from one peptide to another when comparing the non-covalent complexes formed with the same set of four compounds with Aβ1-28/40/42. This allowed the use of peptide Aβ1-40 in this first-line screening assay. It was demonstrated that mutation or scrambling of the Aβ peptide sequence affects the BA of the four targeted LMW compounds with the peptides. A similar effect was observed when the four LMW compounds were tested against Aβ3pE-40. The results indicate that the binding between Aβ and a LMW compound

is likely through charge-charge interaction, and the negative charge motifs in the peptide are important to such. The results obtained using this assay indicated that homotaurine contains the necessary molecular properties for its binding to the soluble Aβ peptide, and could interfere with the amyloidogenic cascade. In this study, BA obtained with the assay can be used as the first-line screening to select LMW compounds for further testing in biological assays.

Chapter 6: Gas phase binding energies for non-covalent Aβ1-40 peptide / small molecule complexes from CID mass spectrometry and RRKM theory [182]

6.1 Introduction

Although RRKM theory can be applied to larger systems, computational facilities may limit its application. The Aβ–substrate complexes investigated here have over 1800 vibrational modes and at internal energies greater than ~9 eV, ρ(E) takes on values $>10^{308}$, the double precision floating point limit for 64-bit compilers. So the exact RRKM solution can only be run on 64-bit with quadruple precision processors.

The modeling of such a large system with RRKM theory is inhibited by the need to calculate the density of states of the system at high enough internal energies. To circumvent this problem we outline an approximate method that involves extrapolating the form of the post-collision internal energy distribution, $P(E,E_{CM})$, and microcanonical rate constant k(E), the latter by fitting a 4-parameter Hill equation to those values which are computationally accessible. The procedure is tested on a polymer system and finally applied to the Aβ1-40/homotaurine complex. The $P(E,E_{CM})$ for the Aβ–substrate complex was calculated at temperatures below 300 K for which it was possible to obtain ρ(E). These distributions were then fit in the Igor program (Igor Pro 4.05A, Wavemetrics Inc., Oregon, USA) to Gaussians of the form:

$$P(E, E_{CM}) = j_0 + j_1 e^{-\left(\frac{E-j_2}{j_3}\right)^2} \quad (6.1)$$

A plot was then made of each 'j' parameter as a function of internal energy and the results fitted with a 2nd order polynomial to yield:

$$j_0 = 5 \times 10^{-13}B^2 - 9 \times 10^{-10}B + 9 \times 10^{-7} \quad (6.2)$$

$$j_1 = 3 \times 10^{-8}B^2 - 6 \times 10^{-5}B + 0.0349 \quad (6.3)$$

$$j_2 = 3 \times 10^{-6}B^2 + 0.0033B + 0.7475 \quad (6.4)$$

$$j_3 = 1 \times 10^{-7}B^2 + 0.0005B + 0.1421 \quad (6.5)$$

where $B = \beta E^{1/2}$. These expressions at least provided visually well-behaved extrapolations to higher T_{eff} values. β is essentially the same as in equation 3.5, governing the increase in $P(E,E_{CM})$ with increasing E_{CM} and is an adjustable parameter in the modeling of the breakdown curves. Finally, once $P(E,E_{CM})$ is calculated it is normalized by dividing by the integrated total population. To extrapolate k(E) above to high internal energies, the chosen method was to fit calculated k(E) values below 9 eV internal energy with a four parameter Hill equation (Igor Pro 4.05A, Wavemetrics Inc., Oregon, USA):

$$k(E) = a_1 + \frac{(a_2 - a_1)}{\left[1 + \left(\frac{a_3}{E}\right)^{a_4}\right]} \quad (6.6)$$

This equation is then used to extrapolate k(E) to higher internal energies. To test the utility of this extrapolation, we chose two smaller systems for which k(E) could be calculated up to 30 eV internal energy completely by equation 3.7. A comparison could then be made for Hill equation extrapolations based on a variable number of fitted k(E) points. The systems chosen have been

studied in our group previously and consist of PMMA oligomers bound in the gas phase to protonated glycine (GH^+), specifically a 7mer and 13mer of PMMA [26].

To develop the approximate RRKM method, the procedure was to calculate 900 k(E) values between 0 and 30 eV internal energy for the dissociation of each complex using equation 3.7. For the complex involving the 7mer, the Hill equations were used to fit the lowest 250, 400, 600 and 900 k(E) values. For the 13mer, the lowest 250, 400, 600 and 900 values were fit. The results are presented in Figure 6.1. Note that equation 6.6 is only used for the extrapolated k(E) values. There are several observations to be made from Figure 6.1. First is that when equation 6.6 is used to fit all of the k(E) points obtained from equation 3.7, the fit is excellent. Agreement gets worse as the number of points on which the fit is established decreases, as is to be expected. When only the lowest energy 250 points are used, there can be a three order of magnitude difference in k(E) at high internal energies. However, the real utility of this procedure does not depend on how well equation 6.6 approximates k(E) over the entire internal energy range calculated, but rather on how well the resulting theoretical breakdown diagram approximates that derived from a rigorous calculation of k(E) with equation 3.7.

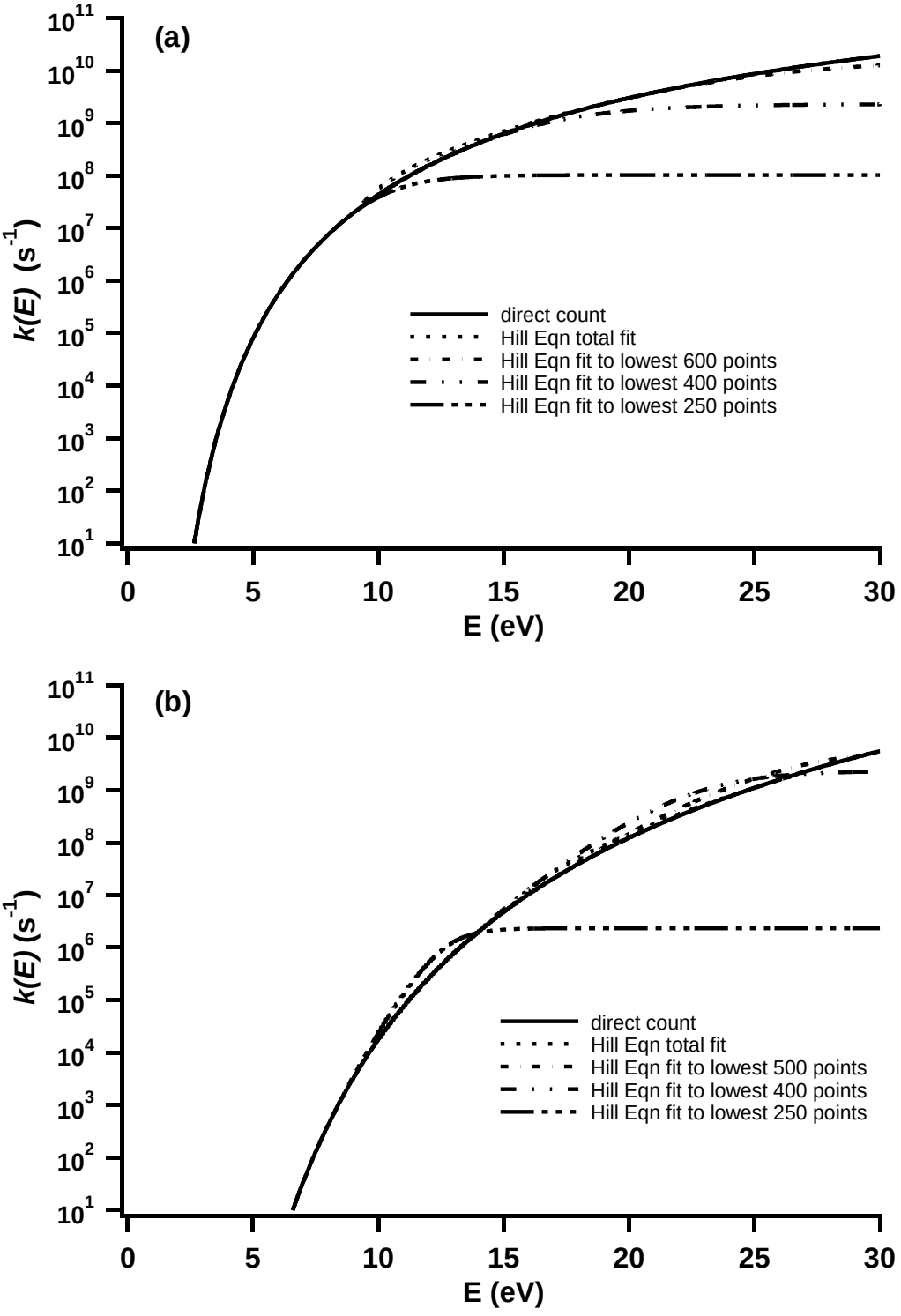


Figure 6.1: Plot of k(E) vs. E based on the direct count of densities and sums of states, and several approximate treatments based on the Hill Equation extrapolation for (a) $(PMMA_7)(GlyH^+)$ and (b) $(PMMA_{13})(GlyH^+)$.

The experimental breakdown curves for the dissociation of the $(PMMA_7)(GlyH)^+$ complex are shown in Figure 6.2. In symbols are the experimental data points and the theoretical curves for the equation 3.7 model, and equation 6.6 models based on the lowest energy 250, 400 and 600 k(E) points are shown. In all cases E_0 = 0.82 eV, $\Delta^{\ddagger}S$ = -14 J K^{-1} mol^{-1} and α = 200. The curves are indistinguishable. This is due to the fact that while there is significant deviation between equations 3.7 and 6.6 results at the higher internal energies in Figure 6.2, the experimental breakdown curves are dominated by k(E) values at much lower internal energies, where the agreement between the two equations is satisfactory. The same result was found for the 13mer complex (Figure 6.2b), although now a small difference can be seen when only 250 points are used for the extrapolation. In the Results and Discussion sections this approximate approach for the dissociation of the Aβ1-40/homotaurine complex in the 4+ charge state will be compared with the rigorous calculation.

6.2 Results

Table 3.6 shows the five aminosulfonate small molecules that were used to complex the Aβ1-40 peptide: homotaurine (M1), 3-(cyclohexylamino)propane-1-sulfonic acid (M2), 3-Cyclohexylamino-2-hydroxy-1-propanesulfonic acid (CAPSO, M3), 3-(1,3,4,9-tetrahydro-2*H*-β-carbolin-2-yl) propane-1-sulfonic acid (M4) and 3-(1,3,4,9-tetrahydro-2*H*-β-carbolin-2-yl)butane1-sulfonic acid (M5).

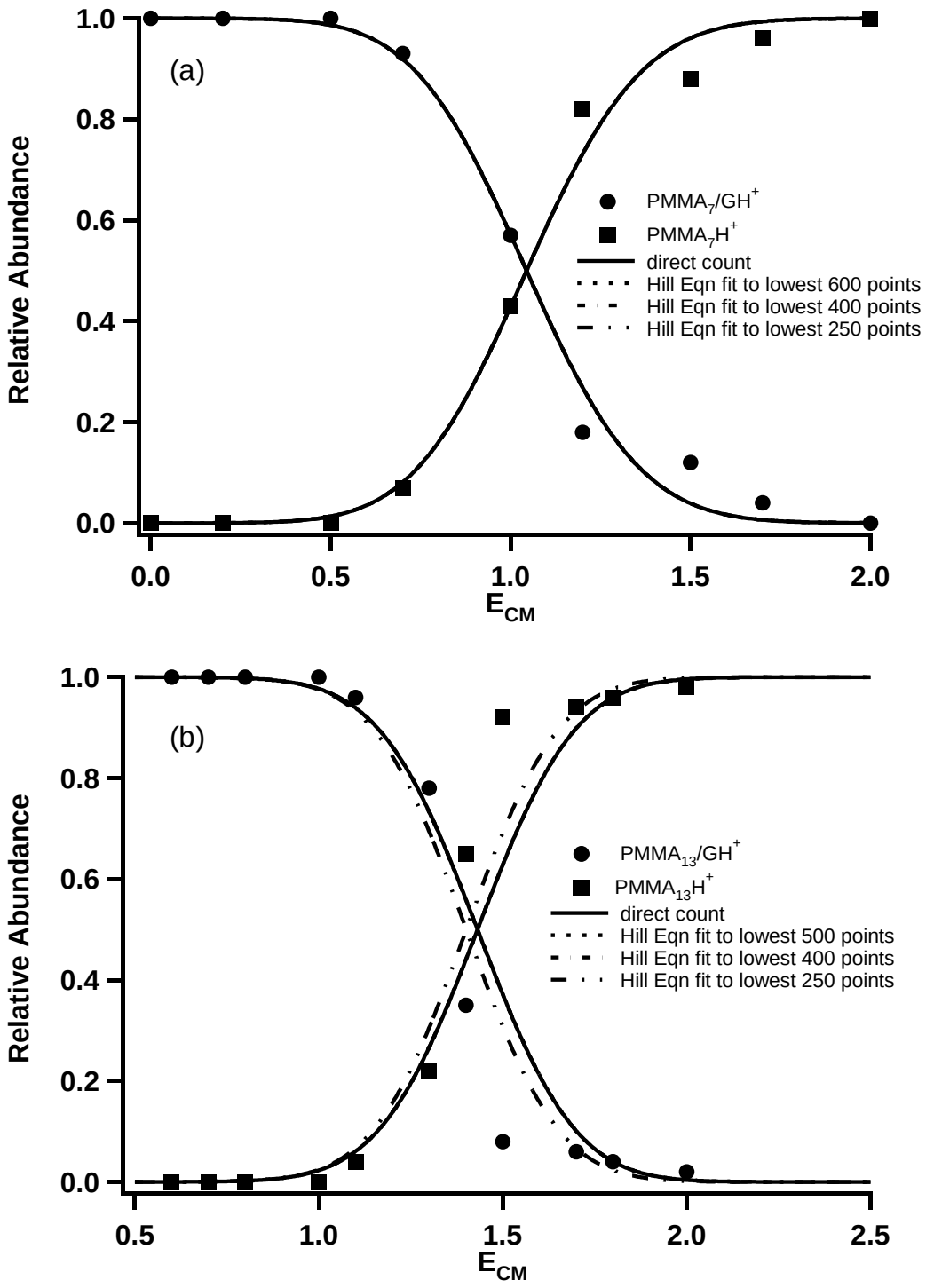


Figure 6.2: Calculated breakdown diagrams for (a) ($PMMA_7$)($GlyH^+$) and (b) ($PMMA_{13}$)($GlyH^+$). In each case the curves were calculated with either the rigorous RRKM treatment or the stated Hill equation extrapolation of k(E).

6.2.1 Dissociation of the Aβ1-40/homotaurine complex

The complex selected for this study consists of the Aβ peptide, one molecule of homotaurine and seven sodium ions, hereafter denoted $[A\beta\text{-M1-7Na}]^{4+}$ (m/z 1156.4). The principal dissociation pathway is the loss of a neutral homotaurine molecule to form $[A\beta\text{-7Na}]^{4+}$ (m/z 1121.4). The experimental breakdown curve is presented in Figure 6.3. The solid lines in Figure 6.3 represent the best theoretical fits to the experimental data. Employing the full direct count calculation for the sums and densities of states yields a best fit of E_0 = 0.681 ± 0.002 eV, $\Delta^{\ddagger}S_{300K}$ = 3.0 ± 0.5 J mol^{-1} K^{-1} and α = 207 ± 1. There is however significant interplay between $\Delta^{\ddagger}S_{300K}$ and α, as they both primarily influence the rate of change of the parent and daughter ion curves. Based on a range of visually acceptable fits, more reasonable error limits are E_0 = 0.681 ± 0.005 eV, $\Delta^{\ddagger}S_{300K}$ = 3 ± 3 J·mol^{-1}·K^{-1} and α = 207 ± 10 (Table 6.1). Again, these error limits relate simply to the fitting of the experimental data and do not take into account inaccuracies resulting from the assumptions in the model used. Apparent from Figure 6.3 is that the fit to the high E_{CM} portion of the breakdown curve is not very good. The dissociation of a molecule from the complex may be governed by a variational transition state which can change character with increasing ion internal energy. Typically this variational transition state moves to a configuration having a tighter geometry as internal energy increases and thus is characterized by a less positive (or more negative) $\Delta^{\ddagger}S_{300K}$. Also presented in Figure 6.3 is the result of the approximate, Hill Equation, fit using the same E_0 and $\Delta^{\ddagger}S_{300K}$ and a β value of 56. Agreement between the two approaches is very good. By coincidence, using the Hill Equation to extrapolate k(E) results in a k(E) vs. E curve that resembles that which one would get by introducing a variational transition state as discussed above (see for example Figure 6.1). As will be seen below, this is the only system that shows this behaviour. All of the other breakdown curves are for dissociations leading to two product ions, a process which

should have associated with it a reverse activation barrier, if only a small one, and thus a discrete transition state.

Table 6.1: Parameters of the best fit for each dissociated complex.[a]

complex	E_0 (± 0.005 eV)	$\Delta^{\ddagger}S$ (± 3 J mol^{-1} K^{-1})	α (± 10)
[Aβ-M1-7Na] $^{4+}$ (m/z 1156)	0.681	3	207
[Aβ-M2-7Na] $^{4+}$ (m/z 1177)	0.655	1	200
[Aβ-M3-8Na] $^{5+}$ (m/z 949.6)	0.675	4	207
[Aβ-M4-7Na] $^{5+}$ (m/z 956.5)	0.625	5	195
[Aβ-M4-8Na] $^{5+}$ (m/z 960.8)	0.625	0	190
[Aβ-M4-6Na] $^{4+}$ (m/z 1189.9)	0.625	-5	180
[Aβ-M5-9Na] $^{5+}$ (m/z 968.5)	0.625	-1	204
[Aβ-M5-7Na] $^{4+}$ (m/z 1199)	0.625	0	200

[a]MX represents the small molecule number as presented in Table 3.6

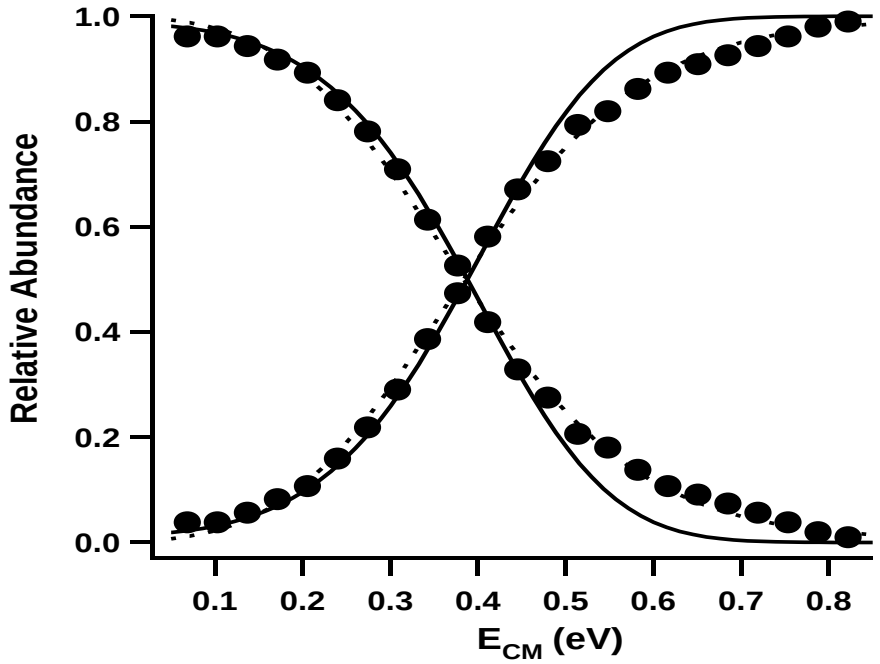


Figure 6.3: Theoretical best fit (solid curves) to the experimental (points) breakdown curve for [Aβ-M1-7Na]$^{4+}$. The result of the Hill equation approximation is shown in dashed lines.

6.2.2 Dissociation of the Aβ1-40/3-(cyclohexylamino) propane-1-sulfonic acid complex

The complex selected for this study consists of the Aβ1-40 peptide, one molecule of M2 and seven sodium ions in the 4+ charge state, hereafter denoted $[Aβ\text{-}M2\text{-}7Na]^{4+}$ (m/z 1177). The primary dissociation is the loss of a sodiated M2 molecule to form $[Aβ\text{-}6Na]^{3+}$ (m/z 1488). Both fragmentation products are thus observed in the mass spectrum. Theoretically, both product ions should have identical relative abundance, but in practice, they do not due to mass discrimination in the time-of-flight detector. In order to minimize mass discrimination effects and ion scattering, we have modeled the breakdown curve using only the $[Aβ\text{-}6Na]^{3+}$ product ion (Figure 6.4). This holds for all breakdown curves in this study where both product ions are observed. M2 introduces a functionalized amino group, presumably making it less available for binding to the peptide. This modification only lowers the binding energy by ~ 3 $kJ \cdot mol^{-1}$ (Table 6.1).

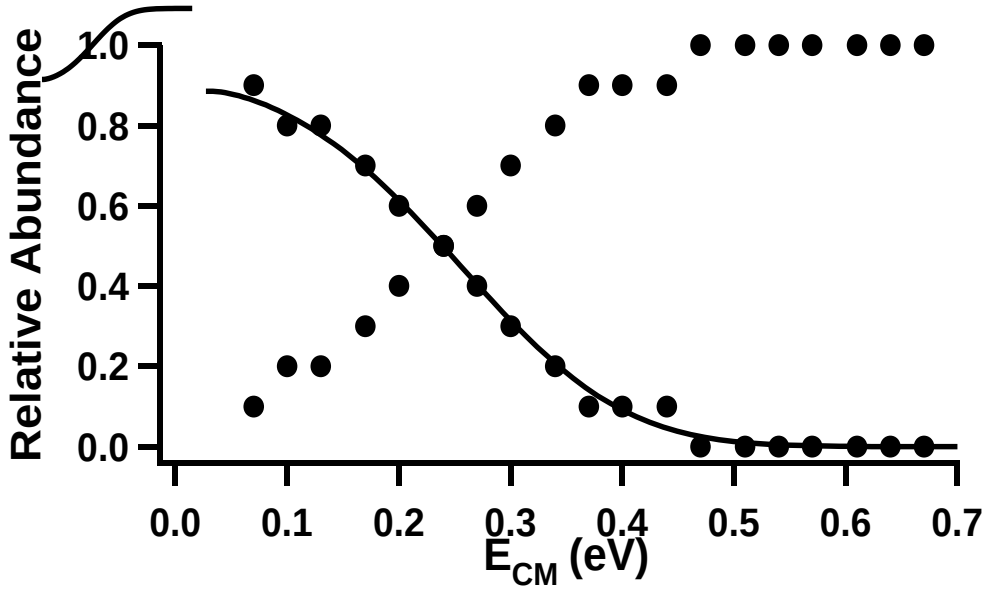


Figure 6.4: Theoretical best fit (solid curves) to the experimental (points) breakdown curve for $[Aβ\text{-}M2\text{-}7Na]^{4+}$.

6.2.3 Dissociation of the Aβ1-40/cyclohexylamino-2-hydroxy-1-propanesulfonic acid (CAPSO) complex

The complex selected for this study consists of the Aβ peptide, one molecule of CAPSO and eight sodium ions in the 5+ charge state, hereafter denoted $[Aβ\text{-}M3\text{-}8Na]^{5+}$ (m/z 949.6). CAPSO introduces a functionalized amino group and a hydroxyl group along the aliphatic chain (Table 1). The principle dissociation pathways is the loss of a sodiated CAPSO molecule to form $[Aβ\text{-}7Na]^{4+}$ (m/z 1121.9), the same product ion obtained from the homotaurine complex in section 6.2.1. The experimental breakdown curve is presented in Figure 6.5 along with the best theoretical fit to the data.

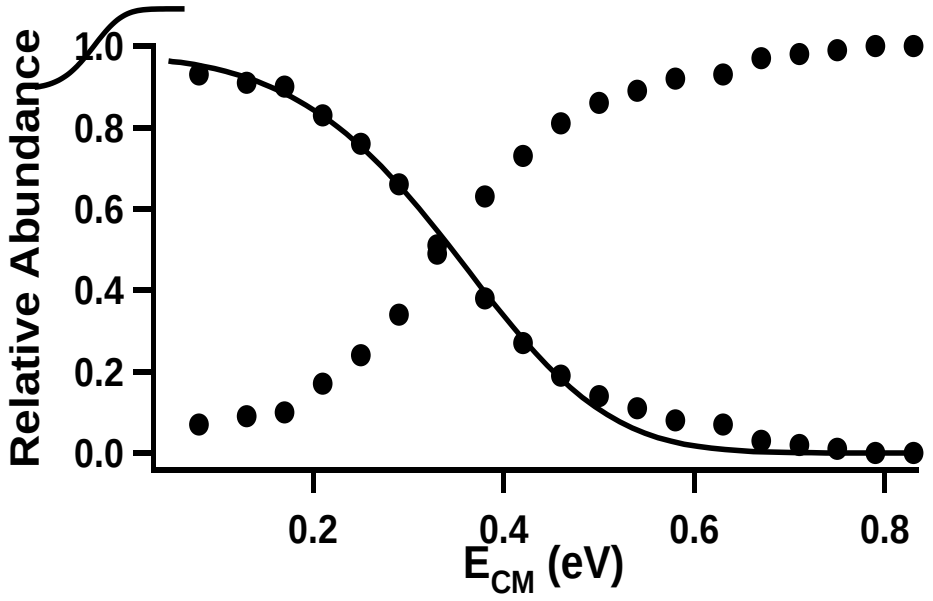


Figure 6.5: Theoretical best fit (solid curves) to the experimental (points) breakdown curve for $[Aβ\text{-}M3\text{-}8Na]^{5+}$.

6.2.4 Dissociation of the Aβ1-40/3-(1,3,4,9-tetrahydro-2*H*-β-carbolin-2-yl) propane-1-sulfonic acid complex

Three complexes containing this substrate were selected consisting of the Aβ peptide, one molecule of M4 and either: seven sodium ions in the 5+ charge state, $[Aβ\text{-}M4\text{-}7Na]^{5+}$ (m/z 956.5), eight sodium ions in the 5+ charge state, $[Aβ\text{-}M4\text{-}8Na]^{5+}$ (m/z 960.8) or six sodium ions in the 4+ charge state $[Aβ\text{-}M4\text{-}6Na]^{4+}$ (m/z 1189.9). Each of these dissociates by loss of a sodiated M4 molecule. The breakdown diagrams and associated fits are presented in Figure 6.6.

6.2.5 Dissociation of the Aβ1-40/3-(1,3,4,9-tetrahydro-2*H*-β-carbolin-2-yl)butane-1-sulfonic acid complex

Two complexes containing this substrate were selected consisting of the Aβ peptide, one molecule of M5 and either: seven sodium ions in the 4+ charge state, $[Aβ\text{-}M5\text{-}7Na]^{4+}$ (m/z 1199.0) or nine sodium ions in the 5+ charge state, $[Aβ\text{-}M5\text{-}9Na]^{5+}$ (m/z 968.1). M5 increases the steric bulk at the amino group relative to homotaurine and CAPSO (Table 3.6). The two complexes dissociate the same way, by loss of a sodiated M5 molecule (forming m/z 1488.2 and 1127.3 in the case of the 4+ and 5+ charge state complexes, respectively). The experimental breakdown curves are presented in Figure 6.7 together with the best theoretical fit to the data.

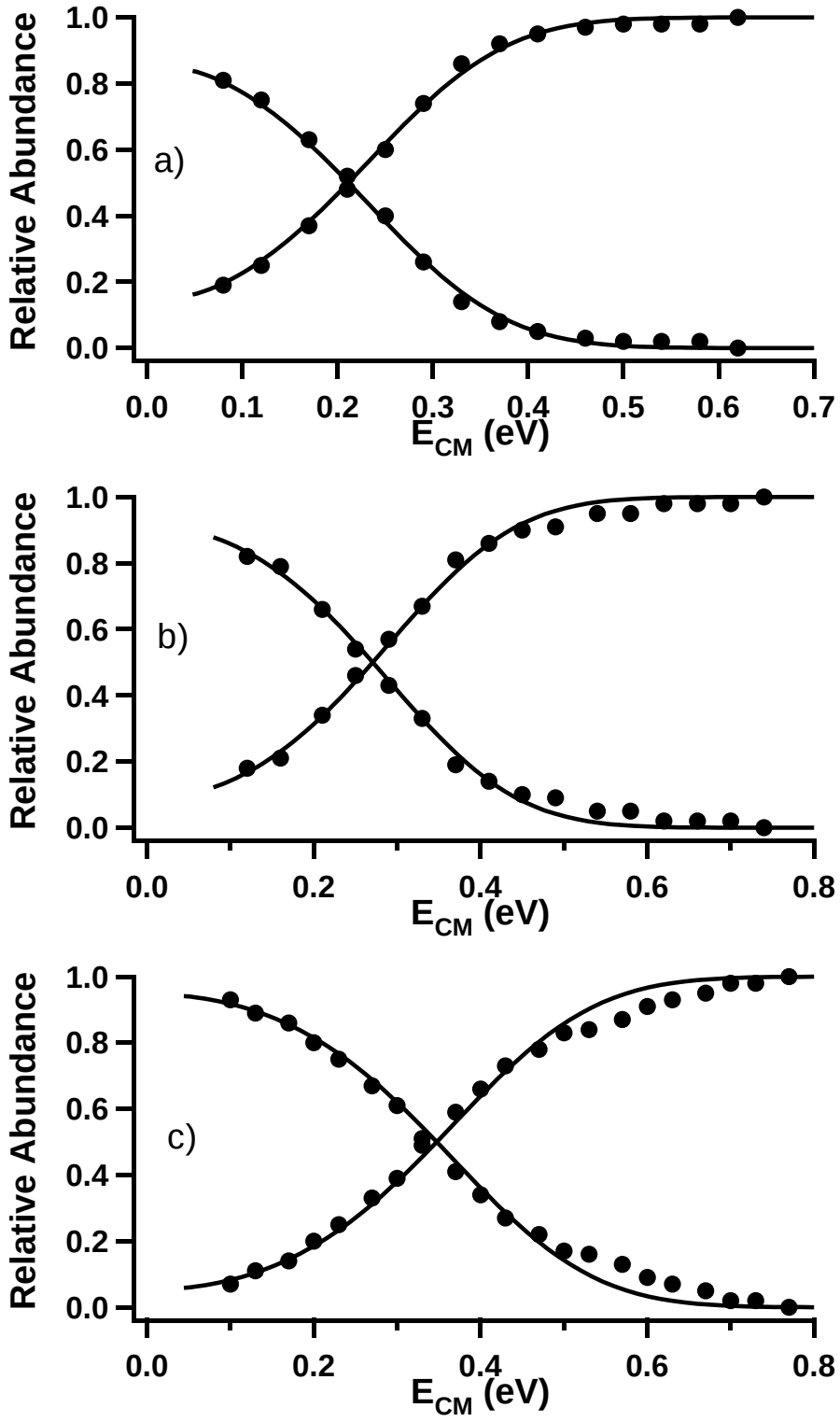


Figure 6.6: Theoretical best fit (solid curves) to the experimental (points) breakdown curve for (a) [Aβ-M4-7Na] $^{5+}$ (m/z 956.5), (b) [Aβ-M4-8Na] $^{5+}$ (m/z 960.8) and (c) [Aβ-M4-6Na] $^{4+}$ (m/z 1189.9).

