

Molecular Recognition of Biomolecules in the Gas Phase

Thesis by

Ryan R. Julian

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2003

(Defended April 28, 2003)

© 2003

Ryan R. Julian

All Rights Reserved

to Jimbo
friend and toolmaster extraordinaire

Acknowledgements

I would like to acknowledge the many people who have contributed, in one way or another, to the completion of this thesis. First, there are the scientists. Mr. Willey and Mrs. Rocha were instrumental as high school teachers that encouraged a love of science. At the University of Utah, excellent courses taught by Bill Epstein and Bob Parry convinced me not to be a biologist. Bill Breckenridge was instrumental in teaching me the ways of physical chemistry and insuring that I didn't end up in law school. It was also in his lab where I got my first real taste of research and finally found something that I really enjoyed. Bill (Breckenridge) also deserves a lot of the credit for helping me to arrive at such a fine institution like Caltech for grad school.

Here at Caltech, I owe many thanks to Bill Goddard for teaching me a lot about chemistry and how to think about molecules and their properties. Collaborating with him in several projects utilizing theory has been very valuable. Jeremy May and Brian Stoltz have been priceless collaborators as well, teaching me about organic chemistry and making marvelous molecules for me to experiment with. A lot of the work in this thesis would not have been possible without their help. I would like to thank all of the Beauchamp group members (Heather Cox, Rob Hodyss, Ron Grimm, Daniel Austin, Jim Smith, Dmitri Kossakovsky, Sang-Won Lee) for putting up with me, working with me, and for providing friendship and some decent games of croquet.

I owe many thanks to Jack Beauchamp, my advisor, for being the largest influence in my development as a scientist and for putting up with my daily visits to his office. Jack has given me enough room to grow on my own while still providing guidance and helping me to think about the most important aspects of a problem. I am indebted to Jack

further for consistently scratching out half of what I write, because it has helped me develop better writing skills (at least when writing papers).

Finally, this list would be incomplete (and probably still is anyway) without acknowledging my parents, who have always encouraged me to achieve great heights. And the biggest thanks goes to my lovely wife, who provides me with a constant foundation of love and support, and who always makes sure to remind me that I am not half as smart as I think I am.

Abstract

The first chapter introduces the most relevant noncovalent forces for gas phase experiments. Chapters 2-5 contain work on small clusters of biologically relevant molecules. In Chapter 2, it is shown that the unusual properties of arginine lead to extensive noncovalent clustering of this amino acid, when sampled by ESI-MS. The stability of the zwitterionic form of arginine for clusters without a net charge is addressed by theoretical methods in Chapter 3.

In Chapter 4, the properties of the unusually abundant serine octamer are examined. Experiments demonstrate that this octamer has a strong preference to be homochiral. A structure for the serine octamer is proposed that is cubic and has a zwitterionic core. The results gathered from the serine octamer demonstrate that a homochiral preference can exist for very small clusters or “nanocrystals.”

The first gas phase synthesis for ATP is given in Chapter 5. ATP is easily synthesized in the gas phase from a cluster of three AMP molecules bound by a sodium salt bridge. Subsequent CAD spectra following the gas phase synthesis are identical to those obtained from an authentic sample of ATP in separate experiments.

Chapters 7-9 deal with the molecular recognition of amino acid side chains in ESI-MS experiments. The ability of 18C6 to recognize and selectively attach to lysine residues is explored. Recognition of arginine side chains is accomplished in a similar manner by utilizing the larger dibenzo-30-crown-10 ether (DB30C10). The two techniques are mutually compatible, allowing for both crowns to be added to the same solution.

Chapters 10-11 combine the recognition of 18C6 with various chemical functionalities in order to mediate peptide chemistry in the gas phase. In Chapter 10, a new class of

molecules termed “molecular mousetraps” is described. The mousetraps combine the recognition of 18C6 with the chemical reactivity of diazo groups. The resulting molecules are capable of noncovalently attaching to any molecule that contains a protonated primary amine. CAD can be utilized to activate the complex. In Chapter 11, the mousetraps are utilized in experiments with peptides. It is shown that covalent attachment can be achieved in a quantitative fashion.

Table of Contents

Chapter 1 All Things Noncovalent	1
Chapter 2 Salt Bridge Stabilization of Charged Zwitterionic Arginine Aggregates in the Gas Phase	16
Chapter 3 Cooperative Salt Bridge Stabilization of Gas Phase Zwitterions in Neutral Arginine Clusters	43
Chapter 4 Nanocrystalline Aggregation of Serine Detected by Electrospray Ionization Mass Spectrometry: Origin of the Stable Homochiral Gas Phase Serine Octamer	56
Chapter 5 Abiotic Synthesis of ATP from AMP in the Gas Phase: Implications for the Origin of Biologically Important Molecules from Small Molecular Clusters	91
Chapter 6 Gas Phase Synthesis of Charged Copper and Silver Fischer Carbenes from Diazomalonates: Mechanistic and Conformational Considerations in Metal Mediated Wolff Rearrangements	115
Chapter 7 Site Specific Sequestering and Stabilization of Charge in Peptides by Supramolecular Adduct Formation with 18-Crown-6 Ether via Electrospray Ionization	146
Chapter 8 Molecular Recognition of Arginine in Small Peptides by Supramolecular Complexation with Dibenzo-30-Crown-10 Ether	173
Chapter 9 The Unusually High Proton Affinity of Aza-18-Crown-6 Ether: Implications for the Molecular Recognition of Lysine in Peptides by Lariat Crown Ethers	193

Chapter 10	212
Molecular Mousetraps: Gas Phase Studies of the Covalent Coupling of Noncovalent Complexes Initiated by Reactive Carbenes Formed by Controlled Activation of Diazo Precursors	
Chapter 11	223
Biomimetic Approaches to Gas Phase Peptide Chemistry: Combining Selective Binding Motifs with Reactive Carbene Precursors to Form Molecular Mousetraps	

List of Figures

Chapter 1

Figure 1.1	2
Figure 1.2	3
Figure 1.3	6
Figure 1.4	6
Figure 1.5	7
Figure 1.6	10
Figure 1.7	11
Figure 1.8	12

Chapter 2

Figure 2.1	22
Figure 2.2	24
Figure 2.3	26
Figure 2.4	29
Figure 2.5	34

Chapter 3

Figure 3.1	48
------------	----

Chapter 4

Figure 4.1	63
Figure 4.2	63
Figure 4.3	65
Figure 4.4	67
Figure 4.5	69
Figure 4.6	71
Figure 4.7	73
Figure 4.8	75
Figure 4.9	76

Chapter 5

Figure 5.1	96
Figure 5.2	98
Figure 5.3	100
Figure 5.4	102
Figure 5.5	104
Figure 5.6	106

Chapter 6

Figure 6.1	123
Figure 6.2	126
Figure 6.3	129

Chapter 7

Figure 7.1	152
Figure 7.2	156
Figure 7.3	158
Figure 7.4	161
Figure 7.5	165
Figure 7.6	168

Chapter 8

Figure 8.1	179
Figure 8.2	180
Figure 8.3	182
Figure 8.4	183
Figure 8.5	186
Figure 8.6	189

Chapter 9

Figure 9.1	198
Figure 9.2	200
Figure 9.3	200
Figure 9.4	204
Figure 9.5	206
Figure 9.6	208

Chapter 10

Figure 10.1	217
Figure 10.2	219

Chapter 11

Figure 11.1	231
Figure 11.2	233
Figure 11.3	236
Figure 11.4	239
Figure 11.5	241
Figure 11.6	243
Figure 11.7	245

List of Tables

Chapter 2

Table 2.1	18
Table 2.2	31

Chapter 3

Table 3.1	49
Table 3.2	49

Chapter 4

Table 4.1	79
Table 4.2	81

Chapter 5

Table 5.1	107
-----------	-----

Chapter 6

Table 6.1	132
Table 6.2	137
Table 6.3	141

Chapter 7

Table 7.1	147
Table 7.2	154
Table 7.3	163
Table 7.4	163

Chapter 1

All Things Noncovalent

1.1 Introduction

Noncovalent interactions are an extremely important subset of the chemical processes occurring in the world. In biology, selectivity and recognition are achieved primarily through noncovalent contacts. In chemistry, noncovalent interactions influence virtually every chemical reaction and the design of building blocks held together by noncovalent bonds comprises an entire field, known as supramolecular chemistry.^{1,2} In both chemistry and biology, the majority of work on this subject to date has focused on interactions in solution, where the results are strongly influenced by the presence of a solvent. However, more recently the study of supramolecular chemistry has taken to the gas phase.³ Similarly, the study of biological systems is increasingly performed with the aid of gas phase techniques.⁴ The results presented in this thesis expand upon the gas phase study of noncovalent interactions, and emphasis is placed on utilizing molecules and principles from supramolecular chemistry to study biological molecules.

Electrospray ionization mass spectrometry (ESI-MS) is the primary experimental method used throughout this work.⁵ With this technique, molecules of low volatility can be introduced into the gas phase from a solution of water and methanol. Ions are typically produced by the addition or removal of a proton. This method of ionization is gentle, allowing for the gas phase observation of noncovalent complexes formed originally in

solution.⁶ Various theoretical methods can be applied to quantitatively assess the energetics of these complexes in the gas phase. Density functional theory and semi-empirical calculations are used extensively in combination with molecular dynamics to evaluate the energetics of and determine structures for the noncovalent complexes studied in this thesis.

Background. All noncovalent interactions were *not* created equally; therefore, we will briefly review those that are most important to the present work. Coulombic interactions are very important in the gas phase because they are both strong and long range. The $1/r$ dependence for coulombic interactions allows them to operate over long distances in the gas phase. In solution on the other hand, solvents with high dielectric constants (particularly water) will substantially mediate the importance of coulombic interactions. Salt bridges are an excellent example of a coulombic interaction that serves to elegantly demonstrate this difference.

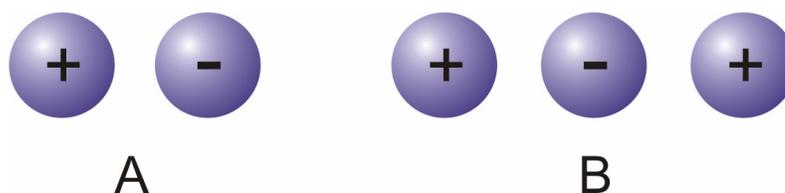


Figure 1.1 Typical salt bridges in proteins (A) and in the gas phase (B).

Minimally, a salt bridge is defined as a favorable interaction between charged functional groups (as shown generically in Figure 1.1). Biologists usually refer to the attractive interaction between (for example) protonated lysine and deprotonated aspartic acid as a salt bridge (see Figure 1.1A). By contrast, in the gas phase a salt bridge will

usually consist of three charged groups, two positive and one negative or vice versa, so that overall there is a net charge allowing the ion to be easily manipulated by electric and magnetic fields. Most of the salt bridges in the present work will be similar to that shown in Figure 1.1B, which can also be thought of as a charge stabilized ion pair.

A straightforward example is given in structure **1.1**. The binding energy of an iodide anion to the doubly charged crown ether is calculated to be ~ 150 kcal/mol in the gas phase (at the PM5 semi-empirical level of theory). This value is larger than most covalent bond strengths, demonstrating the strength of coulombic interactions in the absence of solvent mediation. This high binding energy is confirmed by collisional activation of **1.1**, which leads exclusively to the breaking of covalent bonds and the loss of methyl iodide (see Figure 1.2). However, **1.1** is also an unusual system because there are no labile protons present that could disrupt the salt bridge. Typically in biological systems there will be labile protons present that can neutralize charges by transferring from a protonated basic site to a deprotonated acidic site. For an isolated acid/base pair, charge separation is typically not favored over proton transfer in the gas phase, but many factors can easily stabilize charge separation as explained in Chapter 3.

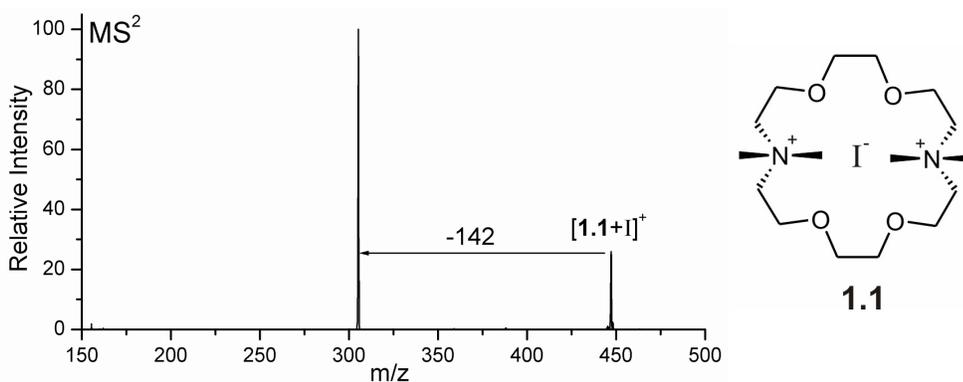


Figure 1.2 CAD of $[1.1+I]^+$ leads exclusively to the loss of MeI (mass 142Da).

Conversely, the stability of charge separation in solution is greatly enhanced, but the binding strength of the interaction is greatly reduced. As a result the removal of iodide anion from **1.1** in aqueous solution does not result in any covalent bond cleavage. It is for this same reason that the role salt bridges play in stabilizing protein structure in solution remains hotly debated.^{7,8} Stabilization free energies vary widely, but even the highest estimates are not more than 5-10 kcal/mol for a single salt bridge in aqueous solution.⁹

Similar arguments can be made for other coulombic interactions such as ion-dipole and dipole-dipole noncovalent bonds. A special case that requires further comment is the hydrogen bond. Hydrogen bond strengths are much higher in the gas phase than in water or other polar solvents. Typical values for hydrogen bonds in water range from 2-10 kcal/mol, whereas hydrogen bond strengths up to 20 kcal/mol are possible in the gas phase.¹⁰ Hydrogen bonds to a charged donor or acceptor can lead to hydrogen bonds up to 45 kcal/mol in the gas phase.¹¹ Therefore, hydrogen bonds are potentially much stronger and more important in the gas phase than in aqueous solution. Furthermore, the combination of just a few hydrogen bonds in the gas phase can easily equal the energy of a typical covalent bond (~85 kcal/mol) under the right circumstances. It should also be noted that all coulombic interactions are directional, with the strongest forces being achieved by the most linear arrangement of charges or partial charges. This directionality can be particularly important for hydrogen bonds.

Finally, another important though poorly understood solution phase interaction that is very relevant to the observation of ions by electrospray ionization mass spectrometry (ESI-MS) is the hydrophobic effect.^{12,13} This is the driving force behind the aggregation of nonpolar molecules in aqueous solution, which is thought to be very important in

protein folding.¹⁴ The strength of this interaction is difficult to define and certainly weaker in the gas phase than it is in solution. This is primarily due to the fact that, formally, there is no “hydrophobic” effect when there is no water. Only weakly binding Van der Waals forces remain after desolvation and introduction into the gas phase. Therefore, complexes that are held together in solution through largely hydrophobic interactions are not likely to be observed in the gas phase by ESI or matrix assisted laser desorption ionization (MALDI) experiments. However, it should be mentioned that the higher the relative hydrophobicity of a charged molecule, the more abundant it will appear in an ESI mass spectrum. This interesting phenomenon will be explained in greater detail in Chapter 7.

1.2 Content of Thesis

Clusters of Biomolecules. Chapters 2-5 contain work on small clusters of biologically relevant molecules. In Chapter 2, it is shown that the unusual properties of arginine lead to extensive noncovalent clustering of this amino acid, when sampled by ESI-MS as shown in Figure 1.3. The clusters can be formed as cations or anions, with a variety of different molecules serving as the charge carrier. Of particular interest are a series of anionic trimers which demonstrate unusual abundance (see Figure 1.4). A structure (**1.2**) in which each arginine interacts in a head-to-tail arrangement while maintaining an intramolecular bond is proposed to explain the unusual abundance of these clusters. Each arginine in the trimer is in the zwitterionic form. The stability of the zwitterionic form of arginine for clusters without a net charge is addressed further by theoretical methods in Chapter 3.

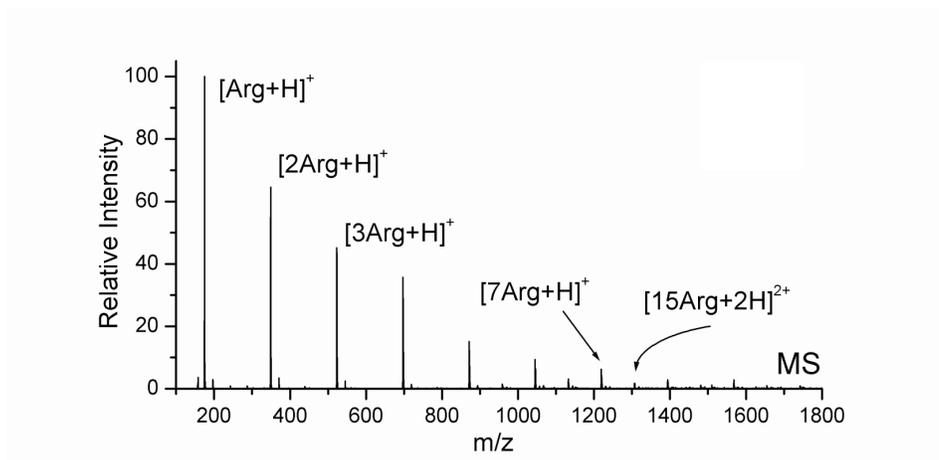


Figure 1.3 Cationic clusters of arginine.

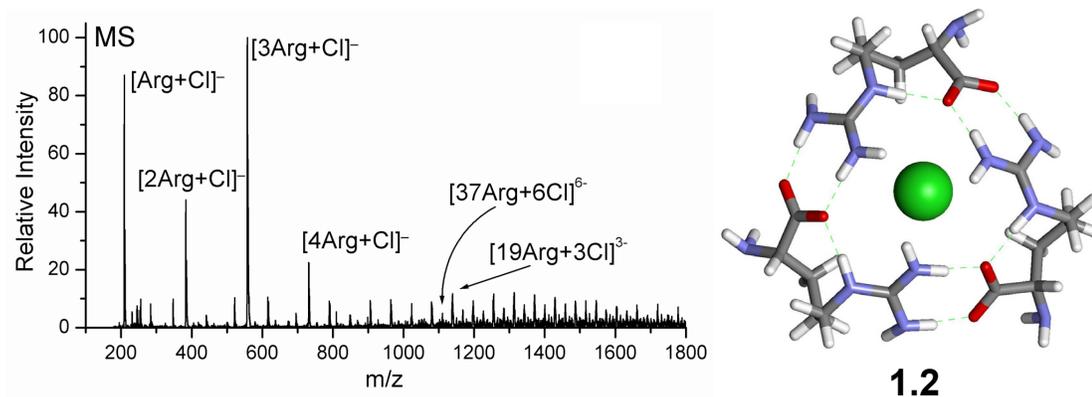


Figure 1.4 Anionic clusters of arginine with chloride.

This work on arginine clusters allows us to draw several important conclusions. The significance of the strength and specificity that can be obtained by the interaction between alkyl-guanidiniums and carboxylates is clearly demonstrated. These interactions have been observed in many crystal structures.^{15,16,17} The cluster work presented in Chapter 2 establishes the importance of these interactions in the gas phase as well. There are several implications for gas phase protein structure as a result of this observation. Salt bridges between arginine and either aspartic or glutamic acid are likely to remain as charge separated salt bridges in the gas phase. Furthermore, this class of salt bridge is

predicted to be more specific and strongly bound in the gas phase than salt bridges involving lysine residues. These factors should be taken into account when molecular modeling is utilized to examine the gas phase structures of proteins.

In Chapter 4, the properties of another unusually abundant cluster are examined. The ESI-MS spectrum of a 0.01M solution of serine reveals an unusually abundant protonated octamer. Further experiments employing isotopic labeling demonstrate that this octamer has a strong preference to be homochiral. These startling observations have attracted the attention of several groups,¹⁸ each with different structures and explanations for the observed characteristics of the serine octamer. Utilizing the hierarchy of interactions outlined in the introduction above, we constructed a serine octamer that maximized favorable coulombic interactions and hydrogen bonds. The resulting structure is cubic and has a zwitterionic core as shown in Figure 1.5. A recent review on the subject critically compared the energetics of the structures from each group and found ours to be the lowest energy conformation by a significant margin.¹⁹

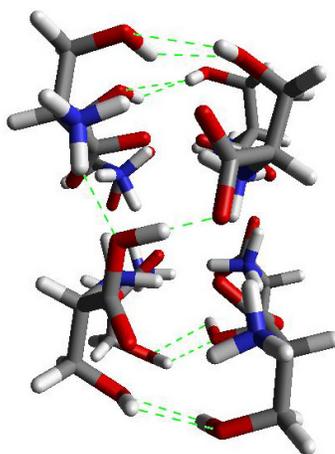


Figure 1.5 Structure for the homochiral protonated serine octamer.

The results gathered from the serine octamer demonstrate that a homochiral preference can exist for very small clusters or “nanocrystals.” Furthermore, ESI may offer a new

experimental technique for investigating the early stages of homogeneous crystal nucleation in solution. A universal theory for the explanation of homogeneous nucleation is still lacking.²⁰ Several spectroscopic methods have been employed to study crystal nucleation,²¹ but mass spectrometry offers the additional ability to sample small clusters and study them in the gas phase. Moreover, methods for symmetry breaking in racemic mixtures are very important to understanding the origin of life.²² The work in Chapter 4 clearly demonstrates that symmetry breaking can be achieved in small molecular clusters. If some means of activation that leads to polymerization of the cluster components can be achieved, then a route to the generation of homochiral polymers is possible. While this may sound rather outlandish, polymerization reactions for biological molecules from small molecular clusters are reported in Chapter 5.

The first gas phase synthesis for ATP (adenosine triphosphate) is given in Chapter 5. This extremely important molecule is easily synthesized in the gas phase from a cluster of three AMP (adenosine monophosphate) molecules bound together by a sodium salt bridge. Subsequent collision activated dissociation (CAD) spectra following the gas phase synthesis are identical to those obtained from an authentic sample of ATP in separate experiments. It is further shown that similar chemistry is possible with phosphate itself, allowing for the gas phase generation of polyphosphate (which is another biologically important molecule).²³

Natural processes such as sea spray could theoretically lead to the occurrence of such desolvated clusters in the Earth's atmosphere.²⁴ The sun provides an ample supply of energy, which could initiate the chemical reactions that lead to polymerization. Given that ATP can be synthesized by such a process and that the serine octamer provides

evidence for symmetry breaking, the chemistry of small clusters of biological molecules and their precursors probably deserves more attention.

Highly Reactive Chemistry in Clusters. Chapter 6 studies the fundamental properties of the chemistry of carbenes with emphasis on the important Wolff rearrangement. Metal ion coordination is utilized to facilitate both the generation of carbenes from diazo malonate precursors and the subsequent multiple Wolff rearrangements that follow. Isotopic labeling is employed to determine the mechanisms for the various reactions and confirm that rearrangement does not proceed through an oxirene intermediate. The influences that conformation and metal ion coordination have on the Wolff rearrangement are studied experimentally and theoretically. Reactions between these carbene species and various noncovalent adducts are also examined. The end result is a more detailed understanding on the fundamental aspects of a process which is very important to synthetic organic chemistry.

Molecular Recognition in Biological Systems. Chapters 7-9 deal with the molecular recognition of amino acid side chains through noncovalent attachment in ESI-MS experiments. In Chapter 7, the ability of 18-crown-6 ether (18C6) to recognize and selectively attach to lysine residues is explored. It is found that the number of lysines can be quantified for small peptides. For proteins, the number of 18C6 ethers that attach is related to the structure of the protein in solution, with more crowns attaching to unfolded proteins. Furthermore, 18C6 is shown to enhance the ESI signal for the ion to which it attaches by effectively desolvating the charge and increasing the surface activity of the ion on the highly charged electrospray droplet. This leads to the observation of $[\text{KKKK}+4(18\text{C}6)+4\text{H}]^{4+}$ as the base peak in the spectrum from a solution containing a

1:1 mixture of KKKK and 18C6. This structure is shown in Figure 1.6. Clearly, the solution phase compositions are not accurately reflected in the ESI-MS data for these systems due to preferential sampling of ions that are coordinated to 18C6.

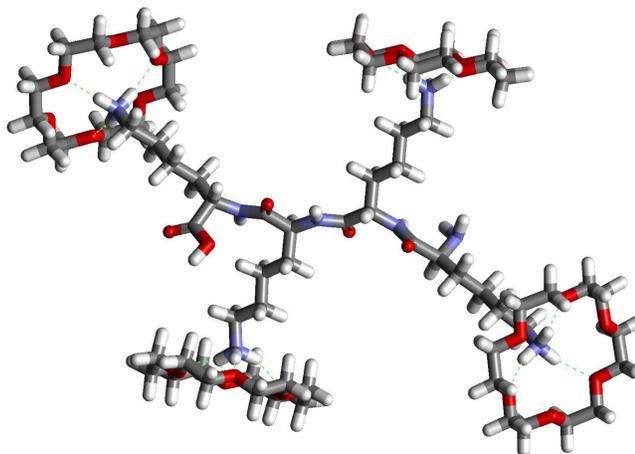


Figure 1.6 Tetralysine with 4 18C6 ethers attached.

In Chapter 8, recognition of arginine side chains is accomplished in a similar manner by utilizing the larger dibenzo-30-crown-10 ether (DB30C10) as shown in Figure 1.7. This crown preferentially recognizes the side chain of arginine, but does not form highly abundant adducts like 18C6 does with lysine. The reason for this is unclear, but the net result is that only one arginine will be reliably identified on a peptide that may contain multiple arginines. It is shown through competitive CAD experiments that the larger crown has a higher binding energy to arginine than 18C6 does to lysine. Furthermore, the techniques in Chapter 7 and 8 are mutually compatible, allowing for both crowns to be added to the same solution. This should allow for the easy separation of a tryptic digest into the lysine and arginine containing fragments without *a priori* knowledge of the sequences of the peptides. Such a technique may be useful in confirming the identity of a

peptide which had undergone post-translational modification and appeared at a mass other than the expected value.

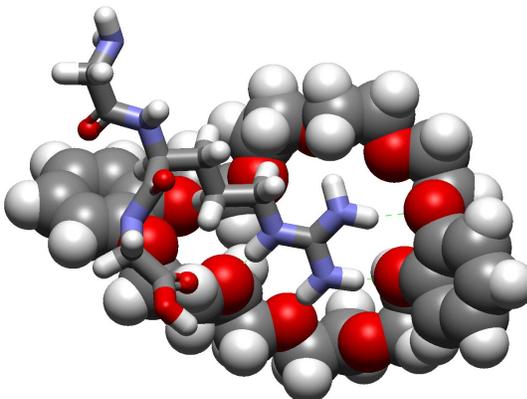


Figure 1.7 Interaction between dibenzo-30-crown-10 and the peptide GRG.

In Chapter 9, the most effective method for adding lariats to 18C6 without disrupting its excellent recognition abilities are systematically explored for reasons that will become obvious in Chapters 10 and 11. The simplest way to obtain a lariat crown ether is to begin with aza-18C6 and attach a functional group to the nitrogen heteroatom. Unfortunately, the high proton affinity of the secondary amine destroys the recognition ability of crown for gas phase experiments. It is shown that conversion of the amine to an amide reduces the effect, but the overall binding energy is still substantially lower when compared to 18C6. Interestingly, aza-18C6 is shown to have a proton affinity ~ 10 kcal/mol higher than any other secondary amine in the NIST database.²⁵ This unusually high proton affinity is due to intramolecular hydrogen bonds within the crown. With regards to the addition of lariat side chains, the results indicate that the side chain must branch off of one of the carbons in 18C6. This hypothesis was confirmed in later experiments as shown below.

De Novo Biomimetic Reagents. Chapters 10 and 11 combine the recognition of 18C6 with various chemical functionalities in order to mediate peptide chemistry in the gas phase. In Chapter 10, a new class of molecules termed “molecular mousetraps” is described (see Figure 1.8). The mousetraps combine the recognition of 18C6 with the chemical reactivity of diazo groups. The resulting molecules are capable of noncovalently attaching to any molecule that contains a protonated primary amine. In the gas phase, CAD can then be utilized to activate the complex, which results in the formation of a highly reactive carbene which then preferentially inserts intermolecularly. The noncovalent complex is transformed into a covalently bound molecule by this process. Importantly, this is an example of a system where collisional activation of a noncovalent complex results in a chemical reaction rather than simple dissociation, which is typically the dominant process in the vast majority of systems.

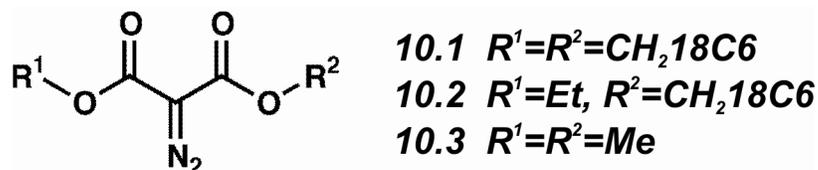


Figure 1.8 Structures of the “molecular mousetraps” in Chapter 10.

In Chapter 11, the mousetraps are utilized in experiments with peptides. It is shown that covalent attachment can be achieved in a quantitative fashion. In addition, the results for two other reagents designed to initiate peptide backbone cleavage are given. Although a directed cleavage process was not observed, progress towards the design of such a reagent was achieved. Primarily, these experiments revealed that any successful reagent must be designed with the proper combination of high binding energy and highly reactive

chemical functionalities. High binding energy can be achieved by using two 18C6 ethers to attach to a peptide containing two lysine residues. Chemical reactions with activation barriers similar to the reaction barrier for converting a diazo into a carbene appear to be optimal. These complexes are stable until triggered by CAD. If the activation energy is too low, then the process cannot be controlled and may occur prior to detection in the mass spectrometer. These initial experiments demonstrate that the de novo design of reagents capable of mediating peptide chemistry in the gas phase is possible and worth exploring further.

1.3 Summary

In conclusion, the study of noncovalent interactions can reveal fundamental information about the chemistry that is happening in the world around us. Noncovalent complexation can reveal properties of biological molecules and offers a glimpse into the possible origin of such molecules. The combination of molecular recognition with additional chemical functionalities is a promising area for the development of de novo reagents that operate in the gas phase.

1.4 References

-
- ¹ *Comprehensive Supramolecular Chemistry*, vol. 1 (Ed.:G. W. Gokel), Pergamon/Elsevier: Oxford, 1996.
- ² *Supramolecular Chemistry* Beer, P. D.; Gale, P. A.; Smith, D. K. Oxford University Press, New York, 1999.
- ³ Schalley, C. A. *Int. J. Mass Spec.* **2000**, *194*, 11-39.
- ⁴ Hoaglund-Hyzer, C.S.; Counterman, A.E.; Clemmer, D.E. *Chem. Rev.* **1999**, *99*, 3037-3079.
- ⁵ Fenn, J.; Rosell, J.; Nohmi, T.; Shen, S.; Banks, F. *Biochemical and Biotechnological Applications of Electrospray Ionization Mass Spectrometry* **1996**, *619*, 60-80.
- ⁶ (a) Veenstra, T. D. *Biophys. Chem.* **1999**, *79*, 63-79. (b) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.
- ⁷ Kumar, S.; Nussinov, R. *J. Mol. Biol.* **1999**, *293*, 1241-1255.
- ⁸ Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3*, 211-226.
- ⁹ Lebbink, J. H. G.; Knapp, S.; van der Oost, J.; Rice, D.; Ladenstein, R.; de Vos, W. M. *J. Mol. Biol.* **1998**, *280*, 287-296.
- ¹⁰ Prins, L. J.; Reinhoudt, D. N.; Timmerman, P. *Angew. Chem. Int. Ed.* **2001**, *40*, 2382-2426.
- ¹¹ *Hydrogen Bonding: A Theoretical Perspective* (Ed. S. Scheiner). Oxford University Press, New York, **1997**.
- ¹² Marmur, A. *J. Am. Chem. Soc.* **2000**, *122*, 2120-2121.
- ¹³ Silverstein, K. A. T.; Haymet, A. D. J.; Dill, K. A. *J. Am. Chem. Soc.* **1998**, *120*, 3166-3175.