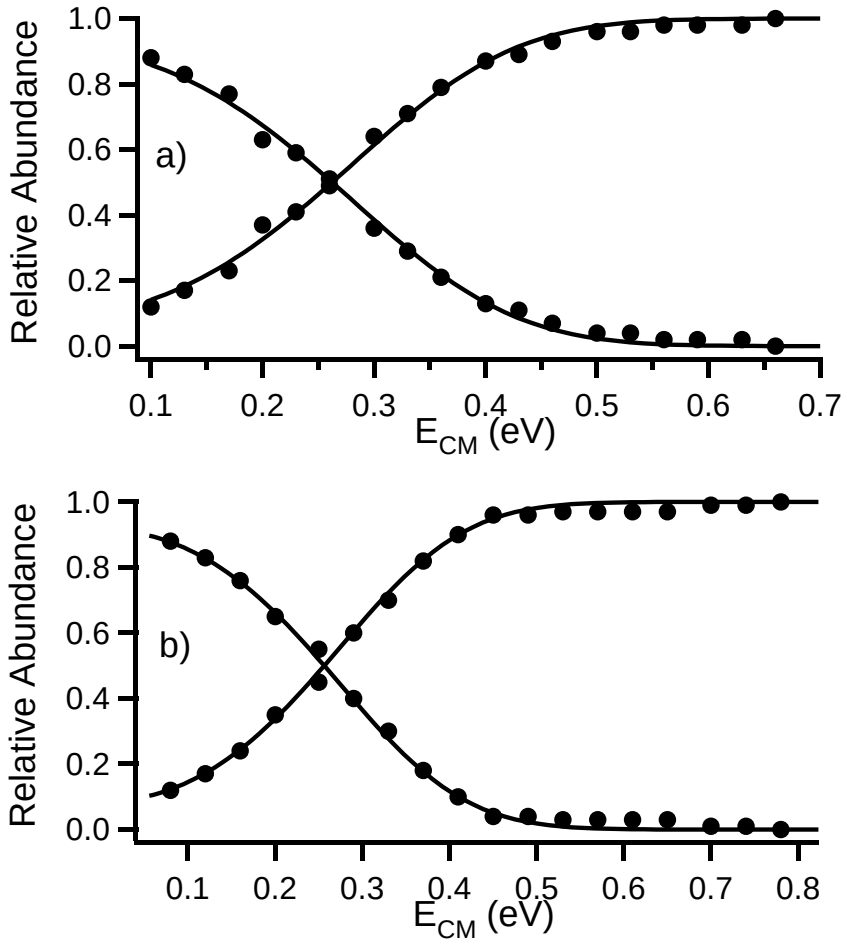


Figure 6.7: Theoretical best fit (solid curves) to the experimental (points) breakdown curve for (a) [Aβ-M5-9Na] $^{5+}$ (m/z 968.5) and (b) [Aβ-M5-7Na] $^{4+}$ (m/z 1199).

6.3 Discussion

The fitting results are summarized in Table 6.1. It is seen that E_0 for all five complexes are within 5 kJ mol^{-1} (0.625 - 0.675 eV). Indeed the same applies to the $\Delta^{\ddagger}S$ which is in the order of -5 to +5 J mol^{-1} K^{-1}. The sulfonate end groups of the five molecules are the same, whereas they exhibit quite distinct, chemically modified amino end groups. Thus, one conclusion from the observed similarity in binding energy is that binding is likely to be primarily through the sulfonate end group

and does not involve the amino terminal group to any significant extent. Given the similarity in the binding of these small molecules, it is consistent that it is the deformation of the peptide during the dissociation process that governs the entropy of activation. If the values for $\Delta^{\ddagger}S$ are taken at face value to be around 0 J mol^{-1} K^{-1}, they suggest that binding of these small molecules does not induce a large conformational change on the ionized peptide. In addition, charge state/number of sodium ions does not appear to play a role in the structure of these complexes, an observation consistent with a binding that can be described primarily as non-specific. For $[A\beta\text{-}M5\text{-}9Na]^{5+}$ (m/z 968.1), the complex having one less sodium was also fit (not shown), $[A\beta\text{-}M5\text{-}8Na]^{5+}$ (m/z 963.9), and the resulting E_0 and $\Delta^{\ddagger}S$ are similar at 0.630 eV and -5 J mol^{-1} K^{-1}. While the value of α does appear to change depending on the nature of the complex, all species could be represented with a value of 200 ± 20, suggesting that there is not that great a 3-dimensional change in shape in the three peptide/substrate complexes studied.

These gas phase results can be compared to our earlier investigation of the relative solution-phase binding affinity of two of these compounds. In solution it was found that M3, CAPSO, had a ~ 66 % higher binding affinity than M1, homotaurine. In the gas phase, this difference disappears and is even slightly reversed (though the values only differ by a few $kJ \cdot mol^{-1}$). This result explicitly points to the difference between gas phase and solution-phase binding. It is possible that, in solution, water plays a role in helping to bind the hydroxyl group in CAPSO to the peptide through hydrogen bonding with water, thus increasing the binding affinity (Figure 6.8). This result points to caution that must be exercised in using mass spectrometry to assess solution-phase binding.

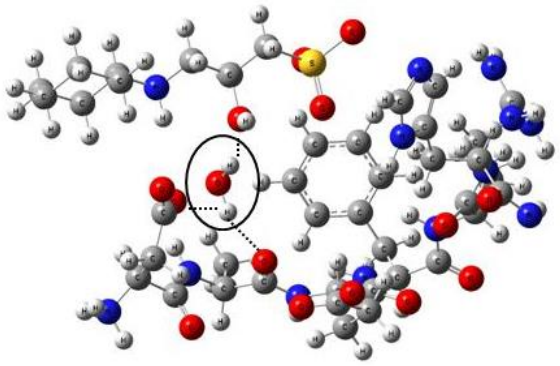


Figure 6.8: Representation of the role water may play in the binding of CAPSO (M3) to the Asp1-His6 segment of the Aβ1-40 peptide.

6.4 Summary

Two approaches to the RRKM modeling of the collision-induced dissociation breakdown curves for the binding of the Aβ peptide to small substrates have been presented: one involving the rigorous calculation of the density and sums of states and another that involves an extrapolation of k(E) to high internal energies using a fitted Hill equation. The approach has been used to evaluate the activation energies for dissociation of five small aminosulfonates bound to the Aβ peptide. The binding energies of the five compounds were found to be very similar, suggesting that binding is primarily through the sulfonate head group of the molecules and does not involve the amino-end of the compounds. The obtained entropies of activation imply little in the way of three-dimensional conformations of the dissociating complexes and only small entropic effects in the gas phase. There is a reversal in relative binding for two of the compounds between solution- and gas phases, indicating a potential role of water in the solution-phase binding dynamics.

Chapter 7: Conclusions

7.1 Modeling the dissociation of $(PMMA_n)(AA)(H^+)$ complexes

CID mass spectrometry was used to dissociate $(PMMA_n)(AA)(H)^+$ complexes the gas phase. The complexes were formed with an oligomer of $PMMA_n$ (7mer, 10mer, or 13mer) and an amino acid (glycine, leucine, or phenylalanine). Breakdown diagrams were acquired and RRKM modeling was used to fit the dissociation curves in order to determine ΔE_0 and $\Delta^{\ddagger}S$. Two major factors were opposing for the dissociation channels: the proton affinity of the amino acid vs. the oligomer, and the length of the oligomer. RRKM modeling of the $(PMMA_7)(AA)(H^+)$ breakdown curves showed that $(PMMA_nH^+ + AA)$ dissociation channels had lower E_0 but higher $\Delta^{\ddagger}S$ whereas the $(PMMA_nH^+ + AA)$ dissociation channels (for leucine and phenylalanine) had higher E_0 but smaller $\Delta^{\ddagger}S$. RRKM modeling of the $(PMMA_{10})(AA)(H^+)$ breakdown curves showed that $(PMMA_nH^+ + AA)$ dissociation channels had higher E_0 but lower $\Delta^{\ddagger}S$ whereas the $(PMMA_nH^+ + AA)$ dissociation channels (for leucine and phenylalanine) had lower E_0 but higher $\Delta^{\ddagger}S$. RRKM modeling of the $(PMMA_{13})(AA)(H^+)$ breakdown curves showed mix results as to the relative E_0 and $\Delta^{\ddagger}S$, however, the α value for these complexes was 300 vs. 200 for the systems involving 7mer and 10mer. Further ion-mobility experiments will be required to understand this switch in the α value.

In order to obtain information on the 3D conformations, as well as their energetic and entropic effect, MD-SA simulations were carried on the 4mer, 7mer and, 13mer systems. Three parameters of the SA protocol were first optimized: (i) 1500 SA cycles for the 4mer and 7mer systems vs. 2000 for the 13mer systems; (ii) the mid-cycle temperature for the 4mer, 7mer and 13mer systems

was adjusted to: 700, 900 K, 1100 K respectively and a slower damping temperature schedule of 10 K ps^{-1} was preferred to allow better conformational sampling. The modeling of the gas phase complexes provided useful insights on the 3D conformations. For each system, it was found that $(PMMA_n)(AAH^+)$ were always having lower energies compared to their respective $(AA)(PMMA_nH^+)$ complexes. In the majority of cases where the oligomer was protonated, it was found that it was more energetically stable for the protonation site to be in the middle of the oligomer. Unfortunately, the level of theory used was not sufficient to obtain a correlation between the experimental and the theoretical results. Solely based on the modeling results, it is not possible to fully explain the competition between both dissociation channels. The only qualitative trend that could be observed is that there seems to exist a competition between the energy levels of $(AA)(PMMA_nH^+)$ and $(PMMA_n + AAH^+)$: as the oligomer length increases, it becomes energetically favourable to dissociate through the protonated oligomer channel.

Using the vibrational frequencies and the rotational constants from the MD-SA or AM1 calculations, it was possible to evaluate the ΔS_{tot} for each dissociation channel. The translational and rotational contributions to ΔS_{tot} canceled out between both dissociation channels. The main contribution to the ΔS_{tot} came from the vibrational contribution. In order to compare the theoretical entropy results obtained, the $\Delta(\Delta S_{tot})$ were calculated for the 7mer and 13mer systems, and reasonable agreement was obtained when compared the $\Delta^{\ddagger}S$ obtained by RRKM theory. These similarities suggest that even if the 3D conformation are not completely accurate, the transition states of these dissociation channels are probably close the structures of the products. Finally, as a follow up on another project initiated in our group, the question was raised as to what is the PA of

a PMMA oligomer?. Using the results of dissociation at threshold collision energy, it was possible to provide an estimate that the PA of the 3mer, 4mer and 7mer was approximately 915 ± 11 kJ mol^{-1}, whereas theoretically, a ΔPA_{4-7} of 3.1 kJ mol^{-1} was calculated. This calculated difference in PA is relatively small, but is included inside the experimental error. For the 13mer, the experimental PA was estimated to be 949 ± 37 kJ mol^{-1} for a ΔPA_{7-13} of 34 kJ mol^{-1}. This difference is greater than the theoretical ΔPA_{7-13} of 7.9 kJ mol^{-1} (even if this result was considering a change from bidentate binding for the 7mer to tridentate binding for the 13mer). In order to obtain better experimental $PMMA_n$ PA values, ion-mobility experiments will have to be conducted on several $(PMMA_n)(AA)(H^+)$ complexes. On the computational level, a higher level of theory will have to be used to obtain refined 3D conformations and hence, better energetics, rotational constants, and vibrational frequencies.

7.2 Investigation of the binding affinity between Aβ peptides and LMW molecules

An efficient and reproducible ESI-MS binding assay was developed for semi-quantitative evaluation of the binding affinity of LMW compounds to Aβ peptides. This method allowed classifying semi-quantitatively small aminosulfonate molecules into four categories (strong, moderate, weak, and non-binders). From this ESI-MS method, we have found that the same small aminosulfonate molecules were binding similarly to Aβ1-28, Aβ1-40, and Aβ1-42. As an alternative for compounds with a lower solubility in water, solutions containing 20% ethanol help the solubilisation while providing similar results as well. The binding of four different compounds (having different BA) was tested on chimeric peptides (which have a different charge motif than the Aβ peptides). This allows to find what were the requirement for the small aminosulfonate

molecules to have, in order to bind strongly to the Aβ peptide. This method was important because, in the amyloidogenic cascade, some of these LMW molecules have the ability to inhibit Aβ fibril formation. However, before having this ability, these LMW must first bind the the soluble Aβ peptide. This is why, this moderate throughput assay was developped.

7.3 Development of an alternate RRKM method for larger molecular systems

Peptides have important biological roles in life and being able to calculate variations of energy and entropy upon their dissociation with small molecules is crucial to understand their role. RRKM is a reliable method to determine ΔE_0 and $\Delta^{\ddagger}S$, as long as the number of degrees of freedom is limited to a relatively small number. With large values for the sum and density of states, it may not be computationally feasible on standard computers to do RRKM modeling. An alternate RRKM method was developed to fit the breakdown curves obtained by CID mass spectrometry for larger systems. This alternate method involves an extrapolation of k(E) to high internal energies using a fitted Hill equation. This reliability of this new method was tested against results previously obtained on $PMMA_{13}$ systems with the classical rigorous RRKM theory. Once it was found that the alternate method could reproduce such fittings, it was applied to the dissociation of gas phase complexes involving Aβ1-40 and five aminosulfonate molecules. The binding of the five LMW molecules was rather simlar and seems to involve the sulfonate function to the molecules. The entropy of activation calculed implies no significant conformational rearrangement of the peptide upon dissociation.

Claims to Original Research

1. Modeling the energetic and entropic effects of the dissociation channels of gas phase $(PMMA_n)(AA)(H^+)$ complexes using CID mass spectrometry, RRKM theory, and molecular dynamics simulated annealing simulations. This work is described in Chapter 4 and is not published yet. Some of the experimental and computational aspects presented in this section will be published in manuscript(s).

2. Development and description of an efficient and reproducible screening method for identifying and describing low molecular weight compounds that bind to amyloid β peptides using ESI-MS. This work was published in reference [179].

3. Development and description of an approximate model for extracting information from the CID breakdown curves for complexes between the Aβ1-40 peptide with small molecule substrates with RRKM theory to obtain their relative gas phase binding energies. Due to the sheer size of this system, an alternative approach to the rigorous sum- and density-of-states calculations required in RRKM theory based on a Hill equation extrapolation of k(E). The performance of this approximate, but faster, approach to k(E) was tested for several polymer-based systems as well as for the Aβ1-40 peptide. This was published in reference [182].

References

[1] Séguin, M.; Villeneuve, B.: *Astronomie & Astrophŷsique*; ERPI: St-Laurent, Canada, 1995.

[2] Feynman, R. P.; Leighton, R. B.; Sands, M. L.: *The Feynman Lectures on Physics: the definitive and extended edition*; Addison-Wesley Publishing Company, USA, 2005.

[3] Atkins, P. W.: *Physical Chemistry*; 5th ed.; W.H. Freeman: New York, USA, 1994.

[4] McQuarrie, D. A.; Simon, J. D.: *Physical Chemistry: a molecular approach*; University Science Books, Sausalito, USA, 1997.

[5] London, F.: The general theory of molecular forces. *Transactions of the Faraday Society* **1937**, *33*, 8-26.

[6] Hill, J. W.; Petrucci, R. H.; McCreary, T. W.; Perry, S. S.: *Chimie Générale*; ERPI, St-Laurent, Canada, 2008.

[7] van der Waals, J. D. Over de Continuiteit van den Gas- en Vloeistoftoestand (*On the continuity of the gas and liquid state)* (Ph.D. thesis), Leiden, The Netherlands, 1873.

[8] Alagona, G.; Ghio, C.: *Ab Initio Theoretical Methods for Studying Intermolecular Forces.*, In: *Introduction to Advanced Topics of Computational Chemistry*, Montero, L. A.; Diaz, L. A.; Bader. R., Eds.; Editorial de la Universidad de La Habana: Havana, 2003; p.137-58.

[9] Leach, A. R.: *Molecular Modelling: Principles and Applications*; 2nd ed.; Prentice Hall, Essex, England, 2001.

[10] Szabo, A.; Ostlund, N. S.: *Modern Quantum Chemistry – Introduction to Advanced Electronic Structure Theory*; Dover editions, Mineola, USA, 1996.

[11] Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P.: AM1: A new general purpose quantum mechanical model. *Journal of the American Chemical Society* **1985**, *107*, 3902-9.

[12] Glusker, J. P.: Dorothy Crowfoot Hodgkin (1910-1994). *Protein Science* **1994**, 2465-9.

[13] The RCSB Protein Data Bank (www.rcsb.org), consulted on December 28th 2012.

[14] Friebolin, H.: *Basic One- and Two-Dimensional NMR Spectroscopy*; 3rd ed.; Wiley-VCH, Weinheim, Germany, 1998.

[15] Wüthrich, K.: Protein structure determination in solution by NMR spectroscopy. *The Journal of Biological Chemistry* **1990**, *265*, 22059-62.

[16] Wüthrich, K.: NMR Studies of Structure and Function of Biological Macromolecules (Nobel Lecture). *Angewandte Chemie International Edition* **2003**, *42*, 3340-63.

[17] Kanu, A. B.; Dwivedi, P.; Tam, M.; Matz, L.; Hill Jr., H. H.: Ion mobility-mass spectrometry. *Journal of Mass Spectrometry* **2008**, *43*, 1-22.

[18] Katta, V.; Chait, B. T.: Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. *Rapid Communications in Mass Spectrometry* **1991**, *5*, 214-7.

[19] Stocks, B. B.; Konermann, L.: Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-Induced Oxidative Labeling and Mass Spectrometry. *Analytical Chemistry* **2009**, *81*, 20-7.

[20] Montaudo, G.; Lattimer, R. P.: *Mass Spectrometry of Polymers*; CRC Press, Boca Raton, USA, 2002.

[21] Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F.: Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry and Ion Processes* **1987**, *78*, 53-68.

[22] Tanaka, K.: The Origin of Macromolecule Ionization by Laser Irradiation (Nobel Lecture). *Angewandte Chemie International Edition* **2003**, *42*, 3860-70.

[23] Yamashita, M.; Fenn, J. B.: Electrospray Ion Source. Another Variation on the Free-Jet Theme. *Journal of Physical Chemistry* **1984**, *88*, 4451-59.

[24] Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M.: Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246*, 64-71.

[25] Fenn, J. B.: Electrospray Wings for Molecular Elephants (Nobel Lecture). **2003**, *42*, 3871-94.

[26] Alhazmi, A. M.; Mayer, P. M.: Protonating polymer oligomers in the gas phase to change fragmentation pathways. *Journal of the American Society for Mass Spectrometry* **2009**, *20*, 60-6.

[27] McLafferty, F. W.; Bente, P. F.; Kornfeld, R.; Shih-Chuan, T.; Howe, I.: Metastable ion characteristics. XXII. Collisional activation spectra of organic ions. *Journal of the American Chemical Society* **1973**, *95*, 2120-9.

[28] Jackson, A. T.; Jennings, K. R.; Scrivens, J. H.: Generation of average mass values and end group information of polymers by means of a combination of matrix-assisted laser desorption/ionization-mass spectrometry and liquid secondary ion-tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry* **1997**, *8*, 76-85.

[29] Jackson, A. T.; Yates, H. T.; Scrivens, J. H.; Green, M. R.; Bateman, R. H.: Utilizing matrix-assisted laser desorption/ionization-collision induced dissociation for the generation of structural information from poly(alkyl methacrylate)s. *Journal of the American Society for Mass Spectrometry* **1997**, *8*, 1206-13.

[30] Gidden, J.; Jackson, A. T.; Scrivens, J. H., Bower, M. T.: Gas phase conformations of synthetic polymers: poly(methylmethacrylate) oligomers cationized by sodium ions. *International Journal of Mass Spectrometry* **1999**, *188*, 121-30.

[31] Scrivens, J. H.; Jackson, A. T.; Yates, H. T.; Green, M. R.; Critchley, G.; Brown, J.; Bateman, R. H.; Bowers, M. T.; Gidden, J.: The Effect of the Variation of Cation in the Matrix-Assisted Laser Desorption/Ionisation-Collision Induced Dissociation (MALDI-CID) Spectra of Oligomeric Systems. *International Journal of Mass Spectrometry and Ion Processes* **1997**, *165/166*, 363-75.

[32] Jackson, A. T.; Yates, H. T.; Scrivens, J. H.; Critchley, G.; Brown, J. Green, M. R., Bateman, R. H.: The Application of Matrix-assisted Laser Desorption/Ionization Combined with Collision-induced Dissociation to the Analysis of Synthetic Polymers. *Rapid Communications in Mass Spectrometry* **1996**, *10*, 1668-74.

[33] Giguère, M-S.; Mayer, P. M.: Climbing the internal energy ladder: the unimolecular decomposition of ionized poly(vinyl acetate). *International Journal of Mass Spectrometry* **2004**, *231*, 59-68.

[34] Renaud, J.; Alhazmi, A. M.; Mayer, P. M.: Comparing the fragmentation chemistry of gas phase adducts of poly(dimethylsiloxane) oligomers with metal and organic ions. *Canadian Journal of Chemistry* **2009**, *87*, 453-9.

[35] Gronert, S.; Simpson, D. C.; Conner, K. K.: A Reevaluation of Computed Proton Affinities for the Common α-amino acids. *Journal of the American Society for Mass Spectrometry* **2009**, *20*, 2116-23.

[36] Creighton, T. E.: *Proteins – Structures and Molecular Properties*; 2nd ed.; W. H. Freeman and Company, New York, USA, 1996.

[37] van Heijenoort, J.: Formation of the glycan chains in the synthesis of bacterial peptidoglycan, *Glycobiology* **2001**, *11*, 25R-36R.

[38] Wolosker, H.; Dumin, E.; Balan, L.; Foltyn, V. N.: D-amino acids in the brain: D-serine in neurotransmission and neurodegeneration. *The FEBS Journal* **2008**, *275*, 3514-26.

[39] Wermuth, C. G.; Ganellin, C. R.; Lindberg, P.; Mitscher, L. A.: Glossary of terms used in medicinal chemistry. *Pure & Applied Chemistry* **1998**, *70*, 1129-43.

[40] Wisniewski, H. M.; Wegiel, J.: The neuropathology of Alzheimer's disease. *Neuroimaging clinics of North America* **1995**, *5*, 45-57.

[41] Esiri, M. M.: *The Neuropathology of Alzheimer's Disease.*, In: *Neurobiology of Alzheimer's Disease, 2nd ed.*, Dawbarn, D.; Allen, S. J., Eds.; Oxford University Press: Oxford, England, 2001; p. 33–53.

[42] Clark, C. M.; Karlawish, J. H.: Alzheimer's disease: current concepts and emerging diagnostic and therapeutic strategies. *Annals of Internal Medicine* **2003**, *138*, 400-10.

[43] Hardy, J.; Selkoe, D. J.: The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **2002**, *297*, 353-6.

[44] Hardy, J. A.; Higgins, G. A.: Alzheimer's disease: the amyloid cascade hypothesis. *Science* **1992**, *256*, 184-5.

[45] Selkoe, D. J.: Towards a comprehensive theory for Alzheimer's disease. Hypothesis: Alzheimer's disease is caused by the cerebral accumulation and cytotoxicity of amyloid β-protein. *Annals of the New York Academy of Sciences* **2000**, *924*, 17-25.

[46] Wolfe, M. S.: Secretase targets for Alzheimer's disease: identification and therapeutic potential. *Journal of Medicinal Chemistry* **2001**, *44*, 2039-60.

[47] Sambamurti, K.; Greig, N. H.; Lahiri, D. K.: Advances in the cellular and molecular biology of the β-amyloid protein in Alzheimer's disease. *Neuromolecular Medicine* **2002**, *1*, 1-31.

[48] Barghorn, S.; Nimmrich, V.; Striebinger, A.; Krantz, C.; Keller, P.; Janson, B.; Bahr, M.; Schmidt, M.; Bitner, R. S.; Harlan, J.; Barlow, E.; Ebert, U.; Hillen, H.: Globular amyloid β-peptide oligomer - a homogenous and stable neuropathological protein in Alzheimer's disease. *Journal of Neurochemistry* **2005**, *95*, 834-47.

[49] Jarrett, J. T.; Berger, E. P.; Lansbury, P. T., Jr.: The carboxy terminus of the β amyloid protein is critical for the seeding of amyloid formation: implications for the pathogenesis of Alzheimer's disease. *Biochemistry* **1993**, *32*, 4693-7.

[50] Vetrivel, K. S.; Thinakaran, G.: Amyloidogenic processing of β-amyloid precursor protein in intracellular compartments. *Neurology* **2006**, *66*, S69-73.

[51] Chen, Y. R.; Glabe, C. G.: Distinct early folding and aggregation properties of Alzheimer amyloid-β peptides Aβ40 and Aβ42: stable trimer or tetramer formation by Aβ42. *The Journal of Biological Chemistry* **2006**, *281*, 24414-22.

[52] Klein, W. L.; Krafft, G. A.; Finch, C. E.: Targeting Small Aβ Oligomers: The Solution to an Alzheimer's Disease Conundrum? *Trends in Neurosciences* **2001**, *24*, 219-24.

[53] Heneka, M. T.: Inflammation in Alzheimer's Disease. *Clinical Neuroscience Research* **2006**, *6*, 247-60.

[54] Kisilevsky, R.; Fraser, P. E.: Aβ amyloidogenesis: unique, or variation on a systemic theme? *Critical Reviews in Biochemistry and Molecular Biology* **1997**, *32*, 361-404.

[55] Verbeek, M. M.; Otte-Holler, I.; van den Born, J.; van den Heuvel, L. P.; David, G.; Wesseling, P.; de Waal, R. M.: Agrin is a major heparan sulfate proteoglycan accumulating in Alzheimer's disease brain. *The American Journal of Pathology* **1999**, *155*, 2115-25.

[56] Kisilevsky, R.; Lemieux, L. J.; Fraser, P. E.; Kong, X.; Hultin, P. G.; Szarek, W. A.: Arresting amyloidosis in vivo using small-molecule anionic sulphonates or sulphates: implications for Alzheimer's disease. *Nature Medicine* **1995**, *1*, 143-48.

[57] Gupta-Bansal, R.; Frederickson, R. C. A.; Bruden, K. R.: Proteoglycan-Mediated Inhibition of Aβ Proteolysis: a Potential Cause of Senile Plaque Accumulation. *The Journal of Biological Chemistry* **1995**, *270*, 18666-71.

[58] LeVine III, H.: Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: detection of amyloid aggregation in solution. *Protein Science* **1993**, *2*, 404-10.

[59] Wang, S. S.; Chen, Y. T.; Chou, S. W.: Inhibition of amyloid fibril formation of β-amyloid peptides via the amphiphilic surfactants. *Biochimica et Biophysica Acta* **2005**, *1741*, 307-13.

[60] Kaneko, I.; Tutumi, J.: Matters Arising: Conformations of β-Amyloid in Solution. *Journal of Neurochemistry* **1997**, *68*, 307-13.

[61] Soto, C.; Castano, E. M.; Kumar, R. A.; Beavis, R. C.; Frangione, B.: Fibrillogenesis of synthetic amyloid-β peptides is dependent on their initial secondary structure. *Neuroscience Letters* **1995**, *200*, 105-8.

[62] Hou, L.; Kang, I.; Marchant, R. E.; Zagorski, M. G.: Methionine 35 oxidation reduces fibril assembly of the amyloid Aβ-(1-42) peptide of Alzheimer's disease. *The Journal of Biological Chemistry* **2002**, *277*, 40173-6.

[63] Nilsson, M. R.; Driscoll, M. D.; Raleigh, D. P.: Low levels of asparagine deamidation can have a dramatic effect on aggregation of amyloidogenic peptides: Implications for the study of amyloid formation. *Protein Science* **2002**, *11*, 342-9.

[64] Eisert, R.; Felau, L.; Brown, L. R.: Methods for enhancing the accuracy and reproducibility of Congo red and Thioflavin T assays. *Analytical Biochemistry* **2006**, *353*, 144-6.

[65] Nilsson, M. R.: Techniques to study amyloid fibril formation in vitro. *Methods* **2004**, *34*, 151-60.

[66] Bruce, J. E.; Anderson, G. A.; Chen, R.; Cheng, X.; Gale, D. C.; Hofstadler, S. A.; Schwartz, B. L.; Smith, R. D.: Bio-affinity characterization mass spectrometry. *Rapid Communications in Mass Spectrometry* **1995**, *9*, 644-50.

[67] Loo, J. A.: Studying non-covalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrometry Reviews* **1997**, *16*, 1-23.

[68] Smith, R. D.; Light-Wahl, K. J.: The observation of non-covalent interactions in solution by electrospray ionisation mass spectrometry. *Biological Mass Spectrometry* **1993**, *22*, 493-501.

[69] Bazoti, F. N.; Tsarbopoulos, A.; Markides, K. E.; Bergquist, J.: Study of the non-covalent interaction between amyloid-β-peptide and melatonin using electrospray ionization mass spectrometry. *Journal of Mass Spectrometry* **2005**, *40*, 182-92.

[70] Camilleri, P.; Haskins, N. J.; Howlett, D. R.: β-Cyclodextrin interacts with the Alzheimer amyloid β-A4 peptide. *FEBS Letters* **1994**, *341*, 256-8.

[71] Cheng, X.; van Breemen, R. B.: Mass spectrometry-based screening for inhibitors of β-amyloid protein aggregation. *Analytical Chemistry* **2005**, *77*, 7012-5.

[72] Skribanek, Z.; Balaspiri, L.; Mak, M.: Interaction between synthetic amyloid-β-peptide (1-40) and its aggregation inhibitors studied by electrospray ionization mass spectrometry. *Journal of Mass Spectrometry* **2001**, *36*, 1226-9.