-
- ¹⁴ (a) Baldwin, R.L. *Science* **2002**, *295*, 1657-1658. (b) Fernandez, A.; Kardos, J.; Goto, Y. *FEBS Letters* **2003**, *536* (1-3), 187-192 and references therein.
- ¹⁵ Karle, J.; Karle, I.L. *Acta Cryst.* **1964**, *17*, 835-841.
- ¹⁶ Bhat, T.N.; Vijayan, M. *Acta Cryst.* **1977**, *B33*, 1754-1759.
- ¹⁷ Chapiro, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J. *J. Am. Chem. Soc.* **1998**, *120*, 12956-12957.
- ¹⁸ The full story is detailed in Chapter 4, including the contributions of other investigators.
- ¹⁹ Schalley, C.A.; Weis, P. *Int. J. Mass Spectrom.* **2002**, *221*, 9-19.
- ²⁰ Granasy, L.; Igloi, F. *J. Chem. Phys.* **1997**, *107*, 3634-3644.
- ²¹ (a) Granasy, L.; James, P. F. *J. Chem. Phys.* **2000**, *113*, 9810-9821, (b) Peng, X.; Wickham, J.; Alivisatos, A. P. *J. Am. Chem. Soc.* **1998**, *120*, 5343-5344.
- ²² (a) Podlech, J. *Cell. Mol. Life Sci.* **2001**, *58*, 44-60. (b) Editor, David B. Cline *Physical origin of homochirality in life : Santa Monica, California, February 1995* Woodbury, New York : American Institute of Physics, **1996**.
- ²³ Kornberg, A.; Rao, N. N.; Ault-Riche, D. *Annu. Rev. Biochem.* **1999**, *68*, 89-125.
- ²⁴ (a) Ellison, G.B.; Tuck, A.F.; Vaida, V. , *J. Geophys. Res.* **1999**, *104*, 11633-11642. (b) Tuck, A. *Surv. Geophys.* **2002**, *23*, 379-409. (c) Gill, P.S.; Graedel, T.E.; Weschler, C.G. *Rev. Geophys.* **1983**, *21*, 903-920.
- ²⁵ E.P. Hunter and S.G. Lias, "Proton Affinity Evaluation" in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. W.G. Mallard and P.J. Linstrom, February 2000, National Institute of Standards and Technology, Gaithersburg MD, 20899 (<http://webbook.nist.gov>)

Chapter 2

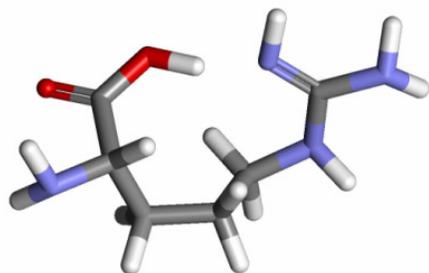
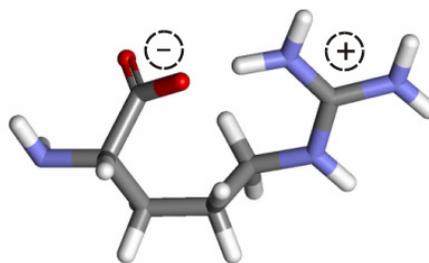
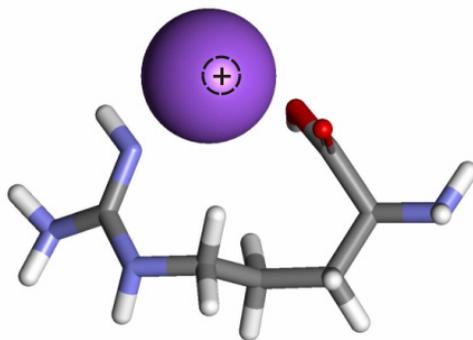
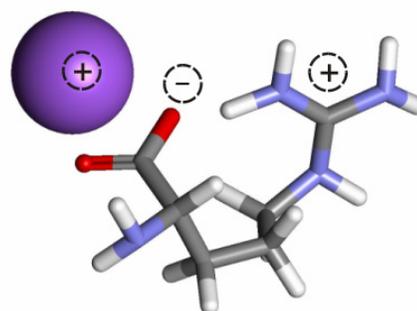
Salt Bridge Stabilization of Charged Zwitterionic

Arginine Aggregates in the Gas Phase

Portions previously published in: Julian R. R.; Hodyss R.; Beauchamp J. L. *J. Am. Chem. Soc.* **2001**, *123*, 3577-3583.

2.1 Introduction

Amino acids are known to exist as zwitterions in solution, but conditions appropriate for stabilization of the zwitterionic form in the gas phase remain debatable. Studies of the gas phase acidity and basicity of glycine with Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry have shown that the glycine zwitterion is unstable by ~ 84 kJ/mol.¹ Many recent calculations confirm that glycine is unlikely to exist in the gas phase as a zwitterion;² furthermore, the zwitterion is not even predicted to be a minima on the potential energy surface. Theory suggests that water molecules can stabilize the zwitterionic form of glycine in the gas phase, with recent calculations suggesting that two water molecules are required.³ In contrast, recent experimental results suggest that the number of water molecules necessary to stabilize glycine as a zwitterion is five.⁴

Arginine **2.1**Zwitterionic Arginine **2.2**Charge Solvated **2.3**Zwitterionic Salt Bridge **2.4**

The gas phase stability of the zwitterionic form of arginine is predicted to be much more favorable. The arginine zwitterion is created by transferring a proton from the C-terminus to the side chain (see **2.1** and **2.2**). The nomenclature indicated for arginine structures **2.1-2.4** will be used in the present work. Arginine is the most basic amino acid (see Table 2.1). This high basicity increases the stability of zwitterionic arginine in the gas phase relative to other amino acids. However, recent experiments have shown that isolated arginine is not a zwitterion. Cavity-ring down laser absorption spectra of jet-cooled arginine do not exhibit a peak corresponding to the calculated carboxylate asymmetric stretch of the zwitterion.⁵ High level computations have confirmed that arginine (**2.1**) is more stable than zwitterionic arginine (**2.2**) in the absence of an

additional charge.⁶ However, theory predicts that the zwitterionic arginine is only less stable than arginine by 4-12 kJ/mol, depending on the level of theory.⁶

Table 2.1 Relative gas phase acidities and basicities.⁷

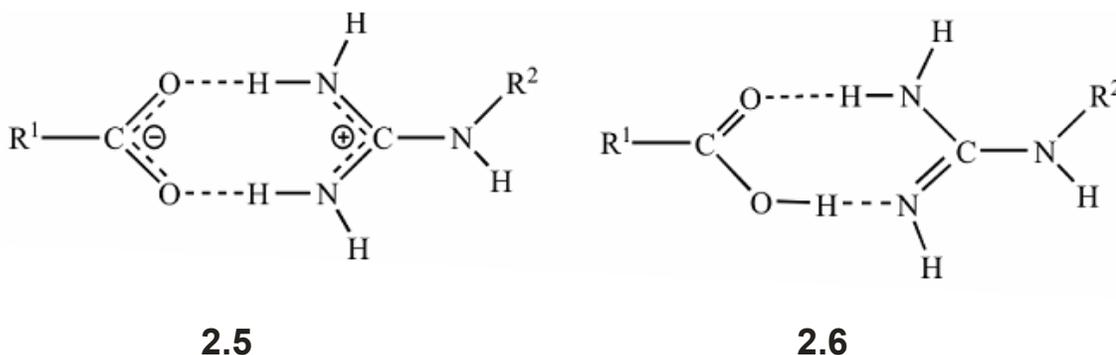
Amino Acid	Gas Phase Acidity (kJ/mol)	Gas Phase Basicity(GB) (kJ/mol)
Glycine	1402(avg.)	852
Lysine	1383	951
Arginine	1360	1007
Histidine	1356	950
HF	1531	--
HCl	1373	--
HBr	1332	--
HI	1294	--
HNO ₃	1330	--

The stability of the zwitterionic forms of amino acids can be substantially modified by specific noncovalent interactions with nearby molecules and ions.⁸ For example, calculations show that copper ions can stabilize zwitterionic glycine.⁹ The addition of a cation to arginine is likely to lead to one of the two structures depicted as **2.3** and **2.4**. Collision induced dissociation spectra obtained with a FT-ICR mass spectrometer suggest that zwitterionic arginine is stabilized by metal cations (K⁺, Rb⁺, and Cs⁺).¹⁰ Related results based on the dissociation of heterodimers also favored structure **2.4** (for K⁺, and Cs⁺).¹¹ Experiments suggest that the protonated arginine dimer exists with one arginine in the zwitterionic state.^{12,13} Ion mobility data for complexes of arginine with Na⁺, Cs⁺, and H⁺ could not discriminate between **2.3** and **2.4** in the case of the metal ions.¹⁴ However, calculations performed in conjunction with these experiments favored the zwitterionic salt-bridge **2.4**.¹⁴

Arginine participates in a variety of specific noncovalent interactions involving ionized functional groups that are easily observed in crystal structures. Crystalline

arginine itself is different from all other amino acids. Typically, amino acids arrange themselves in a peptide-like fashion with the n-terminal amino and c-terminal carboxylate groups aligned with the side chains protruding outward on alternating sides.¹⁵ In contrast, arginine stacks end to end enabling the guanidinium and carboxylate interaction shown in **2.5**.¹⁶ This motif is common in crystal structures containing arginine as illustrated by arginine with acetate⁵ and in the dipeptide Arg-Glu.¹⁷

The stability of **2.5** with various groups R^1 and R^2 in the gas phase has been addressed by several computational approaches. Calculations have been performed on model



systems at various levels of theory, including semi-empirical calculations with PM3 and AM1, density functional theory, and ab initio methods.^{18,19,20} The energetics of **2.5** and **2.6** (where $R^1 = R^2 = \text{CH}_3$) were calculated at the RHF/6-31G** and MP2/6-31G** ab-initio level.¹⁸ Semi-empirical results obtained with AM1 qualitatively reproduced the ab-initio results and both levels of theory predict that **2.5** is not likely to exist in the gas phase, with **2.6** being energetically favored by ~18 kJ/mol at the highest level of theory.

The discovery of several new unusually stable aggregates of arginine that are intermolecularly bound by salt bridges is reported here. Quadrupole ion-trap mass spectrometry provides further evidence for the stability of arginine in the zwitterionic state, coordinated to either a cation or anion. Clusters of arginine with sodium,

potassium, lithium, magnesium, chloride, fluoride, bromide, iodide, and nitrate ions are reported. Computational results at the DFT level are used to assign structures to and assess the energetics of particularly prominent clusters. A new method is introduced to investigate the effect of arginine enantiomeric purity on the stability of several of the more prominent clusters reported in this study. Our results are compared and contrasted to the recent report by Cooks and co-workers²¹ of protonated clusters of arginine alone.

2.2 Experimental Methodology

All spectra were obtained using a Finnigan LCQ ion trap quadrupole mass spectrometer without modification. The signal was tuned using the automatic tuning capabilities of the LCQ on the trimer of the protonated arginine clusters. Subsequent tuning on higher order clusters did not increase the signal or change the profile of the cluster distribution. Tuning on the monomer or dimer favors more harsh conditions that lead to the predissociation of the higher order clusters. The particular settings used for all data collection in this paper included source voltage 4.15 kV, capillary voltage 40.47 V, capillary temperature 199.9 °C, and tube lens offset 10 V. For the negative ion mode, source voltage polarity was simply reversed.

Sample concentrations were kept uniformly at ~300 μ M for arginine and ~100 μ M for the additional species of interest, unless otherwise noted. These concentrations are substantially higher than those used for this instrument for analytical purposes (which will not be our focus here). All samples were electrosprayed in an 80:20 methanol/water mixture. Samples were electrosprayed with a flow of 3-5 μ L/min from a 500 μ L Hamilton syringe for optimal signal. Silica tubing with an inner diameter of 12.7

microns was used as the electrospray tip. No acid was added to any of the samples. All chemicals were purchased from Sigma or Aldrich and used without further purification, unless indicated otherwise.

For the study of chiral effects, samples were electrosprayed at $\sim 200 \mu\text{M}$ concentration for each component (the total concentration being $\sim 400 \mu\text{M}$). D-Arginine HCl was used in conjunction with isotopically labeled L-arginine HCl. The labeled arginine contained two ^{15}N (99+ atom %) at the terminal guanidinium positions. All other conditions were identical to those mentioned previously. The labeled arginine was purchased from Cambridge Isotope and used without further purification.

Initial calculations were performed on the HyperChem 5.1 Professional Suite. Candidate structures were identified with molecular mechanics and submitted to full optimization at the PM3 semi-empirical level. The DFT calculations were carried out using Jaguar 4.0 (Schrödinger, Inc., Portland, Oregon). Full geometry optimization was performed at the B3LYP/6-31G** level of theory.

2.3 Results

Clustering of Electrosprayed Amino Acids and Derivatives. Protonated clusters of arginine are shown in Figure 1.1a. The protonated clusters of arginine fit the distribution $[n\text{Arg}+x\text{H}]^{x+}$ where $n = 1-7$ for $x = 1$, and $n > 11$ for $x = 2$. As shown in the spectrum, the distribution varies uniformly with n , with no indication of clearly favorable structures or “magic numbers”.

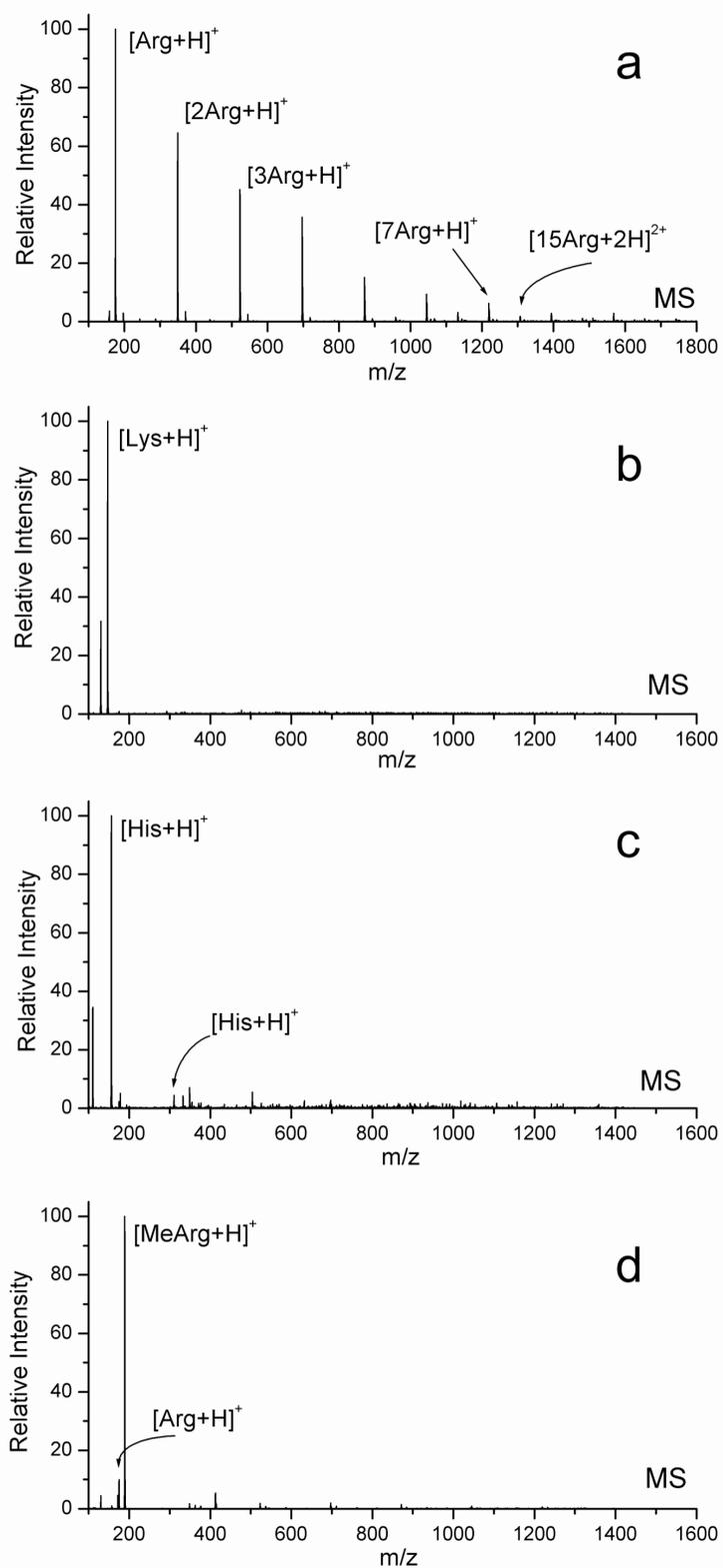


Figure 2.1 Caption on next page.

Figure 2.1 Electrospray mass spectra of basic amino acids and derivatives at 300 μ M concentration. Arginine (a) exhibits extensive noncovalent clusters. Lysine (b) and histidine (c) exhibit no extensive clustering. Arginine methyl ester (MeArg) electrosprayed with arginine (d) shows that MeArg does not cluster with itself or with arginine, demonstrating the importance of the C-terminus in the clustering.

To assure that the arginine clusters were indeed atypical and not the result of electrospray conditions that would lead to the clustering of any small molecule, several other amino acids were run under identical conditions. In particular, lysine and histidine were examined due to the high basicity (Table 2.1)⁷ of these amino acids. The mass spectra of lysine and histidine are shown in Figure 2.1b and 2.1c, respectively. Under conditions where arginine forms extensive clusters, histidine exhibits only a small dimer and no clusters are observed with lysine. Histidine and lysine will cluster at higher concentrations and lower capillary temperatures, but the extent of clustering observed with these milder ion sampling conditions is still less than that illustrated in Figure 2.1a for arginine. The propensity for arginine to cluster more than other amino acids is consistent with past observations.²¹

To probe the origin of the stability of the clusters, a mixture of arginine methyl-ester (MeArg) and arginine was electrosprayed. The spectrum of the mixture is shown in Figure 2.1d. The protonated MeArg peak completely dominates the spectrum, and the intensity of the protonated arginine clusters is dramatically reduced compared to the results in Figure 2.1a. In addition, no clusters are observed for MeArg in a solution

containing no arginine. These results indicate that the C-terminus plays a critical role in the formation of arginine clusters.

Effect of Metal Cations on Clusters of Arginine. In light of previous work^{10,14}, the effect of metal cations on clustering of arginine was examined. The metals were added to the solution in the form of the appropriate chloride salt. The results are summarized in Figure 2.2.

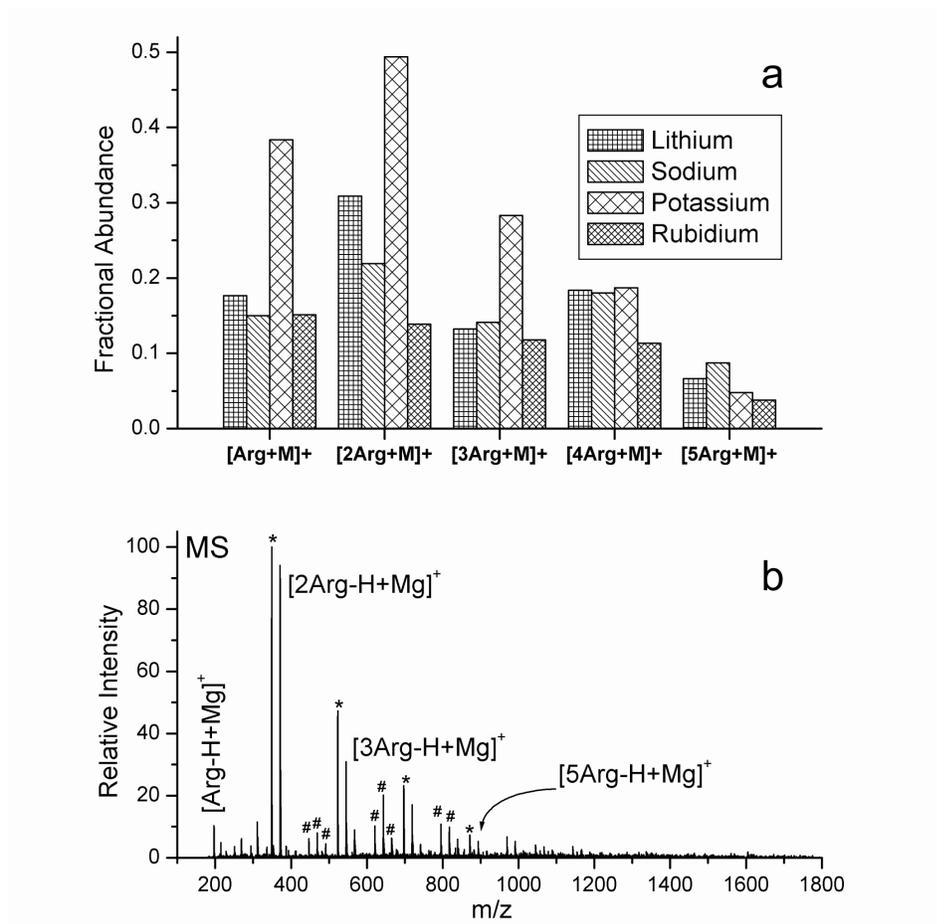


Figure 2.2 (a) Distribution of $[n\text{Arg}+M]^+$ clusters where $M = \text{Li}, \text{Na}, \text{K}, \text{and Rb}$ taken relative to the protonated dimer. (b) Arg/Mg clusters showing the unusual abundance of the dimer. * indicates protonated clusters of Arginine and # indicates clusters of arginine with sulfate.

The results for clusters of arginine with Li, Na, K, and Rb are grouped together in Figure 2.2a. The alkali metal arginine clusters fit a distribution described by $[n\text{Arg}+\text{M}]^+$, where $n = 1-5$ and $\text{M} = \text{Li}, \text{Na}, \text{K}, \text{and Rb}$. Li^+ clusters have a bimodal distribution centered on $n = 2$ and 4. Clusters of arginine with Na^+ fit a similar bimodal distribution. K^+ clusters exhibit a smooth distribution centered on $n = 2$. The distribution changes again for clusters of arginine with Rb^+ , with the intensity of the clusters declining from the monomer. The alkali metal spectra do not suggest that any of the clusters are unusually stable.

The results of mixing arginine with MgSO_4 are shown in Figure 2.2b. The Mg^{2+} clusters are described by $[n\text{Arg}-\text{H}+\text{Mg}]^+$ where $n = 1-5$, resulting in a complex that is singly charged. The dimer ($n = 2$) is particularly favored in this series, exhibiting over twice the intensity of the monomer or the trimer. The structure and stability of this cluster is discussed further below.

Anionic Arginine Clusters. Arginine electrosprayed in negative ion mode yielded the spectrum shown in Figure 2.3a. Anionic clusters of arginine are observed under identical conditions to those producing the cationic clusters (only reversing the polarity). These clusters may be described by $[n\text{Arg}-\text{H}]^-$, where $n = 1-9$, where the net charge on the species is -1 from the loss of one proton. Deprotonated clusters of anionic arginine are observed out to a total of 9 arginines. Apart from the abundant monomer, the distribution of clusters peaks at $n = 4$.

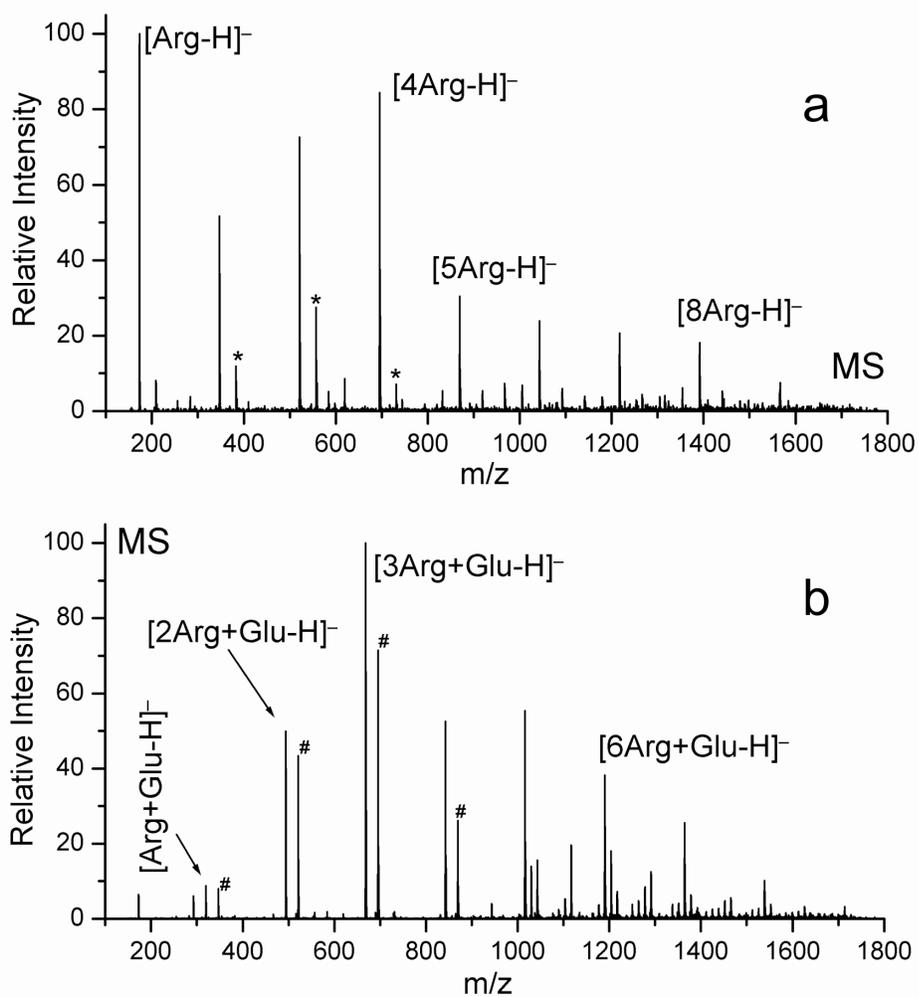


Figure 2.3 Anionic electrospray mass spectra of arginine clusters. (a) Deprotonated arginine clusters. (b) Arg/Glu clusters (150 μM Glu). * indicates clusters of Arg/Cl⁻ (see Figure 2.4) and # indicates deprotonated arginine clusters.

Clusters of arginine with acidic amino acids are generally described by $[n\text{Arg}+x\text{X}-\text{H}]^-$, where $n = 1-9$ for $x = 1$ with $\text{X} = \text{Glu}, \text{Asp}$. The spectrum of a mixture of glutamic acid and arginine is shown in Figure 2.3b. Clusters observed with aspartic acid are nearly identical to the distribution in Figure 2.3b. The presence of an abundant anionic cluster with $n = 3$ is noted and will be discussed further below. Also, there are numerous peaks of low intensity, distributed in the range 1000-1600 m/z . The charge state of these clusters is not within the resolution capabilities of the instrument. However, they appear to be multiply charged higher mass clusters of multiple glutamic acids with multiple arginines.

Anionic clusters of arginine incorporating halide anions were examined by addition of the sodium salt of each halogen to the arginine solution, which was then electrosprayed in negative ion mode. The results are shown in Figure 2.4.

Figure 2.4a shows the spectrum resulting from the addition of fluoride. Interestingly, anionic clusters of arginine incorporating fluoride ions are not observed. The spectrum of clusters observed in the presence of chloride ion is shown in Figure 2.4b. The chloride clusters fit a distribution described by $[n\text{Arg}+x\text{Cl}]^{x-}$, where $n = 1-58$ and $x = 1-6$. Multiply charged clusters are apparent in this case where $x > 1$. The cluster corresponding to the trimer $n = 3$, $x = 1$ is particularly prominent, suggesting unusual stability. Bromide clusters (Figure 2.4c) are described by $[n\text{Arg}+\text{Br}]^-$, where $n = 2-4$. The trimer ($n = 3$) is again the most prominent of the observed clusters. In comparison to chloride, the overall cluster intensity and variety are greatly reduced.

The spectrum for iodide is shown in Figure 2.4d. The iodide clusters are described by $[n\text{Arg}+\text{I}]^-$, where $n = 2, 3$. With iodide, the cluster intensity is reduced again relative to

chloride and bromide. The results for the halogen series demonstrate that chloride has the greatest propensity to form stable clusters with arginine.

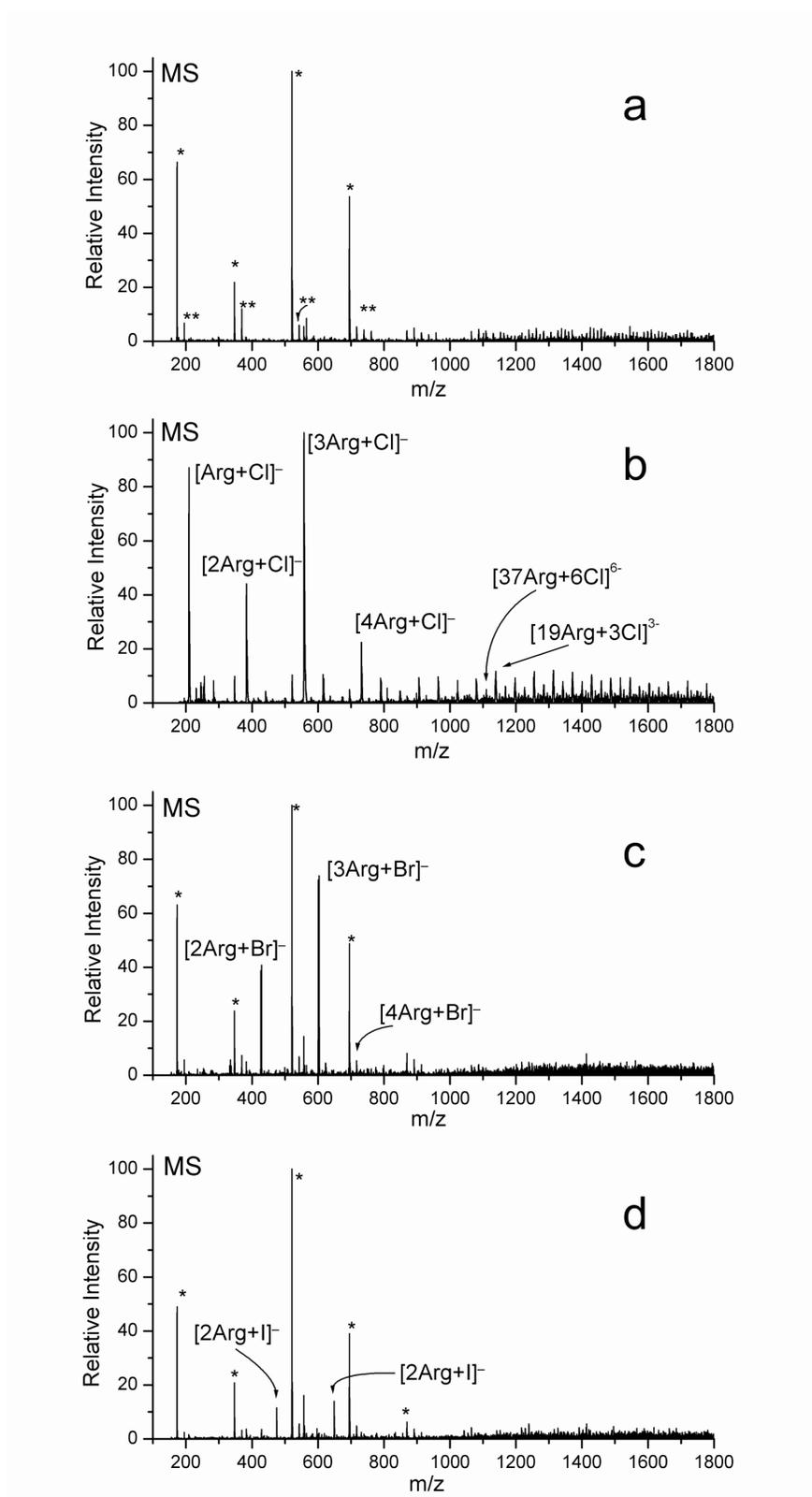


Figure 2.4 Caption on next page.

Figure 2.4 Anionic electrospray mass spectra of Arg/halide clusters. (a) Arginine does not cluster with fluoride. (b) Addition of chloride leads to extensive clustering and “magic” trimer (600 μ M Arg). (c) Arginine clusters to a small degree with the bromide anion, where again the trimer is the primary peak. (d) Arginine forms clusters with iodide, including the trimer, though the clustering is greatly reduced overall. * indicates deprotonated arginine clusters and ** indicates anionic sodiated cluster of Arg.

Chiral Effects. To determine the effect of enantiomeric purity on the stability of the clusters, a mixture of D and L arginine was electrosprayed. The L-arginine was isotopically labeled by two ^{15}N (99+ atom %) at the terminal guanidinium positions to distinguish it from D-arginine in the cluster spectra. The relative concentrations of each enantiomer were approximated by the relative monomer intensities. A binomial distribution was then used to predict the intensities of the observed heterochiral and homochiral species. For example, the dimer should follow the distribution described by LL + DL + LD + DD, where the predicted intensities would be LL=0.25, DL(LD)=0.5, and DD=0.25 for a mixture that was initially 50% of each monomer. Deviation of the observed distribution from this binomial distribution would indicate that cluster stability is dependent on the enantiomeric composition.

Table 2.2 lists the predicted distributions for protonated arginine clusters based on the binomial distribution. The results for the particularly stable NO_3^- trimer are also included (the chloride trimer was not evaluated due to the added isotopic complexity

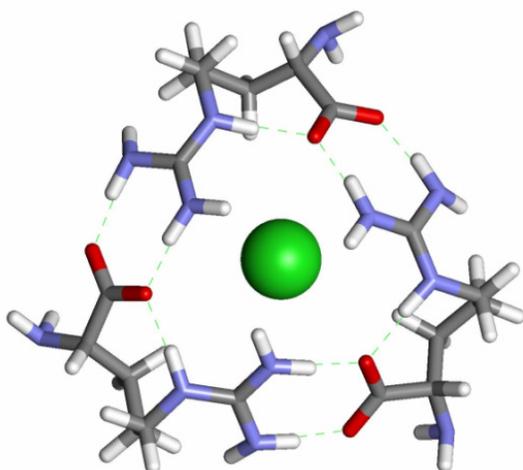
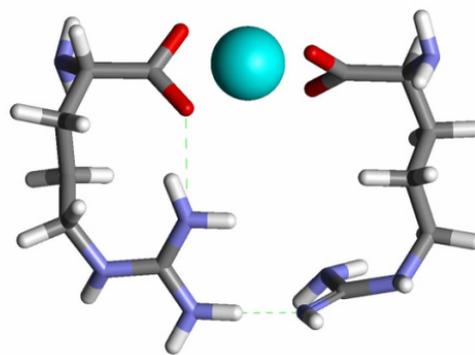
inherent with chloride spectra). The observed distributions do not show a significant deviation from the expected distribution.

Table 2.2 Cluster distributions for an enantiomeric mixture of various protonated arginine clusters and [3Arg+NO₃]⁻. Predicted percentages are based on the expected percent from a binomial distribution of clusters. The initial monomer distribution was 47.7% D-arginine hydrochloride and 52.3% isotopically labeled L-arginine hydrochloride.

Clusters of Arginine from D,L Mixture	Distribution (as % of total)		
	Predicted	Observed	Difference
DD	22.8	21.8	1.0
DL	49.9	49.2	0.7
LL	27.4	28.9	-1.5
DDD	10.9	10.6	0.3
DDL	35.7	33.2	2.5
DLL	39.1	41.5	-2.4
LLL	14.3	14.7	-0.4
DDDD	5.2	5.6	-0.4
DDDL	22.7	19.6	3.1
DDLL	37.3	33.9	3.4
DDL	27.3	29.9	-2.6
LLLL	7.5	10.9	-3.4
DDDDD	2.5	5.0	-2.5
DDDDL	13.5	9.6	3.9
DDDLL	29.7	26.8	2.9
DDLLL	32.5	33.9	1.4
DLLLL	17.8	18.8	-1.0
LLLLL	3.9	5.9	-2.0
DDD*	6.0	5.9	0.1
DDL*	28.0	29.8	-1.8
DLL*	43.5	44.6	-1.1
LLL*	22.5	19.7	2.8

*Results correspond to structure **7** capped with NO₃⁻.

Calculations. DFT calculations at the B3LYP/6-31G** level have been utilized to assess the energetics and structures of the arginine clusters. This level of calculation is feasible for the trimeric clusters, but the size of the remaining observed clusters prohibits the use of higher level calculations. We have found that PM3 calculations will qualitatively reproduce the DFT results, enabling calculations to be performed on larger systems at this level.

**2.7****2.8**

The proposed structure for the anionic trimer (**2.7**) is stabilized by the intermolecular interaction between the guanidinium and carboxylate groups **2.5** that is often observed in crystal structures. The structure allows all three arginines to participate in this highly favored interaction, while allowing for the additional stabilization of an intra-molecular hydrogen bond within each arginine. Higher level DFT calculations carried out at the B3LYP/6-31G** level support structure **2.7**. The charge solvated version of structure **2.7** (with non-zwitterionic arginines) will convert automatically to the zwitterionic version when minimized. Collision induced dissociation (CID) experiments on structure **2.7** demonstrate primarily the sequential loss of neutral arginine.

The calculations on the $[2\text{Arg-H+Mg}]^+$ dimer reveal another geometrical configuration by which the zwitterionic state of arginine is favored. The charge of the dimer is overall +1. This indicates that the one of the arginines is negatively charged and that the other is neutral. The neutral arginine could either be zwitterionic or not. DFT calculations show that the zwitterionic form is more stable by 270 kJ/mol. The geometry is shown in **2.8**, and it can be observed that much of the stabilization comes from the magnesium ion interaction with the carboxylate groups. This highly favorable interaction in which the two carboxylate groups closely approach the cation at nearly right angles to one another is only possible in the dimer and probably accounts for the special stability of this cluster.

Experimental Support for Structure 2.7. It can be noted from the structure for the arginine trimer **2.7** that it has threefold symmetry which might facilitate simultaneous formation of three hydrogen bonds with an appropriate anion. The requirements imposed on this anion would be that it have a size similar to that of the chloride ion and three hydrogen bond acceptor sites that could be aligned with the three available hydrogens from the guanidinium side chains in **2.7**. NO_3^- fits these requirements, with a size ~ 0.1 angstroms larger than Cl^- . The spectrum for the clusters of arginine with NO_3^- is shown in Figure 2.5. The clusters fit the distribution $[n\text{Arg+NO}_3]^-$, where $n = 1-5$. The most dominant $[n\text{Arg+NO}_3]^-$ cluster corresponds to $n=3$.

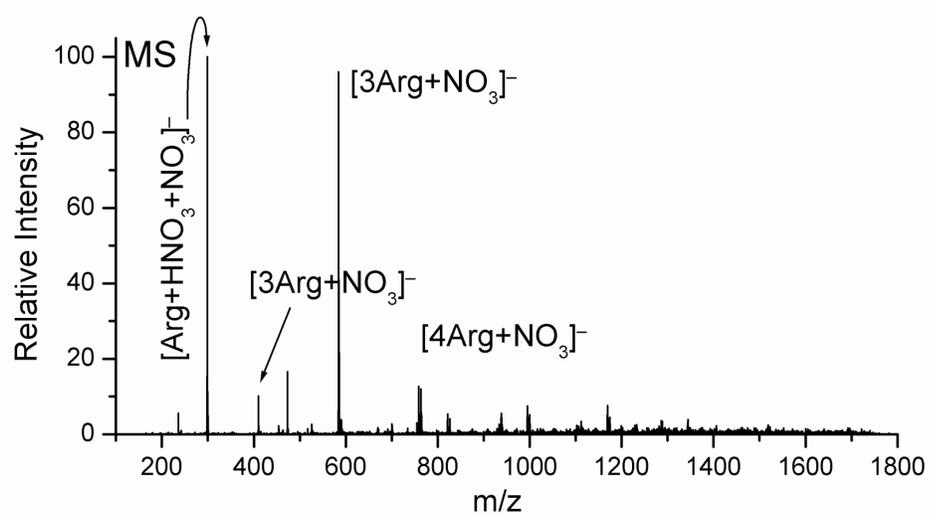


Figure 2.5 Anionic electrospray mass spectra of Arg/NO₃⁻ clusters, exhibiting the trimer as an abundant and unusually stable species, thus supporting the proposed structure (2.7) for the anionic trimer.

2.4 Discussion

Arginine forms extensive supramolecular assemblies in the gas phase, sampled from moderately concentrated solutions by electrospray ionization. The clusters observed in this study are ionic in nature, with the charge bearing group ranging from simple monatomic anions or cations to protonated or deprotonated molecular species. Comparisons with other amino acids show that arginine is unique in its ability to cluster under these conditions. Previous work has shown that arginine may be stable as a zwitterion in the gas phase in the presence of charge.¹⁰⁻¹⁴ It is consistent with all experimental and theoretical results in the present work that salt bridges formed between ions and zwitterionic arginine stabilize the observed clusters.

It has been reported previously that $[4\text{Arg}+\text{H}]^+$ has special stability²¹; however, we find no evidence to support this claim here. The intensity of the cluster corresponding to $[4\text{Arg}+\text{H}]^+$ does not remain unusually abundant with change in concentration, as shown in Figure 2.1a. Although this peak becomes more intense at concentrations several orders of magnitude higher, it still conforms to the surrounding distribution. We attribute the change in the intensity of this cluster to the change in the concentration. We were unable to produce the proposed²¹ planar arrangement of $[4\text{Arg}+\text{H}]^+$ resembling the crystal structure of NaCl with molecular modeling.

The cluster distributions for the alkali metal ions with arginine (Figure 1.2a) are dependent on the particular metal. Na^+ and Li^+ exhibit the same bimodal distribution of clusters, suggesting that these metals form clusters with arginine that are similar in structure. In contrast, clusters of arginine with K^+ exhibit greater intensity, and a different distribution centered on the dimer. This suggests a different series of structures,

and it is consistent with previous observations where K^+ was found to stabilize the zwitterionic state of arginine better than Na^+ or Li^+ .¹⁰ For Rb^+ the intensity of the clusters is lower and a different distribution is observed. Interestingly, the cluster abundances for Rb^+ are nearly constant from the monomer to the tetramer.

Anionic clusters of arginine with the halides are very different in nature from the cationic clusters discussed thus far. Arginine was not observed to cluster with the fluoride anion (Figure 2.4a). Calculations suggest that this is due to the small size and high gas phase basicity of F^- , which deprotonates the guanidinium group of arginine. The resulting formation of HF would yield a spectrum of deprotonated arginine clusters, consistent with the experimental results shown in Figure 2.4a. The chloride anion is able to form abundant clusters with arginine (Figure 2.4b). The lower gas phase basicity and larger size of the chloride anion make it a better candidate for cluster formation. Both bromide and iodide show decreasing cluster intensity, suggesting that these anions are too large for ideal clustering conditions (Figure 2.4c and Figure 2.4d). While the intensity of the arginine clusters with attached halide ions varies with size much like the alkali metal clusters, the similar cluster distribution observed for each anion suggests a common structure.

Two of the clusters (**2.7** and **2.8**) in this work are unusually abundant when they appear in a cluster distribution, suggesting structures with special stability. Both of these clusters exhibit intensities that are at least double the intensities of the $n+1$ and $n-1$ clusters (where $n = \text{Arg}$). This remains evident over a wide range of concentrations, as may be seen by comparison of Figure 2.3a with Figure 2.4b. In Figure 2.3a trace

impurities lead to the formation of $[n\text{Arg}+\text{Cl}]^-$ clusters (where $n = 1-4$). However, even at very low concentrations, the trimer **2.7** is the most prominent peak.

The anionic trimer shown in **2.7** illustrates the features of zwitterionic arginine that facilitate formation of an unusually stable cluster. The cyclic array of molecules accommodates a nearly planar guanidinium and carboxylate interaction between adjacent arginines. Coulombic attraction and two specific strong hydrogen bonds result from this orientation. In addition, an internal hydrogen bond stabilizes each arginine individually (the lowest energy conformations of arginine itself all contain an internal hydrogen bond).⁶ Each positively charged guanidinium group interacts with the central anion. It is possible to form larger cyclic arrays of arginine around a large central anion, however only at the cost of the intramolecular hydrogen bonds. This suggests an explanation for the lack of a magic tetramer with a larger anion such as iodide.



DFT calculations suggest that reaction (2.1) is exothermic by ~ 8.7 kJ/mol. Assuming a negligible entropic change, this reaction should proceed to the right. When the gas phase acidities of both acids are taken into account (Table 2.1), these calculations indicate that the chloride anion binds more strongly (34 kJ/mol) to the neutral arginine trimer than does the nitrate anion. This result is surprising because the nitrate anion can form three hydrogen bonds to the arginine shell. However, the nitrate anion is larger than the chloride anion, leading to steric constraints that mitigate against the formation of three strong hydrogen bonds between the anion and the arginine shell. The result is a lower binding energy. Although such experiments were not conducted as a part of this investigation, it should be possible to observe processes such as reaction (2.1) in the gas

phase and perhaps even measure equilibrium constants that would further quantify the thermochemistry of these species.

The greatly enhanced abundance of $[3\text{Arg}+\text{NO}_3]^-$ relative to the dimer and tetramer of this series can be attributed to the three hydrogen bonds that can be formed with the guanidinium groups of each arginine. This offers evidence that the arginine trimer is cyclic in nature. DFT calculations further show no barrier for conversion to the proposed zwitterionic state when the cluster initially is comprised of non-zwitterionic arginines. Interestingly, the optimization proceeds by conversion of one arginine to the zwitterionic state, which is then followed by conversion of the remaining arginines to zwitterions.

The dimer of arginine coordinated to divalent magnesium (**2.8**) demonstrates another method of stabilizing zwitterionic arginine. It has been shown that one protonated arginine will stabilize another zwitterionic arginine in the protonated dimer.¹² In structure **2.8**, negatively charged arginine with Mg^{2+} is shown to stabilize the second arginine in the zwitterionic form. All divalent metal cations are likely to coordinate to arginine via the carboxylate ends. The interaction with the carboxylate ends also accounts for the unusual stability of the dimer of this series relative to the trimer and the monomer. It is also possible that other amino acids may be stabilized as zwitterions in a similar divalent metal cation bound dimer.

Experiments with the enantiomers of arginine revealed no enhanced stability for homochiral clusters. As clearly demonstrated above, arginine clusters derive special stability from the strong intermolecular interaction between the guanidinium and carboxylate groups of adjacent molecules. Other interactions are secondary in importance, including those involving the amino group, which remains uncharged in

these clusters. This is consistent with the crystal structure for arginine, in which the amino groups are not charged and there is no preference for homochirality.²² This result is consistent with the proposed structure **2.7**, where the amino groups do not participate in any intra or intermolecular bonding.

2.5 Conclusion

Both theoretical and experimental evidence suggests that the zwitterionic form of arginine can be easily stabilized in the gas phase. It has been shown both by experiment and theory that zwitterionic arginine is very near in energy to arginine, and therefore may be stabilized by nearby charged species.

It has been demonstrated here that it is possible to stabilize the zwitterionic form of arginine in cyclic arrays, which take full advantage of the strong guanidinium and carboxylate interaction. Such arrays are neutral, but are detected by their ability to complex with either an anion or cation. This is well illustrated by the cyclic trimer of arginine bound to an anion (structure **2.7**). The abundance of this species suggests that it is unusually stable, especially when coordinated to NO_3^- , which forms three hydrogen bonds to the guanidinium groups and can be regarded as the crown jewel for capping this structure.

We have developed experimental methodology to examine possible preferences for homochirality in molecular clusters which has broad applications beyond the studies of arginine clusters reported in this paper. The arginine clusters show no preference for homochirality, consistent with structures where the amino group is not involved in cluster

bonding. However, we have examined other systems where there is a pronounced preference for homochirality.²³

Assignment of structures for the larger clusters, e.g., $[50\text{Arg}+6\text{Cl}]^{6-}$ is a formidable problem, with the sheer size of the aggregate rendering impossible all calculations except molecular mechanics. Such enormous clusters offer a glimpse of the chemistry that lies at the interface between the solid state and gas phase. Portions of these clusters may approximate arginine in the crystalline state.

2.6 References

- ¹ (a) Locke, M. J.; Hunter, R. L.; McIver, R. T. *J. Am. Chem. Soc.* **1979**, *101*, 272. (b) Locke, M. J.; McIver, R. T. *J. Am. Chem. Soc.* **1983**, *105*, 4226.
- ² Ding, Y.; Kroug-Jespersen, K. *Chem. Phys. Lett.* **1992**, *199*, 261.
- ³ Jensen, J. H.; Gordon, M.S. *J. Am. Chem. Soc.* **1995**, *117*, 8159-8170.
- ⁴ Bowen, K. H. *personal communication*.
- ⁵ Chapo, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J. *J. Am. Chem. Soc.* **1998**, *120*, 12956-12957.
- ⁶ Maksic, Z. B.; Kovacevic, B. *J. Chem. Soc., Perkin Trans. 2*, **1999**, 2623-2629.
- ⁷ (a) E. P. Hunter and S. G. Lias, "Proton Affinity Evaluation" in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. W. G. Mallard and P. J. Linstrom, February 2000, National Institute of Standards and Technology, Gaithersburg MD, 20899 (<http://webbook.nist.gov>). (b) J. F. Liebman, J. L. Holmes, R. D. Levin, and W. G. Mallard, "Ion Energetics Data" in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. W. G. Mallard and P. J. Linstrom, February 2000, National Institute of Standards and Technology, Gaithersburg MD, 20899 (<http://webbook.nist.gov>).
- ⁸ Wytttenbach, T.; Witt, M.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458.
- ⁹ Bertran, J.; Rodriguez-Santiago, L.; Sodupe, M. *J. Phys. Chem. B* **1999**, *103*, 2310-2317.
- ¹⁰ Jockusch, R. A.; William, P. D.; Williams, E. R. *J. Phys. Chem. A* **1999**, *103*, 9266-9274.
- ¹¹ Cerda, B. A.; Wesdemiotis, C. *Analyst* **2000**, *125(4)*, 657-660.

- ¹² Jockusch, R. A.; William, P. D.; Williams, E. R. *J. Am. Chem. Soc.* **1997**, *119*, 11988.
- ¹³ Strittmatter, E. F.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 6069-6076.
- ¹⁴ Wytenbach, T.; Witt, M.; Bowers, M. T. *Int. J. Mass Spec.* **1999**, *182/183*, 243.
- ¹⁵ Suresh, C. G.; Vijayan, M. *Int. J. Peptide Protein Res.* **1983**, *21*, 223-226.
- ¹⁶ Karle, J.; Karle, I. L. *Acta Cryst.* **1964**, *17*, 835-841.
- ¹⁷ Bhat, T. N.; Vijayan, M. *Acta Cryst.* **1977**, *B33*, 1754-1759.
- ¹⁸ Melo, A.; Ramos, M. J.; Floriano, W. B.; Gomes, J. A. N. F.; Leao, J. F. R.; Magalhaes, A. L.; Maignet, B.; Nascimento, M. C.; Reuter, N. *J. Mol. Struct.* **1999**, *463*, 81-90.
- ¹⁹ Zheng, Y. J.; Ornstein, R. L. *J. Am. Chem. Soc.* **1996**, *118*, 11237-11243.
- ²⁰ Barril, X.; Aleman, C.; Orozco, M.; Luque, F. J. *Proteins* **1998**, *32*, 67-79.
- ²¹ Zhang, D. X.; Wu, L. M.; Koch, K. J.; Cooks, R.G. *Eur. Mass. Spec.* **1999**, *5(5)*, 353-361.
- ²² Suresh, J.; Padmanabhan, S.; Vijayan, M. *J. Biomol. Struct. Dyn.* **1994**, *11*, 1425-1435.
- ²³ Julian R. R.; Hodyss R.; Kinnear B.; Jarrold M.F.; Beauchamp J. L. *J. Phys. Chem. B* **2002**, *106*, 1219-1228.

Chapter 3

Cooperative Salt Bridge Stabilization of Gas Phase Zwitterions in Neutral Arginine Clusters

Portions published previously in: Julian R. R.; Beauchamp J. L.; Goddard W. A. *J. Phys. Chem. A* **2002**, *106*, 32-34.

3.1 Introduction

While amino acids are known to exist as zwitterions in solution, the general assumption that zwitterions do not exist in the gas phase has been the subject of recent debate from both experimental and theoretical approaches. Glycine has received much attention because it is the simplest amino acid. Fourier transform-ion cyclotron resonance (FT-ICR) mass spectrometry has demonstrated that glycine is unstable as a zwitterion by ~ 20 kcal/mol.¹ Ab initio calculations confirm that glycine is unlikely to exist in the gas phase as a zwitterion.^{2,3} However, glycine also has the lowest proton affinity of the amino acids, making it the worst candidate for a gas phase zwitterion. On the other hand, the guanidinium group of arginine (Arg) gives it the highest proton affinity of the amino acids, making Arg the best candidate for a gas phase zwitterion.

The first studies on Arg by Williams and coworkers suggested that isolated Arg might exist in the zwitterionic form.⁴ However, further experimental⁵ and theoretical⁶ studies indicate that the isolated Arg monomer is not a zwitterion in the gas phase. On the other

hand several recent experiments, supported by theory, suggest that Arg in the presence of a net charge may exist in the zwitterionic state.⁷ These studies are supported by recent calculations indicating that the attachment of an electron to glycine reduces the instability of the glycine zwitterion from 20 to 9 kcal/mol.⁸

Compared to other amino acids, Arg possesses a high propensity to form abundant clusters when electrosprayed into the gas phase.^{9,10} Recent studies suggest that this clustering ability is due to the salt bridges formed by association of the guanidinium group of one Arg with the carboxylate group of another.¹⁰ The resulting clusters are detected by their ability to associate with either a cation or an anion, suggesting that isolated neutral analogs of such clusters might be stable.

We report here first principles quantum mechanical calculations (density functional theory (DFT) at the B3LYP/CCPVTZ(-F)⁺⁺//B3LYP/6-31G** level) of the structures for the neutral dimer and trimer of Arg and assess the energetics for both the zwitterionic and non-ionic tautomers of each cluster. Predictions for the structures of larger clusters and implications for the gas phase structures of peptides and proteins are made based on these results.

3.2 Methods

Candidate structures were identified by two methods. Extensive MD simulations with simulating annealing were performed to identify candidate structures. The second approach consisted of minimizing structures assembled according to chemical intuition, followed by simulated annealing on each of these structures. In general, the second approach yielded much better results. The lowest energy conformers were submitted to

higher levels of theory. Candidate structures were also evaluated at the PM3 semi-empirical level.

The DFT calculations were carried out using Jaguar 4.0 (Schrödinger, Inc., Portland, Oregon). Full geometry optimization was performed at the B3LYP/6-31G** level of theory, followed by single point calculations at the B3LYP/CCPVTZ(-F)⁺⁺ level for selected structures. Zero point energies and vibrational frequencies were obtained from the analytical Hessian for the optimized structure at the B3LYP/6-31G** level.

The molecular dynamics (MD) calculations were carried out using Cerius² from Molecular Simulations Inc. The MD calculations used the DREIDING force field¹¹ (with the exponential-six form of the van der Waals potentials) and charges from charge equilibration.¹²

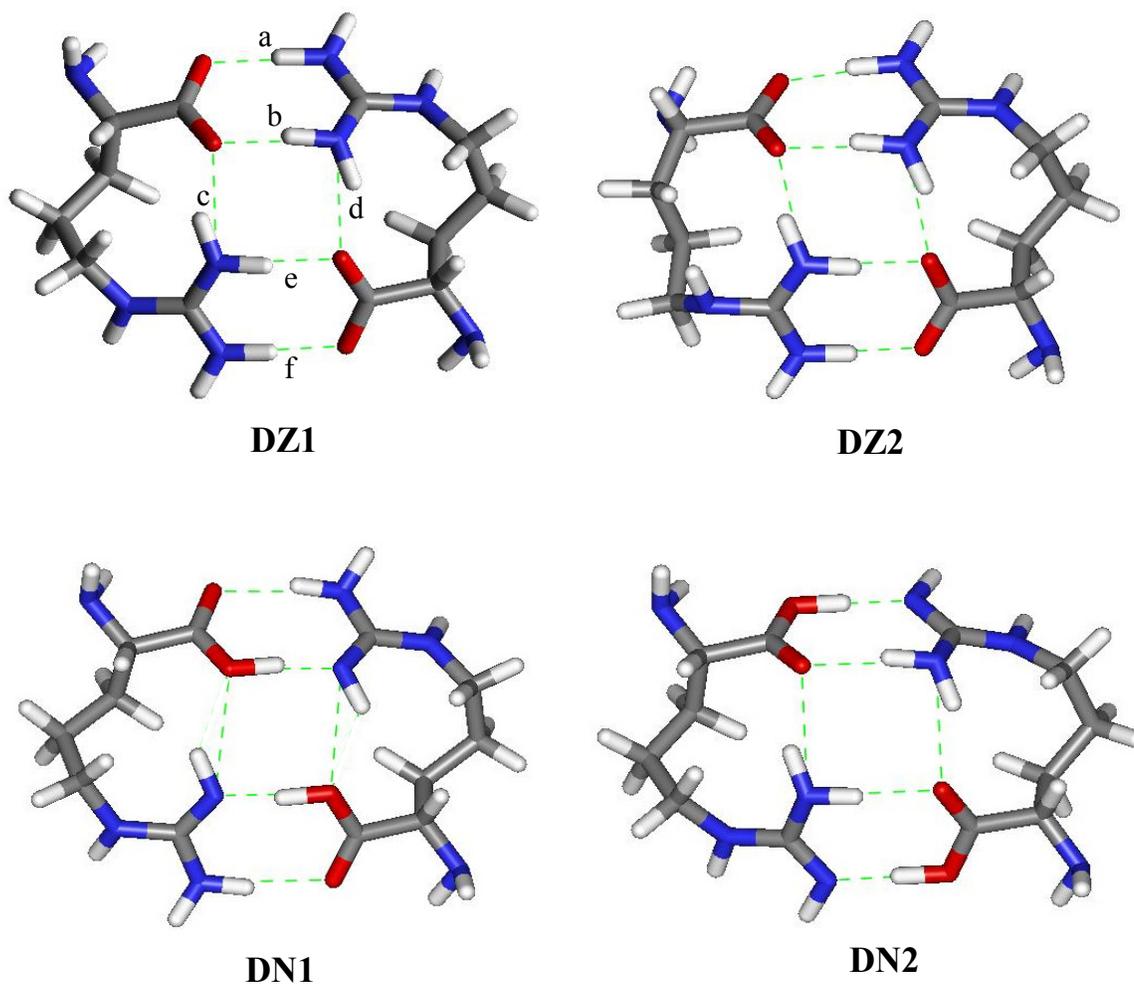
Semi-empirical PM3 MNDO type calculations were carried out using the HyperChem 5.1 Professional Suite (Hypercube, Inc., Gainesville, Florida). Predicted vibrational intensities were calculated at this level of theory.

3.3 Results and Discussion

Dimers of Arginine. After an extensive search of conformational space, the lowest energy conformer for the neutral Arg dimer that we have been able to identify is structure **DZ1**. This structure is composed of two zwitterionic Arg and has C₂ symmetry. The Arg are arranged in a head to tail fashion with the guanidinium group of one molecule interacting with the carboxylate of the neighboring Arg. Two salt bridges are formed by the interacting zwitterions, accompanied by six hydrogen bonds. This structure is not surprising given that guanidinium/carboxylate interactions are observed in both the

crystal structure¹³ of Arg and in the lowest energy conformer^{6b} of the zwitterionic Arg monomer in the gas phase.

A second zwitterionic structure, **DZ2**, exists in a metastable state as a local minimum on the potential energy surface, only ~2 kcal/mol higher in energy (Table 2). This structure is formed from **DZ1** by flipping one guanidinium while leaving the interacting carboxylate unchanged, leading to C_1 symmetry.



There are three metastable non-ionic tautomers of **DZ1**. Two of these (**DN1** and **DN2**) are symmetrical non-ionic tautomers. Structure **DN1** has both interior hydrogens placed on the carboxylate groups, while structure **DN2** has the exterior hydrogens placed on the

carboxylate groups. The relative energetics of the non-ionic and zwitterionic arginine dimers are shown in the reaction coordinate diagram at the B3LYP/6-31G** level in Figure 3.1. All four dimers contain six hydrogen bonds (Table 3.1), but the non-ionic structures lack the coulomb attraction derived from the salt bridges in **DZ1**. The result (Table 3.2) is that the non-ionic tautomers are less stable than **DZ1** by 10 to 12 kcal/mol. Furthermore, starting from a structure with one salt bridge interaction, the dimer minimizes without barrier to **DZ1**. This suggests that the creation of one salt bridge stabilizes the other in a cooperative manner. Thus we conclude that Arg can solvate itself, stabilizing the zwitterionic state that predominates in solution.