[73] Gervais, F.; Morissette, C.; Kong, X.: Proteoglycans and Amyloidogenic Proteins in Peripheral Amyloidosis. *Current Medicinal Chemistry - Immunology Endocrine & Metabolic Agents* **2003**, *3*, 361-70.

[74] Gervais, F.; Paquette, J.; Morissette, C.; Krzywkowski, P.; Yu, M.; Azzi, M.; Lacombe, D.; Kong, X.; Aman, A.; Laurin, J.; Szarek, W. A.; Tremblay, P.: Targeting soluble Aβ peptide with Tramiprosate for the treatment of brain amyloidosis. *Neurobiology of Aging* **2007**, *28*, 537-47.

[75] Martineau, E.; Aman, A. M.; Kong, X.: pH Effect on the Association Between the Amyloid β Peptide and Low Molecular Weight Molecules. *71e congrès de l'Acfas*: Rimouski, Qc, Canada, 2003.

[76] Busman, M.; Knapp, D. R.; Schey, K. L.: Observation of large multimers in the electrospray ionization mass spectrometry of peptides. *Rapid Communications in Mass Spectrometry* **1994**, *8*, 211-6.

[77] Benkestock, K.; Sundqvist, G.; Edlund, P. O.; Roeraade, J.: Influence of droplet size, capillary-cone distance and selected instrumental parameters for the analysis of non-covalent protein-ligand complexes by nano-electrospray ionization mass spectrometry. *Journal of Mass Spectrometry* **2004**, *39*, 1059-67.

[78] Peschke, M.; Verkerk, U. H.; Kebarle, P.: Features of the ESI mechanism that affect the observation of multiply charged non-covalent protein complexes and the determination of the association constant by the titration method. *Journal of the American Society for Mass Spectrometry* **2004**, *15*, 1424-34.

[79] Hardouin, J.; Lange, C. M.: Biological Noncovalent Complexes by Mass Spectrometry. *Current Organic Chemistry* **2005**, *9*, 317-24.

[80] Schilling, S.; Lauber, T.; Schaupp, M.; Manhart, S.; Scheel, E.; Bohm, G.; Demuth, H. U.: On the seeding and oligomerization of pGlu-amyloid peptides (in vitro). *Biochemistry* **2006**, *45*, 12393-9.

[81] Youssef, I.; Florent-Bechard, S.; Malaplate-Armand, C.; Koziel, V.; Bihain, B.; Olivier, J. L.; Leininger-Muller, B.; Kriem, B.; Oster, T.; Pillot, T.: N-truncated amyloid-β oligomers induce learning impairment and neuronal apoptosis. *Neurobiology of Aging* **2008**, *29*, 1319-33.

[82] Miyazawa, K.; Ito, K.; Matsumoto, F.: Aminosulfonic acids in six species of marine algae. *Bulletin of the Japanese Society for the Science of Fish* **1969**, *35*, 1215-9.

[83] Miyazawa, K.; Ito, K.; Matsumoto, F.: Amino acids peptide in seven species of marine green algae. *Journal of the Faculty of Fisheries and Animal Husbandry* **1976**, *15*, 161-9.

[84] Aisen, P. S.; Gauthier, S.; Vellas, B.; Briand, R.; Saumier, D.; Laurin, J.; Garceau, D.: Alzhemed: a potential treatment for Alzheimer's disease. *Current Alzheimer Research* **2007**, *4*, 473-8.

[85] Baer, T.; Hase, W. L.: *Unimolecular Reactions Dynamics - Theory and Experiments*; Oxford University Press: New York, USA, 1996.

[86] Baer, T.; Mayer, P. M.: Statistical Rice-Ramsperger-Kassel-Marcus Quasiequilibrium Theory Calculations in Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **1997**, *8*, 103-15.

[87] Rice, O. K.; Ramsperger, H. C.: Theories of unimolecular gas reactions at low pressures. *Journal of the American Chemical Society* **1927**, *49*, 225-42.

[88] Rice, O. K.; Ramsperger, H. C.: Theories of unimolecular gas reactions at low pressures. II. *Journal of the American Chemical Society* **1928**, *49*, 1617-29.

[89] Kassel, L. S.: Studies in homogeneous gas reactions. I. *Journal of Physical Chemistry* **1928**, *32*, 225-42.

[90] Wigner, E. The transition-state method. *Transactions of the Faraday Society* **1938**, *34*, 29-41.

[91] Lifshitz, C. *Advances in Mass Spectrometry* 7A (ed. Daly, N. R.), London, **1978**, p. 3-18.

[92] Schermann, S. M.; Simmons, D. A.; Konermann, L.: Mass spectrometry-based approaches to protein-ligand interactions. *Expert Review of Proteomics* **2005**, *2*, 475-85.

[93] Hossain, B. M.; Simmons, D. A.; Konermann, L.: Do electrospray mass spectra reflect the ligand binding state of proteins in solution?. *Canadian Journal of Chemistry* **2005**, *83*, 1953-60.

[94] Ferguson, P. L.; Kuprowski, M. C.; Boys, B. L.; Wilson, D. J.; Pan, J.; Konermann, L.: Protein Folding and Protein-Ligand Interactions Monitored by Electrospray Mass Spectrometry. *Current Analytical Chem*istry **2009**, *5*, 186-204.

[95] Sun, J.; Kitova, E. N.; Sun, N.; Klassen, J. S.: Method for Identifying Nonspecific Protein-Protein Interactions in Nanoelectrospray Ionization Mass Spectrometry. *Analytical Chemistry* **2007**, *79*, 8301-11.

[96] Sun, J. X.; Kitova, E. N.; Klassen, J. S.: Method for Stabilizing Protein-Ligand Complexes in Nanoelectrospray Ionization Mass Spectrometry. *Analytical Chemistry* **2007**, *79*, 417-25.

[97] Sun, J. X.; Kitova, E. N.; Wang, W. J.; Klassen, J. S.: Method for Distinguishing Specific and Nonspecific Protein-Ligand Complexes in Nanoelectrospray Ionization Mass Spectrometry. *Analytical Chemistry* **2006**, *78*, 3010-8.

[98] Wang, W. J.; Kitova, E. N.; Klassen, J. S.: Nonspecific Protein-Carbohydrate Complexes Produced by Nanoelectrospray Ionization. Factors Influencing their Formation and Stability. *Analytical Chemistry* **2005**, *77*, 3060–71.

[99] Klassen, J. S.; Kitova, E. N.; Bundle, D. R.: *Abstracts of Papers from the American Chemical Society* **2004**, *227*, 51.

[100] Kitova, E. N.; Bundle, D. R.; Klassen, J. S.: Partitioning of Solvent Effects and Intrinsic Interactions in Biological Recognition. *Angewante Chemie International Edition* **2004**, *43*, 4183-6.

[101] Wang, W. J.; Kitova, E. N.; Klassen, J. S.: Bioactive Recognition Sites May Not Be Energetically Preferred in Protein-Carbohydrate Complexes in the Gas Phase. *Journal of the American Chemical Society* **2003**, *125*, 13630-1.

[102] Wang, W. J.; Kitova, E. N.; Klassen, J. S.: Influence of Solution and Gas Phase Processes on Protein-Carbohydrate Binding Affinities Determined by Nanoelectrospray Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry. *Analytical Chemistry* **2003**, *75*, 4945-55.

[103] Pan, J.; Xu, K.; Yang, X.; Choy, W. Y.; Konermann, L.: Solution-Phase Chelators for Suppressing Nonspecific Protein-Metal Interactions in Electrospray Mass Spectrometry. *Analytical Chemistry* **2009**, *81*, 5008-15.

[104] Shoemaker, G. K.; Kitova, E. N.; Palcic, M. M.; Klassen, J. S.: Equivalency of Binding Sites in Protein-Ligand Complexes Revealed by Time-Resolved Tandem Mass Spectrometry. *Journal of the American Chemical Society* **2007**, *129*, 8674-5.

[105] Kitova, E. N.; Daneshfar, R.; Marcato, P.; Mulvey, G. L.; Armstrong, G.; Klassen, J. S.: Stability of the Homopentameric B Subunits of Shiga Toxins 1 and 2 in Solution and the Gas Phase as Revealed by NanoES-FT-ICR/MS. *Journal of the American Society for Mass Spectrom*etry **2005**, *16*, 1957-68.

[106] Kitova, E. N.; Wang, W. J.; Bundle, D. R.; Klassen, J. S.: Retention of Bioactive Ligand Conformation in a Gaseous Protein-Trisaccharide Complex. *Journal of the American Chemical Society* **2002**, *124*, 13980-1.

[107] Kitova, E. N.; Bundle, D. R.; Klassen, J. S.: Evidence for the Preservation of Specific Intermolecular Interactions in Gaseous Protein-Oligosaccharide Complexes. *Journal of the American Chemical Society* **2002**, *124*, 9340-1.

[108] Kitova, E. N.; Kitov, P.I.; Paszkiewicz, E.; Kim, J.; Mulvey, G. L.; Armstrong, G. D.; Bundle, D. R.; Klassen, J. S.: Affinities of Shiga Toxins 1 and 2 for Univalent and Oligovalent Pk-trisaccharide Analogs Measured by Electrospray Ionization Mass Spectrometry. *Glycobiology* **2007**, *17*, 1127-37.

[109] Sinelnikov, I.; Kitova, E. N.; Armstrong, G. D.; Klassen, J. S.: Effects of Single Amino Acid Substitution on the Dissociation of Multiply Charged Multiprotein Complexes in the Gas Phase. *Journal of the American Society for Mass Spectrometry* **2007**, *18*, 688-92.

[110] Wang, W. J.; Kitova, E. N.; Sun, J. X.; Klassen, J. S.: Blackbody Infrared Radiative Dissociation of Nonspecific Protein-Carbohydrate Complexes Produced by Nanoelectrospray Ionization. The Nature of the Noncovalent Interactions. *Journal of the American Society for Mass Spectrometry* **2005**, *16*, 1583-94.

[111] Daneshfar, R.; Klassen, J. S.: Arrhenius Activation Parameters for the Loss of Neutral Nucleobases from Deprotonated Oligonucleotide Anions in the Gas Phase. *Journal of the American Society for Mass Spectrometry* **2004**, *15*, 55-64.

[112] Felitsyn, N.; Kitova, E. N.; Klassen, J. S.: Thermal Decomposition of the Protein Homodimer Ecotin in the Gas Phase. *Journal of the American Society for Mass Spectrometry* **2002**, *13*, 1432-42.

[113] Felitsyn, N.; Kitova, E. N.; Klassen, J. S.: Thermal Decomposition of a Gaseous Multiprotein Complex Studied by Blackbody Infrared Radiative Dissociation. Investigating the Origin of the Asymmetric Dissociation Behaviour. *Analytical Chemistry* **2001**, *73*, 4647-61.

[114] Dunbar, R. C.: BIRD (blackbody infrared radiative dissociation): Evolution, principles, and applications. *Mass Spectrometry Reviews* **2004**, *23*, 87-159.

[115] Angelelli, F.; Chiavarino, B.; Crestoni, M. E.; Fornarini, S.: Binding of gaseous Fe(III)-heme cation to model biological molecules: Direct association and ligand transfer reactions. *Journal of the American Society for Mass Spectrometry* **2005**, *16*, 589-98.

[116] Schnier, P. D.; Price, W. D.; Jockusch, R. A.; Williams, E. R.: Blackbody Infrared Radiative Dissociation of Bradykinin and Its Analogues: Energetics, Dynamics, and Evidence for Salt-Bridge Structures in the Gas Phase. *Journal of the American Chemical Society* **1996**, *118*, 7178-89.

[117] Jockusch, R. A.; Schnier, P. D.; Price, W. D.; Strittmatter, E. F.; Demirev, P.A.; Williams, E. R.: Effects of Charge State on Fragmentation Pathways, Dynamics, and Activation Energies of Ubiquitin Ions Measured by Blackbody Infrared Radiative Dissociation. *Analytical Chemistry* **1997**, *67*, 1119-26.

[118] Schnier, P. D.; Price, W. D.; Strittmatter, E. F.; Williams, E. R.: Dissociation energetics and mechanisms of leucine enkephalin $(M + H)^+$ and $(2M + X)^+$ ions (X = H, Li, Na, K, and Rb) measured by blackbody infrared radiative dissociation. *Journal of the American Society for Mass Spectrometry* **1997**, *8*, 771-80.

[119] Ge, Y.; Horn, D. M.; McLafferty, F. W.: Blackbody infrared radiative dissociation of larger (42 kDa) multiply charged proteins. *International Journal of Mass Spectrometry* **2001**, *210/211*, 203-14.

[120] Asano, K. G.; Goeringer, D. E.; Butcher, D. J.; McLuckey, S. A.: Bath gas temperature and the appearance of ion trap tandem mass spectra of high-mass ions. *International Journal of Mass Spectrometry* **1999**, *190/191*, 281-93.

[121] Butcher, D. J.; Asano, K. G.; Goeringer, D. E.; McLuckey, S. A.: Thermal Dissociation of Gaseous Bradykinin Ions. *Journal of Physical Chemistry A* **1999**, *103*, 8664-71.

[122] Asano, K. G.; Goeringer, D. E.; McLuckey, S. A.: Thermal dissociation in the quadrupole ion trap: ions derived from leucine enkephalin. *International Journal of Mass Spectrometry* **1999**, *185/187*, 207-219.

[123] Mayer, P. M.; Poon, C.: The mechanisms of collisional activation of ions in mass spectrometry. *Mass Spectrometry Reviews* **2008**, *28*, 608-39.

[124] Asano, K. G.; Goeringer, D. E.; McLuckey, S. A.: *Proceeding of the 46th Conference for the American Society for Mass Spectrometry*, 1998.

[125] Georinger, D. E.; McLuckey, S. A.: Evolution of ion internal energy during collisional excitation in the Paul ion trap: A stochastic approach. *Journal of Chemical Physics* **1996**, *104*, 2214-21.

[126] Hart, K. J.; McLuckey, S. A.: Relative Dissociation Energy Measurements Using Ion Trap Collisional Activation. *Journal of the American Society for Mass Spectrometry* **1994**, *5*, 250-9.

[127] Goeringer, D. E.; Asano, K. G.; McLuckey, S. A.: Ion Internal Temperature and Ion Trap Collisional Activation: Protonated Leucine Enkephalin. *International Journal of Mass Spectrometry* **1999**, *182/183*, 275-88.

[128] Asano, K. G.; Butcher, D.J.; Goeringer, D. E.; McLuckey, S. A.: Effective Ion Internal Temperatures Achieved via Boundary Activation in the Quadrupole Ion Trap: Protonated Leucine Enkephalin. *Journal of Mass Spectrometry* **1999**, *34*, 691-8.

[129] Goeringer, D. E.; Duckworth, D. C.; McLuckey, S. A.: Collision-induced Dissociation in Quadrupole Ion Traps: Application of a Thermal Model to Diatomic Ions. *Journal of Physical Chemistry A* **2001**, *105*, 1882-9.

[130] Goeringer, D. E.; McLuckey, S. A.: Kinetics of Collision-Induced Dissociation in the Paul Trap: A First-Order Model. *Rapid Communication in Mass Spectrometry* **1996**, *10*, 328-34.

[131] Laskin, J.; Futrell, J. H.: Collisional Activation of Peptide Ions in FT-ICR Mass Spectrometry. *Mass Spectrometry Reviews* **2003**, *22*, 158-81.

[132] Laskin, J.; Futrell, J. H.: Activation of Large Ions in FT-ICR Mass Spectrometry. *Mass Spectrometry Reviews* **2005**, *24*,135-167.

[133] Laskin, J.; Byrd, M.; Futrell, J.: Internal energy distributions resulting from sustained off-resonance excitation in FTMS. I. Fragmentation of the bromobenzene radical cation. *International Journal of Mass Spectrometry* **2000**, *195/196*, 285-302.

[134] Laskin, J.; Denisov, E.; Futrell, J. H.: A Comparative Study of Collision-Induced and Surface-Induced Dissociation. I Fragmentation of Protonated Dialanine. *Journal of the American Chemical Society* **2000**, *122*, 9703-14.

[135] Laskin, J.; Denisov, E.; Futrell, J. H.: Fragmentation energetics of small peptides from multiple-collision activation and surface-induced dissociation in FT-ICR MS. *International Journal of Mass Spectrometry* **2002**, *219*, 189-201.

[136] Nesatyy, V. J.; Laskin, J.: Dissociation of non-covalent protein complexes by triple quadrupole tandem mass spectrometry: comparison of Monte Carlo simulation and experiment. *International Journal of Mass Spectrometry* **2003**, *221*, 245-62.

[137] Laskin, J.; Futrell, J.: Internal Energy Distributions Resulting from Sustained Off-Resonance Excitation in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. II. Fragmentation of the 1-Bromonaphthalene Radical Cation. *Journal of Physical Chemistry A* **2000**, *104*, 5484-94.

[138] Bastian, M. J.; Dressler, R. A.; Levandier, D. J.; Murad, E.; Muntean, F.; Armentrout, P. B.: Low energy collision-induced dissociation and photodissociation studies of the $(N_2O,H_2O)^+$ cluster ion. *Journal of Chemical Physics* **1997**, *106*, 9570-9.

[139] Ervin, K. M.; Armentrout, P. B.: Translational energy dependence of $Ar^+ + XY \rightarrow ArX^+ + Y$ ($XY=H_2,D_2,HD$) from thermal to 30 eV c.m. *Journal of Chemical Physics* **1985**, *83*, 166-89.

[140] Muntean, F.; Armentrout, P. B.: Guided ion beam study of collision-induced dissociation dynamics: integral and differential cross sections. *Journal of Chemical Physics* **2001**, *115*, 1213-28.

[141] Knyazev, V. D.; Stein, S. E.: Classical Trajectories and RRKM Modeling of Collisional Excitation and Dissociation of Benzylammonium and *tert*-Butyl Benzylammonium Ions in a Quadrupole-Hexapole-Quadrupole Tandem Mass Spectrometer. *Journal of the American Society for Mass Spectrometry* **2010**, *21*, 425-39.

[142] Knyazev, V. D.; Stein, S. E.: Monte Carlo/RRKM/Classical Trajectories Modeling of Collisional Excitation and Dissociation of *n*-Butylbenzene Ion in Multipole Collision Cells of Tandem Mass Spectrometers. *Journal of Physical Chemistry A* **2010**, *114*, 6384–93.

[143] Herbert, C. G.; Johnstone, R. A. W.: *Mass Spectrometry Basics*; CRC Press, Boca Raton, 2003.

[144] Kebarle, P; Peschke, M.: On the Mechanism by Which the Charged Droplets Produced by Electrospray Lead to Gas Phase Ions. *Analytica Chimica Acta* **2000**, *406*, 11-35.

[145] Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A.: Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004**, *25*, 1157-74.

[146] Allinger, N. L.: Conformational Analysis 130. MM2 A Hydrocarbon Force Field Utilizing V_1 and V_2 Torsional Terms. *Journal of the American Chemical Society* **1977**, *99*, 8127-34.

[147] Allinger, N. L.: Molecular Mechanics. The MM3 Force Field for Alkenes. *Journal of Computational Chemistry* **1990**, *11*, 848-67.

[148] Allinger, N. L.: Molecular Mechanics (MM3) Calculations on Conjugated Hydrocarbons. *Journal of Computational Chemistry* **1990**, *11*, 868-95.

[149] Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I.: Fast, efficient generation of high-quality atomic charges. AM1-BCC Model: I. Method. *Journal of Computational Chemistry* **2000**, *21*, 132-46.

[150] Jakalian, A.; Jack, D. B.; Bayly, C. I.: Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of Computational Chemistry* **2002**, *23*, 1623-41.

[151] Levinthal, C.: How to Fold Graciously. In *Mossbauer Spectroscopy in Biological Systems*; DeBrunner, J. T. P., Munck, E., Eds.; University of Illinois Press: Allerton House, Monticello, Illinois, USA, 1969; pp. 22-24.

[152] Verlet, L.: Computer 'Experiments' on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Reviews* **1967**, *159*, 98-103.

[153] Hockney, R. W.: The Potential Calculation and Some Applications. *Methods in Computational Physics* **1970**, *9*, 136-211.

[154] Wu, X.; Brooks, B. R.: Self-guided Langevin dynamics simulation method. *Chemical Physics Letters* **2003**, *381*, 512-8.

[155] Kirkpatrick, S.; Gelatt, C. D., Jr.; Vecchi, M. P.: Optimization by simulated annealing. *Science* **1983**, *220*, 671-80.

[156] Kottalam, J.; Case, D. A.: Langevin modes of macromolecules: applications to crambin and DNA hexamers. *Biopolymers* **1990**, *29*, 1409-21.

[157] Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, R. C.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A.: *AMBER 9*. University of California, San Francisco, USA, 2006.

[158] Lamm, G.; Szabo, A.: Langevin modes of macromolecules. *Journal of Chemical Physics* **1986**, *85*, 7334-48.

[159] Huang, X.; Atwood, C. S.; Moir, R. D.; Hartshorn, M. A.; Vonsattel, J. P.; Tanzi, R. E.; Bush, A. L.: Zinc-induced Alzheimer's Aβ1-40 aggregation is mediated by conformational factors. *The Journal of Biological Chemistry* **1997**, *272*, 26464-70.

[160] Shen, C.-L.; Murphy, R. M.: Solvent effects on self-assembly of β-amyloid peptide. *Biophysical Journal* **1995**, *69*, 640-51.

[161] Talafous, J.; Marcinowski, K. J.; Klopman, G.; Zagorski, M. G.: Solution structure of residues 1-28 of the amyloid β-peptide. *Biochemistry* **1994**, *33*, 7788-96.

[162] Corcho, F. J.; Filizola, M.; Pérez, J. J.: Evaluation of the iterative simulated annealing technique in conformational search of peptides. *Chemical Physics Letters* **2000**, *319*, 65-70.

[163] Villani, V.: Methodological procedures of Molecular Dynamics Simulated Annealing for the chain folding of the flexible tetrapeptide Boc-Gly-Leu-Gly-Gly-NMe with analysis of the evolution mechanism. *Journal of Molecular Structure (Theochem)* **1996**, *363*, 43-66.

[164] Villani, V.; D'Alessio, L.; Tamburro, A. M.: Conformational non-linear dynamical behavior of the peptide Boc-Gly-Leu-Gly-Gly-NMe. *Journal of the Chemical Society Perkin Transactions* **1997**, *2*, 2375-85.

[165] Frisch, M.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery Jr., J. A.; Vreven, T.; Judin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, W. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Ciolowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.: *Gaussian 03, revision C.02*. Gaussian Inc.: Wallingford, USA, 2004.

[166] Moon, J. H.; Oh, J. Y.; Kim, M. S.: A Systematic and Efficient Method to Estimate the Vibrational Frequencies of Linear Peptide and Protein Ions with Any Amino Acid Sequence for the Calculation of Rice-Ramsperger-Kassel-Marcus Rate Constant. *Journal of the American Society for Mass Spectrometry* **2006**, *17*, 1749-57.

[167] Scott, A. P.; Radom, L.: Harmonic Vibrational Frequencies: An Evaluation of Hartree-Fock, Møller-Plesset, Quadratic Configuration Interaction, Density Functional Theory, and Semiempirical Scale Factors. *Journal of Physical Chemistry* **1996**, *100*, 16502-13.

[168] Kenttämaa, H. I.; Cooks, R. G.: Internal energy distributions acquired through collisional activation at low and high energies. *International Journal of Mass Spectrometry and Ion Processes* **1985**, *64*, 79-83.

[169] Martínez-Núñez, E.; Fernández-Ramos, A.; Vázquez, S. A.; Marques, J. M. C.; Xue, M.; Hase, W. L.: Quasiclassical dynamics simulation of the collision-induced dissociation of $Cr(CO)_6^+$ with Xe. *Journal of Chemical Physics* **2005**, *123*, 154311-9.

[170] Douglas, D. J.: Applications of Collision Dynamics in Quadrupole Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **1998**, *9*, 101-13.

[171] Wysocki, V. H.; Kenttämaa, H. I.; Cooks, R. G.: Internal energy distributions of isolated ions after activation by various methods. *International Journal of Mass Spectrometry and Ion Processes* **1987**, *75*, 181-208.

[172] Beyer, T.; Swinehart, D. F.: Algorithm 448: number of multiply-restricted partitions. *Communications ACM* **1973**, 379.

[173] Alhazmi, A. M. Mass spectrometry of polymers: from synthesis to sequence (Ph.D. thesis), Ottawa, Canada. 2008.

[174] Renaud, J. B.; Martineau, E.; Mironov, G. G.; Berezovski, M.; Mayer, P. M.: The collaborative role of molecular conformation and energetics in the binding of gas phase non-covalent polymer/amine complexes. *Physical Chemistry and Chemical Physics* **2012**, *14*, 165-172.

[175] IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997).

[176] Cooks, R. G.; Patrick, J. S.; Kotiaho, T.; McLuckey, S. A.: Thermodynamical determinations by the kinetics method. *Mass Spectrometry Reviews* **1994**, *13*, 287-339.

[177] Holmes J.L.; Aubry, C.; Mayer, P.M.: *Assigning structures to ions in mass spectrometry*. CRC Press: Boca Raton, USA, 2007.

[178] Hunter, E. P.; Lias, S. G.: Evaluated Gas Phase Basicities and Proton Affinities of Molecules: An Update. *Journal of Physical and Chemical Reference Data* **1998**, *27*, 413-656.

[179] Martineau, E.; de Guzman, J. M.; Rodionova, L.; Kong, X.; Mayer, P. M.; Aman, A. M.: Investigation of Non-covalent Interactions Between Anti-Amyloid Agents and Amyloid β Peptides by ESI-MS. *Journal of the American Society for Mass Spectrometry* **2010**, *21*, 1506-14.

[180] Wortmann, A.; Rossi, F.; Lelais, G.; Zenobi, R.: Determination of zinc to β-peptide binding constants with electrospray ionization mass spectrometry. *Journal of Mass Spectrometry* **2005**, *40*, 777-84.

[181] Dember, L. M.; Hawkins, P. N.; Hazenberg, M. D.; Bouke P. C.; Gorevic, P. D.; Merlini, G.; Butrimiene, I.; Livneh, A.; Lesnyak, O.; Puechal, X.; Lachmann, H. J.; Obicic, L.; Balshaw, R.; Garceau, D.; Hauck, W.; Skinner, M.; Bhargava, A.; Faragasso, A.; Levy, K.; Keown, P.; Kaarela, K.; Grateau, G.; Hachulla, E.; Livneh, A.; Rosner, I.; Merlini, G.; Obici, L.; Butrimiene, I.; Kirdaite, G.; Povilenaite, D.; Venalis, A.; Bijzet, J.; Filipowicz-Sosnowska, A.; Wiland, P.; Chlebicki, A.; Lesnyak, O.; Koryakova, N.; Sibiryakova, A.; Nasonov, E. L.; Stanislav, M.; Jover, J. A.; Munoz, J.; Tena Marsa, X.; Valverde Garcia, J.; Ben Maiz, H.; Ben Mousa, F.; Goucha, R.; Kaaroud, H.; Direskeneli, H.; Temel, M.; Gul, A.; Kamali, S.; Hatemi, G.; Ozdogan, H.; Hawkins, P. N.; Lachman, H. J.; Hunter, J. A.; Benson, M. D.; Dispenzieri, A.: Eprodisate for the treatment of renal disease in AA amyloidosis. Eprodisate for the treatment of renal disease in AA amyloidosis. *The New England Journal of Medicine* **2007**, *356*, 2349-60.

[182] Mayer, P. M.; Martineau, E.: Gas-phase binding energies for non-covalent Aβ-40 peptide/small molecule complexes from CID mass spectrometry and RRKM theory. *Physical Chemistry and Chemical Physics* **2011**, *13*, 5178-86.

Appendix A: Programs

A.1 Script to run the Molecular Dynamics Simulated Annealing simulations on the $PMMA_n$ oligomers and the $(PMMA_n)(AA)(H^+)$ complexes.

This is an example of a script using an annealing schedule: "300 K – 800 K – 300 K, 0 K". The details of the annealing schedules are written separate input files.

```
#! /bin/bash
#$ -S /bin/bash
#$ -cwd
#$ -V

#$ -o job.out
#$ -e job.err

date

maxit="2000"
ntwx="250"
#temp="10000"

echo "Simulated Annealing script, University of Ottawa, Eric Martineau, 2009"
echo "300-800-300-0K"
echo "SA begins"

sander -O -i 01_init_min.in -o 01_init_min.out -p 7p3.prmtop -c 7p3.inpcrd -r 01_init_min.rst
echo "Minimization done"

$AMBERHOME/exe/ambpdb -p 7p3.prmtop < 01_init_min.rst > 01_init_min.pdb
echo "trajin 01_init_min.pdb" > process.trajin

sander -O -i 02_heat_0-300.in -p 7p3.prmtop -c 01_init_min.rst -r 02_heat_0-300.rst -o
02_heat_0-300.out -x 02_heat_0-300.mdcrd
echo "Heating 0-300K done"

sander -O -i 03_equi_300.in -p 7p3.prmtop -c 02_heat_0-300.rst -r 03_equi_300.rst -o
03_equi_300.out -x 03_equi_300.mdcrd
echo "Equilibration @300K done"
```

```
for (( ii = 1 ; ii <= $maxit ; ii++ ))
do

sander -O -i 04_anneal_800.in -p 7p3.prmtop -c 03_equi_300.rst -r 04_anneal_800.rst -o
04_anneal_800.out -x 04_anneal_800.mdcrd
echo $ii "Annealing done"

./process_mdout.perl 04_anneal_800.out
var=$(./min.perl < summary.EPTOT)
echo "grep ' "$var" ' 04_anneal_800.out > out" >> grep_anneal
chmod 777 grep_anneal
./grep_anneal
rm grep_anneal
nstep=$(awk -f awk.nstep out)
frame=$(($nstep/$ntwx))
echo $nstep
echo $frame
echo -e "trajin 04_anneal_800.mdcrd" $frame $frame"\ntrajout lowest_energy.rst restart" >>
extract_frame.trajin

$AMBERHOME/exe/ptraj 7p3.prmtop < extract_frame.trajin > extract_frame.out

mv extract_frame.out ${ii}_lowest_energy.out

mv lowest_energy.rst.* lowest_energy.rst

rm summary* out extract_frame.trajin

#echo $var
#echo $temp

#if [ $var < $temp ]
#then
#cp lowest_energy.rst global_min.rst
#temp=$var
#fi
#echo $temp
#echo $var

#exit

sander -O -i 05_cool_300-0.in -p 7p3.prmtop -c lowest_energy.rst -r 05_cool_300-0.rst -o
05_cool_300-0.out -x 05_cool_300-0.mdcrd
echo $ii "Cooling 300-0K done"

mv lowest_energy.rst ${ii}_lowest_energy.rst.${frame}
```

```
sander -O -i 06_post_min.in -p 7p3.prmtop -c 05_cool_300-0.rst -r 06_post_min.rst -o
06_post_min.out -x 06_post_min.mdcrd
echo $ii "post minimization done"

$AMBERHOME/exe/ambpdb -p 7p3.prmtop < 06_post_min.rst > 06_post_min.pdb

mv 06_post_min.pdb ${ii}_06_post_min.pdb
echo "trajin "${ii}"_06_post_min.pdb" >> process.trajin


if [ $ii -ne $maxit ]
then
  cp 04_anneal_800.rst ${ii}_04_anneal_800.rst
  mv 04_anneal_800.rst 03_equi_300.rst
else
  cp 04_anneal_800.rst ${ii}_04_anneal_800.rst
fi

mv 04_anneal_800.out ${ii}_04_anneal_800.out
mv 04_anneal_800.mdcrd ${ii}_04_anneal_800.mdcrd

mv 05_cool_300-0.out ${ii}_05_cool_300-0.out
mv 05_cool_300-0.mdcrd ${ii}_05_cool_300-0.mdcrd
mv 05_cool_300-0.rst ${ii}_05_cool_300-0.rst

mv 06_post_min.out ${ii}_06_post_min.out
mv 06_post_min.rst ${ii}_06_post_min.rst


date

if [ -f fort.9 ]
then
   echo "core dumped, abort run"
   exit
fi

done

# calculates RMSD of each conformation at the end of a cycle vs the initial conformation
echo "reference 01_init_min.pdb" >> process.trajin
echo "rms reference out rmsd.dat *" >> process.trajin

# calculates the dihedral angles of the polymer side-chains
#echo "dihedral R1.dat :1@C35 :1@O34 :1@C7  :1@C5  out R1.dat" >> process.trajin
```

```
#echo "dihedral R2.dat :1@C47 :1@O43 :1@C13 :1@C11 out R2.dat" >> process.trajin
#echo "dihedral R3.dat :1@C39 :1@O51 :1@C18 :1@C15 out R3.dat" >> process.trajin
#echo "dihedral R4.dat :1@C67 :1@O56 :1@C25 :1@C21 out R4.dat" >> process.trajin

# calculates the "phi and psi" angles of the polymer backbone"
#echo "dihedral psi1.dat :1@C1  :1@C5  :1@C8  :1@C11 out psi1.dat" >> process.trajin
#echo "dihedral phi2.dat :1@C5  :1@C8  :1@C11 :1@C14 out phi2.dat" >> process.trajin
#echo "dihedral psi2.dat :1@C8  :1@C11 :1@C14 :1@C15 out psi2.dat" >> process.trajin
#echo "dihedral phi3.dat :1@C11 :1@C14 :1@C15 :1@C20 out phi3.dat" >> process.trajin
#echo "dihedral psi3.dat :1@C14 :1@C15 :1@C20 :1@C21 out psi3.dat" >> process.trajin
#echo "dihedral phi4.dat :1@C15 :1@C20 :1@C2H12 :1@C21 out phi4.dat" >> process.trajin

# calculates the H-bond angle"
#echo "dihedral HB.dat :1@O34 :1@C7 :1@O33 :1@4H12 out HB.dat" >> process.trajin

# calculates end-to-end distances"
#echo "distance d_4mer.dat :1@C5 :1@C21 out d_4mer.dat" >> process.trajin
#echo "distance d_comp.dat :1@N64 :1@C67 out d_comp.dat" >> process.trajin

# average conformation
echo "average average.pdb pdb" >> process.trajin

# HBD masks
#echo "donor mask :2@O1" >> process.trajin
#echo "donor mask :2@N" >> process.trajin
#echo "donor mask :1@O34" >> process.trajin
#echo "donor mask :1@O42" >> process.trajin
#echo "donor mask :1@O43" >> process.trajin
#echo "donor mask :1@O51" >> process.trajin
#echo "donor mask :1@O52" >> process.trajin
#echo "donor mask :1@O56" >> process.trajin
#echo "donor mask :1@O57" >> process.trajin

# HBA masks
#echo "acceptor mask :1@O33 1@4H12" >> process.trajin
#echo "acceptor mask :2@O 2@H4" >> process.trajin
#echo "acceptor mask :2@N 2@H" >> process.trajin
#echo "acceptor mask :2@N 2@H1" >> process.trajin

#hbond print .05 series hbt

$AMBERHOME/exe/ptraj 7p3.prmtop < process.trajin > process.dat

sander -O -i 07_equi_300.in -p 7p3.prmtop -c 04_anneal_800.rst -r 07_equi_300.rst -o
07_equi_300.out -x 07_equi_300.mdcrd
echo $maxit "Equilibration at 300K done"
```

```
echo -e "trajin 07_equi_300.mdcrd" >> process_300.trajin
echo -e "reference "${maxit}"_06_post_min.pdb" >> process_300.trajin
echo -e "rms reference out rmsd_to_final.dat *" >> process_300.trajin
echo -e "average average_300.pdb pdb" >> process_300.trajin

$AMBERHOME/exe/ptraj 7p3.prmtop < process_300.trajin > process_300.dat

mkdir Analysis_800K
mv *.dat average* 07*rst 07*mdcrd 07*out *trajin Analysis_800K

mkdir Results_800K
mv mdinfo *out *mdcrd *rst* *pdb job*err Results_800K

if [ -f fort.9 ]
then
  echo "Abnormal termination"
  exit
fi

date
```

A.2 Awk and Perl codes called inside the MD-SA script

Script: **Awk.min**. This script was used to..

```
NR == 1 { min = $2 }
    $2 < min { min = $2 }
    END { print min }
```

Script: **Awk.nstep**. This script was used to ...

```
NR == 1 { nstep = $3 }
    END { print nstep }
```

Script: **min.perl**. This script was used to ...

```
#!/bin/perl

my %erg;

while (<>) {
  # match pairs of numbers
  /([\d\.]+)\s+([\d\.]+)/ or next;

  # time => energy map
  $erg{$1} = $2;
}

# build sorted time => energy map, smallest energy first
my @s_erg = sort {$erg{$a} <=> $erg{$b}} (keys %erg);

print $s_erg[0] . "\n";
```

A.3 Program for calculating the sum of states, the density of states, and the unimolecular rate constants

This program was used for the calculation of the density and sum of states and unimolecular rate constants with the use of the Beyer-Swinehart scheme for Direct Count.

```
!  This FORTRAN program was adapted from the BASIC version quoted in
!  P.M. Mayer and T. Baer, J. Am. Soc. Mass Spectrom. 1997, 8, 103-115.
!   by Paul M. Mayer (Chem Dept, U of Ottawa, Ottawa, Canada K1N 6N5)
!    25/7/97
!
!  The variables:
!  Eo   the activation energy
!  MIN and IMIN are the minimum energy for which you want a k printed
!  MAX and IMAX are the maximum energy for which you want a k calc and print
!  NUM and INUM is the interval you want the k printed at
!  Nfreq is the number of vibrational frequencies for both ion and ts
!  w(1000) is the array into which freqs are put
!  p(32000) is the density of states vector in which densities are put
!  N(32000) is the sum of states vector in which sums are put
!  RATE(32000) is the rate constant vector into which k's are put
!  LOGRATE(32000) is the vector into which is put log10(k)
!  QMolc, QTs are the ion and ts partition functions at 600K
!  UMolc, UTs are the ion and ts internal energies at 600K
!
!  The constants:
!   H=3.3364E-11 is Planck's constant in cm-1 sec
!   Kb=1.3807E-23 in Boltzman's constant in J/K
!   Na=6.022E23 is Avogadro's number
!   HJ=6.6262E-34 is Planck's constant in Jsec
!   c=2.9979E10 is the speed of light in cm/sec
!   T is temperature in Kelvin
!
!  STEP 1. Variable declaration:
!
      INTEGER IEo,Nfreq,IMIN,INCR,IMAX,ITMAX,J,I,T,v(3000)
      REAL w(3000), m(3000)
      DOUBLE PRECISION p(480000)
      DOUBLE PRECISION N(480000)
      DOUBLE PRECISION RATE(480000)
      DOUBLE PRECISION QMolc, UMolc, QTs, UTs
      REAL Eo,MAX,MIN,NUM,Sts,H,Kb,Na,HJ,c,Sigma
      REAL LOGRATE(480000)
      CHARACTER Title*80,Name*6
!
```

```
! STEP 2. Reading in the stuff
!
     WRITE (*,'("input file name")')
     READ (*,'(a6)') Name
     OPEN (UNIT=15, FILE=Name, STATUS='OLD')
     READ (15,'(a80)') Title
     READ (15,*) Eo,MIN,MAX,NUM,Sigma
     READ (15,*) Nfreq
     READ (15,*) (w(I),I=1,Nfreq)

!
! STEP 3. Converting input energies to integral wavenumbers
!
     Eo=Eo*10000          !converting to units of 0.1 meV
     IEo=NINT(Eo)         ! converting to an integer
     MAX=MAX*10000        ! converting to units of 0.1meV
     IMAX=NINT(MAX)       ! converting to an integer
     ITMAX=IMAX-IEo       ! defining transition state maximum energy
     MIN=MIN*10000
     IMIN=NINT(MIN)
!
!       Converting input frequencies to units of 0.1meV and then integers
!
     DO 20 J=1,Nfreq,1
     m(J)=w(J)*1.24
     v(J)=NINT(m(J))
20    CONTINUE
!
! STEP 4. Initializing p and N -vectors
!
     p(1)=1
     N(1)=1
     DO 50 J=2,480000,1
        p(J)=0
        N(J)=1
50    CONTINUE
!
! STEP 5. Calculating density of states
!
     DO 200 J=1,Nfreq,1
        DO 100 I=v(J),IMAX,1
          IF ((I-v(J)) .EQ. 0) THEN
             p(I)=p(I)+0
          ELSE
           p(I) = p(I) + p(I - v(J))
          END IF
```

```
100      CONTINUE
200   CONTINUE
!
!     PRINT 201, p(5000),p(12000),p(16000)
!201   FORMAT (1PG10.4E2)
!
! STEP 6. Calculating ion partition function and internal energy
!
     QMolc=1
     UMolc=0
     DO 250 I=1,Nfreq,1
       QMolc=QMolc*(1/(1-EXP(-w(I)/208.4)))     !calcing molc partion fcn
       UMolc=UMolc + (w(I)/(EXP(w(I)/208.4)-1)) ! calcing molc int nrg
250   CONTINUE
!
!     PRINT 260, QMolc,UMolc
!260   FORMAT (1PG11.4E3)
!
! STEP 7. Calculating sum of states
!
     READ (15,*) Nfreq
     READ (15,*) (w(I),I=1,Nfreq)
!
     DO 270 J=1,Nfreq,1
     m(J)=w(J)*1.24
     v(J)=NINT(m(J))
270   CONTINUE
!
     DO 400 J=1,Nfreq,1
       DO 300 I=v(J),ITMAX,1
         N(I) = N(I) + N(I-v(J))
300      CONTINUE
400   CONTINUE
!
! STEP 8. Calculating ts partition function and internal energy at 600K
!
     QTs=1
     UTs=0
     DO 450 I=1,Nfreq,1
       QTs=QTs*(1/(1-EXP(-w(I)/208.4)))     !calcing molc partion fcn
       UTs=UTs + (w(I)/(EXP(w(I)/208.4)-1)) ! calcing molc int nrg
450   CONTINUE
!
!     PRINT 451, QTs,UTs
!451   FORMAT (1PG11.4E3)
!
```

```
! STEP 9. Calculating rate constant at specific intervals
!
     H=(3.3364E-11)*(1.24)
     DO 500 J=1,ITMAX,1
     RATE(J)=(Sigma*N(J))/(H*p(J+IEo))
500   CONTINUE
!
! STEP 10. Calculating entropy of activation at 600K
!
     Kb=1.3807E-23
     Na=6.022E23
     HJ=6.6262E-34
     c=2.9979E10
     T=300
     Sts= ((Kb*LOG(Qts/QMolc)) + (HJ*c*(UTs-UMolc))/T)*Na
!
! STEP 8. Printing it off
!
     PRINT '(''RRKM k(E) vs E using the Direct Count method'')'
     PRINT 550, Title
550    FORMAT (1X,A80)
     PRINT 560, Sts
560   FORMAT ('entropy of activation (J/K/mol) at 600K is', 1PG15.7E2)
     PRINT 600
600   FORMAT ('  E                RATE             density ')
!
     INCR=INT((IMAX-IMIN)/(NUM-1))
     DO 800 J=IMIN-IEo,IMAX-IEo,INCR
     PRINT 700, (J+IEo)/10000.0, RATE(J), p(J+IEo)
700   FORMAT (1PG10.4E2,E20.7E2,1PG20.4E2,E20.7E3)
!
800   CONTINUE
!
!     write the data to a file
     open(unit=10 ,file='k1.dat', status='replace')

     DO 950 J=IMIN-IEo,IMAX-IEo,(IMAX-IMIN)/900
     WRITE (10,900) RATE(J)
900   FORMAT (E13.7E2)
!
950   CONTINUE
     endfile 10
     close (unit=10)

!     write the data to a file
     open(unit=11 ,file='e.dat', status='replace')
```

```
      DO 980 J=IMIN-IEo,IMAX-IEo,(IMAX-IMIN)/900
      WRITE (11,960) (J+IEo)/10000.0
960   FORMAT (E13.7E2)
!
980   CONTINUE
      endfile 11
      close (unit=11)

!     write the data to a file
      open(unit=12 ,file='p.dat', status='replace')

      DO 1050 J=IMIN-IEo,IMAX-IEo,(IMAX-IMIN)/900
      WRITE (12,1000) p(J+IEo)
1000  FORMAT (E14.7E3)
1050  CONTINUE
      endfile 12
      close (unit=12)
      close (unit=15)
      END
```

A.4 Program for fitting the breakdown diagrams

This program was used to fit the experimental breakdown curves.

```
!       Program main
!       Calculate breakdown curve. Convolute breakdown curve with time scale, P(E,T)
!
              integer i,j,l,m,nptsk,nptse,nptecom, npfreq
        real k(1000),ep(1000),R,h,kb,S,F
        real elab(1000),ecom(1000),mw
              real k1(1000),k2(1000),TFracdiss(1000),freq(1000)
        real time(1000),Fracdiss(1000,1000)
              real kobs(1000,1000)
        real Parent(1000), Product(1000)
        double precision Qt(1000),q(1000),Teff(1000)
        double precision p(1000), Eint(1000,1000),ek(1000)
        double precision TEint(1000), NEint(1000,1000)

        do 1,i=1,1000
          k1(i)=0.0
          ecom(i)=0.0
          elab(i)=0.0
          Teff(i)=0.0
          q(i)=0.0
          Qt(i)=1.0
          TEint(i)=0.0
          NEint(i,j)=0.0
          TFracdiss(i)=0.0
          freq(i)=0.0
!         k2(i)=0.0
          freq(i) = 0.0
 1        continue
!       print *,k1(1000)

!       Read in rate constant from RRKM calculation
        open (unit=13, file='k1.dat',status='old')
!     +   form='formatted', access='sequential')
        i=1
 5       read(13,*,end=6) k1(i)
        i=i+1
```

```
      goto 5
 6     nptsk=i-1

!      Read in internal energy from RRKM calculation
      open (unit=14, file='e.dat', status='old',
     +   form='formatted', access='sequential')
      i=1
 7     read(14,*,end=8) ek(i)
      i=i+1
      goto 7
 8     nptse=i-1
!
!      open (unit=17, file='k2.dat',status='old')
!      i=1
!7          read(17,*,end=8) ek(i),k2(i)
!      i=i+1
!      goto 7
!8     nptsk=i-1
!
      print *,nptsk

      do 9, i=1,nptsk
      k(i) = k1(i)
 9     continue

!      Read in MW of parent ion, calculate observation window
      open (unit=15, file='mw.dat', status='old')
      read (15,*,err=10) mw
!10    print *,mw
 10    do 11, i=1,1000
      elab(i) = i/10.0
      ecom(i) = elab(i)*40.0/(mw+40.0)
      time(i) = 1/(SQRT(elab(i)*100000.0*2.0*(1000.0/mw)*100.0))
 11    continue
!     print *, elab(1000)
!     print *, time(1), time(1000)

!      Calculate kobs based on k and time
      do 13, i=1,nptsk
      do 12, j=1,1000
```

```
      kobs(i,j)=1-exp(-k(i)*time(j))
 12    continue
 13    continue
!      print *,kobs(1000,1000)

!      Calculate effective T for Q calc
      open (unit=16, file='slope.dat', status='old')
      read (16,*,err=14) slope
!14    print *,slope
 14    do 15, i=1,1000
      Teff(i)=300+slope*ecom(i)
 15    continue
      print *,Teff(1000)

!      read in vibrational frequencies for Q calc
      i=1
      open (unit=17, file='freq.dat', status='old')
!    +   form='unformatted', access='sequential')
 16    read(17,*,end=17) freq(i)
      i=i+1
      goto 16
 17    npfreq=i-1
     print *,npfreq
!     print *,freq(252)

!      Calculate partition function, Q
      h=6.6262E-34
      kb=1.38E-23
      R=8.314
      S=2.99E+8
      do 28, j=1,1000
      i = 1
 26    q(j)=1.0/(1.0-exp((-freq(i)*100.0*S*h)/(kb*Teff(j))))
      Qt(j)=Qt(j)*q(j)
      i = i+1
      if(i .GT. npfreq) goto 28
      goto 26
 28    continue
      print *,Qt(1000)
```

```
      open (unit=18, file= 'p.dat',status='unknown')
!    +   form='formatted', access='sequential')
      i=1
 29    read(18,*,end=30) p(i)
      i=i+1
      goto 29
 30    nptsp=i-1
      print *,nptsp

!     Calculate Internal Energy Distribution
      F=100.0
      do 42, j=1,1000
      i=1
 41    Eint(i,j)=(p(i)/Qt(j))*(DEXP(-(ek(i)*96480)/(R*Teff(j))))
      i=i+1
      if (i .GT. nptsp) goto 42
      goto 41
 42    continue
      print *, Eint(104,1000)

!     Calculate total population at each Ecom so we can normalize the distribution
      do 44, j=1,1000
      i = 1
 43    TEint(j)=TEint(j)+Eint(i,j)
      i=i+1
      if (i .GT. nptsp) goto 44
      goto 43
 44    continue
      print *,TEint(1000)

!     Normalize the internal energy distributuion
      do 56, j=1,1000
      i = 1
 55    NEint(i,j) = Eint(i,j)/TEint(j)
      i = i+1
      if (i .GT. nptsp) goto 56
      goto 55
 56    continue
!     print *,NEint(1000,1000)
```

```
      close (unit=13)
      close (unit=14)
      close (unit=15)
      close (unit=16)
      close (unit=17)
      close (unit=18)

!     Calc net observation
      do 63 j=1,1000
      i = 1
 62    Fracdiss(i,j)=kobs(i,j)*NEint(i,j)
      i = i+1
      if (i .GT. nptsp) goto 63
      goto 62
 63    continue
!     Calculate total observable at each Ecom
      do 66, j=1,1000
      i = 1
 65    TFracdiss(j) = TFracdiss(j)+Fracdiss(i,j)
      i = i+1
      if (i .GT. nptsp) goto 66
      goto 65
 66    continue
      print *,TFracdiss(1), TFracdiss(1000)

!     Calculate total observable as a fraction of total
      do 68, j=1,1000
 67    Product(j) = TFracdiss(j)
      Parent(j) = 1-Product(j)
 68    continue

!     write the breakdown diagram to a file
      open(unit=8 ,file='breakdown.dat', status='unknown')

      do 200 j=1,1000
      write(8,100) ecom(j),Parent(j),Product(j)
 100   format(1PG20.8E2,1PG20.8E2,1PG20.8E2)
 200    continue

!     write the population to a file
```

```
      open(unit=11 ,file='population.dat', status='replace')
      do 400 i=1,nptsk,1
      j=1000
      write(11,300) ek(i), Eint(i,j)
 300   format(1PG20.8E3,1PG20.8E3)
 400   continue
      close (unit=11)

!     print 50 points to the screen to facilitate fitting
      PRINT 600
600   FORMAT ('  Ecom              Parent             Product  ')
*
      DO 800 J=1,1000,40
      PRINT 700, ecom(j), Parent(j), Product(j)
700   FORMAT (1PG10.4E2,1PG20.4E2,1PG20.4E2)
*
800   CONTINUE
      close (unit=8)
      end
```

A.5 Program for calculating the sum of states, the density of states, and the unimolecular rate constants using the Hill equation approximation

This program was used for the calculation of the density of states, sum of states and unimolecular rate constants with the use of the Beyer-Swinehart scheme for Direct Count and the Hill equation approximation.

```
!  This FORTRAN program was adapted from the BASIC version quoted in
!  P.M. Mayer and T. Baer, J. Am. Soc. Mass Spectrom. 1997, 8, 103-115.
!   by Paul M. Mayer (Chem Dept, U of Ottawa, Ottawa, Canada K1N 6N5)
!    25/7/97
!
!  The variables:
!  Eo   the activation energy
!  MIN and IMIN are the minimum energy for which you want a k printed
!  MAX and IMAX are the maximum energy for which you want a k calc and print
!  NUM and INUM is the interval you want the k printed at
!  Nfreq is the number of vibrational frequencies for both ion and ts
!  w(1000) is the array into which freqs are put
!  p(32000) is the density of states vector in which densities are put
!  N(32000) is the sum of states vector in which sums are put
!  RATE(32000) is the rate constant vector into which k's are put
!  LOGRATE(32000) is the vector into which is put log10(k)
!  QMolc, QTs are the ion and ts partition functions at 600K
!  UMolc, UTs are the ion and ts internal energies at 600K
!
!  The constants:
!   H=3.3364E-11 is Planck's constant in cm-1 sec
!   Kb=1.3807E-23 in Boltzman's constant in J/K
!   Na=6.022E23 is Avogadro's number
!   HJ=6.6262E-34 is Planck's constant in Jsec
!   c=2.9979E10 is the speed of light in cm/sec
!   T is temperature in Kelvin
!
!  STEP 1. Variable declaration:
!

      INTEGER  IEo ,IMAX, ITMAX, Nfreq ,J ,I ,T ,o,k, npts,v(3000)
      REAL   w(3000), slope(10)
      REAL   g(3000)
      real*16  p(480000)
      real*16  N(480000)
      real*16  Rate(480000)
      real*8  QMolc, UMolc, QTs, UTs
```

```
      REAL  Eo ,MAX ,Sts ,H ,Kb ,Na ,HJ ,c
      integer i,j,l,m,nptsk,nptse,nptecom, npfreq,nptsexpt
      integer nptssl
      real ep(1000),R,h,kb,S,F, TDiff(10), Diff(10,100)
      real elab(100),ecom(100),mw
      real TFracdiss(10,100),freq(2000)
      real time(100),Fracdiss(1000,10,100)
      real kobs(1000,100)
      real Parent(10,100), Product(10,100) , Expt(100)
      real*16 Qt(10,100),q(10,100), Teff(10,100)
      real*16 den(1000)
      real*16 TEint(10,100), NEint(1000,10,100)
      real*16 Eint(1000,10,100)
      real*16  k1(1000), ek(1000), kt(1000)


!
! STEP 2. Reading in the stuff sequentially from rrkm.dat
!
!      open (unit=14, file='npts.dat', status='old')
!       read (14,*) npts
!      close (unit=14)

      open(unit=20 ,file='results.dat', status='new',
     +     position='append')

      open(unit=21 ,file='breakdown.dat', status='new',
     +     position='append')

      OPEN (UNIT=15, FILE='rrkm.dat', STATUS='OLD')
!     +   form='formatted', access='sequential')

      o=1
10    READ (15,*,end=6000) Eo
      read (15,*) MAX
      READ (15,*) Nfreq
      READ (15,*) (w(I),I=1,Nfreq)
!
! STEP 3. Converting input energies to integral wavenumbers
!
      Eo=Eo*10000          !converting to units of 0.1 meV
      IEo=NINT(Eo)         ! converting to an integer
      MAX =MAX*10000
      IMAX = NINT(MAX)
      ITMAX=IMAX-IEo       ! defining transition state maximum energy
```

```
!
!       Converting input frequencies to units of 0.1meV and then integers
!
     DO 20 J=1,Nfreq,1
     g(J)=w(J)*1.24
     v(J)=NINT(g(J))
20    CONTINUE
!
! STEP 4. Initializing p and N -vectors
!
     p(1)=1
     N(1)=1
     DO 50 J=2,480000,1
        p(J)=0
        N(J)=1
50    CONTINUE
!
! STEP 5. Calculating density of states
!
     DO 200 J=1,Nfreq,1
        DO 100 I=v(J),IMAX,1
          IF ((I-v(J)) .EQ. 0) THEN
            p(I)=p(I)+0
          ELSE
           p(I) = p(I) + p(I - v(J))
          END IF
100      CONTINUE
200   CONTINUE
!
!     PRINT 201, p(5000)
!201   FORMAT (1PG10.4E2)
!
! STEP 6. Calculating ion partition function and internal energy
!
     QMolc=1
     UMolc=0
     DO 250 I=1,Nfreq,1
        QMolc=QMolc*(1/(1-EXP(-w(I)/208.4)))     !calcing molc partion fcn
        UMolc=UMolc + (w(I)/(EXP(w(I)/208.4)-1)) ! calcing molc int nrg
250   CONTINUE
!
!     PRINT 260, QMolc,UMolc
!260   FORMAT (1PG11.4E3)
!
! STEP 7. Calculating sum of states
!
```

```
      READ (15,*) Nfreq
      READ (15,*) (w(I),I=1,Nfreq)
!
      DO 270 J=1,Nfreq,1
      g(J)=w(J)*1.24
      v(J)=NINT(g(J))
270   CONTINUE
!
      DO 400 J=1,Nfreq,1
        DO 300 I=v(J),ITMAX,1
          N(I) = N(I) + N(I-v(J))
300     CONTINUE
400   CONTINUE

!       print *, N(6000)
!
! STEP 8. Calculating ts partition function and internal energy at 300K
!
      QTs=1
      UTs=0
      DO 450 I=1,Nfreq,1
        QTs=QTs*(1/(1-EXP(-w(I)/208.4)))     !calcing molc partion fcn
        UTs=UTs + (w(I)/(EXP(w(I)/208.4)-1)) ! calcing molc int nrg
450   CONTINUE
!
!     PRINT 451, QTs,UTs
!451   FORMAT (1PG11.4E3)
!
! STEP 9. Calculating rate constant at specific intervals
!
      H=(3.3364E-11)*(1.24)
      DO 500 J=1,ITMAX,1
      RATE(J)=(N(J))/(H*p(J+IEo))
500   CONTINUE
!
!     print *, RATE(900)

! STEP 10. Calculating entropy of activation at 300K
!
      Kb=1.3807E-23
      Na=6.022E23
      HJ=6.6262E-34
      c=2.9979E10
      T=300
      Sts= ((Kb*LOG(Qts/QMolc)) + (HJ*c*(UTs-UMolc))/T)*Na
!       print *,Sts
```

```
!
!     write the data to a file
      open(unit=10 ,file='k.dat', status='replace')

      DO 950 J=0-IEo,IMAX-IEo,(IMAX)/900
      IF (J .LT. 0) goto 849
      goto 850
 849   RATE(J) = 0.0
 850   WRITE (10,900) RATE(J)
 900   FORMAT (E13.5E4)

 950   CONTINUE
      endfile 10
      close (unit=10)
!     write the data to a file
      open(unit=11 ,file='e.dat', status='replace')
      DO 980 J=0-IEo,IMAX-IEo,(IMAX)/900
     WRITE (11,960) (J+IEo)/10000.0
960   FORMAT (E13.5E2)
!
980   CONTINUE
     endfile 11
     close (unit=11)

!     write the data to a file
     open(unit=12 ,file='p.dat', status='replace')

     DO 1050 J=0-IEo,IMAX-IEo,(IMAX)/900
     WRITE (12,1000) p(J+IEo)
1000  FORMAT (E14.7E4)
1050  CONTINUE
     endfile 12
     close (unit=12)

!       Program main
!      Calculate breakdown curve. Convolute breakdown curve with time scale, P(E,T)
!

       do 1200,i=1,1000
         k1(i)=0.0
            kt(i)=0.0
            ecom(i)=0.0
         elab(i)=0.0
            freq(i)=0.0
!        k2(i)=0.0
```

```
1200      continue

!      Read in rate constant from RRKM calculation
       open (unit=13, file='k.dat',status='old')
!     +   form='formatted', access='sequential')
       i=1
1205    read(13,*,end=1210) k1(i)
       i=i+1
       goto 1205
1210    nptsk=i-1
!       print *,nptsk

!      Read in internal energy from RRKM calculation
       open (unit=14, file='e.dat', status='old',
     +   form='formatted', access='sequential')
       i=1
1215    read(14,*,end=1220) ek(i)
       i=i+1
       goto 1215
1220    nptse=i-1
!
!      open (unit=17, file='k2.dat',status='old')
!      i=1
!7           read(17,*,end=8) ek(i),k2(i)
!      i=i+1
!      goto 7
!8      nptsk=i-1
!
       do 1240, i=1,nptsk
       kt(i) = k1(i)
1240    continue

!      Read in MW of parent ion, calculate observation window
       open (unit=19, file='mw.dat', status='old')
       read (19,*,end=1260) mw
!10     print *,mw
1260    do 1270, i=1,100
       elab(i) = (i**2)/100.0
       ecom(i) = elab(i)*40.0/(mw+40.0)
       time(i) = 1/(SQRT(elab(i)*100000.0*2.0*(1000.0/mw)*100.0))
1270    continue
!      print *, elab(1000)
!       print *, time(1), time(1000)

!      Calculate kobs based on k and time
```

```
          do 1275 j=1,100,1
         do 1274 i=1,1000,1
         kobs(i,j) = 0.0
1274      continue
1275    continue

       do 1300, i=1,nptsk
       do 1290, j=1,100
       kobs(i,j)=1-exp(-kt(i)*time(j))
1290    continue
1300    continue
!       print *,kobs(1000,1000)

!       Calculate effective T for Q calc
       open (unit=16, file='slope.dat', status='old')
       k=1
1330    read (16,*,end=1340) slope(k)
!14     print *,slope
       k=k+1
          goto 1330
1340    nptssl=k-1

          do 1346 k=1,10,1
         do 1345 j=1,100,1
         Teff(k,j) = 0.0
1345     continue
1346    continue

       do 1350, k=1,nptssl,1
            do 1349 j=1,100,1
       Teff(k,j)=300+slope(k)*ecom(j)
1349    continue
1350    continue
!        print *,Teff(nptssl,1000)

!       read in vibrational frequencies for Q calc
       i=1
       open (unit=17, file='freq.dat', status='old')
!    +    form='unformatted', access='sequential')
1360    read(17,*,end=1370) freq(i)
       i=i+1
       goto 1360
1370    npfreq=i-1
!       print *,npfreq
!      print *,freq(252)
```

```
!       Calculate partition function, Q
        h=6.6262E-34
        kb=1.38E-23
        R=8.314
        S=2.99E+8

!       initialize the partition function array to avoid zeros
        k=1
1969    do 1970 j=1,100,1
         Qt(k,j) = 1.0
1970    continue
        k=k+1
           if (k.GT.10) goto 2000
           goto 1969
!    print *, Qt(1,1)

2000    k=1
2001    j=1
2002    do 2100 i=1,npfreq,1
2003    q(k,j)=1.0/(1.0-exp((-freq(i)*100.0*S*h)/(kb*Teff(k,j))))
        Qt(k,j)=Qt(k,j)*q(k,j)
2100    continue
        j=j+1
        if (j .GT. 100) goto 2120
        goto 2002
2120    k=k+1
           if (k .GT. nptssl) goto 2249
           goto 2001

2249        print *,Qt(1,100)

2250    open (unit=18, file= 'p.dat',status='unknown')
!    +   form='formatted', access='sequential')
        i=1
2290    read(18,*,end=2300) den(i)
        i=i+1
        goto 2290
2300    nptsp=i-1
!        print *,nptsp

!       Calculate Internal Energy Distribution
!        F=100.0
2400    k=1
2401    j=1
2402    do 2450 i=1,nptsp,1
2403    Eint(i,k,j)=(den(i)/Qt(k,j))*(DEXP(-(ek(i)*96480)/(R*Teff(k,j +   )))
```

```
2450    continue
        j=j+1
        if (j .GT. 100) goto 2460
        goto 2402
2460    k=k+1
          if (k .GT. nptssl) goto 2469
          goto 2401
2469    print *, Eint(200,1,1), Eint(800,1,1)
        print *, Eint(200,1,100), Eint(800,1,100)

!       Calculate total population at each Ecom and slope so we can normalize the distribution
2470    k=1
2480    do 2490 j=1,100,1
         TEint(k,j) = 0.0
2490    continue
        k=k+1
          if (k.GT.10) goto 2500
          goto 2480

2500    k=1
2501    j=1
2502    do 2600 i=1,nptsp,1
2580    TEint(k,j)=TEint(k,j)+Eint(i,k,j)
2600    continue
        j=j+1
        if (j .GT. 100) goto 2620
        goto 2502
2620    k=k+1
          if (k .GT. nptssl) goto 2700
          goto 2501

!        print *,TEint(1000)

!       Normalize the internal energy distributuion
2700    k=1
2701    j=1
2702    do 2800 i=1,nptsp,1
2780    NEint(i,k,j) = Eint(i,k,j)/TEint(k,j)
2800    continue
        j=j+1
        if (j .GT. 100) goto 2820
        goto 2702
2820    k=k+1
          if (k .GT. nptssl) goto 2900
          goto 2701
```

```
!2890       print *,NEint(900,1,1), NEint(900,1,900)

2900    close (unit=13)
        close (unit=14)
        close (unit=16)
        close (unit=17)
        close (unit=18)
        close (unit=19)

!       Calc net observation
        k=1
2988    j=1
2989    do 3000 i=1,nptsp,1
2990    Fracdiss(i,k,j)=kobs(i,j)*NEint(i,k,j)
3000    continue
        j=j+1
        if (j .GT. 100) goto 3020
        goto 2989
3020    k=k+1
          if (k .GT. nptssl) goto 3050
          goto 2988
!3049    print *, Fracdiss(900,1,1), Fracdiss(900,1,900)
!       Calculate total observable at each Ecom and slope

3050    k=1
3051    do 3052 j=1,100,1
         TFracdiss(k,j) = 0.0
3052    continue
        k=k+1
          if (k.GT.10) goto 3150
          goto 3051

3150    k=1
3200    j=1
3201    do 3203 i=1,nptsp,1
3202    TFracdiss(k,j) = TFracdiss(k,j)+Fracdiss(i,k,j)
3203    continue
        j=j+1
        if (j .GT. 100) goto 3220
        goto 3201
3220    k=k+1
          if (k .GT. nptssl) goto 3380
          goto 3200

!3379       print *,TFracdiss(1,1), TFracdiss(1,1000)
```

```
!       Calculate total observable as a fraction of total at each slope
3380    k=1
3389    do 3400, j=1,100
        Product(k,j) = TFracdiss(k,j)
        Parent(k,j) = 1-Product(k,j)
3400    continue
        k=k+1
           if (k .GT. nptssl) goto 3580
           goto 3389

!       write the fit results to a file
3580    Do 3585 j=1,100,1
         TDiff(j)=0.0
3585    continue

        open (unit=9, file='expt.dat', status='old')
        i=1
3590    read(9,*,end=3600) Expt(i)
        i=i+1
        goto 3590
3600    nptsexpt=i-1
        close (unit=9)

        k=1
3700    do 3800 j=1,100
3701       Diff(k,j)=ABS(Parent(k,j)-Expt(j))
          if (Diff(k,j) .GT. 2) Diff(k,j)=0.0
             TDiff(k)=TDiff(k)+Diff(k,j)
3800    continue

        write(20,*) o, slope(k), Eo/10000, Sts, TDiff(k)
           k=k+1
           if (k .GT. nptssl) goto 3900
           goto 3700

!3900    format(1PG20.8E2,1PG20.8E2,1PG20.8E2)

!       write the breakdown diagram to a file
3900    write(21,*) o
        k=1
3990    write(21,*) slope(k)
        do 4000 j=1,100
           write(21,*) ecom(j),Parent(k,j),Product(k,j)
!3990    format(1PG20.8E2,1PG20.8E2,1PG20.8E2)
4000    continue
        k=k+1
```

```
          if (k .GT. nptssl) goto 4100
          goto 3990
4100    o=o+1
          goto 10

!       write the population to a file
!        open(unit=11 ,file='population.dat', status='replace')
!        do 400 i=1,nptsk,1
!        j=1000
!        write(11,300) ek(i), Eint(i,j)
! 300    format(1PG20.8E3,1PG20.8E3)
! 400    continue
!        close (unit=11)
!
!       print 50 points to the screen to facilitate fitting
!        PRINT 600
!600     FORMAT ('  Ecom              Parent             Product  ')
!*
!        DO 800 J=1,1000,80
!        PRINT 700, ecom(j), Parent(j), Product(j)
!700     FORMAT (1PG10.4E2,1PG20.4E2,1PG20.4E2)
!*
!800     CONTINUE
!        close (unit=8)

!5000  continue

6000  close (unit=20)
      close (unit=15)
      close (unit=21)
      END
```