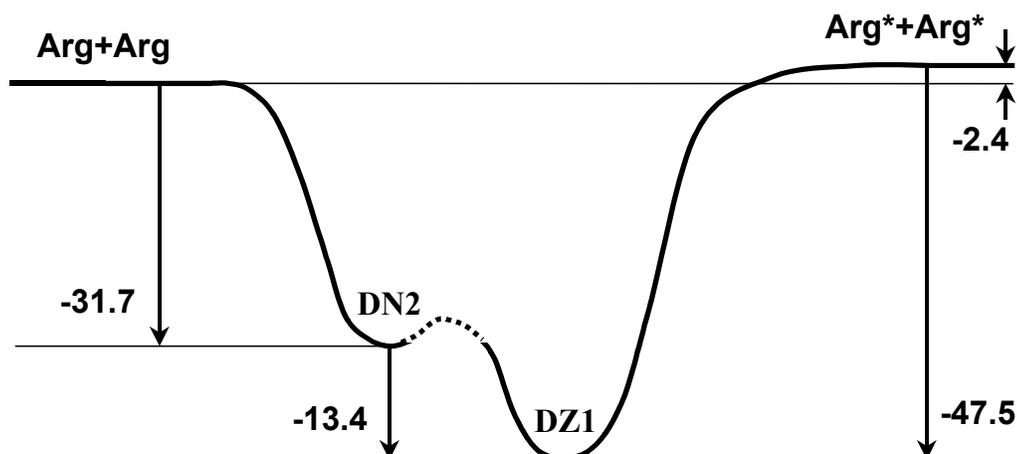


Figure 3.1 Reaction coordinate diagram for the dimerization of arginine as determined at the DFT B3LYP/6-31G** level in kcal/mol. Arg represents the neutral ground state of arginine. Arg* is the lowest energy zwitterionic conformation of arginine. The transition barrier to convert DZ1 to DN2 was not calculated and is represented by a dashed line; however, this barrier is expected to be very small. The monomeric structures were taken from ref 6b (it should be noted that the actual difference between the separated monomers is probably closer to 5.6 kcal/mol as indicated in ref 6b).

Table 3.1 Heteroatom Separation in Hydrogen Bonds of Arginine Dimers.^a

Bond ^b	DZ1	DZ2	DN1	DN2
a	2.74	2.68	2.97	2.59
b	2.72	2.79	2.57	2.88
c	2.73	2.82	2.98	2.94
d	2.73	2.74	2.98	2.94
e	2.72	2.74	2.57	2.88
f	2.74	2.69	2.97	2.59

- a. All lengths given in angstroms
b. As defined by the pattern in structure **DZ1**

Table 3.2 Calculated Electronic Energies for Arginine Dimers.^a

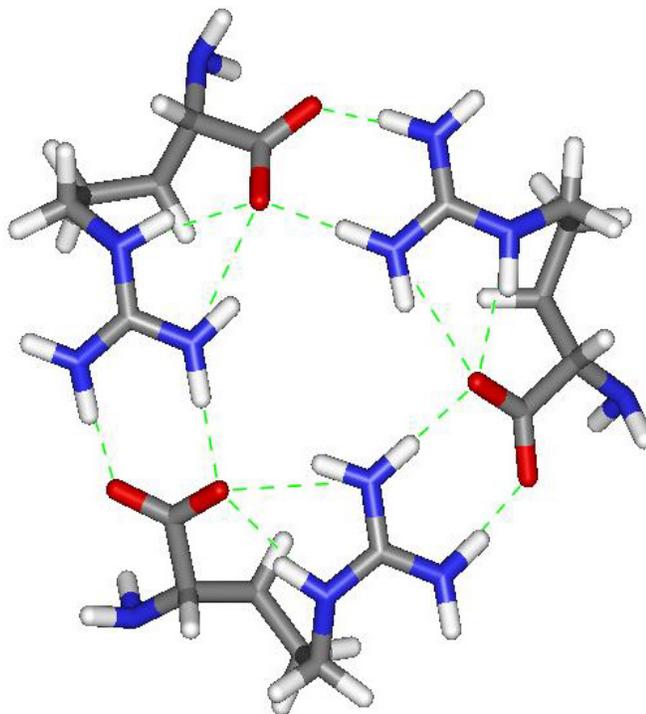
Structure	Zero Point Energy	B3LYP/6-31G** ^b	Zero Point Corrected ^b	B3LYP/CCPVTZ(-F) ^{++b}	Zero Point Corrected ^b
DZ1	281.7	0	0	0	0
DZ2	282.3	2.3	2.9	na	na
DN1	280.7	14.6	13.6	11.6	10.5
DN2	280.6	13.4	12.3	10.4	9.3

- a. All energies given in kcal/mol
b. Energies given relative to structure **DZ1**

For the gas phase, similar calculations on the monomer indicate that the zwitterion is less stable than the nonionic form by 2.8 kcal/mol.⁶ Therefore, the coulomb attraction derived from the two salt bridges provides about 18 kcal/mol to compensate the conversion of neutral Arg to the zwitterionic form. Structure **DZ1** is stabilized by 12 kcal/mol as a result of this coulomb attraction.

Interestingly, it should be possible to confirm the preferred state of the arginine dimer experimentally. The asymmetric CO stretch for the non-ionic tautomers is calculated to fall in the range of 1900 to 2200 cm^{-1} . The predicted intensity of these bands (2049.55 cm^{-1} , 2024.42 cm^{-1} and 2134.39 cm^{-1} , 2056.54 cm^{-1}) should be sufficient for experimental absorption measurements. The zwitterionic tautomers have no predicted vibrations in this range (ranging from 1798 cm^{-1} to 2789 cm^{-1}).

Trimers of Arginine. A particularly stable anionic Arg trimer has previously been identified experimentally.¹⁰ After extensive conformational searching, we find the neutral Arg trimer to have a very similar structure, **TZ1**. Again the cyclic structure is dominated by guanidinium/carboxylate interactions, with each Arg maintaining two intramolecular hydrogen bonds. This structure has C_3 symmetry. There are a variety of non-ionic tautomers of **TZ1**. Only the two symmetrical cases were studied at the DFT level of theory, one with the exterior hydrogens transferred to the carboxylate groups and the other with the interior hydrogens transferred to the carboxylate groups. In both cases, we find that the non-ionic tautomers convert without barrier to the zwitterionic state. PM3 calculations indicate that all other tautomers are bracketed energetically by the two symmetrical cases. Restricting the lowest energy tautomer to be non-ionic yields an energy 31 kcal/mol above **TZ1**.

**TZ1**

The cyclic array of salt bridges in **TZ1** allows for a variety of additional long-range coulomb interactions. **TZ1** demonstrates much greater stability relative to the non-ionic trimer than is the case with the dimers. This observation is a result of the larger cyclic salt bridge in **TZ1**, which allows for more favorable synergistic interactions between the salt bridges and reduced coulomb repulsion between like charges. The results suggest that these long-range interactions help to stabilize the zwitterionic state of Arg. Or put another way, the greater the number of Arg present, the easier it becomes to self-solvate charged groups. The extreme example of this is found in crystal structures, where even glycine is stabilized in the zwitterionic state.¹⁴ This suggests that it should be possible to stabilize all of the amino acids as zwitterions through self-solvation, provided a sufficient cluster size is attained. For Arg, the critical cluster size is the dimer. The results presented here suggest that all Arg clusters larger than the dimer will consist of

zwitterionic Arg and that the primary bonding pattern will be dominated by interactions between guanidinium and carboxylate groups.

Several computational studies have addressed the stability of gas phase salt bridges by analyzing isolated salt bridge pairs in the gas phase. When isolated, the non-ionic hydrogen bound tautomers are always lower in energy.^{15,16,17} The results presented here suggest that in a large protein, the synergistic interaction of an array of salt bridges may enhance their stability in the gas phase. The complex molecular environment surrounding a candidate gas phase salt bridge (for example between the side chains of arginine and aspartic acid) in a protein must be carefully analyzed to identify stabilizing factors. The presence of other charges, salt bridges, hydrogen bonding or any combination of these possibilities may be sufficient to stabilize a salt bridge in a protein in the gas phase.

3.4 Conclusions

Arg prefers the zwitterionic form when clustered together with at least one other Arg even in the absence of solvent or net charge. The structure for the gas phase dimer (**DZ1**) of Arg is defined by two salt bridges, where the guanidinium group of one Arg interacts with the carboxylate of another. Two intramolecular hydrogen bonds and four intermolecular hydrogen bonds additionally bind the juxtaposed arginines. Although the non-ionic tautomers of the dimer are local minima on the potential energy surface and maintain all of the hydrogen bonds, they lack the additional favorable coulombic interactions. The neutral Arg trimer exhibits a similar structural motif. However in this

case, the non-ionic tautomers convert without barrier to the zwitterionic structure shown in **TZ1**.

Molecular clusters provide an interesting bridge between the gas and solid phase properties of molecules. Given sufficient size, the molecular clusters will begin to exhibit solid phase characteristics, such as the stabilization of zwitterionic salt bridges through self-solvation. This principle is not limited to Arg and should be general in nature, although the particular results will vary from molecule to molecule. For instance, similar studies suggest that serine is stabilized as a zwitterion in the gas phase at ~8 serines.¹⁸ Experimentally, these serine octamers also exhibit a chiral preference, also a property commonly associated with the solid phase.

3.5 References

-
- ¹ (a) Locke, M. J.; Hunter, R. L.; McIver, R. T. *J. Am. Chem. Soc.* **1979**, *101*, 272. (b) Locke, M. J.; McIver, R. T. *J. Am. Chem. Soc.* **1983**, *105*, 4226.
- ² Ding, Y.; Kroug-Jespersen, K. *Chem. Phys. Lett.* **1992**, *199*, 261.
- ³ (a) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1995**, *117*, 8159-8170. (b) Bertran, J.; Rodriguez-Santiago, L.; Sodupe, M. *J. Phys. Chem. B* **1999**, *103*, 2310-2317.
- ⁴ Jockusch, R. A.; William, P. D.; Williams, E. R. *J. Am. Chem. Soc.* **1997**, *119*, 11988-11989.
- ⁵ Chapo, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J. *J. Am. Chem. Soc.* **1998**, *120*, 12956-12957.
- ⁶ (a) Maksic, Z. B.; Kovacevic, B. *J. Chem. Soc., Perkin Trans. 2*, **1999**, 2623-2629, (b) Skurski, P.; Gutowski, M.; Barrios, R.; Simons, J. *Chem. Phys. Lett.*, **2001**, *337*, 143-150.
- ⁷ (a) Wyttenbach, T.; Witt, M.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458. (b) Jockusch, R. A.; William, P. D.; Williams, E. R. *J. Phys. Chem. A* **1999**, *103*, 9266-9274. (c) Cerda, B. A.; Wesdemiotis, C. *Analyst* **2000**, *125(4)*, 657-660. (d) Strittmatter, E. F.; Williams, E. R. *J. Phys. Chem. A* **2000**, *104*, 6069-6076. (e) Wyttenbach, T.; Witt, M.; Bowers, M. T. *Int. J. Mass Spec.* **1999**, *182/183*, 243.
- ⁸ Gutowski, M.; Skurski P.; Simons, J. *J. Am. Chem. Soc.* **2000**, *122*, 10159-10162.
- ⁹ Zhang, D. X.; Wu, L. M.; Koch, K. J.; Cooks, R. G. *Eur. Mass. Spec.* **1999**, *5(5)*, 353-361.
- ¹⁰ Julian, R. R.; Hodyss, R.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2001**, *123*, 3577-3583.
- ¹¹ Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem. A* **1990**, *94*, 8897-8909.
- ¹² Rappe A. K.; Goddard W. A. *J. Phys. Chem.* **1991**, *95(8)*, 3358-3363.

-
- ¹³ Karle, J.; Karle, I. L. *Acta Cryst.* **1964**, *17*, 835-841.
- ¹⁴ Marsh, R. E. *Acta Cryst.* **1958**, *11*, 654.
- ¹⁵ (a) Mu, T. W.; Feng, Y.; Liu, L.; Guo, Q. X. *Chinese Chem. Lett.* **2001**, *12*, 219-222.
- (b) Barril, X.; Aleman, C.; Orozco, M.; Luque, F. J. *Proteins* **1998**, *32*, 67-79.
- ¹⁶ Melo, A.; Ramos, M. J.; Floriano, W. B.; Gomes, J. A. N. F.; Leao, J. F. R.; Magalhaes, A. L.; Maignet, B.; Nascimento, M. C.; Reuter, N. *J. Mol. Struct.* **1999**, *463*, 81-90.
- ¹⁷ Zheng, Y. J.; Ornstein, R. L. *J. Am. Chem. Soc.* **1996**, *118*, 11237-11243.
- ¹⁸ Julian R. R.; Hodyss R.; Kinnear B.; Jarrold M. F.; Beauchamp J. L. *J. Phys. Chem. B* **2002**, *106*, 1219-1228.

Chapter 4

Nanocrystalline Aggregation of Serine Detected by Electrospray Ionization Mass Spectrometry: Origin of the Stable Homochiral Gas Phase Serine Octamer

Portions published previously in: Julian R. R.; Hodyss R.; Kinnear B.; Jarrold M. F.; Beauchamp J. L. *J. Phys. Chem. B* **2002**, *106*, 1219-1228.

4.1 Introduction

Did homochirality precede life or was homochirality a consequence of life? This issue¹ is complex enough to warrant study from all avenues. This paper addresses a novel mechanism for the spontaneous generation of homochirality, or homochirogenesis. Homochirogenesis may be achieved by at least three fundamental mechanisms: 1) selective synthesis of only one enantiomer of a chiral molecule², 2) the preferential destruction of one enantiomer of a heterochiral mixture³, and 3) separation of a racemic mixture into distinct homochiral parts.⁴

The spontaneous breaking of symmetry for a racemic mixture of chiral molecules has only been achieved by a handful of methods. Certain molecules will spontaneously separate from a racemic solution into homochiral crystals.⁴ The assembly of these macroscopic homochiral crystals must be orchestrated at the molecular level.

Intermediate to the formation of macroscopic crystals from single molecules, noncovalently bound molecular clusters may be formed. The study of these intermediate molecular clusters is likely to offer insight into the most fundamental requirements for spontaneous symmetry breaking. Molecular clusters may share structural similarities with their parent crystal structures.⁵ Furthermore, homochiral self-assembly into small molecular clusters offers a possible target of opportunity for conversion of the aggregates into a homochiral polymer or macromolecule with well-defined stereochemistry.

Electrospray ionization mass spectrometry (ESI-MS) is an excellent method for observing molecular clusters in the gas phase.⁶ A variety of noncovalent clusters have been studied by ESI-MS, ranging from the molecular recognition of organic molecules⁷ to the structures of salt clusters.⁸ The flexibility and sensitivity of ESI-MS allows for the examination of solutions over a wide range of concentrations.

By itself, mass spectrometry provides only the molecular weight of a detected species and no direct information relating to molecular structure in general or to chirality in particular. This has led to the development of a variety of methods to circumvent this deficiency. These experiments, involving chiral amino acids⁹, amino acid derivatives¹⁰, tartrates¹¹, metals and amino acids¹², and host/guest chemistry involving crown ethers¹³, cyclodextrins¹⁴, cyclofructans¹⁵, and monosaccharides¹⁶ have all utilized mass spectrometry and taken advantage of the different physical properties of diastereomers to successfully discriminate between enantiomeric pairs. Experimental techniques include both kinetic and equilibrium methods. The preferred method uses one of the enantiomers isotopically labeled to distinguish it from the other.⁸⁻¹⁶

Alternatively, mass spectrometry can be coupled with another technique such as ion mobility spectrometry (IMS).¹⁷ IMS gives direct structural information about the gas phase conformation of a molecule or cluster in the form of a collisional cross section. The experimental cross section can then be compared to the theoretical cross section determined computationally for likely structures of a particular molecule or cluster.

We have recently developed a new method to determine the extent to which spontaneous chiral separation occurs in small molecular clusters.^{5,18} Our previous work¹⁸ demonstrated the successful application of this new technique on the serine octamer. The serine octamer was originally reported by Cooks and coworkers,¹⁹ using a different experimental methodology, to have a homochiral preference.

In the current work, we report detailed experiments that elucidate the structure of the serine octamer and offer an explanation for its unusual abundance. Examination of a mixture of D-serine and labeled L-serine using ESI-MS indicates that the octamer strongly prefers a homochiral composition. IMS experiments indicate that the structure of the octamer is consistent with a cubic arrangement of serine. Blockage of the N-terminus or C-terminus of serine leads to no octamer formation, suggesting that amino and carboxylate functionalities play a critical role in the bonding. The spontaneous symmetry breaking and likely involvement of zwitterionic serine aggregates in solution suggest that nanocrystals of serine precede the formation of the gas phase octamer. DFT calculations reveal several low energy structures that are related to the crystal structure of serine. The analogs threonine and homoserine may also form similar structures. The experimental results for these molecules are compared to the results found for serine.

4.2 Methods

Mass spectra were obtained using a Finnigan LCQ ion trap quadrupole mass spectrometer without modification. The signal was optimized using the automatic tuning capabilities of the LCQ on the protonated serine octamer. For serine and analogs, the settings used were source voltage 4.15 kV, capillary voltage 27.30 V, capillary temperature 159.9°C, and tube lens offset 10 V. Collision induced dissociation (CID) was performed on isolated parent ions by applying a resonance excitation RF voltage of between 0.98 V and 2.45 V for a period of 30 ms.

Sample concentrations were varied from $\sim 10^{-5}$ to $\sim 10^{-2}$ M for serine. The maximum octamer formation was observed at ~ 0.01 M. These concentrations are substantially higher than those used for analytical purposes. Samples were electrosprayed using a 50:50 methanol/water mixture containing 0.1% v/v acetic acid at a flow rate of 3 $\mu\text{L}/\text{min}$ from a 500 μL Hamilton syringe. Silica tubing with an inner diameter of 12.7 microns was used as the electrospray tip. Unlabeled compounds were purchased from Sigma or Aldrich and used without further purification. L-serine labeled by replacing the hydrogen in the three C-H bonds with deuterium (99+ atom %, 98+% ee) was purchased from Cambridge Isotope and used without further purification.

Candidate structures were identified with molecular mechanics and submitted to full optimization at the PM3 level. All calculations at the PM3 semi-empirical level were performed using the HyperChem 5.1 Professional Suite. In Table 1, the structures were optimized at the PM3 level, followed by single point calculations at the DFT B3LYP/6-31G level, corrected for zero point energies. All remaining structures (i.e. those in Table 2) were fully optimized and analyzed utilizing DFT. These calculations were carried out

at the B3LYP/6-31G**//B3LYP/6-31G level and corrected for zero point energies using Jaguar 4.0 by Molecular Simulations, Inc. Cross sections for candidate structures were determined using a trajectory method with a Lennard-Jones potential.²⁰

To determine the extent to which chiral separation occurs in molecular aggregates, a solution containing a mixture of enantiomers, with one labeled, is electrosprayed.¹⁸ For each cluster, a distribution of peaks will be observed. For example, three peaks corresponding to the LL, LD(DL), and DD clusters will represent a dimer. If the structure of these noncovalently bound clusters is not sensitive to the chirality of the constituent molecules, then the relative intensities of the peaks will follow a binomial distribution. For the dimer above the predicted intensities would be LL=0.25, LD(DL)=0.5, and DD=0.25 for a mixture that was initially 50% of each monomer. If the stability of the cluster is sensitive to the chirality of the individual molecules, the observed distribution will deviate from the binomial distribution. A preference for homochirality will be indicated by an increase in the relative intensities of the pure L and pure D clusters. A preference for heterochirality might also be observed, with the mixed species being more abundant than predicted by the statistical distribution.

The experimental cross sections of the serine clusters were measured using a high-resolution ion mobility apparatus.²¹ A solution was prepared by dissolving five milligrams of L-serine in 1 ml H₂O, 0.2 ml CH₃COOH, and 0.1 ml CH₃CN. The solution was electrosprayed using a 5 kV potential across a 0.5 cm gap. The ions were guided into a drift tube through an ion gate with a 556 Vcm⁻¹ field against a 2000 sccm flow of helium. The 63 cm long drift tube contains 46 guard rings and was operated at 10,000 V with ~760 torr of helium buffer gas.

Ions exit the drift tube through a 0.125 mm aperture and are focused into a differentially pumped region where they are mass selected by a quadrupole mass spectrometer and detected by an off-axis collision dynode and dual microchannel plates. A multichannel scaler that is synchronized with an electrostatic shutter located between the ion gate and the drift tube records arrival-time distributions. Drift times are determined by correcting the arrival times for the time that the ions spend traveling from the drift tube to the detector. Collision cross sections are calculated using Equation 1.²²

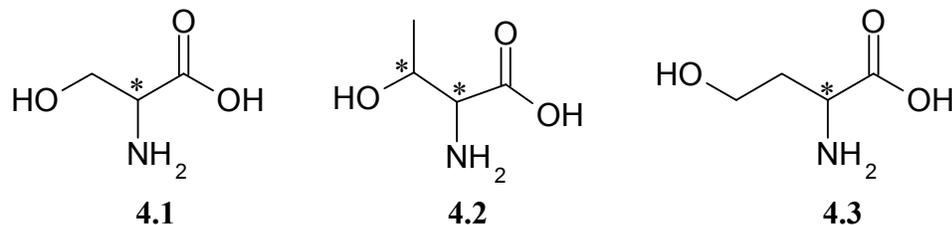
$$\Omega_{\text{avg}}^{(1,1)} = \frac{(18\pi)^2}{16} \left[\frac{1}{m} + \frac{1}{m_b} \right]^2 \frac{z_e}{(k_B T)^2} \frac{t_D E}{L} \frac{1}{\rho} \quad (1)$$

In this expression, m and m_b are the masses of the ion and buffer gas, z_e is the charge on the ion, ρ is the buffer gas number density, L is the length of the drift tube, and E is the drift field. Ions with the same m/z but different structures have different collisional cross sections and can be distinguished.²³ This is particularly useful in distinguishing “multimers” where, for example, the peak for the singly charged monomer is overlapped by the doubly charged dimer.²⁴

4.3 Results

Serine. Serine is the simplest amino acid with a polar side chain, the hydroxy-methyl group. The chiral center is marked with an asterisk in structure **4.1**. In separate ESI-MS experiments, the concentration of serine was varied by factors of ten from 10 μ M to 0.1M. Maximum clustering was observed at 0.01M concentration. The positive ion mass spectrum obtained with 0.01M L-serine is presented in Figure 4.1. The protonated octamer, [8Ser+H]⁺, is the base peak in the spectrum, followed by an

abundant protonated dimer. It is interesting to note that all of the odd clusters, i.e. the trimer and pentamer, prior to the octamer are disfavored. All of the odd clusters after the octamer fit smoothly with the distribution of the even numbered clusters. Furthermore, at masses higher than the hexamer, there are abundant half-integer peaks corresponding to doubly charged clusters such as $[13\text{Ser}+2\text{H}]^{2+}$. At masses higher than the protonated 7mer (equivalent to the $[7\text{Ser}+\text{H}]^+$ notation), one-third-integer peaks arise corresponding to triply charged species such as $[23\text{Ser}+3\text{H}]^{3+}$. In general, the distribution of clusters with m/z higher than $[8\text{Ser}+\text{H}]^+$ exhibits greater relative intensities than those clusters with a lower m/z . Clearly, the most prominent feature of the spectrum is the unusual abundance of the protonated serine octamer.



Ion mobility data (Figure 4.2) confirm that the peak corresponding to the serine octamer is itself comprised of three species, $[8\text{Ser}+\text{H}]^+$, $[16\text{Ser}+2\text{H}]^{2+}$, and $[24\text{Ser}+3\text{H}]^{3+}$. The relative intensities by peak height are 100, 75, and 28, respectively. The cross sections are 187 \AA^2 , 285 \AA^2 and 380 \AA^2 , respectively. The theoretical cross section for a proposed structure should agree to within approximately $\pm 2\%$ of these numbers. This paper will focus on the structure of the octamer itself. In related studies, Clemmer and coworkers have discussed possible structures for the multimers and report a cross section of 191.4 \AA^2 for $[8\text{Ser}+\text{H}]^+$.²⁵

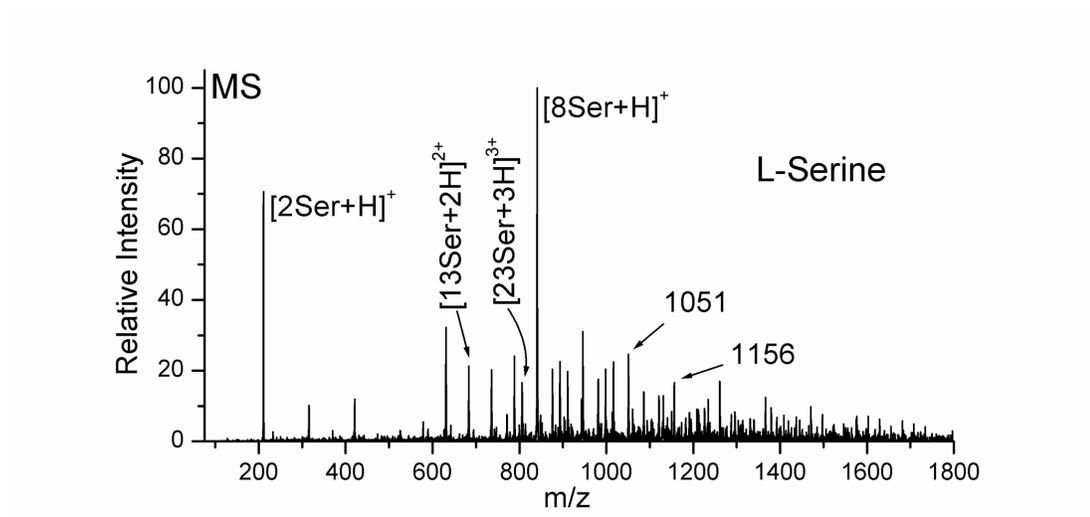


Figure 4.1 Mass spectrum of 0.01 M L-serine demonstrating abundant clustering and an unusually abundant octamer. Multiply charged as well as singly charged species are present.

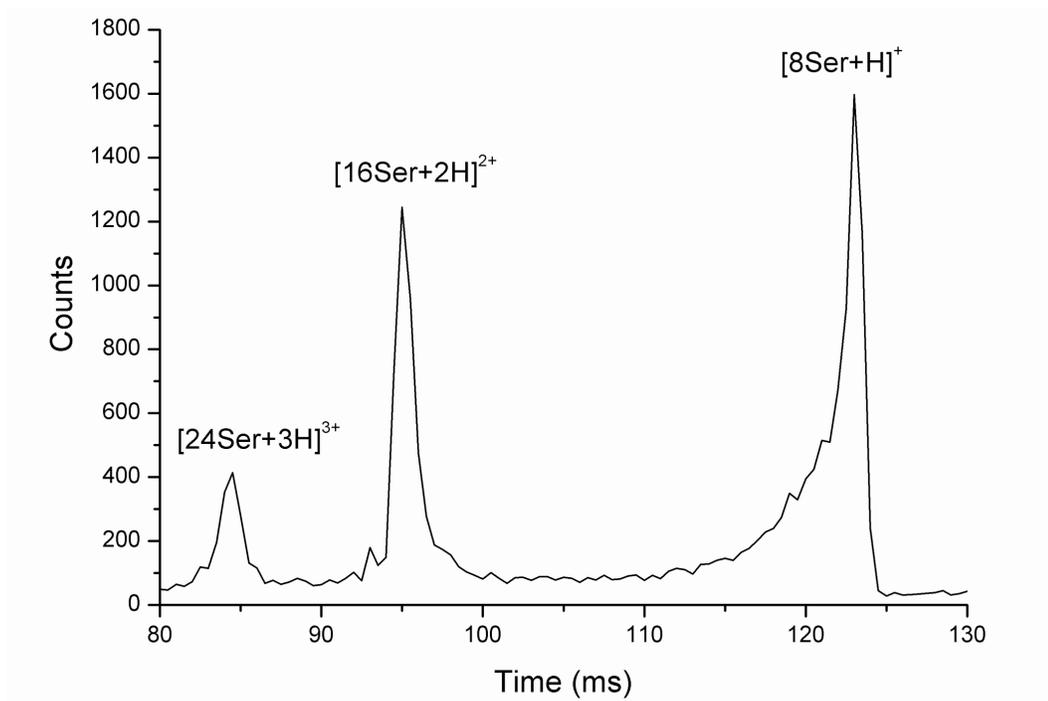


Figure 4.2 Ion mobility spectrum for m/z 841 from 0.036M serine. The three peaks correspond to the singly charged serine octamer and two multiply charged multimers.

The spectrum for the collision induced dissociation (CID) of the isolated singly protonated octamer is presented in Figure 4.3a. In the ion trap, the process of isolating ions with a certain m/z often leads to collisional heating with the helium bath gas. In this case, collisional heating eliminated all of the multiply charged serine clusters as evidenced by ^{13}C distribution. The most abundant fragment corresponds to the loss of two neutral serines, leaving the singly protonated hexamer. To a lesser extent, the singly protonated pentamer and tetramer are formed. The virtual absence of the heptamer is noted, suggesting that the loss of a single neutral serine is not favored.

The CID spectra for two of the higher mass clusters are presented in Figures 4.3b and 3c. The dissociation of the peak at m/z 1051 yields almost exclusively peaks separated by 35 mass units (Figure 4.3b). This corresponds to 1/3 the mass of serine, suggesting that the peak at 1051 is primarily composed of $[\text{30Ser}+3\text{H}]^{3+}$. In Figure 4.3c similar behavior is noted for the CID of the peak at m/z 1156. However, in this case there is an additional series of peaks separated by 35 mass units with m/z ratios higher than 1156. This series could only have resulted from a larger cluster with more than three charges on it, most likely $[\text{44Ser}+4\text{H}]^{4+}$. The relative intensities of the two series indicate that the quadruply charged species comprises a significant portion of the total intensity for the peak at m/z 1156.