Appendix B: Vibrational frequencies obtained by MD-SA for the RRKM modeling of the ($PMMA_n$)(AAH^+) complexes

Table B.1: Calculated vibrational frequencies of the ($PMMA_7$)($GlyH^+$) complex.

Vibrational harmonic frequencies (cm^{-1})
14, 15, 21, 26, 32, 34, 36, 42, 47, 49, 51, 53, 60, 62, 68, 73, 75, 80, 83, 87, 92, 93, 96, 100, 103, 109, 112, 121, 129, 130, 134, 135, 141, 143, 152, 158, 162, 166, 169, 176, 178, 182, 186, 194, 196, 205, 213, 216, 227, 229, 233, 243, 246, 251, 256, 262, 271, 273, 276, 278, 282, 282, 283, 286, 287, 294, 300, 302, 313, 322, 327, 342, 343, 346, 355, 360, 363, 369, 378, 389, 394, 397, 400, 400, 404, 407, 409, 413, 421, 425, 432, 437, 441, 443, 446, 452, 458, 459, 469, 473, 480, 484, 498, 506, 509, 525, 537, 539, 563, 564, 570, 572, 575, 594, 608, 621, 631, 648, 674, 695, 700, 725, 727, 729, 733, 753, 758, 770, 787, 794, 800, 803, 808, 810, 819, 835, 901, 904, 907, 910, 912, 913, 916, 922, 958, 965, 977, 989, 992, 992, 994, 995, 1001, 1005, 1006, 1017, 1025, 1026, 1032, 1042, 1049, 1064, 1066, 1066, 1067, 1068, 1068, 1072, 1075, 1076, 1079, 1080, 1081, 1085, 1087, 1088, 1089, 1108, 1122, 1125, 1134, 1144, 1152, 1169, 1173, 1200, 1205, 1208, 1213, 1216, 1219, 1223, 1260, 1269, 1270, 1276, 1285, 1293, 1314, 1321, 1330, 1355, 1381, 1384, 1399, 1415, 1435, 1460, 1463, 1466, 1469, 1472, 1475, 1479, 1479, 1482, 1484, 1485, 1487, 1488, 1489, 1489, 1490, 1490, 1491, 1491, 1492, 1492, 1493, 1493, 1494, 1495, 1495, 1496, 1497, 1498, 1499, 1501, 1502, 1503, 1506, 1510, 1512, 1518, 1520, 1523, 1531, 1532, 1533, 1534, 1536, 1538, 1538, 1542, 1544, 1548, 1553, 1560, 1564, 1565, 1578, 1588, 1593, 1594, 1598, 1605, 1607, 1610, 1611, 1611, 1617, 1618, 1623, 1623, 1627, 1628, 1629, 1650, 1661, 1662, 1687, 1693, 1699, 1713, 2848, 2848, 2850, 2851, 2853, 2855, 2856, 2863, 2865, 2866, 2869, 2873, 2875, 2879, 2886, 2899, 2900, 2902, 2903, 2906, 2907, 2911, 2935, 2955, 2958, 2959, 2961, 2962, 2964, 2965, 2965, 2965, 2965, 2965, 2965, 2966, 2967, 2968, 2968, 2968, 2968, 2968, 2969, 2970, 2972, 2972, 2973, 2973, 2974, 2974, 2976, 2977, 2977, 2979, 2979, 2983, 2984, 2985, 2991, 2993, 2999, 3016, 3082, 3096

Table B.2: Calculated vibrational frequencies of the $(PMMA_7)(LeuH^+)$ complex.

Vibrational harmonic frequencies (cm^{-1})
13, 14, 17, 19, 24, 28, 32, 35, 43, 48, 51, 53, 54, 59, 61, 64, 71, 75, 76, 85, 89, 93, 95, 102, 105, 110, 116, 124, 128, 133, 137, 144, 150, 153, 158, 165, 170, 173, 177, 181, 188, 192, 193, 198, 207, 208, 221, 227, 231, 236, 237, 240, 242, 245, 251, 253, 259, 261, 264, 268, 270, 273, 276, 277, 278, 285, 285, 292, 294, 297, 300, 306, 315, 324, 329, 333, 340, 341, 342, 360, 368, 370, 376, 381, 387, 388, 397, 400, 404, 408, 412, 415, 417, 424, 432, 437, 449, 451, 459, 460, 464, 478, 482, 490, 497, 502, 502, 518, 528, 541, 553, 569, 571, 576, 579, 584, 585, 595, 600, 615, 623, 650, 663, 664, 735, 739, 742, 746, 747, 750, 752, 753, 759, 768, 775, 776, 782, 790, 795, 806, 810, 816, 822, 822, 832, 850, 902, 902, 906, 910, 911, 915, 917, 924, 944, 955, 961, 969, 976, 985, 990, 993, 995, 999, 1001, 1004, 1005, 1008, 1009, 1010, 1018, 1025, 1027, 1040, 1059, 1065, 1066, 1067, 1070, 1071, 1072, 1077, 1077, 1082, 1082, 1084, 1084, 1086, 1086, 1088, 1088, 1090, 1105, 1114, 1115, 1129, 1134, 1142, 1143, 1154, 1162, 1181, 1188, 1213, 1218, 1225, 1229, 1231, 1238, 1272, 1278, 1286, 1290, 1295, 1304, 1306, 1324, 1328, 1337, 1341, 1359, 1377, 1386, 1390, 1403, 1414, 1419, 1436, 1462, 1464, 1468, 1469, 1470, 1471, 1475, 1478, 1479, 1481, 1484, 1484, 1486, 1488, 1488, 1489, 1489, 1489, 1489, 1490, 1491, 1492, 1493, 1494, 1494, 1495, 1495, 1496, 1496, 1499, 1501, 1502, 1503, 1504, 1505, 1507, 1511, 1513, 1515, 1520, 1524, 1528, 1530, 1534, 1534, 1535, 1537, 1537, 1540, 1540, 1543, 1550, 1554, 1555, 1556, 1568, 1571, 1572, 1582, 1584, 1586, 1593, 1597, 1599, 1604, 1616, 1617, 1622, 1628, 1633, 1663, 1667, 1681, 1694, 1700, 1710, 1717, 1724, 1724, 1726, 1728, 1729, 1730, 1735, 2848, 2848, 2850, 2851, 2852, 2855, 2857, 2863, 2864, 2865, 2865, 2867, 2869, 2872, 2875, 2877, 2883, 2899, 2902, 2904, 2906, 2907, 2909, 2911, 2936, 2936, 2942, 2954, 2958, 2959, 2961, 2961, 2964, 2965, 2965, 2965, 2965, 2965, 2966, 2966, 2966, 2967, 2968, 2968, 2968, 2968, 2969, 2969, 2970, 2971, 2971, 2972, 2973, 2973, 2974, 2975, 2975, 2975, 2977, 2977, 2978, 2978, 2979, 2980, 2983, 2986, 2987, 2989, 2996, 3038, 3081, 3094

Table B.3: Calculated vibrational frequencies of the $(PMMA_7)(PheH^+)$ complex.

Vibrational harmonic frequencies (cm^{-1})
9, 12, 16, 19, 21, 25, 28, 31, 38, 44, 45, 47, 48, 49, 55, 58, 60, 65, 68, 70, 72, 76, 81, 83, 86, 95, 97, 99, 103, 107, 110, 116, 121, 126, 131, 132, 137, 140, 146, 151, 163, 165, 170, 172, 179, 181, 185, 193, 200, 209, 216, 225, 226, 230, 233, 243, 247, 249, 254, 257, 263, 265, 271, 275, 278, 280, 281, 282, 284, 287, 287, 294, 298, 309, 313, 323, 327, 340, 343, 348, 355, 359, 363, 367, 378, 387, 391, 397, 402, 402, 406, 409, 413, 414, 421, 421, 431, 433, 438, 445, 447, 451, 454, 456, 460, 464, 470, 476, 478, 484, 504, 507, 529, 533, 539, 549, 552, 562, 564, 566, 570, 575, 595, 601, 610, 611, 624, 646, 651, 676, 687, 695, 698, 725, 727, 727, 729, 749, 753, 770, 781, 788, 795, 799, 806, 809, 816, 820, 836, 886, 901, 902, 903, 909, 912, 914, 917, 923, 950, 957, 964, 979, 985, 990, 991, 994, 996, 998, 1000, 1005, 1006, 1008, 1017, 1025, 1026, 1029, 1030, 1033, 1049, 1063, 1064, 1066, 1066, 1068, 1070, 1071, 1073, 1075, 1083, 1085, 1086, 1087, 1088, 1090, 1090, 1102, 1108, 1110, 1123, 1126, 1129, 1135, 1145, 1152, 1168, 1175, 1192, 1197, 1205, 1212, 1216, 1217, 1219, 1220, 1254, 1269, 1269, 1280, 1287, 1288, 1296, 1296, 1317, 1322, 1331, 1332, 1357, 1357, 1381, 1384, 1402, 1416, 1440, 1460, 1462, 1468, 1469, 1472, 1475, 1479, 1480, 1485, 1486, 1487, 1487, 1488, 1489, 1489, 1490, 1491, 1491, 1491, 1493, 1493, 1494, 1494, 1495, 1495, 1496, 1497, 1498, 1499, 1500, 1502, 1502, 1504, 1507, 1509, 1511, 1511, 1519, 1521, 1526, 1532, 1533, 1533, 1535, 1536, 1537, 1540, 1541, 1545, 1548, 1556, 1559, 1561, 1565, 1571, 1577, 1578, 1591, 1593, 1598, 1599, 1607, 1612, 1613, 1617, 1620, 1623, 1624, 1625, 1628, 1629, 1630, 1649, 1660, 1661, 1666, 1691, 1697, 1701, 1731, 1752, 1783, 1803, 1848, 2848, 2848, 2849, 2851, 2852, 2854, 2857, 2863, 2865, 2867, 2867, 2870, 2875, 2878, 2881, 2902, 2904, 2905, 2908, 2911, 2916, 2919, 2935, 2942, 2958, 2961, 2963, 2964, 2965, 2965, 2965, 2965, 2965, 2966, 2966, 2966, 2967, 2968, 2968, 2968, 2968, 2969, 2969, 2970, 2970, 2970, 2971, 2971, 2971, 2972, 2972, 2973, 2973, 2973, 2973, 2975, 2976, 2977, 2977, 2978, 2979, 2980, 2983, 2984, 2985, 2990, 2990, 3038, 3080, 3081

Table B.4: Calculated vibrational frequencies of the ($PMMA_{10}$)($GlyH^+$) complex.

Vibrational harmonic frequencies (cm^{-1})
6, 7, 7, 14, 15, 18, 19, 21, 24, 27, 28, 30, 31, 33, 35, 38, 40, 41, 45, 47, 48, 49, 51, 52, 56, 60, 63, 64, 67, 69, 72, 73, 75, 76, 77, 80, 84, 89, 93, 95, 98, 101, 107, 113, 119, 122, 135, 137, 143, 154, 157, 158, 165, 167, 169, 176, 179, 183, 186, 192, 194, 196, 204, 212, 215, 218, 221, 225, 235, 237, 237, 240, 241, 245, 249, 251, 258, 260, 265, 268, 268, 268, 271, 274, 277, 281, 283, 288, 288, 300, 301, 306, 308, 310, 314, 319, 322, 326, 328, 331, 333, 335, 341, 342, 344, 350, 351, 354, 370, 375, 380, 386, 390, 392, 392, 394, 398, 399, 399, 402, 403, 408, 410, 415, 417, 417, 420, 427, 428, 432, 432, 434, 440, 447, 451, 455, 455, 456, 457, 465, 465, 475, 478, 481, 486, 491, 492, 495, 497, 500, 504, 522, 552, 554, 564, 572, 577, 582, 601, 606, 607, 608, 614, 617, 618, 625, 633, 638, 644, 670, 679, 691, 730, 730, 733, 738, 743, 749, 752, 754, 761, 764, 768, 804, 805, 825, 843, 845, 852, 867, 869, 871, 879, 902, 904, 908, 913, 916, 920, 927, 928, 928, 930, 935, 967, 968, 985, 985, 989, 991, 994, 996, 998, 1000, 1004, 1010, 1013, 1019, 1019, 1023, 1025, 1026, 1028, 1031, 1035, 1062, 1065, 1065, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1068, 1070, 1082, 1085, 1085, 1085, 1085, 1085, 1085, 1086, 1086, 1087, 1088, 1090, 1114, 1115, 1119, 1127, 1136, 1139, 1153, 1157, 1161, 1163, 1181, 1188, 1188, 1198, 1204, 1208, 1216, 1219, 1223, 1231, 1251, 1257, 1264, 1266, 1270, 1280, 1297, 1302, 1305, 1324, 1330, 1343, 1353, 1374, 1380, 1381, 1391, 1397, 1401, 1402, 1415, 1425, 1461, 1466, 1469, 1469, 1471, 1474, 1475, 1477, 1479, 1480, 1482, 1484, 1485, 1486, 1486, 1488, 1488, 1488, 1488, 1488, 1488, 1492, 1492, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1494, 1495, 1496, 1499, 1505, 1505, 1507, 1508, 1508, 1510, 1510, 1511, 1517, 1517, 1521, 1521, 1524, 1525, 1526, 1528, 1528, 1528, 1529, 1529, 1530, 1532, 1533, 1534, 1534, 1535, 1536, 1536, 1538, 1540, 1542, 1546, 1549, 1551, 1556, 1558, 1559, 1561, 1567, 1570, 1570, 1572, 1577, 1579, 1583, 1585, 1588, 1594, 1598, 1599, 1602, 1609, 1612, 1620, 1624, 1627, 1628, 1631, 1634, 1636, 1644, 1650, 1652, 1656, 1659, 1665, 1667, 1671, 1681, 1684, 1688, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2849, 2850, 2851, 2864, 2864, 2865, 2865, 2866, 2866, 2866, 2866, 2867, 2868, 2869, 2904, 2906, 2907, 2913, 2916, 2918, 2920, 2923, 2923, 2927, 2936, 2961, 2962, 2964, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2966, 2966, 2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2969, 2969, 2969, 2970, 2972, 2973, 2973, 2973, 2974, 2974, 2975, 2978, 2979, 2980, 2981, 2983, 2985, 2986, 2988, 2994, 2999, 3037, 3089, 3104

Table B.5: Calculated vibrational frequencies of the ($PMMA_{10}$)($LeuH^+$) complex.

Vibrational harmonic frequencies (cm^{-1})
5, 6, 6, 10, 14, 15, 15, 19, 22, 25, 27, 27, 29, 30, 33, 33, 36, 39, 39, 42, 44, 47, 49, 49, 51, 52, 53, 56, 59, 60, 64, 64, 66, 67, 71, 72, 73, 75, 77, 79, 81, 85, 90, 94, 99, 101, 106, 114, 121, 123, 134, 138, 148, 149, 159, 161, 163, 169, 170, 173, 177, 181, 184, 187, 189, 195, 202, 205, 220, 222, 224, 230, 232, 234, 235, 238, 240, 240, 249, 252, 253, 257, 260, 264, 267, 267, 268, 268, 269, 270, 273, 277, 277, 281, 282, 283, 292, 294, 299, 303, 305, 308, 311, 318, 320, 322, 326, 328, 331, 334, 335, 338, 343, 345, 346, 349, 356, 370, 370, 380, 384, 390, 392, 394, 395, 399, 400, 402, 404, 410, 414, 415, 418, 419, 420, 428, 428, 430, 435, 439, 442, 447, 453, 455, 456, 460, 462, 462, 465, 470, 475, 476, 479, 489, 490, 492, 497, 499, 502, 506, 508, 526, 530, 552, 552, 565, 569, 577, 582, 589, 600, 603, 606, 607, 612, 615, 621, 627, 632, 639, 652, 667, 677, 724, 729, 732, 739, 742, 748, 753, 754, 756, 757, 763, 789, 804, 826, 843, 845, 849, 857, 862, 863, 871, 877, 888, 904, 906, 909, 910, 921, 925, 927, 929, 930, 931, 961, 967, 969, 977, 985, 987, 990, 992, 995, 996, 1001, 1002, 1003, 1006, 1012, 1013, 1021, 1021, 1022, 1024, 1026, 1027, 1028, 1031, 1047, 1065, 1066, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1067, 1069, 1083, 1084, 1085, 1085, 1085, 1085, 1086, 1086, 1086, 1086, 1087, 1088, 1097, 1106, 1106, 1116, 1119, 1119, 1128, 1134, 1137, 1147, 1153, 1159, 1163, 1181, 1187, 1189, 1197, 1204, 1207, 1216, 1219, 1220, 1231, 1252, 1255, 1259, 1270, 1272, 1279, 1289, 1294, 1303, 1314, 1324, 1331, 1350, 1367, 1374, 1378, 1385, 1389, 1390, 1398, 1399, 1402, 1414, 1424, 1461, 1467, 1468, 1469, 1470, 1472, 1472, 1474, 1476, 1477, 1477, 1482, 1484, 1486, 1486, 1486, 1487, 1487, 1488, 1488, 1488, 1488, 1489, 1490, 1491, 1492, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1494, 1495, 1497, 1498, 1504, 1505, 1507, 1507, 1507, 1509, 1510, 1513, 1514, 1514, 1518, 1521, 1521, 1523, 1525, 1527, 1528, 1529, 1529, 1530, 1530, 1533, 1533, 1534, 1534, 1535, 1536, 1536, 1538, 1538, 1539, 1541, 1542, 1545, 1547, 1549, 1550, 1555, 1559, 1561, 1563, 1566, 1567, 1569, 1570, 1570, 1572, 1575, 1581, 1584, 1588, 1592, 1594, 1598, 1602, 1609, 1617, 1622, 1625, 1627, 1629, 1630, 1632, 1637, 1644, 1647, 1653, 1656, 1658, 1665, 1670, 1675, 1679, 1681, 1739, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2850, 2850, 2862, 2862, 2864, 2864, 2865, 2865, 2865, 2865, 2866, 2866, 2866, 2870, 2904, 2906, 2907, 2919, 2919, 2919, 2920, 2921, 2923, 2924, 2936, 2940, 2961, 2962, 2964, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2966, 2966, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2969, 2969, 2970, 2971, 2972, 2972, 2972, 2973, 2973, 2974, 2974, 2975, 2975, 2976, 2977, 2980, 2980, 2981, 2984, 2984, 2986, 2987, 2988, 2990, 3038, 3082, 3101

Table B.6: Calculated vibrational frequencies of the ($PMMA_{10}$)($PheH^+$) complex.

Vibrational harmonic frequencies (cm^{-1})
3, 6, 7, 10, 13, 15, 18, 19, 21, 23, 25, 29, 30, 32, 33, 35, 38, 39, 41, 43, 45, 47, 49, 51, 51, 54, 58, 59, 61, 63, 65, 67, 69, 71, 73, 74, 75, 77, 78, 79, 87, 89, 92, 100, 101, 101, 106, 116, 119, 122, 133, 138, 147, 157, 162, 162, 167, 169, 170, 175, 180, 184, 184, 188, 189, 193, 197, 209, 212, 215, 222, 233, 234, 235, 237, 238, 239, 241, 243, 249, 251, 253, 255, 257, 267, 268, 269, 270, 276, 277, 278, 280, 282, 283, 292, 296, 300, 303, 306, 309, 311, 315, 317, 321, 324, 326, 327, 330, 334, 335, 340, 345, 348, 351, 354, 370, 378, 378, 386, 390, 391, 395, 398, 398, 400, 403, 404, 405, 409, 413, 413, 416, 417, 420, 420, 422, 426, 428, 432, 438, 442, 448, 450, 452, 454, 455, 459, 462, 465, 471, 473, 475, 478, 485, 488, 490, 493, 498, 499, 506, 526, 534, 551, 552, 553, 564, 572, 576, 583, 599, 601, 606, 613, 615, 619, 624, 629, 632, 640, 652, 652, 668, 679, 686, 690, 724, 729, 731, 737, 740, 749, 752, 756, 757, 759, 763, 766, 805, 821, 826, 844, 846, 851, 855, 863, 865, 878, 896, 900, 904, 906, 909, 911, 922, 924, 926, 929, 930, 932, 967, 969, 977, 984, 986, 988, 990, 993, 996, 997, 1001, 1004, 1005, 1014, 1017, 1018, 1019, 1021, 1022, 1025, 1027, 1028, 1033, 1045, 1047, 1063, 1065, 1066, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1070, 1084, 1084, 1085, 1085, 1085, 1085, 1086, 1086, 1086, 1087, 1093, 1106, 1114, 1117, 1119, 1125, 1129, 1135, 1140, 1147, 1158, 1161, 1166, 1176, 1182, 1186, 1191, 1198, 1203, 1207, 1215, 1217, 1219, 1222, 1231, 1253, 1256, 1259, 1260, 1264, 1270, 1283, 1286, 1295, 1303, 1307, 1315, 1322, 1334, 1342, 1350, 1380, 1383, 1383, 1390, 1396, 1399, 1407, 1408, 1424, 1461, 1463, 1467, 1468, 1470, 1470, 1472, 1473, 1474, 1476, 1477, 1482, 1484, 1486, 1486, 1487, 1487, 1487, 1488, 1488, 1488, 1490, 1492, 1492, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1494, 1495, 1496, 1496, 1497, 1504, 1504, 1506, 1506, 1507, 1508, 1510, 1510, 1514, 1516, 1518, 1521, 1522, 1524, 1527, 1527, 1528, 1529, 1529, 1529, 1530, 1532, 1532, 1533, 1534, 1535, 1536, 1538, 1539, 1539, 1540, 1543, 1544, 1545, 1555, 1556, 1557, 1563, 1566, 1567, 1568, 1569, 1570, 1571, 1573, 1578, 1584, 1587, 1589, 1593, 1597, 1599, 1601, 1613, 1616, 1622, 1625, 1627, 1627, 1630, 1636, 1637, 1644, 1656, 1658, 1659, 1662, 1668, 1673, 1678, 1686, 1700, 1753, 1803, 1848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2849, 2850, 2851, 2862, 2864, 2865, 2865, 2866, 2866, 2866, 2867, 2868, 2868, 2872, 2905, 2906, 2907, 2919, 2919, 2920, 2924, 2925, 2925, 2936, 2944, 2961, 2962, 2964, 2964, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2966, 2966, 2966, 2967, 2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2969, 2969, 2970, 2970, 2971, 2972, 2972, 2972, 2973, 2973, 2973, 2973, 2974, 2975, 2975, 2977, 2978, 2978, 2979, 2981, 2981, 2983, 2984, 2985, 2988, 2989, 2990, 3002, 3038, 3079, 3083

Table B.7: Calculated vibrational frequencies of the ($PMMA_{13}$)($GlyH^+$) complex.