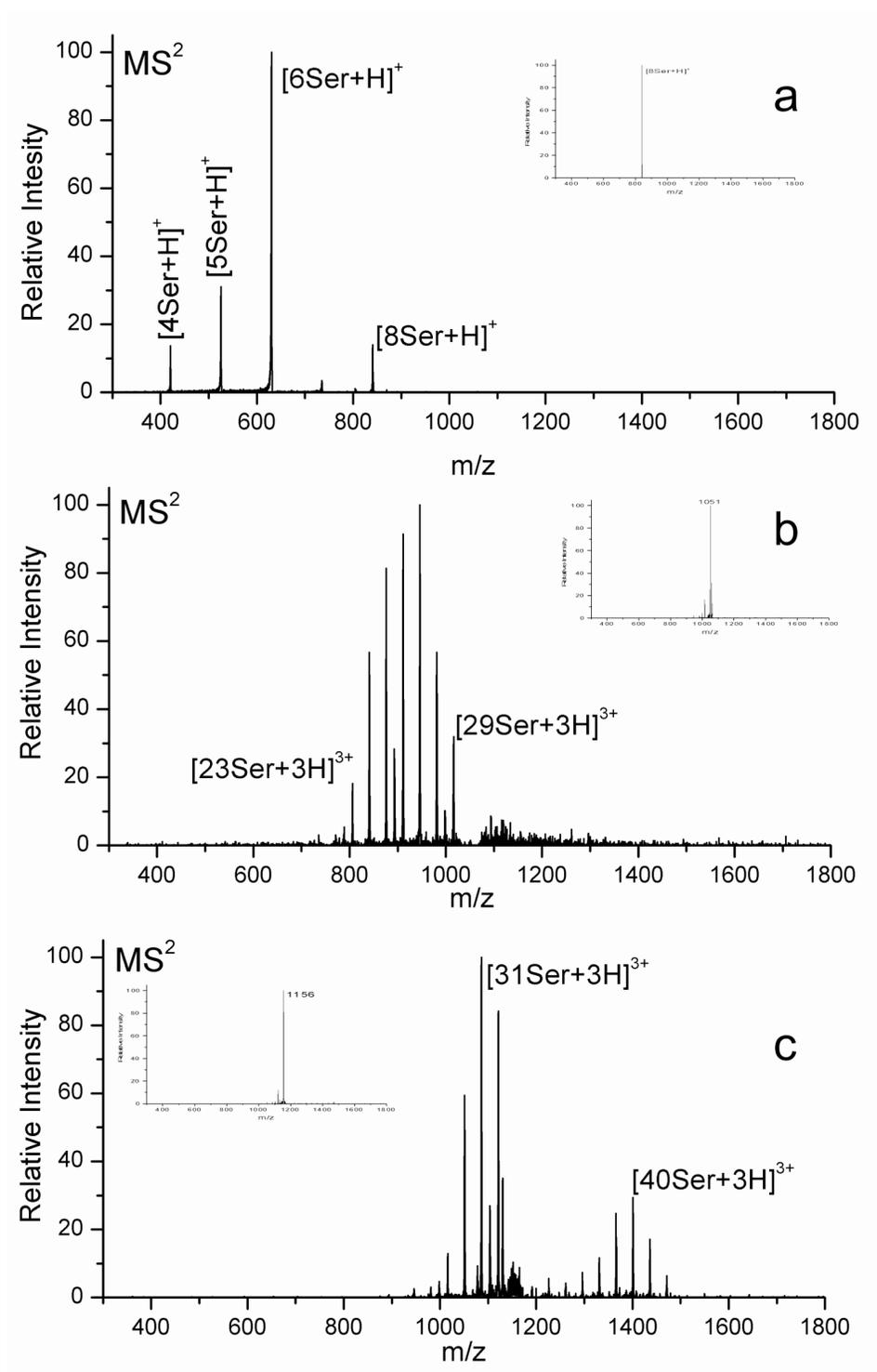


Figure 4.3 (a) CID spectrum for the singly charged serine octamer. The octamer preferentially loses a neutral serine dimer. (b) CID spectrum for m/z 1051. Several fragments are separated by 35 m/z, indicating that the parent ion is composed primarily

of $[30\text{Ser}+3\text{H}]^{3+}$. (c) CID spectrum for m/z 1156. The emergence of a cluster distribution with a separation of 35 m/z at a higher m/z suggests that a significant portion of this cluster is composed of $[44\text{Ser}+4\text{H}]^{4+}$.

Figure 4.4a shows the distribution of serine octamer clusters observed in a 54:46 mixture of D-serine with L-serine, respectively. The L-serine is labeled with deuterium in the three C-H bonds. In these experiments, collisional heating of all ions by mild excitation of the entire mass range was used to eliminate the multiply charged clusters, yielding only singly charged octamer peaks. The resulting experimental distribution of mixed clusters differs significantly from the predicted statistical distribution, included in Figure 4.4a for comparison. These results are in stark contrast to those for the serine dimer. Figure 4.4b illustrates the results of the same experiment for the protonated serine dimer in which no preference for homochirality is observed.

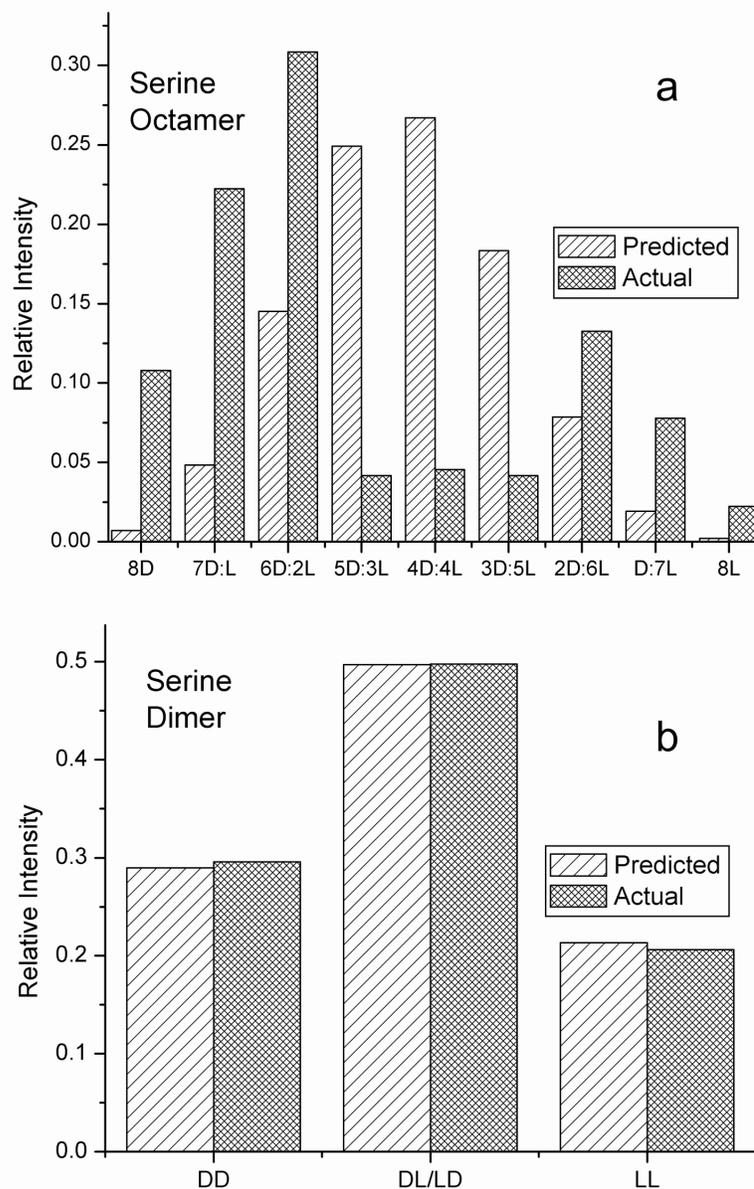


Figure 4.4 (a) Comparison between the predicted statistical distribution and observed distribution for the serine octamer for a 54:46 mixture of D-serine and isotopically labeled L-serine. Full spectrum CID was used to break up the multiply charged aggregates of the octamer. A clear preference for homochirality is indicated by this data. (b) Comparison between the predicted statistical distribution and observed distribution for the serine dimer. No clear preference for chirality is indicated.

At first glance the data in Figure 4.4a suggests that octamers containing a mixture of 6D and 2L or 2D and 6L serines are favored. A more detailed analysis is shown in Figure 4.5. Dividing the observed intensities by the statistical prediction (Figure 4.5a) yields relative intensities for the octamers which better reflect their energetic stabilities. This is more evident in the semilog plots in Figures 4.5b and 4.5c, where starting with either the 8D or 8L octamers, respectively; there is a sequential replacement of serines of one enantiomer with the other. In both instances the limiting case of a single replacement indicates an energetic cost of 2.9 ± 0.3 kJ/mol, calculated from the slope of the semilog plot assuming a Boltzmann analysis is valid, for incorporating the incorrect enantiomer in the homochiral cluster.

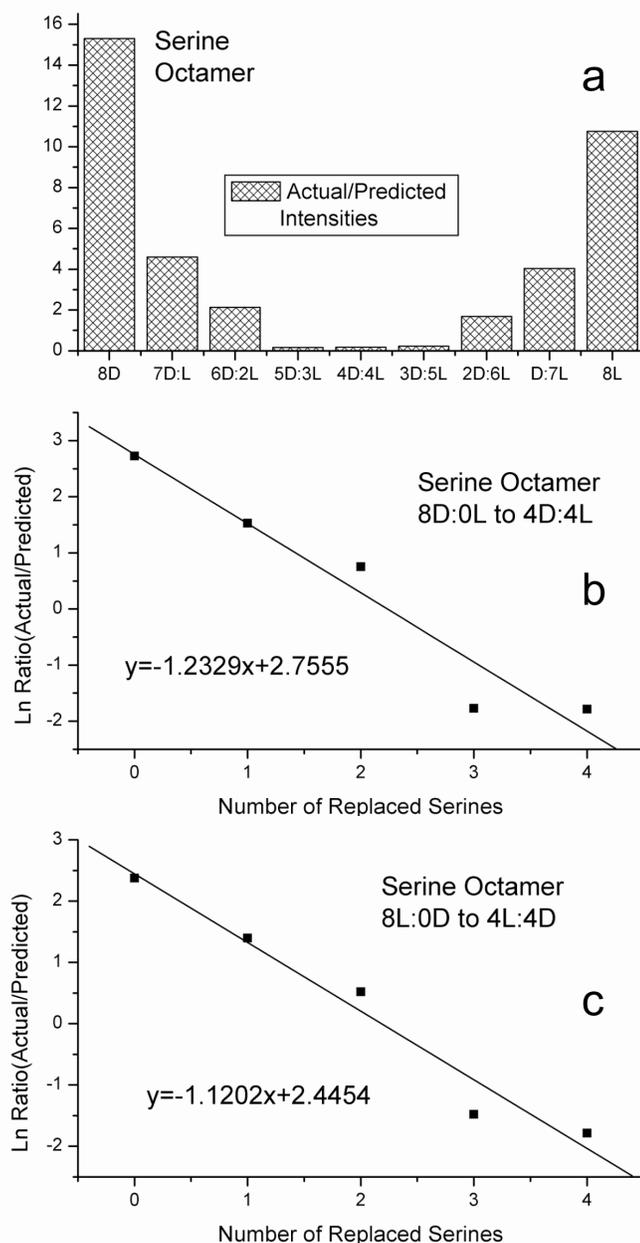


Figure 4.5 (a) Ratio of the observed to statistical intensities for the serine cluster distribution shown in Figure 4a. (b) Plot of the Ln of the ratios in (a) versus the number of L-serines switched for D-serine starting with the D-serine octamer. (c) Plot of the Ln of the ratios in (a) versus the number of D-serines switched for L-serine starting with the homochiral L-serine octamer.

Serine Derivatives. The spectra for L-serine methyl ester and L-N-tertbutoxycarbonyl-serine (L-tboc serine) are shown in Figures 4.6a and 4.6b, respectively. As seen in Figure 4.6a, the C-terminal methyl ester group eliminates the pattern of clustering that is observed for serine. The dimer is the only prominent peak in the spectrum. The small distribution of clusters at higher masses contains multiply charged clusters of L-serine methyl ester with sodium (present as a contaminant). Similarly in Figure 4.6b, the dimer of L-tboc serine is the dominant peak, followed by some higher order clusters. The trimer and higher mass clusters are primarily sodiated. The distribution of clusters is again quite different than that observed for serine itself (Figure 4.1).

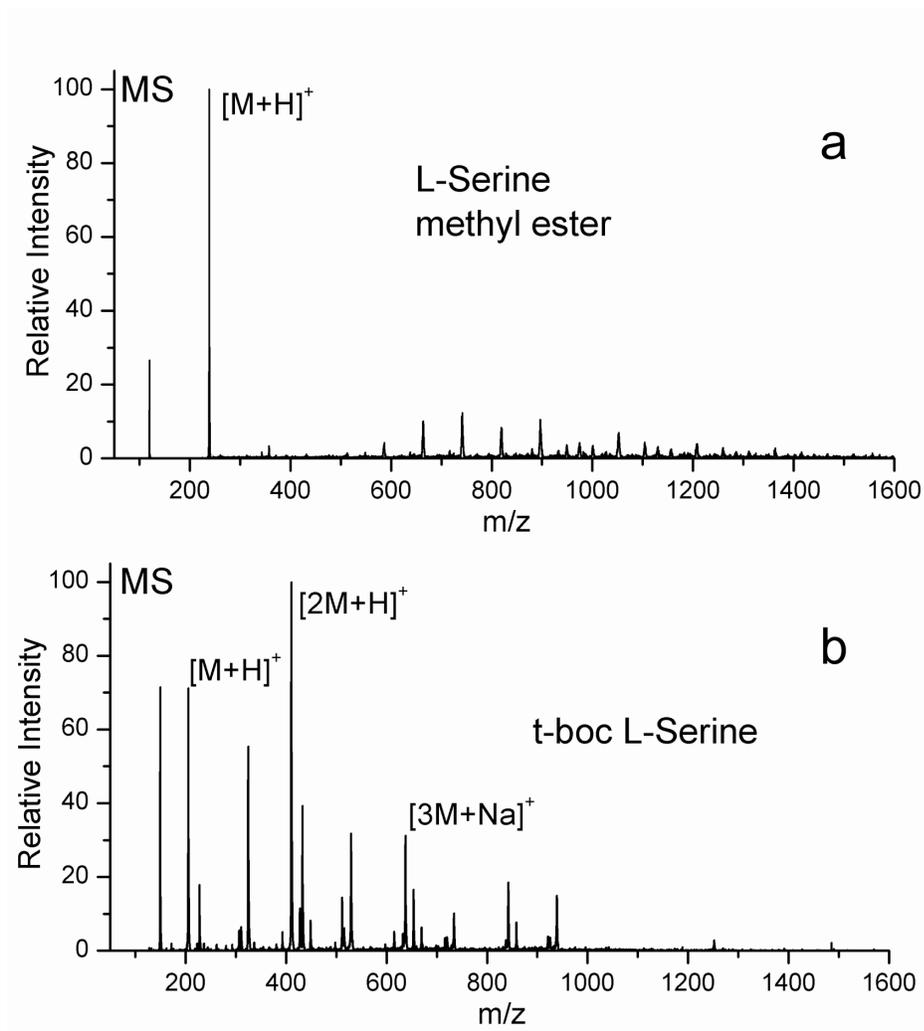


Figure 4.6 (a) Mass spectrum of L-serine methyl ester. There is no evidence for an unusually abundant octamer. (b) Mass spectrum of t-boc serine. Again, there is no evidence for an unusually abundant octamer.

Homoserine and Threonine. Two amino acids closely related to serine are threonine (structure 4.2) and homoserine (structure 4.3). Homoserine contains an additional methylene group, which extends the side chain relative to serine. The mass spectrum for L-homoserine is presented in Figure 4.7a. The monomer and dimer are both abundant, but the singly charged octamer does not exhibit unusual abundance (in contrast with serine). The base peak corresponds to 477 m/z, which might suggest that homoserine forms an unusually abundant tetramer. However upon closer inspection of the carbon-13 peaks, it becomes clear that the peak at 447 m/z is primarily composed of $[8\text{Hser}+2\text{H}]^{2+}$ (where Hser = L-homoserine). Therefore, homoserine also forms an unusually abundant octamer. The peak corresponding to $[11\text{Hser}+2\text{H}]^{2+}$ is also unusually abundant with respect to the surrounding distribution.

Several experiments were performed with mixtures of L-homoserine and L-serine. A 50/50 mixture of L-serine and L-homoserine will yield mixed clusters, but the mixed clusters are not unusually abundant. A 6/2 mixture of L-serine/L-homoserine yields abundant mixed serine octamers with the incorporation of 1 or 2 homoserine molecules into the cluster (Figure 4.7b). Isolation of $[6\text{Ser}+2\text{Hser}+\text{H}]^+$, followed by CID yields the spectrum shown in Figure 4.7c. The mixed cluster preferentially loses two neutral serines. Homoserine is always retained.

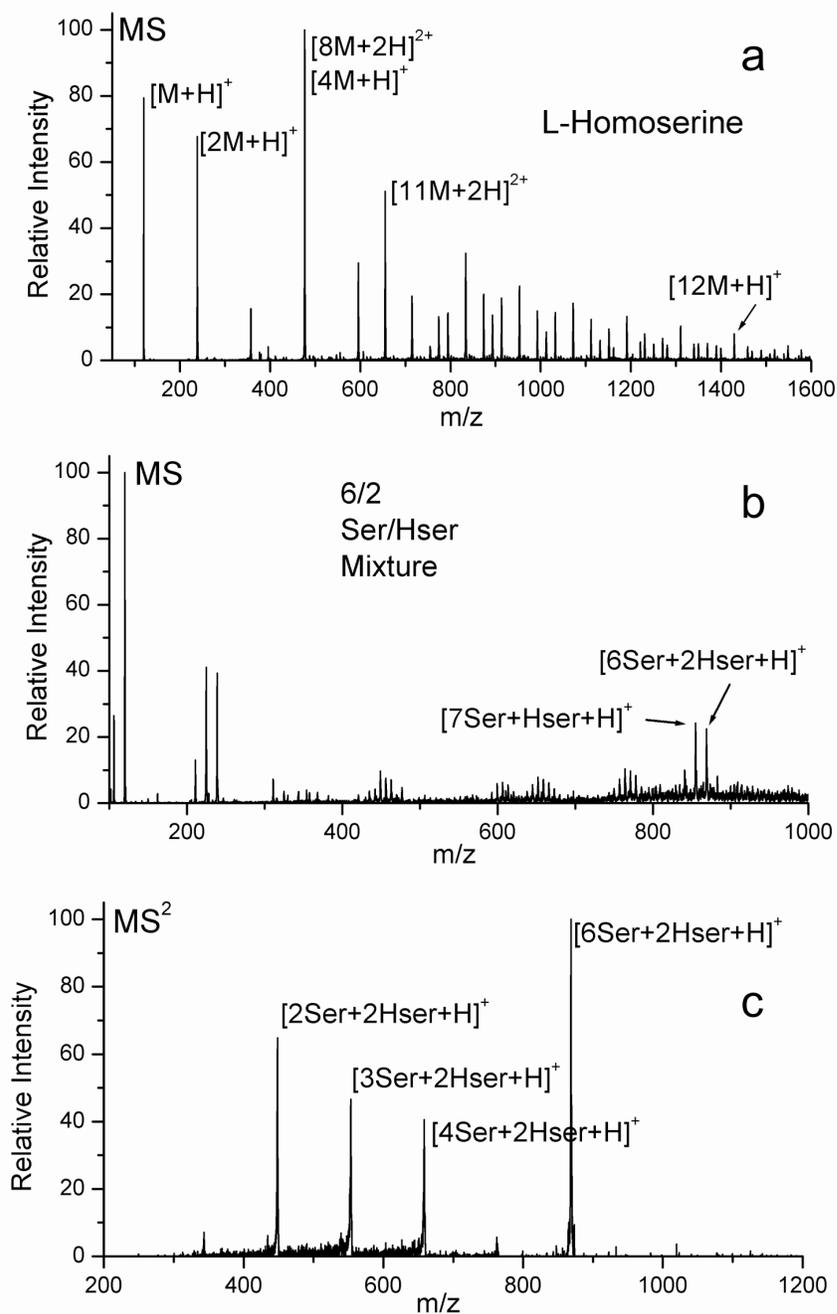


Figure 4.7 (a) Mass spectrum of L-homoserine showing extensive clustering and an abundant octamer. (b) Spectrum for a 6:2 mixture of serine and homoserine. (c) Collision induced dissociation on the mixed octamer yields the loss of at least two serines. Homoserine is always retained.

The mass spectrum for D-threonine is shown in Figure 4.8 (allo-threonine isomers gave very similar results). This spectrum shares many similarities with the spectrum of homoserine. As in the case with homoserine, $[8\text{Thr}+\text{H}]^+$ and $[8\text{Thr}+2\text{H}]^{2+}$ are both present. $[7\text{Thr}+\text{H}]^+$ is only prominent in the threonine spectrum. The peak corresponding to $[11\text{Thr}+2\text{H}]^{2+}$ is unusually abundant. However, the combined intensity of the octameric peaks is not as prominent as it is in the case of serine or homoserine.

A 50/50 mixture of D-threonine and D-serine was electrosprayed and is shown in Figure 4.9a. The singly and doubly charged mixed octamers are labeled in Figure 4.9a, where it is observed that they exhibit unusual abundance when compared to the other clusters. Figure 4.9b shows the distribution for the singly charged mixed threonine/serine octamer. Mixed clusters ranging from $[6\text{Ser}+2\text{Thr}+\text{H}]^+$ to $[2\text{Ser}+6\text{Thr}+\text{H}]^+$ are easily observed. The abundance of the two remaining mixed clusters and the two pure clusters cannot be discerned from noise. A similar distribution is observed for the doubly charged mixed octamers ($[6\text{Ser}+2\text{Thr}+2\text{H}]^{2+}$ to $[2\text{Ser}+6\text{Thr}+2\text{H}]^{2+}$) in Figure 4.9c. The different distribution, particularly the extra prominence of $[4\text{Ser}+4\text{Thr}+2\text{H}]^{2+}$, is likely due to the overlap with mixed tetramers such as $[2\text{Ser}+2\text{Thr}+\text{H}]^+$.

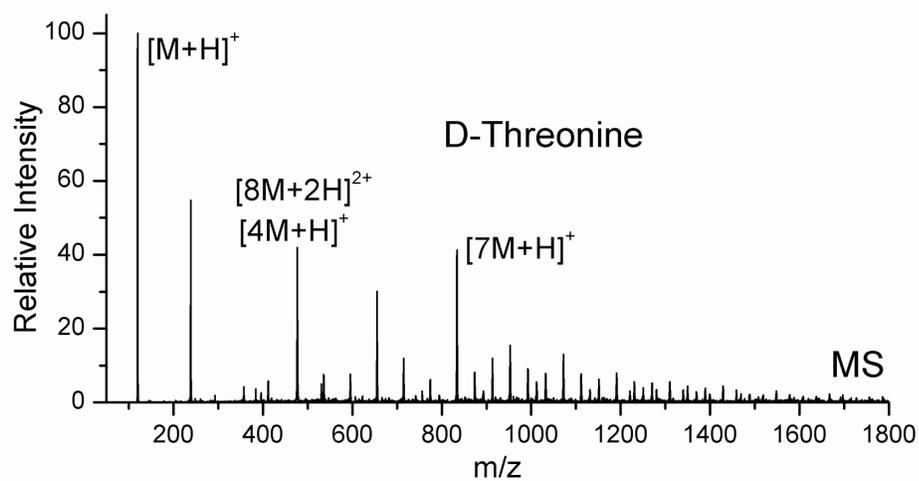


Figure 4.8 Mass spectrum for D-threonine. The octamer of this species is primarily doubly charged.

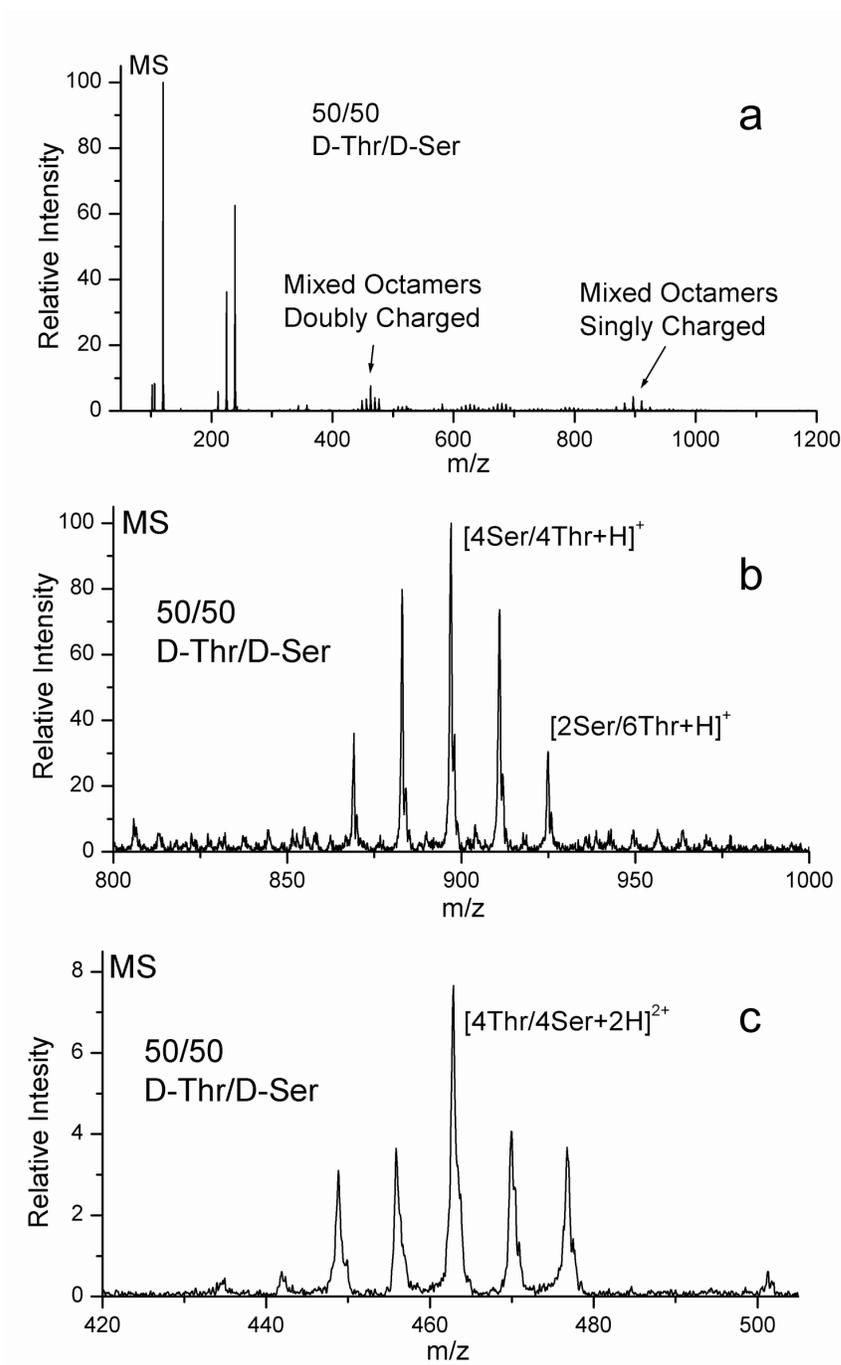
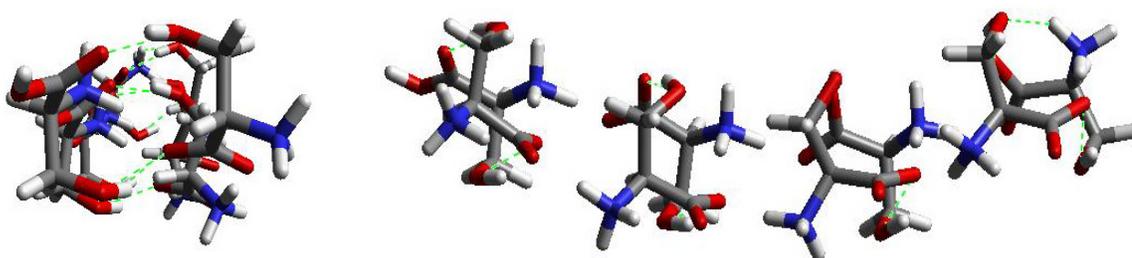


Figure 4.9 (a) Mass spectrum for a 50:50 mixture of D-threonine with D-serine. The mixed clusters still form octamers with unusual abundance. (b) Mass spectrum for the singly charged mixed octamers. (c) Mass spectrum for the doubly charged mixed octamers.

4.4 Discussion

The experiments presented thus far place several constraints on the structure of the serine octamer. (1) The unusual abundance of the octamer suggests a very stable structure. (2) The ESI-MS experiments in Figure 4.4 suggest that the octamer must have strong preference for homochirality. (3) Ion mobility data is consistent with a single structure with a cross section of 187 \AA^2 . (4) Figures 4.9b and 4.9c suggest that the structure should allow for the mixing of serine and threonine to give a stable octamer. Any structure that is proposed for $[8\text{Ser}+\text{H}]^+$ must be consistent with and account for these results.

Fundamentally, crystal formation begins in solution when molecules assemble themselves in a highly ordered fashion. This molecular organization is selective, and can lead to spontaneous symmetry breaking.⁴ The serine octamer demonstrates selectivity which leads to homochiral molecular clusters. Might the homochiral preference of the serine octamer be derived from solution aggregates that are precursors to the formation of crystalline serine?



4.4

Crystalline L-serine is bound through zwitterionic salt bridges and by rows of hydrogen bonds between the hydroxy-methyl side chains.²⁶ Structure 4.4 is derived from the crystal structure of L-serine, where four of the hydroxy-methyl groups have been

rotated by $\sim 120^\circ$ relative to the crystal structure. The resulting structure has two rows of salt-bridges formed by zwitterionic serine in addition to hydrogen bonding between each hydroxy-methyl side chain and the carboxylate group of the serine in the opposite row. The minimized protonated gas phase structure as determined by DFT optimization is shown in **4.4**. The addition of a proton causes the structure to curl relative to the more linear neutral structure.

Structure **4.4** is formed from zwitterionic serine, which is the predominant form of serine in both solution and the solid phase. Would a zwitterionic structure also be stable in the gas phase? Table 4.1 demonstrates the progressive stabilization of zwitterionic serine with increased cluster size for a model structure similar to **4.4**, in the absence of any net charge. The structures are stabilized in the zwitterionic state by the coulombic attraction that is derived from the two rows of salt bridges. This coulombic attraction compensates for the energy required to generate zwitterionic serine, which is not the lowest energy structure for the isolated serine molecule in the gas phase. If the coulombic attraction overcomes the “zwitterion penalty,” the zwitterionic state of the cluster will be favored. A net charge has been shown previously to enhance the stability of zwitterions in the gas phase.²⁷ Our calculations indicate that structures similar to **4.4**, with varying numbers of serines in the chain, would be stable as zwitterions with eight or more serine molecules, even with errors of 5-10 kcal/mol in the calculations.

Table 4.1 Progressive Stabilization of Structure 4.4 at Various Chain Lengths.

Structure	Zwitterion ^a (kcal/mol)	Neutral ^a (kcal/mol)	Δ (kcal/mol)	Δ/Ser (kcal/mol)
2mer	-500335	-500342	7	3.5
4mer	-1000716	-1000726	10	2.5
6mer	-1501103	-1501094	-10	-1.6
8mer	-2001491	-2001468	-23	-2.9
10mer	-2501870	-2501825	-45	-5.7

a.) Total energy as determined by single point calculations on the PM3 minimized neutral structures at the 6-31G//B3LYP level, corrected for zero point energies.

Thus **4.4** is a reasonable candidate structure for the serine octamer because it can exist both in solution and the gas phase. Furthermore, **4.4** is a derivative of the crystal structure of L-serine, and as a result displays a strong preference for homochirality. It is not possible to change the chirality of four of the serines and generate a cluster composed of 4 D-serines with 4 L-serines (a 4D/4L cluster). The highly ordered bonds of **4.4** suggest that the structure may be very stable. This is confirmed in Table 4.2, which lists the calculated gas phase cross sections with the accompanying solution and gas phase energetics of several possible structures for the serine octamer. In solution, **4.4** has the best binding energy of the three structures. However, the calculated cross section for the protonated structure is 224 \AA^2 , which is much greater than the experimental value of 187 \AA^2 . Therefore **4.4** must be ruled out as the *gas phase* structure for the serine octamer.