Vibrational harmonic frequencies (cm^{-1})
6, 7, 7, 14, 15, 16, 18, 19, 21, 23, 24, 27, 28, 29, 30, 31, 33, 34, 35, 38, 40, 41, 44, 45, 47, 48, 49, 51, 52, 56, 56, 60, 63, 64, 65, 67, 69, 72, 73, 74, 75, 76, 77, 78, 80, 84, 89, 89, 93, 95, 98, 101, 107, 113, 119, 122, 123, 135, 137, 143, 146, 154, 157, 158, 162, 165, 167, 169, 170, 176, 179, 183, 186, 187, 192, 194, 196, 198, 204, 212, 215, 218, 221, 225, 235, 235, 237, 237, 240, 241, 242, 245, 249, 251, 253, 258, 260, 265, 267, 268, 268, 268, 271, 271, 274, 277, 281, 281, 283, 288, 288, 292, 300, 301, 306, 307, 308, 310, 314, 316, 319, 322, 326, 328, 329, 331, 333, 335, 338, 341, 342, 344, 347, 350, 351, 354, 370, 370, 375, 380, 386, 390, 392, 392, 394, 394, 398, 399, 399, 400, 402, 403, 408, 410, 410, 415, 417, 417, 419, 420, 427, 428, 430, 432, 432, 434, 435, 440, 447, 451, 452, 455, 455, 456, 457, 459, 465, 465, 475, 477, 478, 481, 486, 490, 491, 492, 495, 496, 497, 500, 504, 522, 536, 552, 554, 564, 567, 572, 577, 582, 589, 601, 606, 607, 608, 614, 617, 618, 625, 632, 633, 638, 644, 653, 670, 679, 691, 725, 730, 730, 733, 738, 743, 749, 752, 754, 756, 761, 764, 768, 774, 804, 805, 825, 839, 843, 845, 852, 859, 867, 869, 871, 879, 888, 902, 904, 908, 910, 913, 916, 920, 923, 927, 928, 928, 929, 930, 935, 967, 968, 976, 985, 985, 989, 990, 991, 994, 996, 997, 998, 1000, 1004, 1006, 1010, 1013, 1019, 1019, 1020, 1023, 1025, 1026, 1028, 1028, 1031, 1035, 1047, 1062, 1065, 1065, 1066, 1066, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1067, 1068, 1070, 1082, 1083, 1085, 1085, 1085, 1085, 1085, 1085, 1086, 1086, 1086, 1087, 1088, 1090, 1106, 1114, 1115, 1119, 1125, 1127, 1136, 1139, 1146, 1153, 1157, 1161, 1163, 1167, 1181, 1188, 1188, 1191, 1198, 1204, 1208, 1211, 1216, 1219, 1223, 1226, 1231, 1251, 1257, 1264, 1265, 1266, 1270, 1280, 1283, 1297, 1302, 1305, 1316, 1324, 1330, 1343, 1353, 1374, 1380, 1381, 1391, 1394, 1397, 1401, 1402, 1408, 1415, 1425, 1461, 1463, 1466, 1469, 1469, 1471, 1471, 1474, 1475, 1477, 1477, 1479, 1480, 1482, 1483, 1484, 1485, 1486, 1486, 1486, 1488, 1488, 1488, 1488, 1488, 1488, 1488, 1490, 1492, 1492, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1494, 1495, 1496, 1497, 1499, 1505, 1505, 1506, 1507, 1508, 1508, 1509, 1510, 1510, 1511, 1514, 1517, 1517, 1521, 1521, 1523, 1524, 1525, 1526, 1527, 1528, 1528, 1528, 1528, 1529, 1529, 1530, 1531, 1532, 1533, 1534, 1534, 1535, 1535, 1536, 1536, 1536, 1538, 1540, 1542, 1543, 1546, 1549, 1551, 1554, 1556, 1558, 1559, 1561, 1565, 1567, 1570, 1570, 1571, 1572, 1577, 1579, 1583, 1583, 1585, 1588, 1591, 1594, 1598, 1599, 1602, 1609, 1609, 1612, 1620, 1620, 1624, 1627, 1628, 1630, 1631, 1634, 1636, 1636, 1644, 1650, 1652, 1656, 1659, 1665, 1667, 1671, 1675, 1681, 1684, 1688, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2849, 2850, 2851, 2863, 2864, 2864, 2865, 2865, 2865, 2866, 2866, 2866, 2866, 2867, 2868, 2869, 2871, 2904, 2906, 2907, 2907, 2913, 2916, 2918, 2919, 2920, 2923, 2923, 2924, 2927, 2936, 2961, 2962, 2964, 2964, 2965, 2966, 2966, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2969, 2969, 2969, 2969, 2970, 2970, 2972, 2973, 2973, 2973, 2973, 2974, 2974, 2974, 2975, 2978, 2979, 2979, 2980, 2981, 2983, 2985, 2985, 2986, 2988, 2988, 2994, 2999, 3037, 3089, 3104

Table B.8: Calculated vibrational frequencies of the ($PMMA_{13}$)($LeuH^+$) complex.

Vibrational harmonic frequencies (cm^{-1})
5, 6, 6, 10, 14, 15, 15, 19, 22, 23, 25, 27, 27, 29, 30, 33, 33, 34, 36, 39, 39, 42, 44, 47, 49, 49, 50, 51, 52, 53, 55, 56, 59, 60, 62, 64, 64, 66, 67, 69, 71, 72, 73, 75, 77, 79, 81, 82, 85, 90, 94, 99, 99, 101, 106, 114, 117, 121, 123, 134, 136, 138, 148, 149, 154, 159, 161, 163, 165, 169, 170, 173, 177, 181, 184, 187, 189, 193, 195, 202, 205, 213, 220, 222, 224, 226, 230, 232, 234, 235, 237, 238, 240, 240, 243, 249, 252, 253, 255, 257, 260, 264, 267, 267, 268, 268, 269, 270, 273, 277, 277, 280, 281, 282, 283, 289, 292, 294, 299, 302, 303, 305, 308, 311, 314, 318, 320, 322, 324, 326, 328, 331, 332, 334, 335, 338, 340, 343, 345, 346, 349, 351, 356, 370, 370, 378, 380, 384, 390, 391, 392, 394, 395, 397, 399, 400, 402, 404, 408, 410, 414, 415, 416, 418, 419, 420, 426, 428, 428, 430, 432, 435, 439, 442, 447, 451, 453, 455, 456, 458, 460, 462, 462, 465, 470, 475, 476, 478, 479, 489, 490, 492, 497, 497, 499, 502, 506, 506, 508, 526, 528, 530, 552, 552, 560, 565, 569, 577, 582, 588, 589, 600, 603, 606, 607, 612, 615, 619, 621, 627, 632, 637, 639, 652, 667, 677, 690, 724, 729, 732, 735, 739, 742, 748, 753, 754, 756, 757, 760, 763, 789, 804, 826, 840, 843, 845, 849, 852, 857, 862, 863, 866, 871, 877, 888, 903, 904, 906, 909, 910, 916, 921, 925, 927, 928, 929, 930, 931, 932, 961, 967, 969, 973, 977, 985, 987, 990, 991, 992, 995, 996, 999, 1001, 1002, 1003, 1004, 1006, 1012, 1013, 1019, 1021, 1021, 1022, 1024, 1024, 1026, 1027, 1028, 1029, 1031, 1047, 1065, 1065, 1066, 1066, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1067, 1067, 1069, 1070, 1083, 1084, 1085, 1085, 1085, 1085, 1086, 1086, 1086, 1086, 1087, 1088, 1091, 1097, 1106, 1106, 1113, 1116, 1119, 1119, 1123, 1128, 1134, 1137, 1141, 1147, 1153, 1159, 1163, 1166, 1181, 1187, 1189, 1191, 1197, 1204, 1207, 1211, 1216, 1219, 1220, 1223, 1231, 1252, 1255, 1259, 1264, 1270, 1272, 1279, 1282, 1289, 1294, 1303, 1306, 1314, 1324, 1331, 1342, 1350, 1367, 1374, 1378, 1380, 1385, 1389, 1390, 1393, 1398, 1399, 1402, 1406, 1414, 1424, 1461, 1463, 1467, 1468, 1469, 1470, 1471, 1472, 1472, 1474, 1474, 1476, 1477, 1477, 1479, 1482, 1484, 1486, 1486, 1486, 1487, 1487, 1488, 1488, 1488, 1488, 1488, 1488, 1489, 1490, 1491, 1492, 1492, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1494, 1495, 1495, 1497, 1498, 1501, 1504, 1505, 1507, 1507, 1507, 1507, 1509, 1510, 1510, 1513, 1514, 1514, 1515, 1518, 1521, 1521, 1523, 1523, 1525, 1527, 1528, 1528, 1529, 1529, 1530, 1530, 1530, 1533, 1533, 1533, 1534, 1534, 1535, 1536, 1536, 1538, 1538, 1539, 1539, 1541, 1542, 1545, 1546, 1547, 1549, 1550, 1554, 1555, 1559, 1561, 1563, 1563, 1566, 1567, 1567, 1569, 1570, 1570, 1572, 1572, 1575, 1581, 1584, 1587, 1588, 1592, 1594, 1595, 1598, 1602, 1609, 1613, 1617, 1622, 1625, 1627, 1628, 1629, 1630, 1632, 1636, 1637, 1644, 1647, 1650, 1653, 1656, 1658, 1660, 1665, 1670, 1675, 1679, 1681, 1739, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2849, 2850, 2850, 2862, 2862, 2864, 2864, 2864, 2865, 2865, 2865, 2865, 2865, 2866, 2866, 2866, 2866, 2868, 2870, 2904, 2906, 2907, 2913, 2919, 2919, 2919, 2919, 2920, 2921, 2923, 2923, 2924, 2936, 2940, 2943, 2961, 2962, 2964, 2965, 2966, 2966, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2968, 2969, 2969, 2969, 2970, 2971, 2971, 2972, 2972, 2972, 2973, 2973, 2973, 2974, 2974, 2974, 2975, 2975, 2976, 2976, 2977, 2980, 2980, 2981,2982, 2984, 2984, 2986, 2987, 2988, 2990, 3038, 3082, 3101

Table B.9: Calculated vibrational frequencies of the ($PMMA_{13}$)($PheH^+$) complex.

Vibrational harmonic frequencies (cm^{-1})
3, 6, 7, 10, 13, 15, 18, 19, 21, 22, 23, 25, 29, 29, 30, 32, 33, 34, 35, 38, 39, 41, 42, 43, 45, 47, 48, 49,
51, 51, 53, 54, 58, 59, 59, 61, 63, 65, 67, 69, 71, 73, 74, 75, 77, 78, 79, 83, 87, 89, 92, 95, 100, 101,
101, 106, 114, 116, 119, 122, 124, 133, 138, 147, 152, 157, 162, 162, 166, 167, 169, 170, 175, 177,
180, 184, 184, 188, 189, 193, 197, 201, 209, 212, 215, 220, 222, 233, 234, 235, 237, 238, 239, 241,
242, 243, 249, 251, 252, 253, 255, 257, 263, 267, 268, 269, 270, 272, 276, 277, 278, 279, 280, 282,
283, 288, 292, 296, 300, 303, 306, 309, 311, 315, 316, 317, 321, 324, 325, 326, 327, 330, 333, 334,
335, 340, 342, 345, 348, 351, 354, 370, 370, 378, 378, 381, 386, 390, 391, 394, 395, 398, 398, 400,
403, 404, 405, 409, 413, 413, 416, 417, 418, 420, 420, 422, 423, 426, 428, 432, 433, 438, 442, 448,
450, 450, 452, 454, 455, 458, 459, 462, 465, 469, 471, 473, 475, 477, 478, 485, 488, 490, 492, 493,
498, 499, 505, 506, 526, 534, 542, 551, 552, 553, 557, 564, 572, 576, 583, 589, 599, 601, 606, 607,
613, 615, 619, 620, 624, 629, 632, 637, 640, 652, 652, 668, 679, 686, 690, 724, 729, 731, 737, 740,
744, 749, 752, 756, 756, 757, 759, 763, 766, 787, 805, 821, 826, 834, 844, 846, 851, 854, 855, 863,
865, 871, 878, 896, 900, 904, 905, 906, 909, 911, 915, 922, 924, 926, 929, 929, 930, 932, 939, 967,
969, 977, 984, 986, 986, 988, 990, 992, 993, 996, 997, 999, 1001, 1004, 1005, 1012, 1014, 1017, 1018,
1019, 1020, 1021, 1022, 1025, 1026, 1027, 1028, 1033, 1044, 1045, 1047, 1063, 1064, 1065, 1066,
1066, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1067, 1070, 1074, 1084, 1084, 1085, 1085, 1085,
1085, 1085, 1086, 1086, 1086, 1086, 1087, 1092, 1093, 1106, 1114, 1116, 1117, 1119, 1125, 1127,
1129, 1135, 1140, 1147, 1153, 1158, 1161, 1166, 1176, 1182, 1186, 1191, 1191, 1198, 1203, 1207,
1211, 1215, 1217, 1219, 1222, 1225, 1231, 1253, 1256, 1259, 1260, 1264, 1270, 1280, 1283, 1286,
1295, 1297, 1303, 1307, 1315, 1322, 1325, 1334, 1342, 1350, 1374, 1380, 1383, 1383, 1387, 1390,
1396, 1399, 1405, 1407, 1408, 1424, 1461, 1462, 1463, 1467, 1468, 1468, 1470, 1470, 1472, 1473,
1474, 1476, 1477, 1479, 1482, 1484, 1486, 1486, 1486, 1487, 1487, 1487, 1488, 1488, 1488, 1488,
1488, 1490, 1492, 1492, 1492, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1494, 1494,
1495, 1496, 1496, 1496, 1497, 1504, 1504, 1506, 1506, 1507, 1507, 1508, 1510, 1510, 1513, 1514,
1516, 1518, 1520, 1521, 1522, 1524, 1524, 1527, 1527, 1528, 1529, 1529, 1529, 1529, 1530, 1530,
1532, 1532, 1533, 1534, 1534, 1535, 1536, 1537, 1538, 1539, 1539, 1540, 1542, 1543, 1544, 1545,
1547, 1555, 1556, 1557, 1563, 1563, 1566, 1567, 1568, 1568, 1569, 1570, 1571, 1572, 1573, 1578,
1584, 1587, 1587, 1589, 1593, 1594, 1597, 1599, 1601, 1611, 1613, 1616, 1622, 1625, 1627, 1627,
1627, 1630, 1634, 1636, 1637, 1644, 1648, 1656, 1658, 1659, 1661, 1662, 1668, 1673, 1678, 1681,
1686, 1700, 1753, 1783, 1803, 1848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2849,
2850, 2850, 2851, 2862, 2864, 2864, 2865, 2865, 2866, 2866, 2866, 2866, 2867, 2867, 2868, 2868,
2872, 2904, 2905, 2906, 2907, 2918, 2919, 2919, 2920, 2921, 2924, 2925, 2925, 2927, 2936, 2944,
2961, 2962, 2964, 2964, 2964, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965,
2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2966, 2966, 2966, 2966, 2967, 2967,
2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2968, 2969, 2969, 2970, 2970, 2970, 2971, 2972,
2972, 2972, 2973, 2973, 2973, 2973, 2974, 2974, 2975, 2975, 2977, 2977, 2978, 2978, 2979, 2980,
2981, 2981, 2983, 2984, 2985, 2988, 2989, 2990, 3002, 3038, 3079, 3083

Appendix C: Vibrational frequencies and rotational constants calculated for the determination of ΔS_{vib} and ΔS_{rot}

Table C.1: Calculated vibrational frequencies of glycine, leucine, and phenylalanine at the AM1 level of theory for the determination of ΔS_{vib}.

Amino acid	Vibrational harmonic frequencies (cm^{-1})
Gly	44, 230, 265, 478, 502, 579, 626, 936, 969, 1036, 1137, 1296, 1338, 1340, 1412, 1464, 1546, 1729, 2077, 2978, 3047, 3416, 3423, 3464
$GlyH^+$	28, 110, 282, 473, 521, 554, 602, 931, 983, 1093, 1177, 1237, 1259, 1332, 1375, 1408, 1581, 1628, 1642, 1644, 2084, 2940, 3008, 3166, 3204, 3324, 3334
Leu	33, 50, 88, 147, 168, 177, 221, 271, 287, 349, 408, 428, 498, 550, 564, 658, 715, 897, 967, 981, 992, 1020, 1034, 1044, 1131, 1157, 1211, 1219, 1243, 1271, 1273, 1328, 1374, 1391, 1392, 1393, 1397, 1402, 1408, 1417, 1421, 1432, 1446, 1460, 1540, 1722, 2075, 2983, 3004, 3024, 3062, 3063, 3065, 3066, 3091, 3156, 3157, 3420, 3423, 3465
$LeuH^+$	34, 57, 81, 117, 146, 161, 190, 255, 324, 356, 425, 452, 475, 519, 554, 605, 698, 906, 969, 987, 997, 1021, 1036, 1082, 1094, 1165, 1191, 1222, 1226, 1245, 1252, 1292, 1327, 1356, 1380, 1384, 1387, 1388, 1399, 1407, 1412, 1414, 1431, 1452, 1576, 1630, 1634, 1646, 2080, 2937, 2989, 3010, 3051, 3058, 3063, 3066, 3073, 3151, 3157, 3191, 3209, 3330, 3340
Phe	26, 38, 81, 109, 206, 211, 292, 316, 340, 373, 396, 489, 497, 525, 616, 624, 658, 682, 728, 803, 891, 912, 937, 961, 991, 996, 1009, 1053, 1093, 1122, 1152, 1175, 1189, 1197, 1228, 1236, 1270, 1315, 1343, 1377, 1379, 1394, 1424, 1439, 1464, 1534, 1570, 1634, 1723, 1765, 1779, 2066, 2965, 3010, 3083, 3177, 3179, 3185, 3189, 3199, 3409, 3431, 3457
$PheH^+$	38, 44, 61, 88, 140, 186, 265, 306, 358, 374, 421, 483, 519, 553, 598, 606, 657, 674, 708, 808, 890, 931, 944, 959, 992, 994, 1018, 1066, 1096, 1107, 1149, 1169, 1184, 1195, 1199, 1227, 1240, 1299, 1314, 1317, 1366, 1385, 1391, 1410, 1452, 1567, 1577, 1629, 1631, 1645, 1651, 1762, 1769, 2080, 2955, 2999, 3074, 3141, 3167, 3171, 3175, 3179, 3187, 3194, 3322, 3342

Table C.2: Calculated vibrational frequencies of $PMMA_4$ at the AM1 level of theory for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
14, 24, 26, 31, 36, 44, 47, 60, 65, 71, 82, 92, 101, 105, 111, 122, 128, 131, 137, 149, 155, 164, 172, 181, 186, 199, 206, 219, 227, 243, 248, 261, 279, 293, 307, 316, 329, 348, 356, 366, 370, 372, 379, 402, 411, 420, 450, 507, 533, 539, 571, 598, 635, 645, 662, 670, 733, 736, 760, 790, 904, 911, 921, 930, 940, 961, 977, 994, 1002, 1012, 1016, 1038, 1043, 1052, 1055, 1082, 1106, 1158, 1174, 1179, 1179, 1180, 1180, 1193, 1203, 1214, 1218, 1224, 1227, 1242, 1245, 1269, 1292, 1296, 1299, 1302, 1303, 1330, 1352, 1352, 1352, 1353, 1354, 1356, 1359, 1367, 1372, 1377, 1377, 1378, 1378, 1381, 1384, 1385, 1386, 1387, 1389, 1389, 1391, 1393, 1400, 1401, 1405, 1406, 1411, 1415, 1419, 1421, 1426, 1429, 1432, 1434, 1438, 1438, 1440, 1446, 1450, 1465, 1518, 1518, 1520, 1524, 2066, 2068, 2071, 2072, 2996, 3001, 3002, 3015, 3050, 3053, 3055, 3060, 3062, 3064, 3065, 3065, 3065, 3066, 3068, 3069, 3069, 3074, 3075, 3075, 3075, 3090, 3091, 3091, 3091, 3149, 3150, 3150, 3153, 3154, 3154, 3155, 3156, 3157

Table C.3: Calculated vibrational frequencies of $PMMA_4H^+$ at the AM1 level of theory for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
20, 32, 39, 46, 69, 78, 91, 98, 101, 108, 120, 126, 128, 131, 136, 144, 147, 150, 166, 169, 173, 176, 187, 194, 202, 202, 215, 222, 232, 245, 264, 268, 280, 290, 305, 309, 315, 355, 368, 370, 382, 391, 398, 406, 416, 422, 461, 495, 535, 555, 570, 606, 620, 624, 641, 663, 670, 690, 742, 756, 782, 885, 895, 923, 924, 945, 960, 970, 979, 996, 1016, 1019, 1038, 1041, 1045, 1063, 1071, 1078, 1145, 1164, 1168, 1175, 1176, 1182, 1185, 1199, 1210, 1216, 1219, 1223, 1226, 1242, 1251, 1272, 1273, 1285, 1292, 1296, 1300, 1338, 1338, 1344, 1346, 1348, 1350, 1350, 1355, 1358, 1359, 1368, 1371, 1372, 1373, 1377, 1378, 1379, 1381, 1381, 1384, 1387, 1390, 1394, 1395, 1395, 1402, 1405, 1408, 1415, 1421, 1427, 1429, 1433, 1437, 1439, 1444, 1445, 1446, 1450, 1455, 1467, 1537, 1552, 1589, 1707, 1759, 1990, 2026, 2058, 2975, 2992, 2995, 3008, 3051, 3051, 3053, 3054, 3057, 3060, 3063, 3063, 3066, 3066, 3068, 3073, 3073, 3075, 3077, 3078, 3079, 3082, 3087, 3089, 3090, 3139, 3149, 3150, 3152, 3152, 3153, 3154, 3157, 3158, 3219

Table C.4: Calculated vibrational frequencies of $(PMMA_4)(GlyH^+)$ at the AM1 level of theory for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
15, 25, 31, 33, 37, 43, 46, 58, 62, 67, 72, 78, 84, 97, 101, 105, 111, 117, 120, 130, 137, 139, 145, 147, 159, 159, 175, 180, 186, 189, 195, 200, 211, 217, 220, 225, 226, 244, 262, 270, 292, 301, 314, 324, 332, 339, 362, 373, 373, 385, 391, 397, 416, 423, 448, 482, 490, 509, 517, 525, 554, 555, 561, 591, 624, 633, 649, 666, 694, 736, 740, 759, 784, 888, 912, 924, 933, 948, 952, 958, 978, 986, 997, 1002, 1019, 1023, 1045, 1048, 1056, 1065, 1074, 1103, 1126, 1154, 1169, 1174, 1175, 1178, 1179, 1190, 1193, 1204, 1210, 1212, 1219, 1225, 1242, 1254, 1271, 1274, 1280, 1283, 1286, 1287, 1293, 1301, 1323, 1342, 1345, 1349, 1349, 1351, 1354, 1355, 1363, 1369, 1371, 1373, 1373, 1376, 1376, 1377, 1380, 1382, 1385, 1385, 1386, 1390, 1391, 1394, 1396, 1398, 1400, 1403, 1414, 1415, 1421, 1422, 1424, 1425, 1428, 1433, 1436, 1438, 1438, 1441, 1454, 1456, 1464, 1470, 1543, 1546, 1555, 1561, 1579, 1643, 1645, 1685, 2022, 2036, 2042, 2055, 2079, 2934, 2979, 2990, 2994, 3019, 3020, 3045, 3050, 3050, 3054, 3061, 3062, 3064, 3064, 3065, 3065, 3066, 3069, 3069, 3069, 3071, 3072, 3072, 3078, 3087, 3087, 3087, 3088, 3148, 3148, 3149, 3149, 3151, 3152, 3152, 3155, 3157, 3160, 3302, 3374

Table C.5: Calculated vibrational frequencies of $(PMMA_4H^+)(Gly)$ at the AM1 level of theory for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
15, 18, 26, 27, 30, 36, 58, 63, 66, 71, 74, 79, 93, 96, 100, 106, 107, 110, 121, 131, 133, 136, 137, 141, 147, 156, 163, 168, 176, 191, 196, 204, 207, 211, 226, 230, 235, 239, 257, 268, 282, 288, 304, 307, 324, 330, 358, 365, 374, 378, 389, 400, 411, 426, 432, 444, 469, 485, 513, 532, 539, 589, 601, 620, 627, 648, 662, 670, 681, 697, 728, 734, 746, 770, 877, 896, 915, 918, 928, 934, 941, 951, 974, 982, 997, 1008, 1013, 1014, 1036, 1045, 1051, 1055, 1081, 1102, 1132, 1144, 1171, 1173, 1174, 1175, 1176, 1179, 1188, 1198, 1209, 1218, 1223, 1234, 1241, 1243, 1261, 1274, 1280, 1286, 1292, 1293, 1304, 1326, 1331, 1333, 1338, 1341, 1348, 1348, 1350, 1355, 1359, 1361, 1369, 1370, 1371, 1372, 1375, 1376, 1380, 1382, 1383, 1386, 1386, 1387, 1391, 1394, 1396, 1398, 1402, 1407, 1415, 1416, 1419, 1419, 1422, 1431, 1432, 1438, 1440, 1441, 1444, 1448, 1451, 1456, 1461, 1479, 1540, 1559, 1564, 1592, 1702, 1727, 1757, 2023, 2024, 2033, 2059, 2969, 2970, 2977, 3000, 3017, 3041, 3042, 3049, 3050, 3054, 3056, 3062, 3062, 3063, 3066, 3067, 3068, 3070, 3072, 3074, 3078, 3079, 3080, 3084, 3087, 3087, 3089, 3139, 3151, 3152, 3152, 3153, 3154, 3155, 3155, 3159, 3226, 3356, 3419, 3465

Table C.6: Calculated vibrational frequencies of $(PMMA_4)(LeuH^+)$ at the AM1 level of theory for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
15, 17, 23, 26, 30, 35, 38, 42, 48, 54, 56, 69, 78, 86, 89, 90, 97, 104, 111, 116, 126, 128, 132, 137, 138, 148, 155, 161, 165, 167, 173, 175, 178, 183, 184, 187, 190, 198, 209, 215, 218, 226, 229, 241, 255, 263, 275, 286, 297, 319, 322, 334, 349, 363, 370, 376, 381, 395, 405, 412, 423, 425, 426, 448, 477, 502, 513, 527, 544, 548, 557, 561, 575, 622, 641, 656, 666, 700, 707, 735, 739, 761, 798, 895, 905, 906, 923, 931, 944, 959, 975, 979, 982, 995, 995, 1012, 1014, 1016, 1025, 1038, 1043, 1043, 1047, 1060, 1081, 1101, 1120, 1150, 1168, 1175, 1176, 1178, 1179, 1181, 1185, 1193, 1200, 1204, 1209, 1214, 1215, 1222, 1227, 1237, 1242, 1251, 1260, 1262, 1277, 1284, 1284, 1293, 1294, 1296, 1326, 1339, 1346, 1349, 1349, 1350, 1351, 1352, 1354, 1362, 1369, 1372, 1374, 1377, 1377, 1379, 1380, 1381, 1383, 1384, 1384, 1386, 1388, 1388, 1389, 1390, 1393, 1393, 1394, 1395, 1397, 1398, 1400, 1403, 1404, 1413, 1414, 1414, 1419, 1422, 1425, 1426, 1431, 1432, 1434, 1436, 1437, 1440, 1442, 1446, 1448, 1451, 1457, 1481, 1540, 1545, 1555, 1564, 1572, 1631, 1649, 1683, 2020, 2034, 2039, 2049, 2079, 2938, 2988, 2990, 2995, 2996, 3009, 3017, 3050, 3050, 3055, 3058, 3058, 3059, 3060, 3061, 3064, 3064, 3065, 3065, 3065, 3066, 3066, 3066, 3069, 3070, 3070, 3071, 3075, 3079, 3085, 3087, 3087, 3088, 3112, 3145, 3148, 3148, 3149, 3150, 3153, 3155, 3156, 3157, 3157, 3161, 3163, 3296, 3374

Table C.7: Calculated vibrational frequencies of $(PMMA_4H^+)(Leu)$ at the AM1 level of theory for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
8, 16, 22, 27, 30, 36, 42, 42, 45, 51, 52, 65, 68, 71, 86, 88, 96, 99, 104, 111, 116, 125, 127, 136, 142, 145, 151, 158, 159, 162, 167, 171, 180, 184, 187, 190, 200, 210, 224, 234, 245, 260, 267, 268, 270, 273, 282, 310, 316, 330, 337, 346, 352, 353, 365, 367, 373, 384, 387, 392, 408, 413, 419, 446, 457, 465, 489, 509, 531, 566, 581, 605, 625, 644, 648, 656, 665, 712, 723, 740, 744, 752, 779, 831, 891, 917, 924, 927, 930, 936, 954, 958, 971, 984, 989, 989, 1011, 1015, 1017, 1023, 1035, 1038, 1041, 1042, 1049, 1062, 1074, 1077, 1113, 1137, 1143, 1160, 1175, 1176, 1177, 1177, 1185, 1193, 1201, 1209, 1215, 1215, 1219, 1226, 1236, 1237, 1242, 1253, 1254, 1276, 1276, 1285, 1289, 1290, 1302, 1309, 1324, 1333, 1337, 1338, 1343, 1349, 1349, 1351, 1351, 1354, 1357, 1359, 1368, 1373, 1374, 1374, 1379, 1379, 1382, 1386, 1389, 1390, 1391, 1391, 1392, 1393, 1394, 1396, 1399, 1399, 1400, 1403, 1404, 1410, 1411, 1411, 1415, 1419, 1421, 1422, 1427, 1429, 1432, 1432, 1434, 1436, 1438, 1441, 1441, 1444, 1447, 1451, 1454, 1459, 1529, 1539, 1541, 1555, 1716, 1721, 1768, 2031, 2042, 2053, 2063, 2937, 2956, 2967, 2998, 3002, 3009, 3027, 3039, 3042, 3054, 3054, 3055, 3057, 3059, 3059, 3063, 3063, 3064, 3064, 3065, 3066, 3067, 3068, 3076, 3076, 3076, 3077, 3077, 3078, 3086, 3088, 3091, 3095, 3140, 3144, 3147, 3149, 3151, 3151, 3152, 3153, 3155, 3157, 3158, 3158, 3382, 3415, 3462

Table C.8: Calculated vibrational frequencies of ($PMMA_4$)($PheH^+$) at the AM1 level of theory for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
16, 21, 26, 29, 33, 37, 40, 44, 49, 53, 59, 61, 74, 76, 79, 83, 88, 98, 101, 108, 110, 113, 123, 127, 137, 143, 146, 151, 158, 165, 176, 180, 183, 190, 191, 202, 206, 212, 224, 237, 240, 245, 256, 266, 273, 279, 284, 298, 315, 321, 328, 354, 359, 362, 371, 375, 381, 386, 394, 402, 410, 424, 452, 474, 481, 503, 508, 519, 548, 549, 583, 593, 598, 608, 631, 645, 657, 658, 659, 695, 712, 734, 739, 766, 794, 802, 891, 902, 908, 917, 926, 929, 932, 940, 965, 969, 974, 988, 995, 996, 1002, 1014, 1015, 1018, 1039, 1041, 1050, 1056, 1083, 1085, 1106, 1109, 1120, 1129, 1166, 1166, 1175, 1175, 1176, 1177, 1179, 1183, 1189, 1197, 1206, 1208, 1214, 1215, 1224, 1231, 1236, 1238, 1249, 1267, 1269, 1281, 1284, 1286, 1289, 1296, 1305, 1320, 1322, 1349, 1349, 1350, 1350, 1351, 1353, 1354, 1358, 1361, 1372, 1372, 1373, 1375, 1376, 1378, 1379, 1381, 1382, 1383, 1386, 1389, 1389, 1390, 1392, 1395, 1397, 1399, 1400, 1404, 1406, 1414, 1420, 1423, 1426, 1429, 1432, 1434, 1437, 1439, 1442, 1444, 1451, 1455, 1456, 1458, 1467, 1532, 1544, 1553, 1565, 1567, 1570, 1625, 1633, 1644, 1657, 1763, 1773, 2024, 2031, 2040, 2060, 2082, 2928, 2995, 2996, 2997, 3003, 3023, 3049, 3049, 3049, 3059, 3060, 3060, 3061, 3062, 3063, 3066, 3066, 3069, 3074, 3077, 3077, 3079, 3080, 3087, 3088, 3088, 3089, 3090, 3106, 3146, 3148, 3149, 3151, 3152, 3153, 3153, 3154, 3157, 3157, 3165, 3175, 3180, 3180, 3190, 3311, 3375