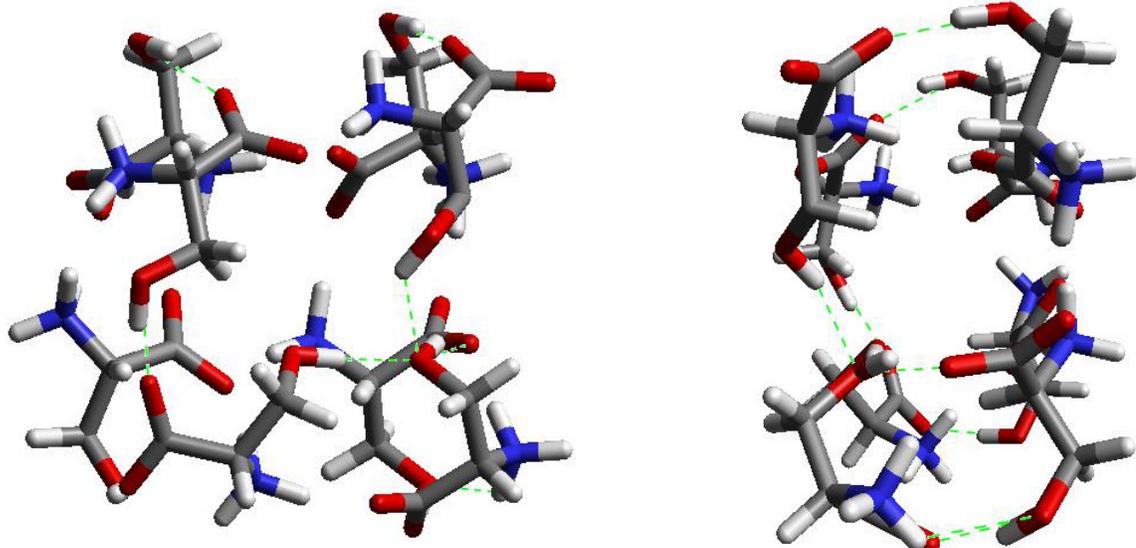
Structure **4.5** is a cubic arrangement of eight serines that is closely related to **4.4**. Structure **4.5** is formed by taking **4.4** and essentially folding it in half to form a cube like structure. The structure has four rows of salt bridges, two interior and two exterior. The structure shown is the minimized unprotonated structure. The cross section calculated for

the neutral structure **4.5** is in reasonable agreement with the experimental results (Table 4.2). However, structure **4.5** undergoes significant reorganization in the gas phase when protonated. The protonated structure will not properly minimize, suggesting that it is significantly unstable in the gas phase. Furthermore, the two interior rows of zwitterions are not well exposed, leading to unfavorable solvation energetics which substantially decrease the solution phase binding energy (Table 4.2). Additionally, a similar heterochiral 4D/4L structure exists. Therefore, it is unlikely that **4.5** is the structure for the serine octamer, but it may be an important intermediate structure as will be explained below.

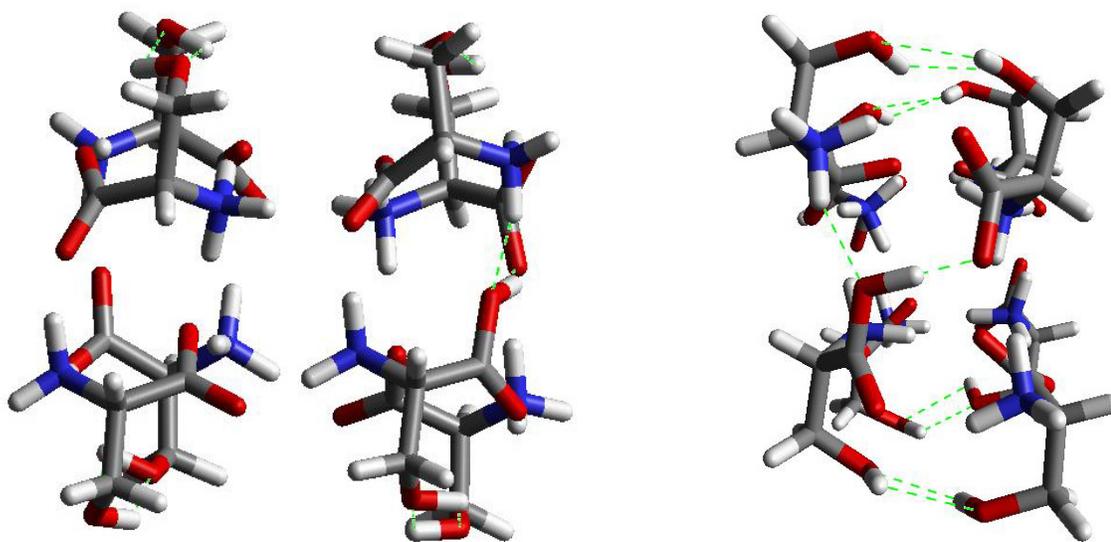
Table 4.2 Calculated Energetics and Cross Sections for Structures 4-6.

Structure	Relative Gas Phase Energy (kcal/mol) ^a	Computed Cross Section (Å ²)	Relative Gas Phase Binding Energy (kcal/mol) ^{a,d}	Relative Solution Phase Binding Energy (kcal/mol) ^e
4	76	224 ^b	112	-55
5	-- ^f	196 ^c	--	-16
6	0	189 ^b	0	0

- a) Energies listed relative to **6** for protonated octamers
- b) Cross sections for protonated structures to be compared with an experimental value of 187 Å²
- c) Cross section calculated for neutral structure (the neutral cross sections for **4.4** and **4.6** are 227 and 189 Å², respectively)
- d) Binding energy calculated relative to eight separated zwitterionic serines
- e) Energies calculated for neutral octamers in water with a dielectric constant of 80.37 and a probe radius of 1.4
- f) This structure is not well behaved and does not fully minimize.



4.5



4.6

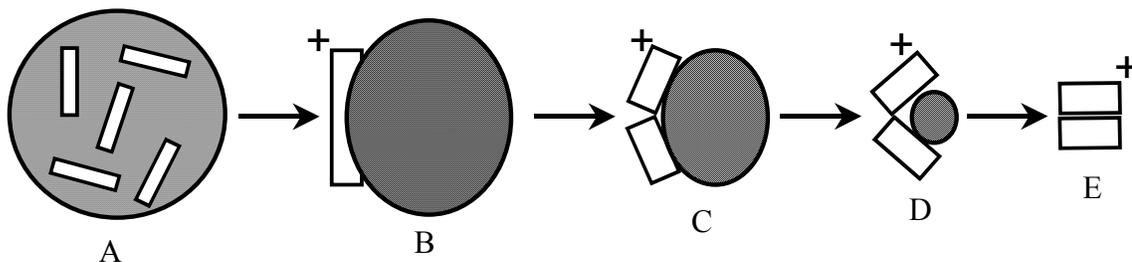
Reorganization of structure 4.5 in the gas phase leads to structure 4.6.¹⁸ The minimized gas phase protonated structure of 4.6 is shown, but it is virtually identical to the neutral structure. All eight serines are bound via a central zwitterionic core. The resulting cubic structure is further stabilized by hydrogen bonding between facing pairs of serines. The computed cross section for 6, 189 \AA^2 , is in excellent agreement with the

experimental value of 187 \AA^2 (Table 4.2). The gas phase binding energy of this structure is the best of any of the candidates that we have tested (Table 4.2). Given that the zwitterionic core of structure **4.6** would be essentially buried from any solvent, it has the worst solution phase binding energy of the three structures. Furthermore, a 4D/4L cluster can be assembled with a zwitterionic core and binding energy similar to that for structure **4.6**. For these reasons, **4.6** is not likely to be the solution phase structure for the serine octamer. However, **4.6** represents the most likely gas phase structure, particularly in light of the excellent gas phase energy and agreement with the experimental cross section.

Of the experimental constraints noted above for the structure of the serine octamer, (2) and (3) are the most difficult to simultaneously satisfy. A three-point interaction between each serine and the rest of the cluster would constitute the minimal requirement for chiral recognition.^{28,29} It is easy to generate a variety of clusters that satisfy this requirement, but such clusters do not necessarily satisfy requirement (2). Given that eight serines must be arranged in an unusually stable structure with a cross section of 187 \AA^2 , cubic motifs (such as **4.5** and **4.6**) seem the most consistent with the ion mobility data. However, the symmetry of a cube generates additional problems that must be taken into account. For example, a cubic octamer connected at four points through the neutral carboxylic acids (a structure actually proposed in a concurrent work by Cooks and coworkers³⁰ during the preparation of this manuscript) can be assembled from homochiral serine or from 4D and 4L serines. Although each serine is held in a three-point interaction in both clusters and the energetics are nearly identical, requirement (2) is clearly not satisfied by such a structure. Furthermore the data presented in Figure 4.4 clearly demonstrates the extremely low abundance of 4D/4L clusters. In fact, it is entirely likely that all cubic

structures (the only reasonable geometry that satisfies the experimental cross section) will have an accompanying 4D/4L cluster.

The combination of requirements (2) and (3) suggests that a non-cubic structure, such as **4.4**, exists in solution which becomes cubic in the gas phase. This is depicted in cartoon fashion in Scheme 4.1. In general terms, **A** depicts various non-cubic nanocrystals of serine forming in bulk solution. In **B**, a charged serine octamer is driven to the surface of a highly charged droplet (other charges not shown for clarity). In **C** and **D**, as the number of solvent molecules is reduced (either through ion evaporation or evaporation of the solvent), the octamer begins to collapse on itself. The fully desolvated protonated serine octamer with a cubic structure is shown in **E**. In this manner the requirements for homochirality and a cubic structure can be satisfied simultaneously.



Scheme 4.1

Singly protonated structures similar to **4.4** (i.e., 6mers and 10mers) may also form in solution, however their individual transfer to the gas phase is not as favorable because only the octamer can form a cube. Aggregation of these singly charged species into multiply charged multimers (such as the multimers of the octamer shown in Figure 4.2) allows for enhanced stabilization in the gas phase. Thus larger, multiply charged clusters such as those shown in Figure 4.1 and Figure 4.3 are also present in the gas phase. However, the stability of these multimers is derived from their size, not from their

structure. Of the *singly* charged species the octamer is king. In the case of the octamer, the unusual abundance is attributed to the special stability of the cubic structure.

Both homoserine and threonine can form structures analogous to **4.4**. The crystal structure of homoserine allows for a homoserine analog of **4.4** with very little rearrangement.³¹ However, in the case of threonine, substantial rearrangement from the crystal structure is necessary.³² This may account for the reduced relative abundance of the octamer for threonine (Figure 4.8). Notwithstanding, the facile incorporation of threonine into the octamer as shown in Figure 4.9b and Figure 4.9c is compatible with the solution phase structure given in **4.4** and the gas phase structure shown **4.6**. It is anticipated that the clusters producing the spectra for homoserine and perhaps threonine are largely analogous to the structures that lead to the serine spectrum. The somewhat higher proton affinity of threonine and homoserine may be related to the observation that the cluster distributions of the analogs sometimes exhibit higher charge states than those observed for serine itself.

4.5 Conclusion

We have conducted a number of experiments designed to elucidate the structure of the serine octamer, and explain its abundance in electrospray ionization mass spectrometry. The octamer demonstrates a strong preference for homochirality. Experiments using L-serine methyl ester and L-N-tertbutoxycarbonyl-serine (L-tboc serine) show no cluster formation, indicating the amino and carboxylate functionalities are necessary for formation of the cluster. Ion mobility data indicates the cross section of the octamer to be 187 \AA^2 ($\pm 2\%$). DFT calculations and the experimental data lead to a structure that is

derived from solution aggregates that precede formation of crystalline serine (Structure 4.4). Structure 4.4 provides the necessary enantiomeric discrimination observed in the serine octamer and the folding of 4.4 into 4.5, which relaxes to 4.6, enables the octamer to conform to the measured cross-section with an energetically favored structure. Both threonine and homoserine may form structures analogous to the serine octamer. Threonine may incorporate freely into serine clusters because the additional methyl group does not interfere with the bonding of the cluster.

The combination of the solution and gas phase properties of the serine octamer leads to its unusual abundance when sampled by mass spectrometry. We have taken advantage of this unusual abundance to demonstrate clearly that small molecular clusters can demonstrate a preference for homochirality. This offers a molecular cluster parallel to the macroscopic observation of chiral symmetry breaking through crystal formation, and a possible pathway for the establishment of prebiotic homochirality. In appropriate energetic or chemical environments, covalent coupling of the components of the clusters could occur, creating homochiral polymers or macromolecules with well-defined stereochemistry. Such stereoregular molecules could then serve as templates in reproductive chemical systems. No inherent preference is given to a particular enantiomer in the formation of these clusters, and therefore an additional mechanism would be required to explain the eventual elimination of one enantiomer.

On a different note, ESI may offer a new experimental technique for investigating the early stages of homogeneous crystal nucleation in solution. A universal theory for the explanation of homogeneous nucleation is still lacking.³³ Several spectroscopic methods have been employed to study crystal nucleation,³⁴ but mass spectrometry offers the

additional ability to sample small clusters and study them in the gas phase. As appears to be the case in the present study, the solution phase structures of the clusters may not be retained in the gas phase. The difference between the solution and gas phase cluster distributions will depend on the detailed mechanism by which they are transferred from solution to the gas phase in the ESI process. The ESI process may increase the monomer concentration through evaporation prior to droplet fissioning. This could conceivably induce greater cluster formation in a super saturated and cooled droplet. If ESI does enhance the formation of nanocrystals, then one could conceivably devise a crystal-seeding source designed to initiate crystallization.

4.6 References

-
- ¹ (a) Podlech, J. *Cell. Mol. Life Sci.* **2001**, *58*, 44-60. (b) Editor, David B. Cline *Physical origin of homochirality in life : Santa Monica, California, February 1995* Woodbury, New York : American Institute of Physics, 1996.
- ² (a) Avalos, M. A.; Babiano, R.; Cintas, P.; Jimenez, J. L.; Palacios, J. C. *Chem. Commun.* **2000**, *11*, 887-892. (b) Soloshonok, V. A.; Svedas, V. K.; Kukhar, V. P.; Galaev, I. Yu.; Kozlova, E. V.; Svistunova, N. Yu. *Bioorg. Khim*, **1993**, *19(4)*, 478-483.
- ³ Buschmann, H.; Thede, R.; Heller, D. *Angew. Chem. Int. Ed.* **2000**, *39(22)*, 4033-4036.
- ⁴ Enantiomers, Racemates, and Resolutions, Jacques, J., Wiley: New York, 1981, 217-434.
- ⁵ Julian, R. R.; Hodyss, R.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2001**, *123*, 3577-3583.
- ⁶ Julian, R. R.; Beauchamp, J. L. *Int. J. Mass Spectrom.* **2001**, *210*, 613-623.
- ⁷ Schalley, C. A. *Int J. Mass Spectrom.* , **2000**, *194*, 11-39.
- ⁸ Hao, C. Y.; March, R. E.; Croley, T. R.; Smith, J. C.; Rafferty, S. P. *J. Mass Spectrom.* **2001**, *36*, 79-96.
- ⁹ Vekey, K.; Czira, G. *Anal. Chem.* **1997**, *69*, 1700-1705.
- ¹⁰ Yao, Z.-P.; Wan, T. S.; Kwong, K. -P.; Che, C. T. *Chem Comm.* **1999**, 2119-2120.
- ¹¹ Nikolaev EN, Denisov EV, Rakov VS, Futrell JH *Int J. Mass Spectrom.* **1999**, *183*, 357-368.
- ¹² Tao, A. T.; Zhang, D.; Wang, F.; Thomas, P. D.; Cooks, R. G. *Anal. Chem.* **1999**, *71*, 4427-4429.
- ¹³ (a) Sawada, M.; Takai, Y.; Yamada, H.; Nishida, J.; Kaneda, T.; Arakawa, R.; Okamoto, M.; Hirose, K.; Tanaka, T.; Naemura, K. *J. Chem. Soc. Perkin Trans. 2* **1998**,

701-710. (b) Sawada, M.; Takai, Y.; Yamada, H.; Hirayama, S.; Kaneda, T.; Tanaka, T.; Kamada, K.; Mizooku, T.; Takeuchi, S.; Ueno, K.; Hirose, K.; Tobe, Y.; Naemura, K. *J. Am. Chem. Soc.* **1995**, *117*, 7726-7736. (c) Sawada, M.; Takai, Y.; Yamada, H.; Kaneda, T.; Kamada, K.; Mizooku, T.; Hirose, K.; Tobe, Y.; Naemura, K. *Chem. Commun.* **1994**, 2497-2498.

¹⁴ (a) So, M. P.; Wan, T. S. M.; Chan, T.-W. D. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 692-695. (b) Ramirez, J.; He, F.; Lebrilla, C. B. *J. Am. Chem. Soc.* **1998**, *120*, 7387-7388.

¹⁵ Sawada, M.; Shizuma, M.; Takai, Y.; Adachi, H.; Takeda, T.; Uchiyama, T. *Chem. Commun.* **1998**, 1453-1454.

¹⁶ Krishna, P.; Prabhakar, S.; Manoharan, M.; Jemmis, E. D.; Vairamani, M. *Chem. Commun.* **1999**, 1215-1216.

¹⁷ Clemmer, D. E.; Jarrold, M. F. *J. Mass Spectrom.* **1997**, *32*, 577-592.

¹⁸ Hodyss, R.; Julian, R. R.; and Beauchamp, J. L. *Chirality* **2001**, *13*, 703-706.

¹⁹ Cooks, R.G. et al. Clustering of amino acids in the gas phase by electrospray ionization mass spectrometry. Proceedings of the 48th ASMS conference on mass spectrometry and allied topics. **2000**, p.1361-1362.

²⁰ Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Phys. Chem. A* **1996**, *100*, 16082-16086.

²¹ Dugourd, P.; Hudgins, R. R.; Clemmer, D. E.; Jarrold, M. F. *Rev. Sci. Instrum.* **1997**, *68*, 1122-1129.

²² Mason, E. A.; McDaniel, E. W. *Transport Properties of Ions in Gases*; Wiley: New York, 1988.

-
- ²³ Weis, P.; Kemper, P. R.; Bowers, M. T. *J. Phys. Chem. A* **1997**, *101*, 8207-8213.
- ²⁴ Counterman, A.E.; Valentine S. J.; Srebalus, C. A.; Henderson, S. C.; Hoaglund, C. S.; Clemmer, D. E. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 743-759.
- ²⁵ Counterman, A. E.; Clemmer, D. E. *J. Phys. Chem. B* **2001**, *105*, 3646.
- ²⁶ Kistenmacher, T. J.; Rand, G. A.; Marsh, R. E. *Acta Cryst.* **1974**, *B30*, 2573-2578.
- ²⁷ (a) Jockusch, R. A.; William, P. D.; Williams, E. R. *J. Phys. Chem. A* **1999**, *103*, 9266-9274. (b) Cerda, B. A.; Wesdemiotis, C. *Analyst* **2000**, *125(4)*, 657-660. (c) Jockusch, R. A.; William, P. D.; Williams, E. R. *J. Am. Chem. Soc.* **1997**, *119*, 11988. (d) Wyttenbach, T.; Witt, M.; Bowers, M. T. *Int. J. Mass Spec.* **1999**, *182/183*, 243.
- ²⁸ Tao W.A.; Zhang D. X.; Nikolaev E. N.; and Cooks R. G. *J. Am. Chem. Soc.* **2000**, *122*, 10598-10609.
- ²⁹ Ahn, S.; Ramirez, J.; Grigorean, G.; Lebrilla, C. B. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 278-287.
- ³⁰ Cooks, R. G.; Zhang, D.; Koch, K. J.; Gozzo, F. C.; Eberlin, M. N. *Anal. Chem.* **2001**, *73*, 3646-3655.
- ³¹ Chacko, K. K.; Swaminathan, S.; Veena, K. R. *Crys. Struct. Commun.* **1982**, *11*, 2057.
- ³² Janczak, J.; Zobel, D.; Luger, P. *Acta Cryst. Sec. C.* **1997**, *53*, 1901.
- ³³ Granasy, L.; Igloi, F. *J. Chem. Phys.* **1997**, *107*, 3634-3644.
- ³⁴ (a) Granasy, L.; James, P. F. *J. Chem. Phys.* **2000**, *113*, 9810-9821, (b) Peng, X.; Wickham, J.; Alivisatos, A. P. *J. Am. Chem. Soc.* **1998**, *120*, 5343-5344.

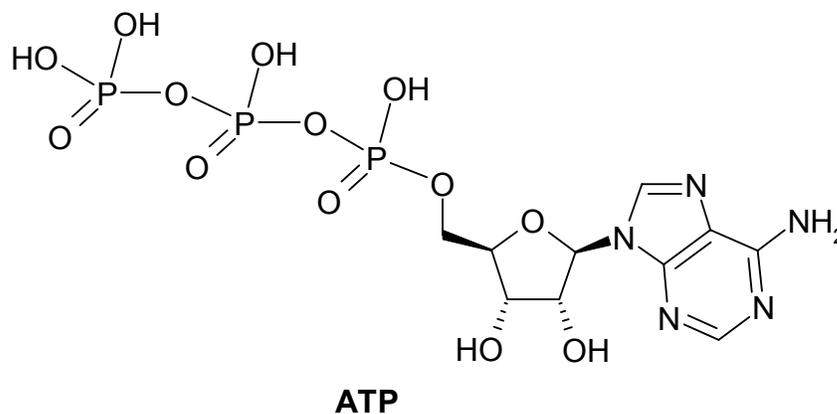
Chapter 5

Abiotic Synthesis of ATP from AMP in the Gas Phase: Implications for the Origin of Biologically Important Molecules from Small Molecular Clusters

Published previously in: Julian, R. R.; Beauchamp, J. L. *Int. J. Mass Spectrom.* **2003**, 227, 147-159.

5.1 Introduction

Adenosine 5'-triphosphate (ATP) is a ubiquitous biological molecule used for storing energy in the form of chemical bonds. The cellular machinery for creating ATP has been found in the oldest forms of life, yet it is still found in all groups of present day organisms.¹ Despite this ancient lineage and the extreme importance of ATP in many biological processes, it is surprising that the molecular mechanism for the synthesis of ATP is still open to debate.^{2,3} Even more puzzling is how ATP originally became the primary energy currency for all of life.⁴ In order to develop a more complete picture of the role of ATP in the origin of life, it is of interest to identify prebiotic routes that may have possibly contributed to the generation of ATP.



Several explanations for the terrestrial synthesis of thermodynamically unstable phosphoanhydride bonds have been given^{5,6} and disputed.⁷ The primary difficulty associated with the production of ATP on earth is that this type of chemistry is unlikely to occur in water. Not only is the hydrolysis of ATP exothermic,⁸ but the oceans of the primitive world contained pyrite, which catalyses the hydrolysis of phosphoanhydride bonds.⁹ Rigorously non-aqueous organic solvents may be used for the synthesis of phosphoanhydride bonds,¹⁰ but the availability of such environments may have been limited in the early earth. Perhaps the best approach to phosphoanhydride synthesis involves chemistry in which no solvent is present, a suggestion which will be explained shortly.

Space is an environment which can easily accommodate conditions which are devoid of solvent. The presence of organic molecules in comets and meteorites and the possible role of these organics in the origin of life have drawn much attention over the years.¹¹ Related work has shown that it is possible to synthesize nucleotides such as adenosine 5'-monophosphate (AMP) by simply exposing the precursor molecules on a thin film to open space.¹² Interestingly, all of the components necessary to synthesize ATP have been detected in meteorites. A wide variety of sugars and their acid and alcohol counterparts

have been detected in the Murchison and Murray meteorites.¹³ Phosphates and alkyl phosphonic acids in many forms have been detected in various meteorites.^{14,15} Nucleobases and other nitrogen heterocycles, including adenine, have been detected in meteorites as well,¹⁶ and they have been synthesized in experiments approximating the conditions in space.¹⁷

Terrestrial phenomena can also account for the extraction of key chemical components into a solvent free state. For example, it has been suggested that aerosol particles produced by sea spray may have formed cell like reactors in which geochemistry may have evolved into biochemistry.¹⁸ Interestingly, organics from meteorites have been observed to transfer into these aerosol particles in high concentrations.¹⁹ However, if the organic content of an aerosol particle is low and thus insufficient to form a protective monolayer, then as the particle travels to a region of low humidity the water will evaporate away. This could eventually lead to small, desolvated salt clusters. This process is fundamentally analogous to sonic spray ionization (SSI) experiments which can be monitored by mass spectrometry in a controlled environment.^{20,21} It is interesting to note that SSI experiments lead to a great enhancement in the formation of clusters when compared to other ionization methods.²²

The gas phase study of non-covalently bound clusters is a rapidly expanding area in the field of mass spectrometry.^{23,24,25} The chemistry of small, noncovalently bound clusters and the role that they may have played in the prebiotic origin of biopolymers is the subject of the present work. It is demonstrated that trimers of AMP will yield ATP exclusively under the appropriate conditions. Collision activated dissociation (CAD) is utilized to initiate the reactions.²⁶ Furthermore, this chemistry is general in nature and can

be applied to the generation of other polyphosphates which are also of biological significance. It is postulated that other linear polymers of biological importance could be generated from small clusters in a similar manner.

5.2 Experimental Methodology

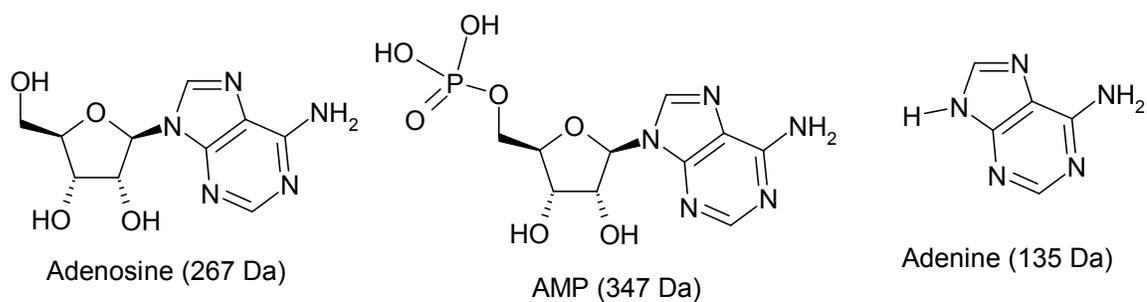
Mass spectra were obtained using a Finnigan LCQ ion trap quadrupole mass spectrometer without modification. The signal was optimized using the automatic tuning capabilities of the LCQ. To enhance cluster formation, the following settings were used for cationic clusters: source voltage 4.15 kV, capillary voltage 5.0 V, capillary temperature 200 °C, and tube lens offset -40 V. Collision activated dissociation was performed on isolated ions by applying a resonance excitation RF voltage of between 0.98 V and 2.45 V for a period of 30 ms.

Sample concentrations were in the range of ~50 μM . Samples were electrosprayed using a 80:20 methanol/water mixture at a flow rate of 3 $\mu\text{L}/\text{min}$ from a 250 μL Hamilton syringe. Silica tubing with an inner diameter of 12.7 microns inside a stainless steel sheath was used for electrospraying the mixture. Compounds were purchased commercially and used without further purification.

5.3 Results

Anionic Clusters. The negative ion ESI-MS spectrum for the sodium salt of AMP is shown in Figure 5.1a. The two major peaks correspond to $[\text{AMP-H}]^-$ and $[\text{2AMP-H}]^-$. Clusters incorporating one or more sodium ions replacing protons are present in smaller abundance. Isolation of $[\text{3AMP-H}]^-$ followed by CAD leads exclusively to the loss of

neutral AMP, as shown in Figure 5.1b. The results differ marginally when one hydrogen is replaced by sodium. The CAD spectrum for $[3\text{AMP}-2\text{H}+\text{Na}]^-$ is shown in Figure 5.1c. The primary product is the loss of neutral AMP. However, there is also a minor loss of 267 Da, which corresponds to the loss of an adenosine nucleoside. In other words, one AMP has been fragmented, releasing neutral adenosine and leaving PO_3 with the cluster. The structures for AMP and two important fragments are given below.



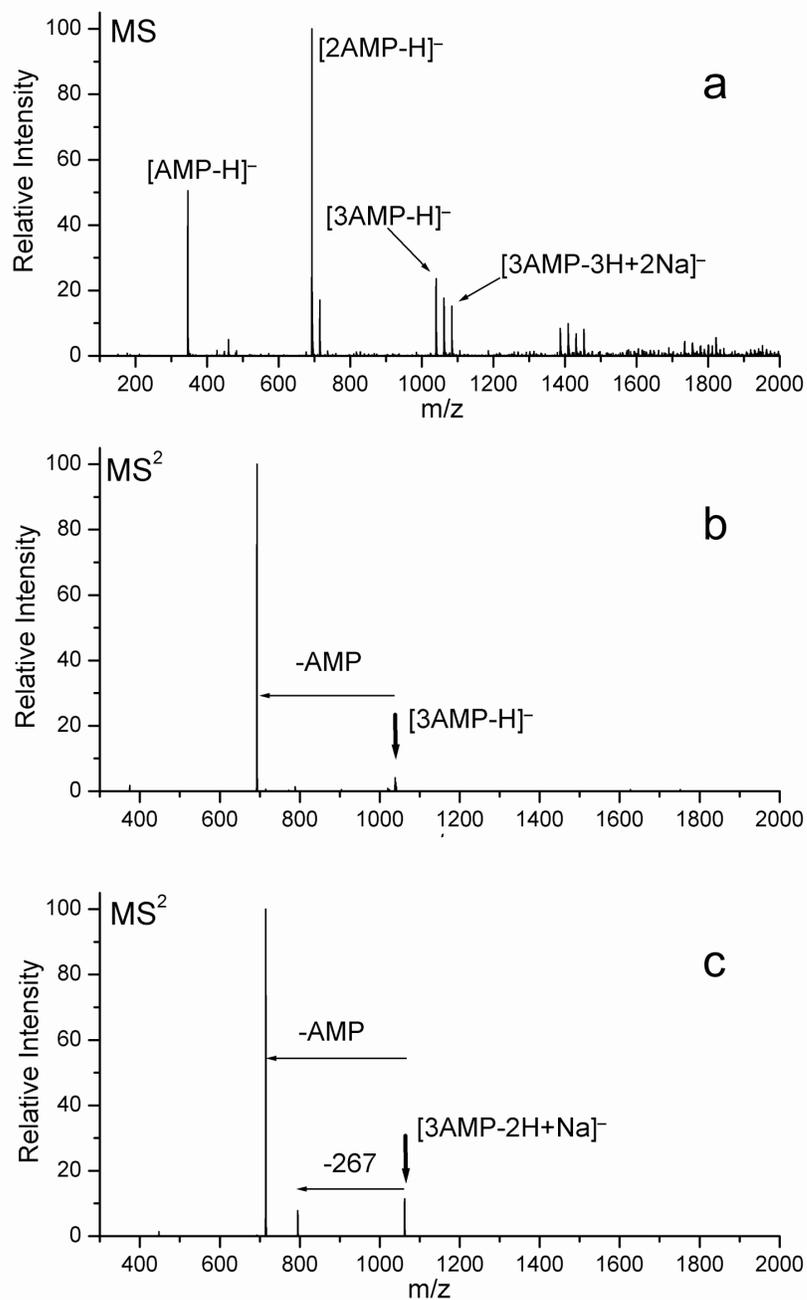


Figure 5.1 (a) Anionic clusters of AMP with sodium. (b) CAD spectrum of the deprotonated AMP trimer showing that no covalent bonds are broken. (c) CAD spectrum of AMP trimer with inclusion of one sodium ion, in which a minor peak corresponding to the fragmentation of one AMP (namely, loss of adenosine) is observed.

Figure 5.2 shows the results for the sequential CAD of the $[3\text{AMP-3H}+2\text{Na}]^-$. The loss of adenosine leads to the base peak in Figure 5.2a. Some additional loss of a second adenosine is observed in Figure 5.2a as well. The primary product is isolated in Figure 5.2b. Further CAD in Figure 5.2c leads exclusively to the loss of the second adenosine. Tentatively, the results in these spectra correspond to the transformation of 3AMP into AMP+ADP, which is then converted into ATP, as will be discussed in greater detail below.

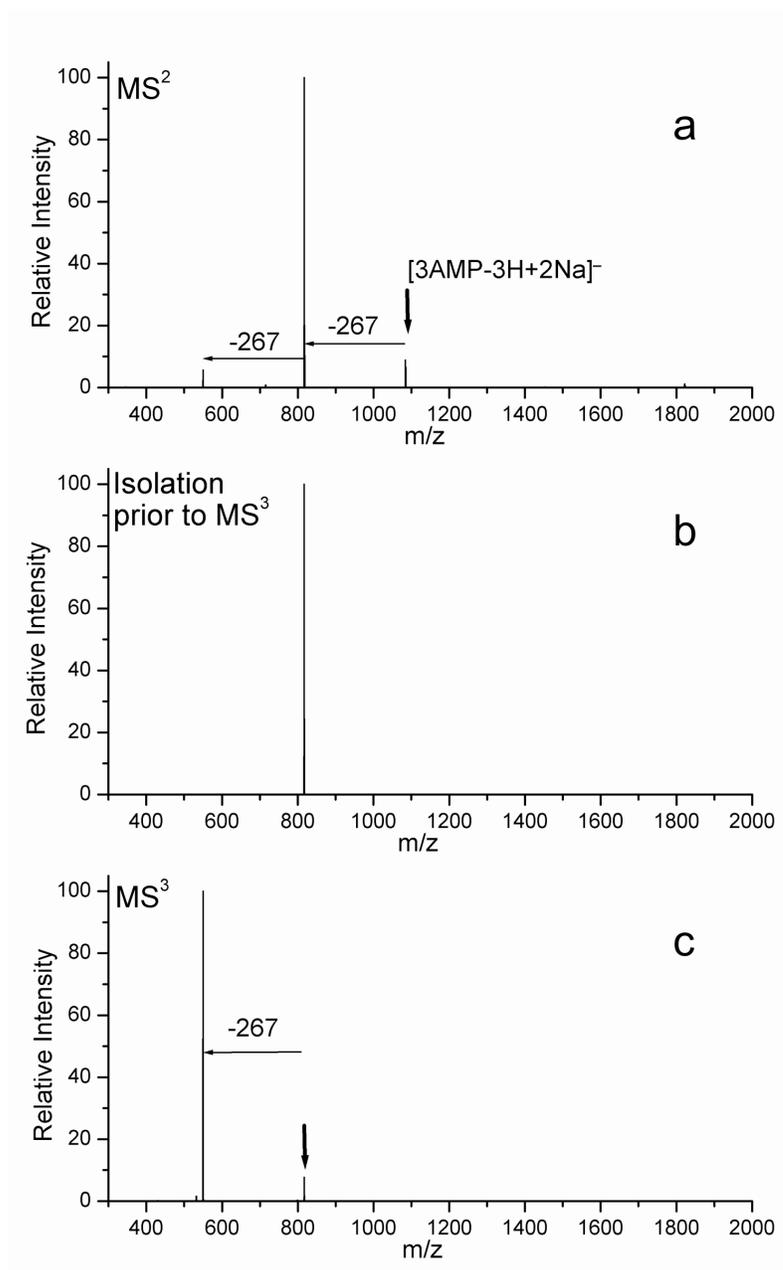


Figure 5.2 (a) CAD spectrum of the anionic sodiated salt cluster of the AMP trimer. The loss of adenosine is predominant and no dissociation of the cluster by loss of AMP is observed. (b) Isolation of the product peak. (c) MS³ spectrum demonstrating the loss of a second adenosine and the generation of ATP.

Three additional collisional activation steps are performed on the primary product from Figure 5.2c in Figure 5.3. The primary peak in Figure 5.3a results from the loss of 18 Da (water). An additional loss of 102 Da is also observed. In Figure 5.3b, the adenine base is lost, leading to the base peak in the spectrum. This loss is accompanied by the pickup of water and methanol, yielding peaks with masses that are 18 Da and 32 Da higher than the daughter fragment, respectively.²⁷ Isolation of the daughter fragment does not eliminate these peaks, confirming that they are picked up from residual solvent in the trap. A competitive loss of 102 Da is also observed. Isolation of the most abundant product, followed by collisional activation yields primarily a loss 102 Da in the MS⁶ spectrum shown in Figure 5.3c. An additional minor peak corresponding to the loss of a second 102 Da is also present in Figure 5.3c. This 102 Da most likely corresponds to neutral loss of NaPO₃.

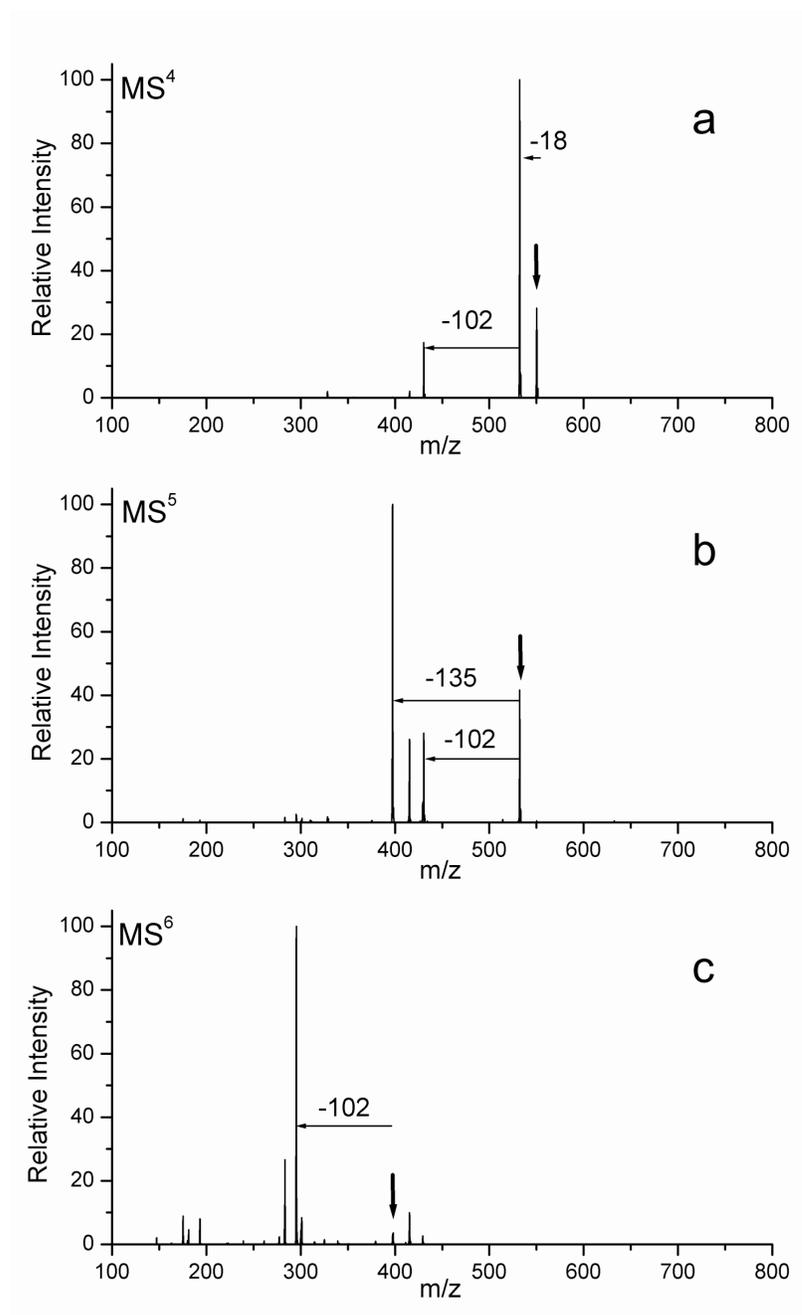


Figure 5.3 Three further collisional activation steps on the product (ATP) generated in Figure 5.2c. (a) The lowest energy pathway leads to the loss of water. (b) Adenine is lost next. (c) Sequential losses of 102 Da or NaPO₃ are the final losses that are observed in this MS⁶ spectrum.

The sequential CAD of an authentic sample of ATP is shown in Figure 5.4. In Figure 5.4a, isolation of $[\text{ATP-3H+2Na}]^-$ is accompanied by the isobaric doubly charged dimer $[\text{2ATP-6H+4Na}]^{2-}$. Therefore, several of the daughter peaks in the CAD spectrum in Figure 4a result from the dissociation of the doubly charged dimer and are marked with a double asterisk. The loss of water leads to the base peak, just as in Figure 5.3a. In Figure 5.4b, interference from doubly charged peaks has been eliminated, and the spectrum is nearly identical to that shown in Figure 5.3b. Similarly, Figure 5.4c is very comparable to Figure 5.3c. In fact, even the majority of the minor peaks in both Figures 5.4b and 5.4c are identical to those found in Figures 5.3b and 5.3c.

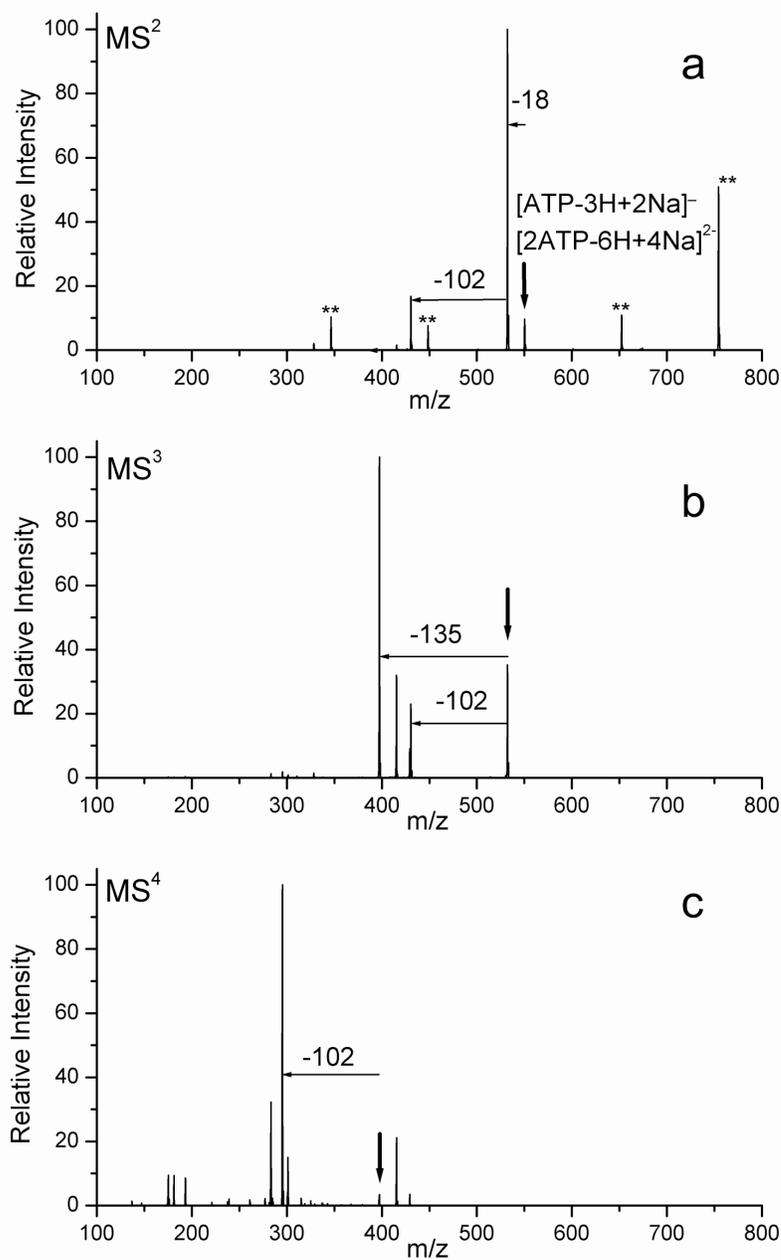


Figure 5.4 Three CAD steps on authentic ATP, $[\text{ATP-3H+2Na}]^-$. (a) Some interference from the doubly charged dimer leads to additional peaks. (b) The loss of adenine is very abundant, as is also observed in Figure 5.3b. (c) Loss of 102 Da is observed in a similar fashion as Figure 5.3c, suggesting that the product produced in Figure 5.2c is in fact ATP. ** Daughter ions from the doubly charged dimer are indicated by a double asterisk.

Cationic Clusters. Figure 5.5 shows the positive ion ESI spectrum for the sodium salt of AMP. Clusters with a sodium cation providing the net charge are easily generated. Clusters incorporating up to five AMP molecules are observed. The maximum intensity for each cluster distribution corresponds to the peak fitting $[x\text{AMP}-x\text{H}+(x+1)\text{Na}]^+$, where $x = 1-5$. Formally then, each AMP is deprotonated and coupled with a sodium cation, forming salt clusters.

In Figure 5.5b, results are presented for the collisional activation of the cluster $[3\text{AMP}-3\text{H}+4\text{Na}]^+$. The loss of adenosine yields the most abundant product. Isolation of this product, followed by further collisional activation, leads primarily to the loss of another adenosine as seen in Figure 5.5c. These losses are identical to those observed in Figure 5.2 for the anionic clusters. The base peak produced in Figure 5.5c is isobaric with $[\text{ATP}-3\text{H}+4\text{Na}]^+$. Further CAD of this peak yields products that are indistinguishable from those produced by CAD of $[\text{ATP}-3\text{H}+4\text{Na}]^+$ using an authentic ATP sample. The results for the cationic clusters closely parallel those obtained for the anionic clusters.

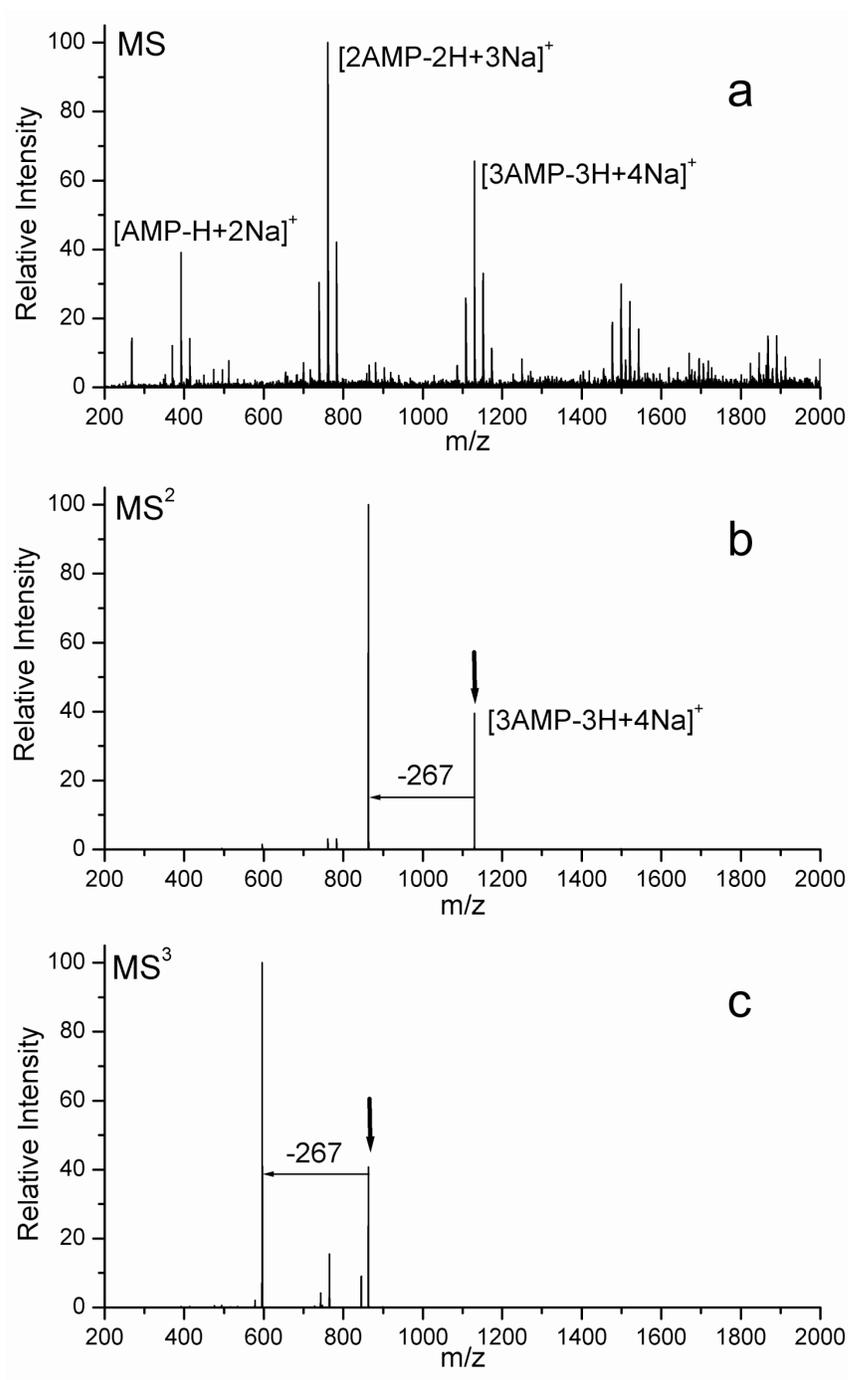


Figure 5.5 (a) Sodiated cationic salt clusters of AMP. (b) CAD of the trimer leads to the formation of ADP+AMP. (c) A second adenosine is lost, implying that ATP is generated in high yield by a reaction similar to that observed for the anions.

Phosphate Clusters. The CAD spectra for clusters of phosphate are shown in Figure 5.6. The sodiated cluster $[3\text{H}_2\text{PO}_4+\text{Na}]^-$ loses water exclusively as shown in Figure 5.6a. Subsequent activation of the daughter ion results in the loss of another water and the formation of triphosphate as shown in Figure 5.6b. This loss is accompanied by the pick up of water and methanol in low abundance. It is unclear whether the added solvent molecules react with the cluster or are merely coordinated to it. In Figure 5.6c, $[3\text{H}_3\text{PO}_4-\text{H}]^-$ is subjected to CAD. In the absence of the sodium ions, the only observed product corresponds to a loss of H_3PO_4 . These results closely parallel those observed for the AMP clusters.

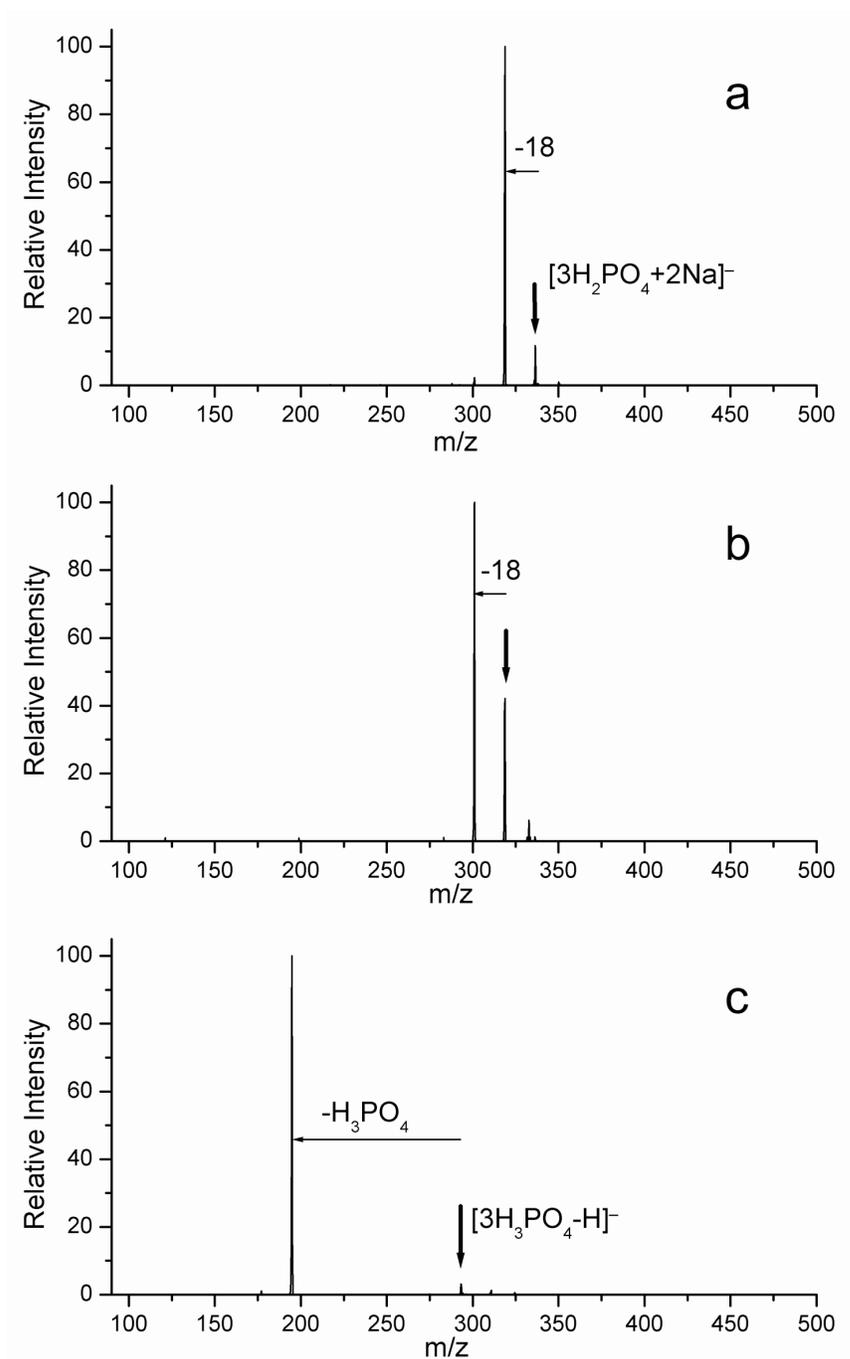


Figure 5.6 (a) CAD spectrum of $[3H_2PO_4 + Na]^-$ which loses water to form a phosphoanhydride. (b) CAD of the daughter ion from (a), yielding triphosphate from the loss of second water. (c) CAD of $[3H_3PO_4 - H]^-$. The hydrogen bound cluster simply dissociates by the loss of H_3PO_4 .

Calculations. The enthalpic contributions to reaction 5.1 were calculated at the PM5 semi-empirical level and utilizing density functional theory at the BLYP/DZVP and B3LYP/6-31G** levels. The results are given in Table 5.1. Methylphosphate is used instead of AMP to evaluate the binding energies of the clusters in order to allow for higher levels of theory to be utilized in a reasonable time. This approximation is supported by the experimental results shown in Figure 5.6 which suggest that most of the chemistry takes place in the phosphate/sodium salt portion of the clusters. The calculations reveal that metal bound phosphate salt clusters are more strongly bound in the gas phase than the proton bound counterparts. The PM5 methodology is sufficient to qualitatively predict the results, although the energetic ordering among the alkali metals is not correct.

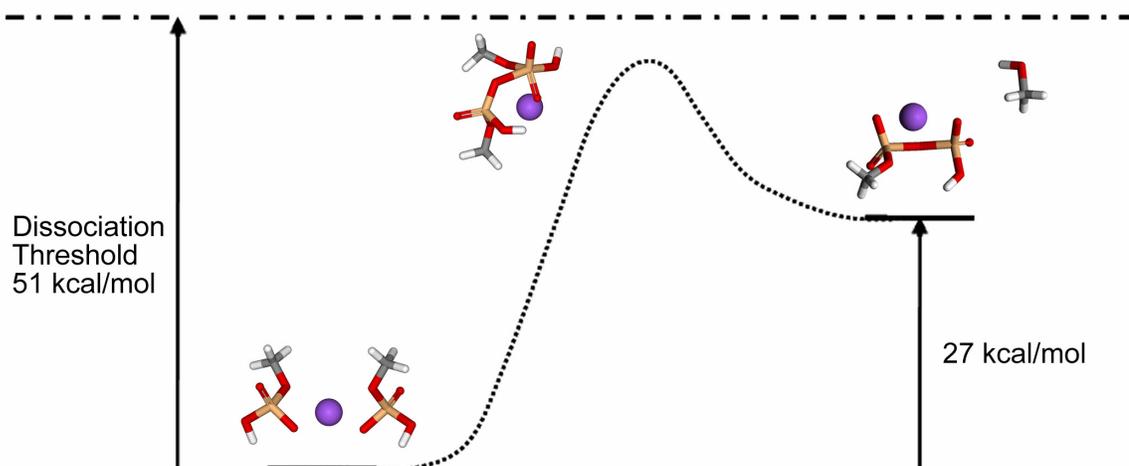


Table 5.1 Calculated Energetics in kcal/mol

Cluster	Binding Energy PM5	Binding Energy BLYP/DZVP	Binding Energy B3LYP/6-31G**
$(CH_3OPO_3H)_2H^-$	-22.36	-34.2309	-40.2867
$(CH_3OPO_3H)_2Li^-$	-46.37	-48.1216	-54.807
$(CH_3OPO_3H)_2Na^-$	-44.95	-44.9093	-50.7288
$(CH_3OPO_3H)_2K^-$	-46.71	-46.0386	-46.6662 ^a
$(CH_3OPO_3H)_2Rb^-$	-59.17	-41.0822	-38.2897 ^a

^a Calculated using the LACVP basis set

The calculated energetics and structures for the collision induced reaction of $[2CH_3HPO_4+Na]^-$ to yield $[CH_3HPO_3PO_4+Na]^-$ and methanol are given in Scheme 5.1 in kcal/mol utilizing DFT at the B3LYP/6-31G** level. The reaction is uphill, and proceeds via a metastable penta-coordinate phosphorous intermediate.

Scheme 5.1 Reaction Coordinate Diagram for Model Chemical System*

*Calculated at the B3LYP/6-31G** level, ΔH given in kcal/mol

5.4 Discussion.

The binding energy of a noncovalent cluster must be greater than the activation barrier for a reaction to occur by collisional activation. In CAD experiments with clusters, rearrangement processes involving the cleavage of covalent bonds are usually not competitive with dissociation of the complex.²⁸ This is well illustrated in Figure 5.1, where the hydrogen bonded clusters are shown to simply dissociate into the constituent molecules. The addition of a single sodium ion increases the binding energy sufficiently to allow some covalent bond cleavage, as seen in Figure 5.1c. However, the formation of a true salt cluster in Figure 5.2 lends greater binding energy to the cluster, enabling reactions with higher activation barriers. The greater binding energy is quantified by the computational results, which predict the salt cluster to be more strongly bound than the hydrogen bound cluster by roughly ~ 10 kcal/mol for each substitution. As a result, in Figure 5.2 the sequential loss of two adenosine nucleosides which results from the

fragmentation of covalent bonds is observed almost exclusively. The final peak produced in Figure 5.2c appears to be ATP, but the fact that it has the same mass is not necessarily proof of structure.

In order to address that question, further CAD experiments and calculations were performed. The results for the calculations are shown in Scheme 5.1 as a reaction coordinate diagram. Theory predicts that the ΔH for the reaction of a sodium bound dimer is significantly uphill, which is to be expected. The exact energy for the transition state was not calculated, but the transition state must lie below the dissociation threshold for the cluster and probably involves a penta-coordinate phosphorous as shown in Scheme 5.1.

The quasimolecular ion produced in Figure 5.2c and actual ATP were both subjected to further steps of CAD for comparison. The results from these experiments are presented in Figure 5.3 and Figure 5.4. With the exception of the peaks originating from the doubly charged cluster of ATP in Figure 5.4a, the spectra in both Figures are nearly identical. The peaks that are produced and the relative intensity of those peaks are very similar, including peaks that only constitute minor contributions to the total ion intensity. Although this is not absolute proof that the structure of the product in Figure 5.2c is ATP, it provides strong support for the conjecture.

Further support for this notion is obtained from the cationic clusters. Despite a different number of sodium ions and presence of net positive charge, the same chemistry is observed in sodiated clusters of AMP. Figure 5.5 clearly shows a favorable route to what appears to be ATP. Again, comparison of the subsequent CAD of this molecule with a verified sample of ATP reveals that the two yield identical daughter ions. There is

no interference from a dimer in this case, and the peaks and relative intensities of the daughter ions are nearly identical. This strongly suggests that ATP is indeed formed from the salt cluster trimer. The notion that two different products would coincidentally produce spectra that overlap exactly with ATP is unlikely.

Other important observations are related to the fact that the same chemistry operates in cationic and anionic clusters. This suggests that the net charge does not play a significant role in the mechanism that leads to the generation of ATP. Neutral clusters would probably yield similar results, but unfortunately neutral clusters are not well suited for the experiments performed in the present work. This does not preclude the formation of such clusters in the atmosphere, either from meteorites or sea spray, or the subsequent chemistry that follows. The results obtained in the present work suggest a viable prebiotic route to the synthesis of ATP.

Furthermore, the chemistry is not limited to the formation of ATP. In fact, it was possible to synthesize adenosine pentaphosphate, although the structure of this compound was not verified for lack of a sample for comparison. Other nucleotides and deoxynucleotides were basically interchangeable with AMP, allowing for the synthesis of other polyphosphates. Sodiated clusters of 5'-ribose-phosphate also yielded ribose polyphosphate upon collisional activation. Polyphosphate, another important and ubiquitous biopolymer,⁵ can be synthesized from sodium salt clusters of phosphate itself. The facile generation of polyphosphates by this mechanism is therefore applicable to a wide range of biologically relevant molecules.

5.5 Conclusion.

ATP is easily generated by collisional activation of a sodiated salt cluster containing three AMP precursors. The sequential loss of two adenosines leads to the formation of ATP. Comparison of several subsequent steps of CAD dissociation of ATP generated by this method with that of actual ATP reveals that the two ions produce identical CAD spectra. This is true for ATP generated from both cationic and anionic sodiated salt clusters of AMP. Polyphosphates can also be generated by collisional activation of sodiated clusters of phosphate, ribose phosphate, and other nucleotides and deoxynucleotides.

This type of chemistry represents a feasible pathway for the generation of biologically important polymers. A potential prebiotic source for ATP or other polyphosphates is outlined in the present work, but conceivably other biopolymers may be generated by similar reactions. It is interesting to note that small clusters of biomolecules can have an inherent preference for homochirality, as is demonstrated by the serine octamer.²⁹ It appears quite possible that the chemistry of these small clusters may be more important than has been previously realized.