Table C.9: Calculated vibrational frequencies of ($PMMA_4H^+$)(Phe) at the AM1 level of theory for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
5, 9, 12, 16, 19, 24, 25, 32, 39, 45, , 58, 61, 66, 81, 84, 86, 97, 101, 102, 109, 112, 121, 127, 133, 140, 142, 145, 154, 168, 172, 185, 186, 189, 202, 215, 225, 226, 230, 231, 237, 243, 245, 254, 270, 272, 280, 289, 311, 321, 332, 341, 346, 350, 367, 373, 378, 382, 383, 407, 409, 417, 419, 423, 456, 463, 477, 499, 515, 534, 550, 573, 596, 609, 611, 620, 625, 648, 657, 659, 678, 701, 728, 735, 744, 756, 759, 809, 876, 892, 896, 918, 921, 933, 933, 947, 956, 962, 977, 978, 984, 992, 999 ,1010, 1010, 1017, 1021, 1033, 1045, 1053, 1060, 1075, 1102, 1108, 1147, 1153, 1162, 1165, 1172, 1175, 1175, 1178, 1178, 1180, 1189, 1189, 1195, 1198, 1208, 1209, 1213, 1219, 1230, 1244, 1249, 1261, 1271, 1279, 1286, 1290, 1294, 1297, 1305, 1317, 1327, 1328, 1346, 1348, 1349, 1349, 1351, 1351, 1356, 1360, 1362, 1372, 1373, 1374, 1378, 1379, 1379, 1380, 1380, 1383, 1386, 1387, 1388, 1388, 1392, 1394, 1398, 1401, 1402, 1406, 1406, 1410, 1416, 1418, 1421, 1424, 1427, 1432, 1435, 1436, 1440, 1442, 1444, 1455, 1457, 1457, 1460, 1468, 1528, 1534, 1549, 1566, 1570, 1635, 1661, 1719, 1765, 1779, 1782, 2021, 2031, 2053, 2065, 2955, 2975, 2991, 3003, 3012, 3019, 3050, 3052, 3052, 3052, 3053, 3056, 3059, 3062, 3063, 3063, 3066, 3066, 3067, 3069, 3069, 3074, 3075, 3080, 3080, 3088, 3089, 3089, 3129, 3147, 3148, 3151, 3152, 3153, 3154, 3155, 3156, 3170, 3179, 3184, 3188, 3190, 3197, 3419, 3420, 3465

Table C.10: Calculated vibrational frequencies of $PMMA_7$ obtained by MD-SA for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
15, 17, 23, 28, 39, 44, 46, 52, 55, 60, 63, 66, 69, 74, 81, 84, 95, 98, 102, 103, 110, 113, 120, 126, 130, 133, 135, 140, 147, 151, 161, 165, 170, 177, 181, 184, 193, 198, 211, 216, 226, 227, 235, 240, 243, 247, 253, 260, 264, 272, 273, 279, 280, 283, 286, 287, 292, 300, 304, 313, 324, 327, 340, 343, 345, 356, 360, 366, 369, 377, 388, 397, 401, 402, 406, 410, 412, 413, 420, 426, 433, 438, 445, 447, 451, 457, 459, 469, 473, 479, 483, 505, 508, 522, 542, 561, 563, 564, 570, 575, 588, 595, 610, 622, 645, 675, 692, 697, 725, 727, 727, 728, 753, 767, 787, 792, 797, 802, 807, 818, 836, 901, 903, 908, 910, 913, 917, 923, 957, 964, 981, 989, 990, 994, 995, 997, 1000, 1005, 1008, 1015, 1024, 1026, 1030, 1048, 1065, 1065, 1066, 1069, 1070, 1072, 1074, 1075, 1082, 1082, 1085, 1087, 1088, 1088, 1089, 1108, 1121, 1126, 1134, 1144, 1151, 1168, 1198, 1205, 1213, 1215, 1219, 1224, 1268, 1269, 1277, 1285, 1294, 1316, 1321, 1331, 1356, 1381, 1385, 1400, 1415, 1437, 1460, 1462, 1468, 1469, 1472, 1476, 1479, 1482, 1485, 1485, 1487, 1488, 1489, 1489, 1490, 1490, 1491, 1492, 1492, 1493, 1493, 1494, 1494, 1496, 1496, 1497, 1498, 1499, 1502, 1503, 1504, 1508, 1510, 1512, 1518, 1521, 1524, 1531, 1532, 1533, 1534, 1535, 1535, 1537, 1539, 1545, 1549, 1553, 1561, 1566, 1572, 1592, 1593, 1597, 1605, 1611, 1611, 1617, 1618, 1622, 1623, 1627, 1628, 1629, 1635, 1661, 1663, 1670, 1691, 1694, 1718, 2848, 2848, 2849, 2851, 2852, 2854, 2855, 2863, 2864, 2867, 2869, 2870, 2874, 2876, 2882, 2898, 2902, 2903, 2906, 2907, 2911, 2935, 2954, 2958, 2959, 2961, 2962, 2964, 2965, 2965, 2965, 2965, 2965, 2966, 2966, 2966, 2967, 2968, 2968, 2968, 2968, 2969, 2970, 2971, 2973, 2973, 2973, 2974, 2974, 2976, 2976, 2976, 2977, 2979, 2979, 2984, 2986, 2992

Table C.11: Calculated vibrational frequencies of $PMMA_7H^+$ obtained by MD-SA for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
16, 19, 24, 32, 35, 44, 46, 50, 56, 62, 64, 64, 68, 70, 82, 86, 92, 96, 101, 109, 112, 115, 123, 126, 132, 137, 141, 144, 148, 154, 164, 167, 174, 181, 188, 193, 200, 209, 213, 219, 225, 231, 234, 236, 242, 243, 251, 257, 262, 266, 269, 274, 280, 283, 285, 288, 288, 291, 304, 308, 325, 328, 342, 346, 350, 357, 361, 364, 370, 373, 377, 391, 393, 404, 406, 407, 410, 414, 420, 427, 431, 443, 447, 452, 456, 459, 463, 469, 472, 479, 483, 506, 509, 524, 542, 550, 559, 560, 564, 570, 576, 592, 593, 609, 625, 641, 681, 691, 701, 718, 726, 728, 737, 756, 767, 784, 796, 798, 804, 809, 819, 831, 899, 902, 909, 910, 914, 917, 923, 959, 964, 980, 987, 990, 992, 995, 998, 1000, 1005, 1011, 1015, 1024, 1026, 1030, 1046, 1064, 1065, 1066, 1068, 1070, 1074, 1074, 1078, 1082, 1082, 1083, 1085, 1088, 1089, 1089, 1112, 1122 ,1126, 1132, 1143, 1151, 1168, 1198, 1205, 1207, 1215, 1219, 1226, 1245, 1267, 1274, 1277, 1286, 1293, 1316, 1320, 1327, 1356, 1381, 1388, 1402, 1421, 1436, 1458, 1460, 1464, 1467, 1469, 1473, 1478, 1479, 1481, 1483, 1484, 1485, 1485, 1486, 1487, 1488, 1489, 1490, 1491, 1491, 1492, 1493, 1493, 1494, 1495, 1496, 1497, 1497, 1499, 1500, 1501, 1504, 1509, 1511, 1521, 1522, 1528, 1531, 1532, 1533, 1534, 1536, 1537, 1541, 1543, 1547, 1552, 1555, 1561, 1562, 1565, 1570, 1594, 1597, 1603, 1606, 1607, 1611, 1618, 1621, 1623, 1627, 1628, 1633, 1643, 1661, 1666, 1668, 1691, 1695, 1714, 2848, 2849, 2849, 2851, 2852, 2854, 2855, 2863, 2865, 2866, 2867, 2870, 2875, 2880, 2881, 2898, 2901, 2903, 2906, 2908, 2911, 2935, 2954, 2958, 2959, 2961, 2962, 2964, 2964, 2964, 2965 ,2965, 2965 ,2965, 2966, 2966, 2967, 2967, 2968, 2968, 2968, 2969, 2969, 2969, 2970, 2971, 2973, 2973, 2973, 2973, 2974, 2976, 2979, 2979, 2981, 2981, 2987, 2992, 2995

Table C.12: Calculated vibrational frequencies of ($PMMA_7H^+$)(Gly) obtained by MD-SA for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
14, 18, 19, 21, 24, 33, 35, 44, 46, 48, 51, 58, 62, 63, 66, 68, 71, 75, 84, 87, 92, 95, 98, 102, 110, 113, 116, 120, 128, 134, 138, 145, 157, 159, 164, 169, 178, 182, 185, 187, 196, 198, 204, 212, 221, 223, 228, 235, 240, 241, 246, 252, 254, 256, 264, 268, 271, 272, 278, 279, 283, 288, 293, 295, 303, 308, 311, 314, 328, 331, 339, 348, 357, 365, 368, 368, 370, 377, 386, 387, 388, 391, 399, 401, 402, 406, 411, 416, 419, 437, 439, 442, 447, 455, 460, 462, 472, 476, 486, 499, 501, 530, 535, 549, 567, 569, 571, 582, 584, 598, 610, 621, 628, 651, 656, 672, 710, 734, 739, 744, 745, 749, 750, 757, 769, 773, 775, 780, 794, 797, 804, 810, 814, 819, 830, 845, 877, 887, 904, 908, 911, 915, 916, 924, 943, 961, 968, 983, 988, 994, 994, 995, 1000, 1004, 1007, 1012, 1019, 1025, 1028, 1033, 1035, 1060, 1064, 1066, 1066, 1069, 1070, 1071, 1076, 1077, 1079, 1081, 1085, 1086, 1089, 1089, 1089, 1090, 1114, 1121, 1135, 1137, 1152, 1155, 1160, 1180, 1192, 1208, 1219, 1224, 1227, 1230, 1233, 1239, 1280, 1287, 1290, 1293, 1305, 1308, 1316, 1328, 1335, 1358, 1388, 1390, 1406, 1424, 1439, 1462, 1466, 1468, 1470, 1473, 1476, 1479, 1480, 1483, 1485, 1485, 1486, 1487, 1488, 1488, 1489, 1490, 1490, 1491, 1492, 1492, 1493, 1493, 1494, 1494, 1495, 1497, 1499, 1501, 1503, 1504, 1506, 1507, 1512, 1514, 1516, 1521, 1524, 1533, 1535, 1536, 1537, 1538, 1539, 1543, 1543, 1548, 1549, 1552, 1557, 1558, 1570, 1575, 1582, 1596, 1600, 1611, 1614, 1617, 1619, 1619, 1621, 1639, 1665, 1670, 1677, 1691, 1698, 1710, 1716, 1723, 1724, 1725, 1727, 1729, 2848, 2848, 2850, 2851, 2854, 2856, 2858, 2863, 2865, 2867, 2869, 2870, 2875, 2879, 2882, 2894, 2902, 2903, 2906, 2907, 2912, 2918, 2936, 2958, 2959, 2961, 2961, 2964, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2966, 2966, 2967, 2968, 2968, 2969, 2970, 2970, 2970, 2971, 2971, 2973, 2975, 2975, 2975, 2977, 2977, 2977, 2977, 2979, 2979, 2982, 2986, 2990, 2993, 3036, 3040, 3119, 3177

The calculated vibrational frequencies of ($PMMA_7$)($GlyH^+$) obtained by MD-SA for the determination of ΔS_{vib} are available in Table B.1.

Table C.13: Calculated vibrational frequencies of ($PMMA_7H^+$)(Leu) obtained by MD-SA for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
8, 8, 15, 18, 19, 24, 28, 30, 36, 41, 43, 45, 53, 53, 54, 63, 64, 68, 68, 70, 73, 83, 87, 91, 100, 102, 108, 113, 116, 120, 127, 134, 136, 141, 156, 160, 164, 171, 176, 177, 185, 187, 195, 197, 203, 211, 220, 225, 228, 229, 234, 240, 242, 245, 247, 253, 255, 259, 263, 267, 269, 270, 273, 278, 279, 282, 288, 294, 295, 298, 303, 311, 312, 328, 331, 338, 339, 343, 347, 355, 357, 365, 368, 371, 377, 386, 387, 390, 399, 401, 402, 406, 412, 414, 418, 431, 436, 440, 446, 448, 456, 461, 463, 474, 480, 489, 500, 501, 505, 529, 538, 550, 551, 565, 569, 572, 582, 585, 587, 599, 610, 622, 627, 650, 656, 673, 710, 734, 739, 745, 745, 749, 750, 756, 770, 775, 776, 778, 782, 795, 810, 814, 820, 824, 830, 832, 847, 887, 904, 908, 910, 914, 915, 916, 925, 938, 948, 961, 968, 975, 984, 988, 994, 994, 996, 999, 1001, 1004, 1007, 1012, 1019, 1025, 1028, 1029, 1034, 1060, 1065, 1066, 1066, 1069, 1070, 1070, 1071, 1077, 1077, 1079, 1081, 1083, 1085, 1089, 1089, 1089, 1090, 1096, 1106, 1115, 1124, 1135, 1137, 1153, 1155, 1163, 1180, 1183, 1192, 1209, 1219, 1225, 1228, 1231, 1233, 1281, 1286, 1289, 1290, 1295, 1305, 1310, 1311, 1330, 1335, 1348, 1362, 1388, 1391, 1395, 1407, 1417, 1425, 1443, 1463, 1466, 1468, 1471, 1471, 1473, 1477, 1478, 1479, 1481, 1484, 1485, 1486, 1487, 1487, 1488, 1489, 1489, 1490, 1491, 1491, 1492, 1492, 1493, 1493, 1494, 1494, 1495, 1497, 1498, 1500, 1500, 1503, 1504, 1505, 1506, 1510, 1513, 1516, 1518, 1521, 1525, 1532, 1534, 1535, 1537, 1537, 1538, 1540, 1542, 1543, 1547, 1552, 1553, 1559, 1561, 1570, 1575, 1580, 1581, 1598, 1600, 1611, 1613, 1617, 1617, 1621, 1624, 1643, 1663, 1670, 1672, 1694, 1698, 1716, 1721, 1723, 1724, 1725, 1727, 1735, 1736, 2848, 2848, 2850, 2851, 2854, 2855, 2857, 2861, 2863, 2865, 2865, 2867, 2869, 2870, 2875, 2880, 2881, 2900, 2902, 2903, 2904, 2906, 2908, 2914, 2936, 2938, 2948, 2956, 2958, 2959, 2961, 2962, 2965, 2965, 2965, 2965, 2965, 2965, 2965, 2966, 2966, 2967, 2967, 2968, 2968, 2968, 2968, 2969, 2969, 2970, 2970, 2970, 2971, 2972, 2973, 2975, 2975, 2975, 2975, 2976, 2977, 2977, 2978, 2979, 2981, 2986, 2990, 2992, 3037, 3040, 3119, 3177

The calculated vibrational frequencies of ($PMMA_7$)($LeuH^+$) obtained by MD-SA for the determination of ΔS_{vib} are available in Table B.2.

Table C.14: Calculated vibrational frequencies of ($PMMA_7H^+$)(Phe) obtained by MD-SA for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
9, 12, 17, 18, 19, 24, 27, 28, 32, 43, 44, 44, 50, 52, 57, 59, 61, 65, 67, 72, 78, 84, 86, 93, 94, 101, 109, 112, 113, 119, 129, 130, 135, 148, 155, 159, 165, 174, 177, 179, 186, 188, 194, 201, 202, 209, 216, 225, 225, 236, 238, 240, 241, 246, 248, 256, 259, 261, 264, 269, 271, 275, 277, 279, 282, 283, 287, 289, 291, 296, 298, 313, 324, 325, 327, 341, 343, 345, 353, 359, 369, 375, 382, 388, 395, 399, 401, 404, 405, 408, 416, 418, 420, 422, 426, 437, 449, 457, 458, 459, 462, 467, 471, 484, 496, 500, 505, 511, 523, 530, 543, 554, 563, 569, 573, 577, 580, 585, 589, 599, 607, 614, 624, 645, 651, 661, 663, 689, 722, 735, 741, 744, 748, 750, 752, 754, 759, 768, 777, 779, 782, 789, 795, 808, 809, 816, 819, 819, 831, 849, 900, 901, 904, 907, 911, 912, 916, 917, 925, 939, 961, 966, 985, 987, 991, 994, 995, 998, 999, 1001, 1009, 1010, 1012, 1019, 1021, 1028, 1032, 1035, 1036, 1059, 1066, 1067, 1068, 1071, 1071, 1072, 1073, 1077, 1078, 1083, 1083, 1084, 1087, 1088, 1090, 1095, 1106, 1114, 1129, 1129, 1135, 1142, 1150, 1151, 1157, 1166, 1174, 1183, 1209, 1216, 1216, 1220, 1225, 1230, 1238, 1258, 1260, 1277, 1286, 1288, 1288, 1290, 1294, 1301, 1322, 1328, 1337, 1343, 1358, 1375, 1384, 1390, 1401, 1418, 1438, 1458, 1464, 1467, 1470, 1471, 1474, 1475, 1480, 1481, 1484, 1486, 1486, 1487, 1488, 1488, 1489, 1489, 1490, 1492, 1492, 1492, 1493, 1495, 1495, 1496, 1496, 1496, 1497, 1499, 1501, 1503, 1503, 1505, 1507, 1512, 1515, 1515, 1519, 1521, 1524, 1525, 1529, 1534, 1535, 1537, 1538, 1540, 1541, 1542, 1549, 1552, 1559, 1571, 1578, 1581, 1586, 1599, 1600, 1607, 1608, 1617, 1621, 1624, 1634, 1639, 1656, 1665, 1669, 1694, 1698, 1701, 1717, 1719, 1724, 1725, 1726, 1729, 1733, 1754, 1783, 1802, 1848, 2848, 2849, 2850, 2851, 2854, 2854, 2857, 2863, 2865, 2867, 2869, 2875, 2877, 2878, 2881, 2899, 2902, 2904, 2906, 2907, 2910, 2911, 2931, 2936, 2954, 2958, 2959, 2960, 2961, 2963, 2964, 2964, 2965, 2965, 2965, 2966, 2967, 2967, 2967, 2967, 2968, 2968, 2969, 2969, 2970, 2970, 2971, 2971, 2971, 2971, 2973, 2973, 2973, 2973, 2973, 2976, 2977, 2977, 2978, 2978, 2979, 2981, 2982, 2986, 2988, 2991, 2992, 3036, 3118, 3178

The calculated vibrational frequencies of ($PMMA_7$)($PheH^+$) obtained by MD-SA for the determination of ΔS_{vib} are available in Table B.3.

Table C.15: Calculated vibrational frequencies of $PMMA_{13}$ obtained by MD-SA for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
6, 6, 7, 12, 14, 15, 18, 20, 23, 24, 27, 29, 30, 31, 31, 32, 34, 34, 39, 41, 44, 46, 48, 49, 51, 52, 53, 57, 59, 60, 63, 64, 65, 69, 70, 72, 73, 74, 75, 76, 78, 80, 82, 89, 94, 98, 101, 107, 115, 122, 122, 137, 138, 147, 152, 158, 159, 162, 163, 167, 169, 171, 174, 177, 183, 185, 188, 190, 193, 197, 203, 205, 210, 215, 222, 234, 234, 235, 236, 237, 239, 240, 241, 243, 249, 251, 253, 255, 260, 265, 267, 268, 269, 271, 271, 272, 273, 277, 280, 281, 282, 288, 290, 294, 298, 301, 304, 308, 310, 313, 318, 322, 324, 325, 328, 329, 332, 332, 333, 336, 340, 341, 342, 345, 347, 351, 370, 370, 376, 379, 383, 390, 392, 392, 394, 395, 396, 398, 401, 401, 403, 407, 408, 414, 414, 416, 418, 418, 420, 424, 428, 429, 431, 432, 442, 447, 449, 452, 454, 454, 455, 456, 458, 461, 462, 465, 475, 478, 478, 484, 486, 490, 491, 492, 494, 497, 499, 500, 506, 526, 552, 552, 564, 566, 571, 578, 583, 590, 600, 602, 605, 606, 612, 614, 619, 621, 627, 631, 637, 639, 652, 667, 676, 690, 724, 728, 732, 735, 739, 743, 748, 752, 754, 756, 756, 757, 764, 805, 826, 843, 845, 850, 852, 858, 862, 866, 871, 878, 890, 903, 904, 906, 910, 910, 915, 920, 924, 927, 928, 929, 929, 930, 967, 969, 977, 984, 986, 989, 991, 992, 995, 995, 998, 1000, 1001, 1005, 1005, 1012, 1013, 1019, 1021, 1021, 1022, 1023, 1024, 1026, 1026, 1027, 1028, 1047, 1065, 1066, 1066, 1066, 1066, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1067, 1070, 1084, 1085, 1085, 1085, 1085, 1085, 1086, 1086, 1086, 1086, 1086, 1086, 1087, 1106, 1113, 1115, 1118, 1122, 1128, 1134, 1140, 1147, 1153, 1158, 1162, 1166, 1181, 1186, 1191, 1197, 1203, 1207, 1211, 1216, 1218, 1221, 1222, 1231, 1252, 1255, 1259, 1264, 1270, 1279, 1282, 1294, 1303, 1306, 1314, 1323, 1331, 1343, 1350, 1374, 1379, 1380, 1385, 1389, 1393, 1397, 1400, 1402, 1406, 1424, 1461, 1463, 1466, 1468, 1469, 1470, 1471, 1472, 1472, 1474, 1474, 1477, 1479, 1482, 1486, 1486, 1487, 1487, 1488, 1488, 1488, 1488, 1488, 1488, 1488, 1488, 1488, 1490, 1492, 1492, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1494, 1494, 1494, 1495, 1496, 1497, 1505, 1506, 1507, 1507, 1507, 1508, 1508, 1508, 1509, 1510, 1513, 1514, 1518, 1521, 1521, 1523, 1526, 1527, 1527, 1528, 1528, 1528, 1528, 1529, 1529, 1529, 1530, 1531, 1532, 1533, 1533, 1534, 1534, 1535, 1535, 1536, 1538, 1539, 1539, 1540, 1542, 1544, 1547, 1549, 1554, 1555, 1558, 1561, 1562, 1565, 1567, 1568, 1569, 1570, 1570, 1571, 1571, 1572, 1583, 1587, 1591, 1594, 1598, 1602, 1609, 1613, 1619, 1623, 1626, 1627, 1628, 1630, 1632, 1636, 1636, 1644, 1648, 1652, 1653, 1656, 1658, 1660, 1665, 1671, 1676, 1680, 1682, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2851, 2862, 2864, 2864, 2865, 2865, 2865, 2865, 2865, 2865, 2865, 2866, 2866, 2866, 2868, 2904, 2906, 2907, 2919, 2919, 2919, 2919, 2920, 2920, 2920, 2923, 2924, 2936, 2961, 2962, 2965, 2966, 2966, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2969, 2969, 2970, 2971, 2972, 2972, 2973, 2973, 2973, 2973, 2973, 2974, 2974, 2975, 2975, 2977, 2980, 2981, 2981, 2982, 2982, 2983, 2984, 2986, 2988

Table C.16: Calculated vibrational frequencies of $PMMA_{13}H^+$ obtained by MD-SA for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
6, 7, 7, 12, 14, 15, 18, 19, 22, 23, 25, 28, 29, 31, 32, 33, 34, 35, 39, 41, 44, 47, 49, 50, 51, 52, 53, 57, 58, 60, 61, 63, 64, 65, 70, 71, 73, 74, 75, 76, 78, 79, 82, 89, 94, 97, 100, 107, 115, 122, 123, 137, 138, 146, 151, 158, 159, 162, 164, 167, 169, 171, 175, 177, 183, 183, 186, 189, 192, 197, 201, 205, 210, 213, 222, 232, 235, 235, 236, 237, 238, 239, 243, 249, 251, 252, 253, 255, 260, 263, 267, 268, 268, 270, 271, 272, 274, 277, 280, 282, 283, 283, 288, 294, 297, 300, 304, 307, 310, 314, 315, 320, 322, 324, 327, 328, 331, 333, 334, 336, 336, 339, 343, 343, 348, 350, 357, 370, 370, 376, 379, 383, 389, 391, 392, 393, 394, 395, 397, 400, 401, 402, 404, 408, 413, 414, 416, 418, 419, 420, 423, 428, 429, 430, 432, 442, 447, 447, 451, 452, 453, 455, 456, 458, 462, 464, 465, 476, 476, 479, 480, 486, 490, 492, 493, 494, 498, 499, 500, 506, 526, 551, 553, 561, 565, 571, 577, 582, 585, 592, 601, 602, 606, 609, 613, 619, 620, 626, 629, 637, 638, 651, 667, 675, 690, 708, 724, 729, 733, 738, 741, 745, 751, 753, 755, 756, 757, 764, 805, 826, 843, 844, 848, 852, 859, 861, 865, 871, 878, 890, 900, 904, 906, 909, 910, 914, 919, 923, 926, 927, 929, 929, 930, 967, 969, 977, 984, 986, 989, 991, 991, 995, 995, 997, 1000, 1001, 1005, 1006, 1012, 1013, 1019, 1019, 1021, 1022, 1023, 1024, 1025, 1027, 1028, 1029, 1047, 1065, 1066, 1066, 1066, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1067, 1067, 1070, 1075, 1084, 1085, 1085, 1085, 1085, 1086, 1086, 1086, 1086, 1086, 1086, 1087, 1106, 1110, 1114, 1117, 1122, 1126, 1133, 1136, 1143, 1147, 1153, 1158, 1161, 1166, 1184, 1186, 1192, 1197, 1205, 1207, 1212, 1216, 1219, 1221, 1223, 1231, 1255, 1255, 1260, 1264, 1270, 1279, 1283, 1294, 1303, 1306, 1314, 1323, 1332, 1343, 1350, 1374, 1378, 1380, 1386, 1388, 1394, 1397, 1401, 1405, 1406, 1424, 1461, 1463, 1466, 1468, 1469, 1469, 1471, 1471, 1474, 1474, 1477, 1479, 1482, 1482, 1485, 1486, 1486, 1487, 1487, 1487, 1488, 1488, 1488, 1488, 1488, 1488, 1490, 1492, 1492, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1495, 1495, 1496, 1497, 1504, 1505, 1506, 1507, 1508, 1508, 1508, 1509, 1510, 1510, 1513, 1514, 1517, 1518, 1521, 1522, 1523, 1526, 1527, 1528, 1528, 1528, 1529, 1529, 1529, 1529, 1530, 1530, 1531, 1533, 1533, 1534, 1534, 1535, 1536, 1536, 1537, 1538, 1539, 1540, 1542, 1543, 1544, 1547, 1550, 1555, 1555, 1559, 1563, 1564, 1565, 1567, 1567, 1570, 1572, 1572, 1573, 1573, 1583, 1587, 1591, 1593, 1597, 1598, 1601, 1609, 1613, 1622, 1622, 1627, 1627, 1629, 1631, 1636, 1637, 1644, 1647, 1650, 1655, 1657, 1658, 1661, 1666, 1671, 1677, 1679, 1684, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2851, 2862, 2864, 2864, 2865, 2865, 2865, 2865, 2866, 2866, 2866, 2866, 2866, 2866, 2868, 2904, 2906, 2907, 2918, 2919, 2920, 2920, 2921, 2922, 2923, 2924, 2924, 2936, 2961, 2962, 2963, 2964, 2965, 2966, 2966, 2966, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2969, 2969, 2970, 2971, 2972, 2972, 2973, 2973, 2973, 2973, 2974, 2974, 2974, 2975, 2975, 2978, 2979, 2981, 2982, 2983, 2984, 2984, 2985, 2987, 2988, 3037

Table C.17: Calculated vibrational frequencies of ($PMMA_{13}H^+$)(Gly) obtained by MD-SA for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
6, 6, 6, 9, 11, 14, 15, 16, 19, 21, 23, 24, 26, 28, 29, 31, 32, 33, 35, 38, 40, 41, 44, 46, 47, 49, 50, 51, 52, 53, 54, 57, 58, 60, 60, 62, 63, 64, 65, 68, 71, 73, 74, 76, 77, 79, 81, 83, 88, 90, 94, 97, 100, 107, 114, 120, 123, 124, 138, 138, 138, 147, 151, 158, 160, 162, 164, 167, 169, 171, 174, 177, 182, 183, 186, 189, 192, 197, 201, 205, 211, 213, 222, 232, 235, 235, 236, 237, 238, 241, 243, 249, 251, 253, 253, 255, 260, 263, 267, 268, 268, 271, 272, 273, 277, 277, 280, 282, 283, 285, 288, 294, 297, 300, 304, 307, 310, 313, 316, 319, 322, 324, 326, 328, 331, 333, 334, 335, 336, 341, 343, 344, 348, 349, 355, 357, 370, 370, 377, 379, 383, 384, 389, 391, 392, 393, 394, 395, 397, 399, 401, 401, 403, 405, 409, 413, 414, 416, 418, 419, 420, 424, 428, 430, 431, 432, 442, 447, 447, 452, 454, 454, 455, 456, 456, 458, 462, 464, 466, 476, 478, 479, 483, 487, 490, 492, 493, 495, 498, 499, 500, 506, 526, 532, 551, 553, 560, 565, 570, 577, 582, 586, 592, 601, 603, 607, 611, 614, 619, 620, 626, 630, 637, 638, 651, 667, 675, 690, 708, 724, 729, 733, 739, 741, 746, 751, 754, 755, 757, 758, 764, 778, 805, 826, 843, 846, 848, 852, 860, 861, 865, 871, 871, 878, 891, 900, 904, 906, 909, 910, 915, 919, 923, 926, 927, 929, 929, 930, 944, 967, 969, 977, 984, 986, 989, 991, 992, 995, 995, 997, 1000, 1001, 1005, 1006, 1012, 1013, 1019, 1020, 1021, 1022, 1022, 1024, 1025, 1027, 1028, 1028, 1046, 1047, 1065, 1066, 1066, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1067, 1067, 1067, 1070, 1075, 1084, 1085, 1085, 1085, 1086, 1086, 1086, 1086, 1086, 1086, 1087, 1087, 1106, 1109, 1114, 1117, 1122, 1126, 1128, 1133, 1136, 1143, 1147, 1153, 1158, 1159, 1162, 1166, 1183, 1186, 1192, 1197, 1205, 1207, 1211, 1216, 1218, 1221, 1223, 1226, 1231, 1254, 1255, 1260, 1264, 1270, 1279, 1282, 1294, 1303, 1306, 1314, 1323, 1326, 1331, 1343, 1350, 1374, 1378, 1380, 1385, 1388, 1394, 1397, 1400, 1404, 1406, 1424, 1461, 1463, 1467, 1468, 1469, 1470, 1471, 1472, 1474, 1475, 1477, 1479, 1482, 1482, 1486, 1486, 1486, 1487, 1487, 1487, 1488, 1488, 1488, 1488, 1488, 1488, 1490, 1492, 1492, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1494, 1494, 1495, 1495, 1496, 1497, 1498, 1504, 1505, 1506, 1507, 1507, 1508, 1508, 1510, 1510, 1510, 1513, 1514, 1517, 1518, 1521, 1522, 1523, 1527, 1527, 1528, 1528, 1528, 1529, 1529, 1529, 1530, 1530, 1530, 1531, 1533, 1533, 1534, 1534, 1535, 1536, 1536, 1537, 1538, 1539, 1540, 1542, 1543, 1545, 1546, 1547, 1551, 1555, 1555, 1559, 1563, 1564, 1565, 1567, 1567, 1570, 1571, 1572, 1572, 1574, 1583, 1587, 1591, 1593, 1597, 1598, 1601, 1609, 1611, 1613, 1622, 1622, 1626, 1627, 1627, 1629, 1630, 1636, 1637, 1644, 1647, 1650, 1653, 1657, 1658, 1661, 1666, 1671, 1677, 1679, 1683, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2849, 2851, 2862, 2864, 2864, 2865, 2865, 2865, 2865, 2866, 2866, 2866, 2866, 2866, 2868, 2868, 2892, 2904, 2906, 2907, 2918, 2919, 2920, 2920, 2921, 2921, 2922, 2924, 2924, 2936, 2961, 2962, 2963, 2963, 2964, 2965, 2966, 2966, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2969, 2969, 2970, 2970, 2971, 2972, 2972, 2973, 2973, 2973, 2974, 2974, 2974, 2975, 2975, 2975, 2978, 2979, 2981, 2981, 2983, 2984, 2984, 2984, 2986, 2988, 3036, 3037, 3120, 3179

The calculated vibrational frequencies of ($PMMA_{13}$)($GlyH^+$) obtained by MD-SA for the determination of ΔS_{vib} are available in Table B.7.

Table C.18: Calculated vibrational frequencies of ($PMMA_{13}H^+$)(Leu) obtained by MD-SA for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
5, 6, 6, 12, 13, 14, 16, 17, 20, 24, 24, 26, 27, 29, 30, 31, 31, 33, 34, 36, 40, 42, 43, 47, 48, 49, 52, 53, 54, 56, 57, 59, 60, 61, 64, 65, 67, 68, 71, 71, 73, 73, 74, 75, 77, 80, 81, 84, 85, 87, 88, 93, 98, 98, 101, 105, 109, 117, 120, 123, 136, 138, 148, 152, 160, 161, 163, 165, 167, 171, 172, 173, 174, 176, 185, 186, 189, 191, 197, 201, 206, 208, 216, 222, 222, 230, 232, 234, 234, 236, 238, 240, 240, 241, 243, 244, 249, 252, 253, 255, 258, 259, 265, 267, 268, 271, 272, 273, 275, 277, 278, 280, 281, 282, 285, 287, 289, 295, 299, 300, 303, 308, 310, 311, 318, 321, 324, 326, 329, 329, 331, 333, 333, 335, 336, 338, 340, 342, 344, 345, 347, 349, 353, 358, 370, 370, 377, 379, 381, 385, 391, 392, 393, 395, 396, 398, 399, 401, 404, 407, 408, 411, 414, 414, 416, 418, 420, 423, 425, 428, 430, 431, 432, 442, 446, 448, 449, 452, 453, 454, 455, 457, 458, 462, 463, 464, 475, 478, 478, 485, 488, 490, 492, 494, 495, 498, 499, 500, 500, 506, 514, 526, 548, 550, 552, 561, 563, 565, 572, 581, 588, 589, 599, 602, 603, 607, 611, 615, 618, 621, 627, 630, 636, 639, 648, 665, 676, 690, 715, 724, 728, 733, 737, 742, 748, 751, 754, 756, 757, 758, 767, 776, 807, 826, 831, 840, 843, 849, 850, 856, 860, 865, 867, 876, 888, 903, 904, 905, 907, 912, 914, 916, 921, 925, 927, 928, 929, 930, 931, 934, 949, 963, 969, 974, 977, 984, 986, 989, 991, 993, 995, 997, 998, 1000, 1002, 1002, 1004, 1006, 1012, 1017, 1020, 1020, 1021, 1022, 1022, 1024, 1024, 1026, 1028, 1029, 1032, 1043, 1066, 1066, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1067, 1067, 1068, 1069, 1070, 1071, 1077, 1084, 1085, 1085, 1085, 1085, 1085, 1086, 1086, 1086, 1086, 1086, 1090, 1098, 1104, 1107, 1113, 1116, 1119, 1123, 1125, 1128, 1134, 1140, 1147, 1153, 1153, 1158, 1162, 1164, 1166, 1181, 1186, 1190, 1191, 1197, 1204, 1207, 1212, 1217, 1219, 1221, 1222, 1231, 1253, 1255, 1259, 1264, 1269, 1279, 1283, 1288, 1290, 1294, 1305, 1313, 1314, 1323, 1332, 1345, 1349, 1350, 1373, 1380, 1381, 1385, 1389, 1394, 1397, 1397, 1402, 1404, 1406, 1418, 1424, 1463, 1464, 1468, 1468, 1469, 1470, 1471, 1472, 1473, 1474, 1474, 1475, 1476, 1478, 1479, 1481, 1484, 1486, 1487, 1487, 1487, 1487, 1487, 1488, 1488, 1488, 1488, 1488, 1489, 1489, 1490, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1494, 1494, 1494, 1494, 1495, 1496, 1497, 1497, 1501, 1505, 1506, 1507, 1507, 1507, 1508, 1508, 1509, 1510, 1510, 1512, 1513, 1514, 1517, 1519, 1521, 1523, 1525, 1526, 1527, 1527, 1528, 1528, 1528, 1529, 1530, 1530, 1531, 1531, 1532, 1533, 1533, 1534, 1535, 1535, 1535, 1536, 1538, 1538, 1539, 1541, 1541, 1541, 1543, 1547, 1548, 1552, 1552, 1553, 1555, 1557, 1560, 1562, 1566, 1567, 1567, 1569, 1570, 1570, 1572, 1576, 1581, 1584, 1587, 1589, 1593, 1595, 1601, 1605, 1612, 1613, 1613, 1618, 1622, 1624, 1626, 1627, 1628, 1630, 1633, 1636, 1638, 1647, 1652, 1654, 1656, 1657, 1660, 1665, 1670, 1675, 1679, 1682, 1735, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2849, 2849, 2849, 2850, 2851, 2861, 2862, 2864, 2864, 2865, 2865, 2865, 2865, 2865, 2866, 2866, 2866, 2867, 2868, 2868, 2869, 2903, 2904, 2906, 2907, 2918, 2918, 2919, 2919, 2919, 2920, 2922, 2923, 2927, 2935, 2941, 2950, 2961, 2961, 2963, 2964, 2965, 2966, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2969, 2969, 2969, 2969, 2970, 2970, 2970, 2970, 2971, 2972, 2972, 2973, 2973, 2973, 2973, 2974, 2974, 2975, 2975, 2975, 2976, 2977, 2980, 2980, 2981, 2982, 2982, 2984, 2985, 2987, 2992, 3037, 3039, 3118, 3177

The calculated vibrational frequencies of ($PMMA_{13}$)($LeuH^+$) obtained by MD-SA for the determination of ΔS_{vib} are available in Table B.8.

Table C.19: Calculated vibrational frequencies of ($PMMA_{13}H^+$)(Phe) obtained by MD-SA for the determination of ΔS_{vib}.

Vibrational harmonic frequencies (cm^{-1})
1, 6, 6, 8, 13, 14, 17, 18, 21, 23, 23, 25, 28, 29, 30, 32, 33, 34, 35, 36, 38, 40, 42, 43, 46, 48, 49, 50, 51, 53, 53, 54, 58, 60, 61, 62, 63, 64, 65, 68, 70, 72, 73, 75, 77, 78, 78, 82, 85, 86, 92, 93, 95, 100, 105, 108, 114, 121, 122, 124, 128, 137, 139, 150, 153, 159, 162, 164, 165, 166, 170, 171, 175, 177, 183, 183, 183, 187, 189, 192, 196, 200, 205, 207, 212, 222, 233, 235, 236, 236, 237, 239, 244, 246, 250, 251, 253, 253, 256, 260, 263, 266, 268, 268, 269, 271, 273, 275, 277, 277, 281, 283, 284, 284, 286, 288, 295, 300, 301, 304, 308, 311, 312, 314, 317, 320, 322, 324, 326, 329, 332, 333, 333, 334, 336, 337, 342, 343, 344, 348, 350, 357, 369, 370, 376, 378, 383, 389, 390, 391, 393, 395, 396, 397, 399, 400, 402, 403, 404, 408, 412, 414, 417, 417, 419, 420, 422, 425, 428, 430, 430, 432, 443, 447, 448, 452, 452, 455, 456, 456, 458, 463, 464, 465, 465, 475, 476, 480, 480, 486, 490, 492, 493, 495, 498, 500, 502, 503, 509, 528, 537, 553, 554, 561, 566, 571, 576, 577, 582, 585, 593, 601, 605, 607, 607, 613, 614, 619, 621, 623, 626, 629, 638, 639, 652, 653, 667, 675, 685, 690, 708, 725, 729, 733, 738, 743, 745, 746, 753, 754, 756, 756, 761, 764, 784, 805, 819, 826, 844, 846, 851, 855, 860, 863, 871, 872, 880, 892, 896, 901, 901, 904, 906, 910, 910, 915, 921, 923, 927, 927, 929, 929, 931, 938, 967, 969, 977, 984, 986, 987, 989, 990, 991, 994, 995, 997, 1000, 1001, 1004, 1005, 1013, 1014, 1014, 1019, 1019, 1022, 1023, 1023, 1024, 1026, 1027, 1028, 1031, 1032, 1034, 1047, 1065, 1066, 1066, 1066, 1066, 1066, 1066, 1067, 1067, 1067, 1067, 1067, 1068, 1070, 1073, 1075, 1084, 1085, 1085, 1085, 1085, 1086, 1086, 1086, 1086, 1086, 1087, 1091, 1106, 1109, 1110, 1113, 1117, 1122, 1126, 1129, 1133, 1136, 1143, 1147, 1151, 1153, 1159, 1162, 1167, 1175, 1183, 1186, 1192, 1197, 1204, 1207, 1211, 1215, 1216, 1217, 1218, 1221, 1222, 1231, 1253, 1255, 1260, 1264, 1267, 1270, 1280, 1282, 1288, 1293, 1297, 1303, 1306, 1314, 1321, 1331, 1343, 1348, 1351, 1374, 1377, 1379, 1380, 1384, 1387, 1393, 1396, 1401, 1405, 1407, 1425, 1461, 1463, 1466, 1469, 1469, 1470, 1471, 1474, 1475, 1475, 1477, 1480, 1482, 1482, 1486, 1486, 1486, 1487, 1487, 1488, 1488, 1488, 1488, 1488, 1488, 1489, 1491, 1491, 1492, 1492, 1492, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1493, 1494, 1494, 1495, 1496, 1496, 1497, 1500, 1504, 1505, 1505, 1507, 1507, 1507, 1508, 1508, 1509, 1510, 1512, 1514, 1514, 1517, 1519, 1521, 1522, 1526, 1526, 1528, 1529, 1529, 1529, 1529, 1529, 1530, 1531, 1531, 1531, 1532, 1534, 1534, 1535, 1536, 1536, 1536, 1538, 1539, 1539, 1541, 1542, 1543, 1546, 1547, 1550, 1552, 1555, 1558, 1562, 1565, 1565, 1565, 1567, 1568, 1569, 1571, 1572, 1573, 1573, 1583, 1586, 1589, 1591, 1596, 1597, 1598, 1601, 1607, 1609, 1614, 1621, 1622, 1626, 1627, 1627, 1628, 1631, 1636, 1637, 1644, 1648, 1650, 1655, 1657, 1658, 1661, 1667, 1672, 1678, 1680, 1683, 1702, 1754, 1784, 1803, 1849, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2848, 2849, 2850, 2851, 2862, 2864, 2864, 2865, 2865, 2866, 2866, 2866, 2866, 2866, 2866, 2866, 2868, 2877, 2904, 2906, 2906, 2907, 2914, 2915, 2917, 2921, 2922, 2923, 2924, 2924, 2924, 2931, 2936, 2962, 2963, 2963, 2964, 2964, 2965, 2966, 2966, 2966, 2966, 2967, 2967, 2967, 2967, 2967, 2967, 2967, 2968, 2968, 2968, 2969, 2969, 2970, 2971, 2971, 2971, 2971, 2972, 2972, 2973, 2973, 2973, 2973, 2973, 2974, 2974, 2974, 2975, 2975, 2976, 2977, 2978, 2980, 2981, 2982, 2982, 2984, 2985, 2987, 2988, 2992, 3014, 3037, 3118, 3178

The calculated vibrational frequencies of ($PMMA_{13}$)($PheH^+$) obtained by MD-SA for the determination of ΔS_{vib} are available in Table B.9.

Table C.20: Rotational constants (A, B, C) for the amino acids and 4mer systems, calculated at the AM1 level of theory.

chemical species	Rotational constants (GHz)		
	A	B	C
Gly	10.3818	3.7780	2.8648
$GlyH^+$	10.0730	3.6843	2.7927
Leu	2.2302	1.0217	0.8914
$LeuH^+$	2.7455	0.8626	0.7678
Phe	1.9103	0.5586	0.4804
$PheH^+$	2.3937	0.4599	0.4204
$PMMA_4$	0.2891	0.1084	0.0904
$PMMA_4H^+$	0.2081	0.1785	0.1272
$(PMMA_4)(GlyH^+)$	0.1310	0.1155	0.1133
$(PMMA_4H^+)(Gly)$	0.1537	0.1017	0.0884
$(PMMA_4)(LeuH^+)$	0.1035	0.0931	0.0822
$(PMMA_4H^+)(Leu)$	0.1264	0.0791	0.0676
$(PMMA_4)(PheH^+)$	0.1029	0.0737	0.0621
$(PMMA_4H^+)(Phe)$	0.1082	0.0554	0.0518

Table C.21: Moments if inertia (I_A, I_B, I_C) for the 7mer and 13mer systems, calculated by Nmode.

chemical species	Moments of inertia (amu Å^2)		
	I_A	I_B	I_C
$PMMA_7$	3330	23168	23955
$PMMA_7H^+$	3333	23373	23856
$(PMMA_7)(GlyH^+)$	5340	23580	25859
$(PMMA_7H^+)(Gly)$	4656	23225	25095
$(PMMA_7)(LeuH^+)$	6174	28322	31441
$(PMMA_7H^+)(Leu)$	6155	24311	27061
$(PMMA_7)(PheH^+)$	6351	24686	26691
$(PMMA_7H^+)(Phe)$	9569	24229	30566
$PMMA_{13}$	8575	107346	108236
$PMMA_{13}H^+$	8587	107569	108446
$(PMMA_{13})(GlyH^+)$	11821	106711	108811
$(PMMA_{13}H^+)(Gly)$	10384	110533	111660
$(PMMA_{13})(LeuH^+)$	11630	109791	111655
$(PMMA_{13}H^+)(Leu)$	12117	113140	115302
$(PMMA_{13})(PheH^+)$	12376	109360	112618
$(PMMA_{13}H^+)(Phe)$	12869	115585	119603

Appendix D: Vibrational frequencies used for the RRKM modeling of the Aβ1-40/LMW molecule complexes

Table D.1: Calculated vibrational frequencies of LMW molecules at the AM1 level of theory.

ID	Vibrational harmonic frequencies (cm^{-1})
M1	20, 53, 89, 127, 191, 241, 269, 292, 300, 386, 423, 493, 515, 756, 777, 809, 867, 963, 967, 1012, 1077, 1086, 1165, 1193, 1209, 1216, 1239, 1303, 1326, 1349, 1373, 1375, 1387, 1400, 1465, 1731, 2978, 2985, 3019, 3048, 3072, 3089, 3422, 3466, 3542
M2	12, 20, 34, 60, 70, 98, 118, 158, 174, 209, 227, 289, 308, 321, 355, 405, 422, 450, 482, 489, 526, 556, 633, 756, 770, 801, 812, 837, 887, 900, 949, 961, 967, 1002, 1037, 1040, 1064, 1086, 1111, 1133, 1142, 1183, 1186, 1194, 1203, 1218, 1226, 1227, 1235, 1241, 1243, 1247, 1254, 1299, 1307, 1313, 1350, 1369, 1373, 1378, 1389, 1392, 1394, 1399, 1401, 1405, 1411, 1412, 1413, 1424, 1426, 1460, 1469, 1578, 2930, 2940, 2983, 3014, 3018, 3020, 3029, 3032, 3033, 3038, 3072, 3089, 3092, 3093, 3097, 3100, 3102, 3404, 3545
M3	23, 26, 41, 43, 74, 132, 153, 189, 197, 216, 230, 284, 287, 327, 365, 379, 419, 432, 460, 466, 486, 489, 507, 566, 591, 664, 760, 762, 824, 842, 885, 909, 945, 949, 958, 999, 1003, 1035, 1062, 1068, 1076, 1099, 1112, 1134, 1163, 1187, 1188, 1209, 1218, 1225, 1239, 1242, 1248, 1254, 1279, 1293, 1298, 1315, 1341, 1351, 1358, 1368, 1373, 1385, 1390, 1401, 1405, 1409, 1412, 1413, 1424, 1425, 1426, 1448, 1467, 1545, 1577, 2937, 2973, 2979, 3014, 3019, 3029, 3033, 3038, 3040, 3048, 3072, 3092, 3093, 3097, 3100, 3102, 3389, 3427, 3537
M4	11, 14, 31, 50, 68, 89, 111, 127, 149, 179, 188, 198, 200, 243, 279, 281, 322, 330, 346, 385, 423, 427, 434, 488, 518, 528, 554, 590, 613, 644, 667, 742, 751, 773, 782, 806, 825, 846, 882, 903, 915, 956, 959, 962, 988, 992, 1029, 1045, 1071, 1084, 1091, 1134, 1148, 1161, 1178, 1181, 1193, 1193, 1207, 1208, 1222, 1226, 1234, 1245, 1279, 1300, 1306, 1322, 1341, 1363, 1370, 1375, 1378, 1381, 1384, 1389, 1396, 1399, 1422, 1444, 1480, 1495, 1505, 1527, 1542, 1642, 1676, 1745, 1810, 1835, 2936, 2948, 2955, 2994, 3016, 3019, 3021, 3027, 3038, 3076, 3087, 3093, 3182, 3189, 3198, 3207, 3536, 3546
M5	10, 15, 25, 33, 65, 77, 92, 108, 117, 146, 168, 182, 190, 201, 215, 280, 286, 311, 331, 348, 355, 391, 418, 427, 478, 486, 518, 528, 557, 589, 614, 647, 670, 751, 752, 769, 780, 805, 808, 827, 845, 903, 912, 936, 959, 962, 968, 988, 992, 1040, 1050, 1072, 1085, 1091, 1137, 1148, 1162, 1179, 1187, 1191, 1193, 1202, 1204, 1208, 1217, 1228, 1238, 1244, 1246, 1279, 1299, 1300, 1323, 1341, 1351, 1364, 1371, 1378, 1380, 1384, 1389, 1396, 1401, 1404, 1416, 1424, 1449, 1481, 1495, 1506, 1527, 1542, 1642, 1677, 1744, 1810, 1835, 2937, 2949, 2955, 2983, 3008, 3016, 3022, 3026, 3035, 3039, 3069, 3085, 3087, 3103, 3182, 3189, 3198, 3207, 3543, 3547

Table D.2: Vibrational harmonic frequencies of the 20 common amino acids obtained by DFT at the B3LYP/6-31G** level of theory [166].†

	Vibrational harmonic frequencies (cm^{-1})
Ala (A)	60, 222, 235, 268, 331, 417, 492, 560, 635^{N}, 733, 804, 858, 947, 1030, 1111, 1150, 1194^{N}, 1244, 1341^{N}, 1361, 1417, 1431, 1508, 1514, 1644^{c}, 1839, 2914, 3058, 3130, 3153, 3508, 3602^{c}, 3753^{N}
Cys (C)	53, 101, 178, 206, 249, 276, 338, 429, 485, 567, 639N, 691, 752, 789, 848, 873, 946, 1061, 1132, 1182N, 1220, 1259, 1283, 1348^{N}, 1376, 1417, 1472, 1635^{c}, 1846, 2676, 2904, 3101, 3168, 3522, 3619^{c}, 3751^{N}
Asp (D)	45, 70, 87, 154, 239, 291, 335, 365, 441, 488, 523, 596, 621N, 660, 679, 731, 860, 875, 906, 950, 1055, 1137, 1172, 1185^{N}, 1223, 1251, 1288, 1348^{N}, 1372, 1409, 1440, 1459, 1636^{c}, 1832, 1846, 2944, 3076, 3123, 3518, 3615^{c}, 3751, 3754^{N}
Glu (E)	20, 41, 65, 102, 112, 215, 240, 245, 309, 387, 432, 509, 575, 599^{N}, 630, 655, 717, 758, 800, 846, 893, 939, 968, 1024, 1059, 1124, 1144, 1174^{N}, 1220, 1256, 1281, 1306, 1343, 1366^{N}, 1383, 1395, 1424, 1492, 1510, 1683^{c}, 1840, 1850, 3038, 3051, 3061, 3122, 3164, 3490, 3568^{c}, 3745^{N}, 3753
Phe (F)	29, 54, 58, 94, 181, 241, 290, 295, 315, 357, 417, 476, 535, 592, 600, 624N, 636, 715, 759, 770, 799, 837, 860, 865, 924, 957, 972, 998, 1017, 1035, 1057, 1082, 1107, 1154^{N}, 1170, 1188, 1209, 1229, 1249, 1327, 1348, 1359^{N}, 1363, 1371, 1424, 1493, 1498, 1540, 1641, 1656^{c}, 1663, 1865, 3038, 3086, 3105, 3169, 3175, 3186, 3194, 3206, 3491, 3582^{c}, 3754^{N}
Gly (G)	66, 240, 290, 455, 544, 603^{N}, 707, 833, 853, 1024, 1088, 1144^{N}, 1237, 1301, 1366^{N}, 1436, 1505, 1662^{c}, 1857, 2932,3117, 3506, 3596^{c}, 3756^{N}
His (H)	27, 74, 80, 127, 219, 276, 309, 323, 343, 417, 505, 577, 632, 675^{N}, 721, 747, 757, 803, 844, 865, 903, 938, 979(2), 994, 1024, 1119, 1134, 1152, 1156, 1184^{N}, 1265, 1267, 1290, 1308, 1362^{N}, 1378, 1393, 1402, 1441, 1493, 1515, 1619, 1674^{c}, 1842, 2997, 3045, 3130, 3248, 3257, 3493, 3562, 3576^{c}, 3749^{N}

Table D.2: continued

Ile (I)	52, 59, 82, 149, 216, 224, 256, 277, 290, 314, 362, 405, 433, 539, 633^{N}, 660, 759, 790, 844, 856, 898, 968, 986, 998, 1049, 1077, 1125, 1152^{N}, 1158, 1188, 1236, 1278^{N}, 1307, 1334, 1373, 1378, 1389, 1425, 1427, 1435, 1502, 1511, 1514, 1523, 1525, 1675^{c}, 1831, 3010, 3023, 3040, 3048, 3075, 3095, 3109, 3116(2), 3128, 3482, 3566^{c}, 3734^{N}
Lys (K)	43, 61, 63, 86, 135, 150, 187, 227, 243, 297, 311, 375, 390, 462, 525, 614^{N}, 655, 745, 753, 782, 851, 875, 889, 951, 964, 1019, 1039, 1054, 1074, 1097, 1118, 1144N, 1174, 1247, 1250, 1286, 1306, 1334, 1347, 1350^{N}, 1373, 1398, 1402, 1411, 1425, 1496, 1501, 1509, 1525, 1673, 1682^{c}, 1838, 3003, 3018, 3028, 3032, 3038, 3058, 3074, 3080, 3100, 3475, 3486, 3559, 3565^{c}, 3745^{N}
Leu (L)	14, 39, 100, 139, 226, 239, 245, 259, 299, 341, 392, 426, 440, 533, 610, 646^{N}, 741, 807, 850, 871, 934, 945, 956, 970, 985, 1084, 1131, 1172, 1185^{N}, 1197, 1226, 1288, 1315, 1342, 1366^{N}, 1387, 1394, 1403, 1417, 1435, 1495, 1505, 1511, 1521, 1527, 1676^{c}, 1835, 3016, 3031, 3036, 3054, 3063, 3086, 3098, 3105, 3109, 3117, 3496, 3578^{c}, 3748^{N}
Met (M)	14, 38, 68, 100, 119, 178, 185, 252, 259, 292, 359, 406, 523, 588, 654N, 712, 742, 756, 788, 844, 928, 948, 975, 987, 1006, 1072, 1109, 1155, 1187^{N}, 1249, 1266, 1312, 1334, 1369^{N}, 1371, 1388, 1402, 1483, 1498, 1505, 1512, 1677^{c}, 1837, 3033, 3043, 3050, 3052, 3108, 3126, 3133, 3152, 3498, 3581^{c}, 3747^{N}
Asn (N)	14, 26, 93, 151, 206, 238, 289, 325, 377, 424, 502, 546, 581, 613^{N}, 642, 696, 756, 846, 877, 917, 920, 1043, 1085, 1122^{N}, 1159, 1214, 1259, 1291, 1356, 1366^{N}, 1398, 1441, 1472, 1626, 1668^{c}, 1786, 1848, 3038, 3080, 3095, 3493, 3583^{c}, 3604, 3745, 3754^{N}
Pro (P)	21, 56, 200, 269, 299, 418, 550, 582, 624, 639^{N}, 728, 800, 850, 890, 928, 932, 973, 991, 1084, 1100, 1129, 1174^{N}, 1196, 1224, 1264, 1276, 1318^{N}, 1333, 1349, 1385, 1420, 1442^{c}, 1500, 1518, 1538, 1848, 2939, 2965, 3071, 3079, 3084, 3122, 3143, 3563^{c}, 3752^{N}
Gln (Q)	27, 35, 70, 87, 120, 156, 222, 247, 268, 324, 383, 431, 529, 542, 592, 617^{N}, 629, 695, 766, 786, 829, 845, 900, 954,1048, 1067, 1098, 1102, 1151, 1195^{N}, 1250, 1266, 1311, 1340, 1367^{N}, 1373, 1414, 1431, 1482, 1506, 1625, 1657^{c},1802, 1852, 3030, 3051, 3072, 3077, 3113, 3502, 3596^{c}, 3605, 3747, 3755^{N}

Table D.2: continued

Arg (R)	25, 40, 52, 73, 94, 109, 171, 189, 208, 254, 317, 360, 415, 431, 442, 533, 561, 607, 641^{N}, 682, 706, 722, 812, 824, 878, 896, 919, 940, 1000, 1026, 1061, 1129, 1146, 1160, 1207, 1246, 1253, 1275, 1295^{N}, 1308, 1374^{N}, 1385, 1426, 1439, 1459, 1489, 1523, 1561, 1562, 1580, 1622, 1633, 1635, 1662
Ser (S)	36, 102, 196, 206, 266, 291, 349^{N}, 456, 521, 602, 624, 767, 795, 880, 979, 1032, 1098^{N}, 1110, 1148, 1227, 1258, 1343, 1379, 1395^{N}, 1406, 1436, 1522, 1671^{c}, 1842, 2967, 2980, 3132, 3493, 3578^{c}, 3754, 3807^{N}
Thr (T)	51, 96, 208, 233, 266, 312, 324, 364, 401, 476, 491, 546, 574, 642^{N}, 718, 838, 869, 907, 960, 1077, 1092, 1137, 1143^{N}, 1184, 1234, 1282, 1306, 1339^{N}, 1381, 1392, 1409, 1461, 1511, 1512, 1650^{c}, 1842, 2947, 2987, 3052, 3126, 3157, 3518, 3616^{c}, 3746^{N}, 3808
Val (V)	44, 67, 167, 210, 223, 251, 284, 339, 356, 387, 425, 529, 611^{N}, 631, 762, 818, 838, 898, 930, 953, 973, 1084, 1116^{N}, 1125, 1182^{N}, 1193, 1244, 1296, 1359, 1362, 1378, 1411, 1423, 1438, 1499, 1507, 1518, 1525, 1656^{c}, 1853, 3024, 3036, 3044, 3079, 3104, 3111, 3116, 3126, 3504, 3601^{c}, 3753^{N}
Trp (W)	24, 40, 50, 70, 172, 195, 226, 232, 290, 312, 328, 399, 404, 435, 464, 501, 552, 568, 587, 595, 629^{N}, 653, 732, 756(2), 772, 781, 801, 857, 861, 882, 889, 928, 944, 971, 1014, 1041, 1083, 1113, 1130, 1154, 1165^{N}, 1178, 1195, 1250, 1268, 1288, 1308, 1329^{N}, 1350, 1369, 1386, 1390, 1437, 1459, 1485, 1496, 1535, 1603, 1630, 1651^{c}, 1676, 1846, 2922, 3067, 3114, 3177, 3183, 3193, 3206, 3262, 3500, 3592^{c}, 3689, 3751^{N}
Tyr (Y)	11, 32, 70, 83, 157, 233, 264, 274, 322, 338, 352, 361, 423, 430, 504, 534, 567, 570, 634^{N}, 657, 723, 766, 787, 794, 818, 844, 863, 906, 942, 959, 966, 1024, 1033, 1087, 1126, 1163^{N}, 1198, 1202, 1225, 1237^{N}, 1253, 1308, 1321,1355, 1367, 1377, 1394, 1425, 1481, 1499, 1562, 1645, 1656^{c}, 1676, 1841, 2985, 3033, 3120, 3164, 3170, 3206, 3212, 3487, 3574^{c}, 3746^{N}, 3823

† Superscripts C and N denote frequencies to be deleted for residues at the C and N termini, respectively.