5.6 References

-
- ¹ Gogarten, J. P.; Taiz, L. *Photosynthesis Research* **1992**, *33*, 137-146.
- ² Senior, A. L.; Nadanaciva, S.; Weber, J. *Biochim. Biophys. Acta* **2002**, *1553*, 188-211.
- ³ J. E. Walker (Ed.), *Biochim. Biophys. Acta* **2000**, *1458(2-3)*, entire special issue.
- ⁴ Muller, A. W. J. *Prog. Biophys. Molec. Biol.* **1995**, *63*, 193-231.
- ⁵ Kornberg, A.; Rao, N. N.; Ault-Riche, D. *Annu. Rev. Biochem.* **1999**, *68*, 89-125.
- ⁶ Yamagata, Y.; Watando, H.; Saitoh, M.; Namba, T. *Nature*, **1991**, *352*, 516-519.
- ⁷ Keefe, A. D.; Miller, S. L. *J. Mol. Evol.* **1995**, *41*, 693-702.
- ⁸ St Martin, H.; Ortega-Blake, I.; Les, A.; Adamowicz, L. *Biochim. Biophys. Acta* **1994**, *1207*, 12-23.
- ⁹ Tassis, A. C.; Penteado-Fava, A.; Pontes-Buarque, M.; Amorim, H. S. De; Bonapace, J. A. P.; Souza-Barros, F. De; Vieyra, A. *Origins Life Evol. B.* **1999**, *29*, 361-374.
- ¹⁰ Machado, V. G. M.; Bunton, C. A.; Zucco, C.; Nome, F. *J. Chem. Soc., Perk. Trans. 2*, **2000**, 169-173.
- ¹¹ (a) Oro, J. *Nature*, **1961**, *190*, 389-390. (b) Delsemme, A. H. *Orgins Life* **1984**, *14*, 51-60. (c) Kissel, J.; Krueger, F. R. *Nature*, **1987**, *326*, 753-760. (d) Engel, M. H.; Macko, S. A. *Nature*, **1997**, *389*, 265-268. (e) Chyba, C. F.; Sagan, C. *Nature*, **1992**, *355*, 125-132.
- ¹² Kuzicheva, E. A.; Simakov, M. B. *Adv. Space Res.* **1999**, *23*, 387-391.
- ¹³ Cooper, G.; Kimmich, N.; Belisle, W.; Sarinana, J.; Brabham, K.; Garrel, L.; *Nature* **2001**, *414*, 879-883.
- ¹⁴ Mautner, M. N.; Sinaj, S. *Geochim. Cosmochim. Acta* **2002**, *66*, 3161-3174.
- ¹⁵ Cooper, G. W.; Onwo, W. M.; Cronin, J. R. *Geochim. Cosmochim. Acta* **1992**, *56*, 4109-4115.

-
- ¹⁶ Sephton, M. A. *Nat. Prod. Rep.* **2002**, *19*, 292-311.
- ¹⁷ Levy, M.; Miller, S. L.; Oro, J. *J. Mol. Evol.* **1999**, *49*, 165-168.
- ¹⁸ (a) Ellison, G. B.; Tuck, A. F.; Vaida, V. , *J. Geophys. Res.* **1999**, *104*, 11633-11642.
(b) Tuck, A. *Surv. Geophys.* **2002**, *23*, 379-409. (c) Gill, P.S.; Graedel, T.E.; Weschler, C.G. *Rev. Geophys.* **1983**, *21*, 903-920.
- ¹⁹ Murphy, D.M.: 2001, Extraterrestrial material and stratospheric aerosols, In: B. Peucker-Ehrenbrink (ed.), *Accretion of Extraterrestrial Matter Throughout Earth's History*: New York, Kluwer Academic, 2001.
- ²⁰ Hirabayashi, A.; Sakairi, M.; Koizumi, H. *Anal. Chem.* **1995**, *67*, 2878-2882.
- ²¹ Shiea, J.; Chang, D. -Y.; Lin, C. -H.; Jiang, S. -J. *Anal. Chem.* **2001**, *73*, 4983-4987.
- ²² Takats, Z.; Nanita, S. C.; Cooks, R. G.; Schlosser, G.; Vekey, K. *Anal. Chem.* **2003**, *75*, 1514-1523.
- ²³ (a) Schalley, C. A. *Mass Spectrom. Rev.* **2001**, *20*, 253-309. (b) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.
- ²⁴ Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.
- ²⁵ Vekey, K.; Czira, G. *Anal. Chem.* **1997**, *69*, 1700-1705.
- ²⁶ (a) Kuntz, A. F.; Boynton, A. W.; David, G. A.; Colyer, K. E.; Poutsma, J. C. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 72-81. (b) Armentrout, P. B. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 371-379. (c) Vékey, K.; Czira, G. *Anal. Chem.* **1997**, *69*, 1700-1705. (d) Tao, W. A.; Zhang, D.; Wang, F.; Thomas, P. D.; Cooks, R. G. *Anal. Chem.* **1999**, *71*, 4427-4429. (e) Vitale, G.; Valina, A. B.; Huang, H.; Amunugama, R.; Rodgers, M. T. *J. Phys. Chem. A* **2001**, *105*, 11351-11364.
- ²⁷ This phenomenon is often observed in ion traps, particularly with transition metal ions.

²⁸ (a) Wan, K. X.; Gross, M. L.; Shibue, T. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 450-457. (b) Gabelica, V.; De Pauw, E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 91-98. (c) Penn, S. G.; He, F.; Green, M. K.; Lebrilla, C. B. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 244-252. (d) Julian, R. R.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L. *Angew. Chem. Int. Ed.* **2003**, *42(9)*, 1012-1015.

²⁹ (a) Julian, R. R.; Hodyss, R.; Kinnear, B.; Jarrold, M. F.; Beauchamp, J. L. *J. Phys. Chem. B.* **2002**, *106*, 1219-1228. (b) Counterman, A. E.; Clemmer, D. E. *J. Phys. Chem. B.* **2001**, *105*, 8092-8096. (c) Schalley, C. A.; Weis, P. *Int. J. Mass. Spectrom.* **2002**, *221*, 9-19. (d) Cooks, R. G.; Zhang, D.; Kock, K. J.; Gozzo, F. C.; Eberlin, M. N. *Anal. Chem.* **2001**, *73*, 3646-3655.

Chapter 6

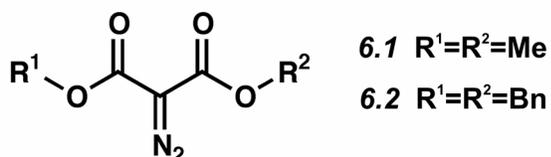
Gas Phase Synthesis of Charged Copper and Silver Fischer Carbenes from Diazomalonates: Mechanistic and Conformational Considerations in Metal Mediated Wolff Rearrangements

Published previously in: Julian, R. R.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L.
J. Am. Chem. Soc. **2003**, *125(15)*, 4478-4486.

6.1 Introduction

The gas phase synthesis of metallo-carbenoid compounds has a surprisingly long history.¹ In early studies, the metallo-carbenoids were created by high energy methods utilizing electron impact ionization. The resulting ions were then used in bimolecular reactions with various neutral molecules to study the chemistry of and determine metal-carbon bond strengths for these charged Fischer carbenes. The importance of studying gas phase species in determining the mechanistic pathways for olefin metathesis was also realized early on,^{1b} and more recent studies have taken advantage of electrospray ionization (ESI) to continue this work.² Gas phase experiments have also enabled mechanistic studies with other highly reactive metallo-carbenoid species.³

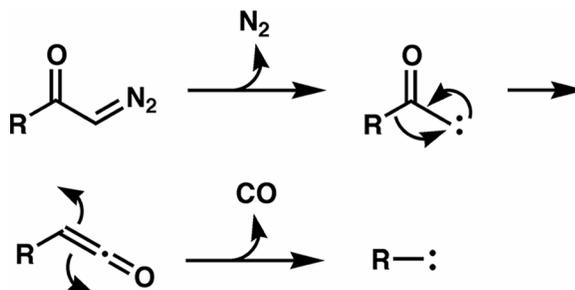
The utility of diazo compounds as carbene and metallo-carbenoid precursors has been heavily exploited in solution phase chemistry.⁴ The use of diazo compounds as precursors in gas phase mass spectrometry (MS) remains less explored. We have recently reported that diazo compounds are excellent carbene precursors for MS.⁵ In these experiments, a carbene is produced through the low-energy collision activated dissociation (CAD)⁶ of a diazo compound. The resulting loss of N₂ and generation of the carbene is achieved under conditions sufficiently mild that non-covalently bound complexes are not dissociated in the process. The highly reactive carbene then inserts into the guest and converts the noncovalent complex into a covalently bound molecule. All of the molecules in this previous study were based on the diazomalonate core (with at least one 18-crown-6 crown ether attached). These reagents have been appropriately named “molecular mousetraps.”⁵



Diazomalonates⁷ (such as **6.1** and **6.2** above) and related diazo ketones⁸ are known to undergo Wolff rearrangement in the gas phase as shown in Scheme 6.1. Since its discovery in 1902, the Wolff rearrangement⁹ has been the subject of numerous studies.¹⁰ In the present work we use the benefits of gas phase experiments to study the mechanism of multiple, consecutive Wolff rearrangements observed in diazomalonates.⁷ The effects that various coordinated metal ions and other charged groups have on Wolff rearrangements are discerned from gas phase MS experiments. Theory is used to quantitatively assess each intermediate for the proposed mechanism. Although the solution phase synthesis of stable copper (I) and silver (I) Fischer carbenes has been

known for some time,¹¹ here we report the first gas phase synthesis of copper (I) and silver (I) Fischer carbenes. The results for several intermolecular reactions of these carbenes with various adduct molecules are presented.

Scheme 6.1



7.2 Experimental Methodology

All mass spectra were acquired on an LCQ Classic instrument from Finnigan. Solutions of the reagents in the ~ 30 - $80 \mu\text{M}$ range were electrosprayed from a $\sim 80/20$ (v/v) solution of methanol/water with a minimum of 0.1% MeCN added. Soft ionization settings were used to maximize the intensity of non-covalently bound complexes.¹² Ions of interest were isolated and subjected to collisional activation until product peaks were observed. For each MS^n step, the peak of interest was re-isolated prior to further dissociation. All chemicals were purchased from Sigma-Aldrich and used without further purification unless otherwise noted. Metal ion complexes were formed by adding an appropriate salt to the solution. No counter-ion effects were noted. For studies of intermolecular reactions, the desired ligand (such as 5-hexynenitrile) was added directly to the solution in several-fold excess.

Calculations. Candidate structures were evaluated initially at the PM3 semi-empirical level. Following minimization at the lower level of theory, structures were optimized

using density functional theory (DFT). The DFT calculations were carried out using Jaguar 4.1 (Schrödinger, Inc., Portland, Oregon). Full geometry optimization was performed at the B3LYP/LACVP** level of theory. Semi-empirical PM3 MNDO type calculations were carried out using the HyperChem 5.1 Professional Suite (Hypercube, Inc., Gainesville, Florida).

Synthesis. Reactions were performed in flame-dried glassware under a nitrogen atmosphere. Solvents were dried and purified using activated alumina columns. All other reagents were used as received from commercial sources. Reaction temperatures were controlled by an IKAmag temperature modulator. Thin-layer chromatography (TLC) was performed using E. Merck silica gel 60 F254 precoated plates (0.25 mm) and visualized by UV and *p*-anisaldehyde staining. ICN Silica gel (particle size 0.032-0.063 mm) was used for flash chromatography. ¹H NMR spectra were recorded on a Varian Mercury 300 spectrometer (at 300 MHz) in CDCl₃ and are internally referenced to the residual chloroform peak (7.27 ppm) relative to Me₄Si. Data for ¹H NMR spectra are reported as follows: chemical shift (δ ppm), multiplicity, coupling constant (Hz), and integration. IR spectra were recorded on a Perkin Elmer Paragon 1000 spectrometer and are reported in frequency of absorption (cm⁻¹). Preparatory reversed phase HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25 x 100 mm, 100 μm C18 column equipped with a guard.

2-Diazodimethylmalonate (6.1). **6.1** was prepared according to previously established methods.¹³ The product was isolated as a yellow oil (2.58 g, 16.29 mmol, 93% yield) with the same physical properties as previously reported.

2-Diazodibenzylmalonate (6.2). A round-bottomed flask (10 mL) was charged with dibenzylmalonate (77 μL, 0.308 mmol), MeCN (3 mL), and *p*-acetamidobenzylsulfonyl

azide (116 mg, 0.485 mmol). TEA (150 μ L, 1.08 mmol) was then added, and the reaction was stirred at room temperature for 10 h. TLC analysis (3:1 hexanes:EtOAc eluent, $R_F=0.46$) showed the reaction to be complete. The solvent was removed by evaporation under reduced pressure, and the crude mixture was subjected to flash chromatographic purification (5:1 hexanes:EtOAc eluent) to afford **6.2** as a yellow oil (75 mg, 0.242 mmol, 78% yield) with the same physical properties as previously reported.¹⁴

Diazomalonate 6.25. A round-bottomed flask (5 mL) was charged with CH_2Cl_2 (1 mL), CD_3OD (500 μ L, 11.3 mmol), and TEA (100 μ L, 0.717 mmol). This mixture was stirred rapidly while malonyldichloride (20 μ L, 0.206 mmol) was added dropwise at room temperature. TLC analysis (3:1 hexanes:EtOAc eluent, $R_F=0.45$) showed complete conversion to the malonate ester within 2 h. The solvent and excess reagents were removed by evaporation under reduced pressure, and then MeCN (2 mL) and *p*-acetamidobenzylsulfonyl azide (103.5 mg, 0.431 mmol) were added to the flask. TEA (100 μ L, 0.717 mmol) was added, and the solution was stirred for 24 h. The solvent was then removed by evaporation under reduced pressure. The product was purified by dissolving the residue in a minimal amount of CH_2Cl_2 (500 μ L) and then precipitating the salts by addition of Et_2O (5 mL). Filtration through celite removed the salts, and the solvent was removed by evaporation under reduced pressure to afford **6.25** as a yellow oil (11.9 mg, 0.073 mmol, 35% yield).

Diazomalonate 6.26. To a stirred solution of $\text{H}_2\text{C}^{13}\text{CO}_2\text{H}$ (19.7 mg, 0.179 mmol) in Et_2O (2 mL) in a scratch-free flask (25 mL) was added ethereal diazomethane solution (0.2 M, 4.0 mL, 0.800 mmol). TLC analysis (3:1 hexanes:EtOAc eluent, $R_F=0.45$) showed the reaction to be complete. The solvent and excess reagents were removed by

evaporation under reduced pressure, and then MeCN (1 mL) and *p*-acetamidobenzylsulfonyl azide (25.9 mg, 0.108 mmol) were added to the flask. TEA (41 μ L, 0.294 mmol) was added, and the solution was stirred for 24 h. The solvent was then removed by evaporation under reduced pressure. The product was purified by dissolving the residue in a minimal amount of CH₂Cl₂ (500 μ L) and then precipitating the salts by addition of Et₂O (5 mL). Filtration through celite removed the salts, and the solvent was removed by evaporation under reduced pressure to afford **6.26** as a yellow oil (17.4 mg, 0.108 mmol, 60% yield).

Diazomalonate 6.27. A round-bottomed flask (10 mL) was charged with CH₂Cl₂ (2 mL), EtOH (25 μ L, 0.687 mmol), and 2-bromoethanol (30 μ L, 0.421 mmol). This mixture was stirred rapidly while malonyldichloride (20 μ L, 0.206 mmol) was added dropwise at room temperature. ESMS showed complete conversion to the 2-bromoethyl ethylmalonate within 2 h. The solvent and excess reagents were removed by evaporation under reduced pressure, and then MeCN (2 mL) and pyridine (50 μ L) were added to the flask. After 24 h, *p*-acetamidobenzylsulfonyl azide (94.1 mg, 0.392 mmol) and then TEA (150 μ L, 1.076 mmol) were added. The solution was then stirred for another 12 h. The solvent was then removed by evaporation under reduced pressure to afford **6.27**; the yield was qualitatively estimated to be ~50% by MS (*m/z* 264.3).

Diazomalonate 28. To a stirred, dry solution of penta(ethylene glycol) (44.4 μ L, 0.210 mmol), CH₂Cl₂ (20 mL), and TEA (150 μ L, 1.076 mmol) was added malonyl dichloride (24 μ L, 0.247 mmol). The mixture was heated to reflux for 5.5 h, cooled, and then the solvent was removed by evaporation under reduced pressure. The product was recovered by extraction of the crude residue with refluxing hexanes (20 mL). The

benzene was then removed by evaporation under reduced pressure to leave the malonate crown ether. This malonate ester¹⁵ was then dissolved in MeCN (3 mL). Stepwise addition of *p*-acetamidobenzenesulfonyl azide (151.0 mg, 0.629 mmol) and TEA (200 μ L, 1.435 mmol) provided **6.28** after stirring for 10 h. The solvent was removed *in vacuo*, the residue was dissolved in a minimal amount of CH₂Cl₂ (500 μ L), and the undesired salts precipitated out of solution with the addition of Et₂O (5 mL). Filtration through celite and removal of solvent by evaporation under reduced pressure yielded **6.28** (59.3 mg, 0.179 mmol, 85% yield).

Diazomalonate 6.29. To a stirred, dry solution of 18-crown-6-methanol (50.0 μ L, 0.159 mmol), CH₂Cl₂ (1.5 mL), and TEA (33 μ L, 0.237 mmol) was added ethyl malonyl chloride (28 μ L, 0.219 mmol). The mixture was heated to reflux for 8 h, the solution was cooled, and then the solvent was removed by evaporation under reduced pressure. The residue was dissolved in MeCN (750 μ L), and treated with TEA (30 μ L, 0.215 mmol). To this solution was added *p*-acetamidobenzenesulfonyl azide (53.1 mg, 0.221 mmol), and the mixture was stirred for 10 h. The solvent was removed by evaporation under reduced pressure, the residue dissolved in a minimal amount of CH₂Cl₂ (500 μ L), and the undesired salts were precipitated out of solution with the addition of Et₂O (5 mL). Filtration through celite and removal of solvent by evaporation under reduced pressure yielded **6.29** (59.8 mg, 0.138 mmol, 87% yield) as a light yellow oil. A small sample (~15 mg) was chromatographed to analytical purity by HPLC (0.1% (wt/v) TFA in water, 8.0 mL/min, 0.30% acetonitrile/min, 82-85 min). FTIR (thin film) 2879, 2142, 1755, 1689; ¹H NMR (300 MHz, CDCl₃) 4.45 (dd, *J* = 3.85, 12.1 Hz, 1H), 4.31 (q, *J* = 7.14 Hz,

2H), 4.27 (m, 1H), 3.85 (t, $J = 4.95$), 3.80 (br s, 1H), 3.67 (br s, 21H), 1.32 (t, $J = 7.14$ Hz, 3H); MS m/z 435.2 (H^+).

6.3 Results and Discussion.

Copper (I) and Silver (I). The ESI-MS spectrum for a mixture of copper (I) and **6.1** is shown in Figure 6.1a. The base peak corresponds to $[6.1+MeCN+Cu]^+$ or **6.3** (for simplicity, the corresponding structures are to the right of the mass spectra in Figure 6.1). It can also be seen from Figure 6.1a that **6.1** has a high affinity for Na^+ (present as an impurity) and that Cu (I) has a high affinity for acetonitrile. As seen in Figure 6.1b, isolation of **6.3** followed by CAD leads to sequential losses of 28 Da. The first loss of 28 Da corresponds to the loss of N_2 from the diazo functionality leading to structure **6.4**. Spontaneous Wolff rearrangement of **6.4** is accompanied by the second loss of 28 Da, which is attributed to be a loss of CO yielding **6.5**. In Figure 6.1c, structure **6.5** is isolated and subjected to further CAD resulting in another loss of 28 Da. This is also attributed to Wolff rearrangement of **6.5** followed by the loss of a second CO yielding the stable copper Fischer carbene **6.6**. Figure 6.1d shows the results of CAD of structure **6.6**. In this case, the noncovalent acetonitrile adduct is lost, followed by the pickup of either water or methanol. This pickup is not surprising given the vacant copper (I) coordination site and the fact that the spectra were acquired from a water/methanol solution.

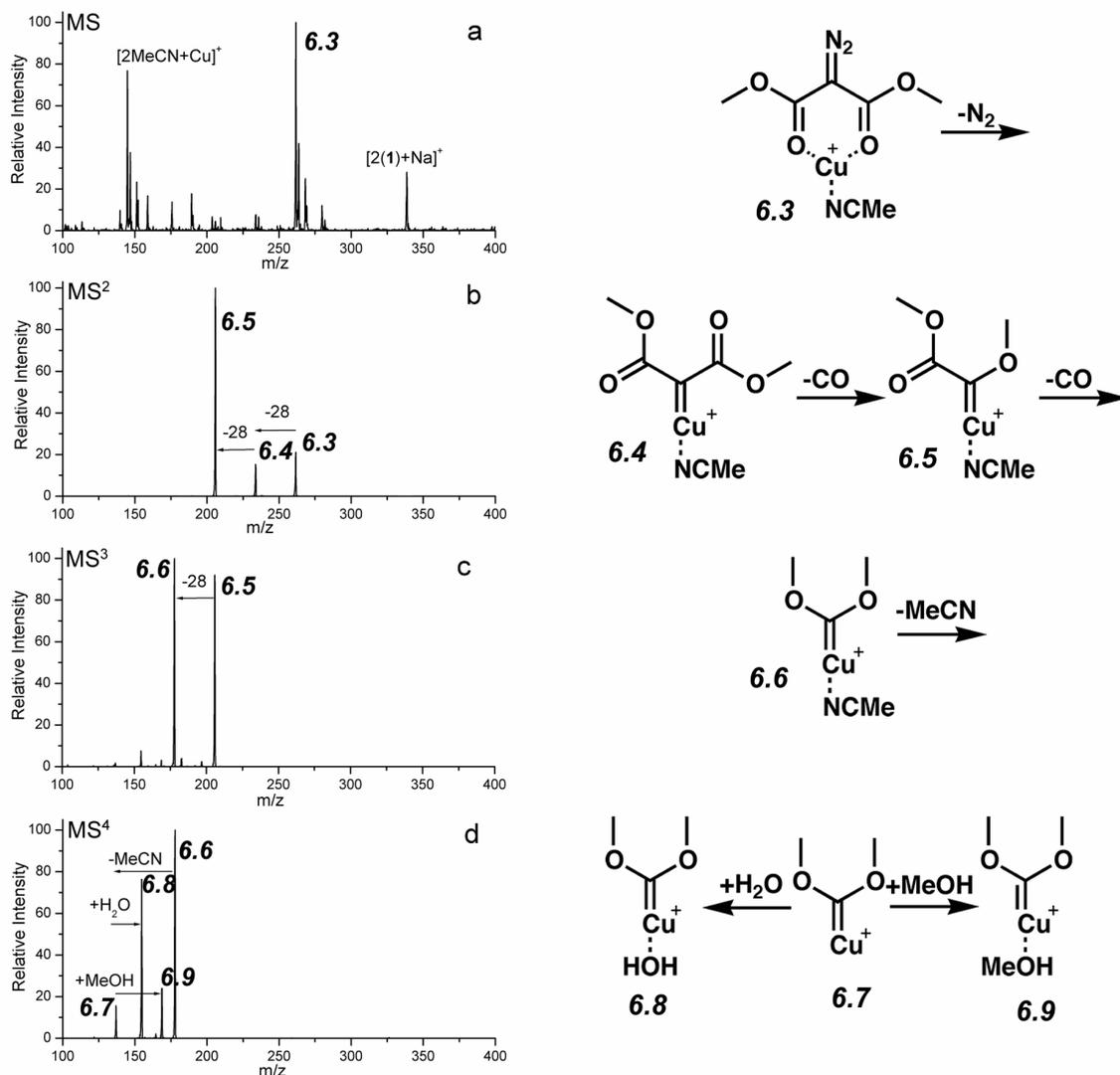


Figure 6.1 Structures for the peaks in the mass spectra on the left are provided in the reaction scheme to the right. (a) MS of a mixture of **6.1** and Cu (I) in a 20/80 (v/v) water/methanol solution with ~0.1% MeCN. (b) MS² on complex **6.3**. N₂ and CO are lost sequentially. (c) MS³ on complex **6.5** resulting in the loss of an additional CO. (d) MS⁴ on **6.6** resulting in the exchange of the MeCN ligand for either water or methanol.

DFT calculations at the B3LYP level using the LACVP** basis set were performed to determine structures and relative energetics for the products and probable intermediates in Figure 6.1. The structures and energetics are presented in the reaction coordinate diagram in Figure 6.2. When a ligand such as N₂ or CO is lost, the energy of the minimized separated molecule is added to compare with the complex prior to dissociation. Structure **6.3** is found to be the global minimum, and there is support for this binding mode from crystal structures.¹⁶ Direct dissociation of N₂ from **6.3** is unlikely because the copper (I) ligand restrains the geometry of the molecule in a triplet-like conformation, whereas the thermal dissociation of N₂ must yield the singlet state.¹⁷ Furthermore, **6.1** has a singlet ground state⁵, and attempts to minimize structure **6.3** with the N₂ removed were unsuccessful. Matrix isolation studies on similar systems suggest that copper (I) insertion into the C-N bond of **6.3** is a more likely alternative.¹⁸

Further experimental and theoretical evidence also suggests that the copper (I) ion mediates the generation of the carbene from the diazo precursor. CAD experiments on [6.1+Na]⁺ result in the complete loss of signal without producing any observable peaks, suggesting dissociation of the sodium ion (which has m/z ratio that is too small to detect). Similar experiments with rubidium yield a peak corresponding to Rb⁺ exclusively from the collisional activation of [6.1+Rb]⁺. The calculated ΔH for reaction 6.1 is -59 kcal/mol, which can be regarded as the binding energy of the sodium ion to **6.1**.



Similarly, calculations reveal a ΔH of -43 kcal/mol for reaction 6.2, which is the binding energy of MeCN to the **6.1**·Cu complex. This value is in reasonable agreement with similar experimental and theoretical results.¹⁹ The MeCN ligand is retained

throughout the entire series of reactions shown in Figure 6.2 without dissociating. Therefore the reaction barriers for each step must be below 43 kcal/mol, otherwise the MeCN ligand would simply dissociate. Furthermore, the binding energy of the sodium ion is much greater at 59 kcal/mol, yet the ion dissociates prior to the loss of N₂ and the generation of the carbene. This demonstrates that copper (I) lowers the activation barrier for the generation of a carbene by at least 16 kcal/mol relative to complexation with sodium.

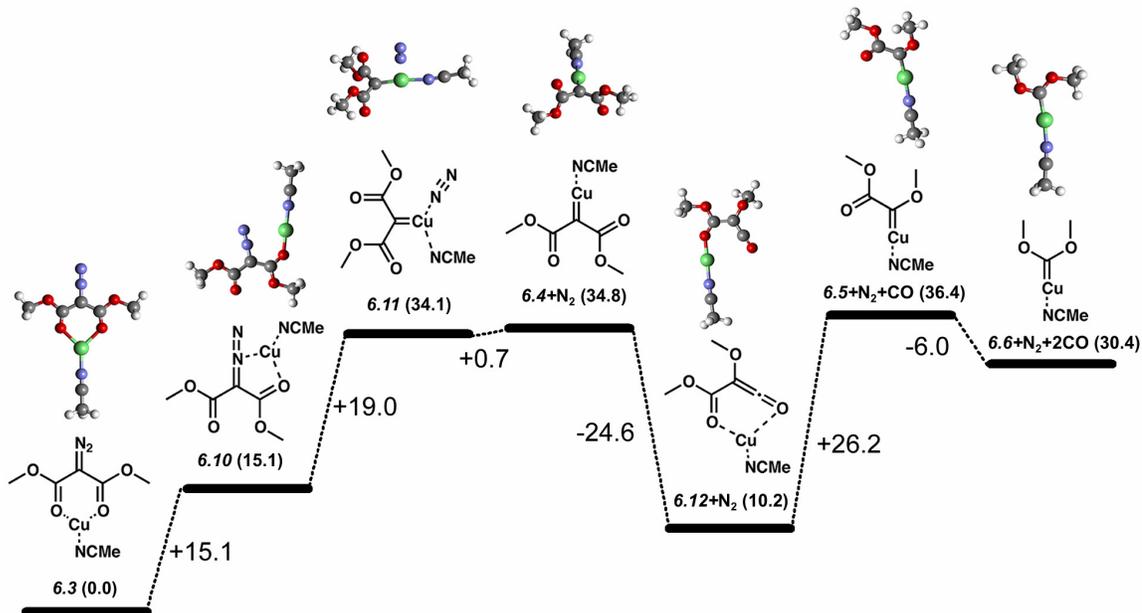


Figure 6.2 Reaction coordinate diagram including intermediate structures for the reactions in Figure 6.1. The energetics are calculated relative to structure **6.3** and are given in parentheses in kcal/mol. When a neutral gas molecule is lost the energy of the minimized, separated molecule is added to that of the remaining structure. See the text for a discussion of the barriers at each step.

The reaction shown in Figure 6.2 proceeds by rearrangement of structure **6.3** to structure **6.10** followed by Cu^+ insertion as shown in **6.11**. Crystal structures of copper (I) Fischer carbenes have similar C-Cu bond lengths to those in **6.11**, and in some cases demonstrate a high affinity for MeCN ligands.²⁰ Copper insertion is postulated as a prelude to the Wolff rearrangement from **6.4** to **6.12**, which must proceed through a singlet state. There is poor agreement between theoretically and experimentally determined barriers to Wolff rearrangement, with theory predicting higher values than those observed experimentally.²¹ Theory would predict a large barrier for the conversion of **6.4** to **6.12** because of the exothermicity of the reaction.^{21b} The experimental evidence in Figure 6.1b suggests that both rearrangement of **6.4** to **6.12** and subsequent loss of CO to produce **6.5** proceed with minimal barriers. It should be pointed out that, given a low barrier to rearrangement, the structure for the observed peak in Figure 6.1b may be the rearranged product (**6.12**) rather than carbene (**6.4**).²² In either case, the observed products suggest that the presence of copper lowers the barrier to Wolff rearrangement in addition to facilitating the generation of the initial carbene.

The Wolff rearrangement product of **6.5** undergoes copper insertion without barrier upon minimization, suggesting that the transition state may occur prior to rearrangement. Therefore, the loss of the second CO likely occurs in a concerted fashion without a true ketene intermediate. The data in Figures 6.1b and 6.1c shows that the initial two losses occur simultaneously, which suggests a higher barrier to the loss of the second CO. However, it should be noted that the non-covalently bound MeCN ligand is retained throughout the entire process, limiting the reaction barriers to the bond dissociation energy of the Cu-NCMe bond (or 43 kcal/mol). Finally, a stable Fischer carbene (**6.6**) is produced upon loss of the second CO. Subsequent excitation of this complex results in

the loss of the MeCN ligand followed by attachment of either H₂O or MeOH from the residual solvent vapor present in the ion trap (Figure 6.1d).

In the presence of silver (I), structure **6.13** (Figure 6.3a) is formed in high abundance. However upon collisional activation in Figure 6.3b, the MeCN ligand is lost instead of N₂ as was the case with copper (I). This difference is due to the weak binding of the third ligand to the silver (I) cation.²³ Loss of the MeCN adduct is followed by collisional cooling of the product (**6.14**), which leads to some pickup of residual methanol (**6.15**) or water (**6.16**) from the trap as seen in Figure 6.3b. Further CAD of re-isolated **6.14** leads to the exclusive loss of 56 Da, comprising N₂ and CO with no intermediate loss of N₂ being observed. Again these losses are followed by the pickup of methanol or water. The absence of an intermediate loss of N₂ suggests that silver (I) is more efficient at facilitating the Wolff rearrangement and subsequent loss of CO than copper (I), consistent with the predominant use of silver (I) as the catalyst where Wolff rearrangement is required for a synthetic route.^{4a} Isolation and activation of **6.17** leads to the expected loss of the second CO, yielding structure **6.20** and ensuing pickup of water or methanol. The silver Fischer carbenes **6.20**, **6.21**, and **6.22** are the final observable carbene products in this sequence. Further CAD of **6.20** leads to the loss of the Ag⁺ cation, presumably generating the neutral dimethoxy carbene in the process.

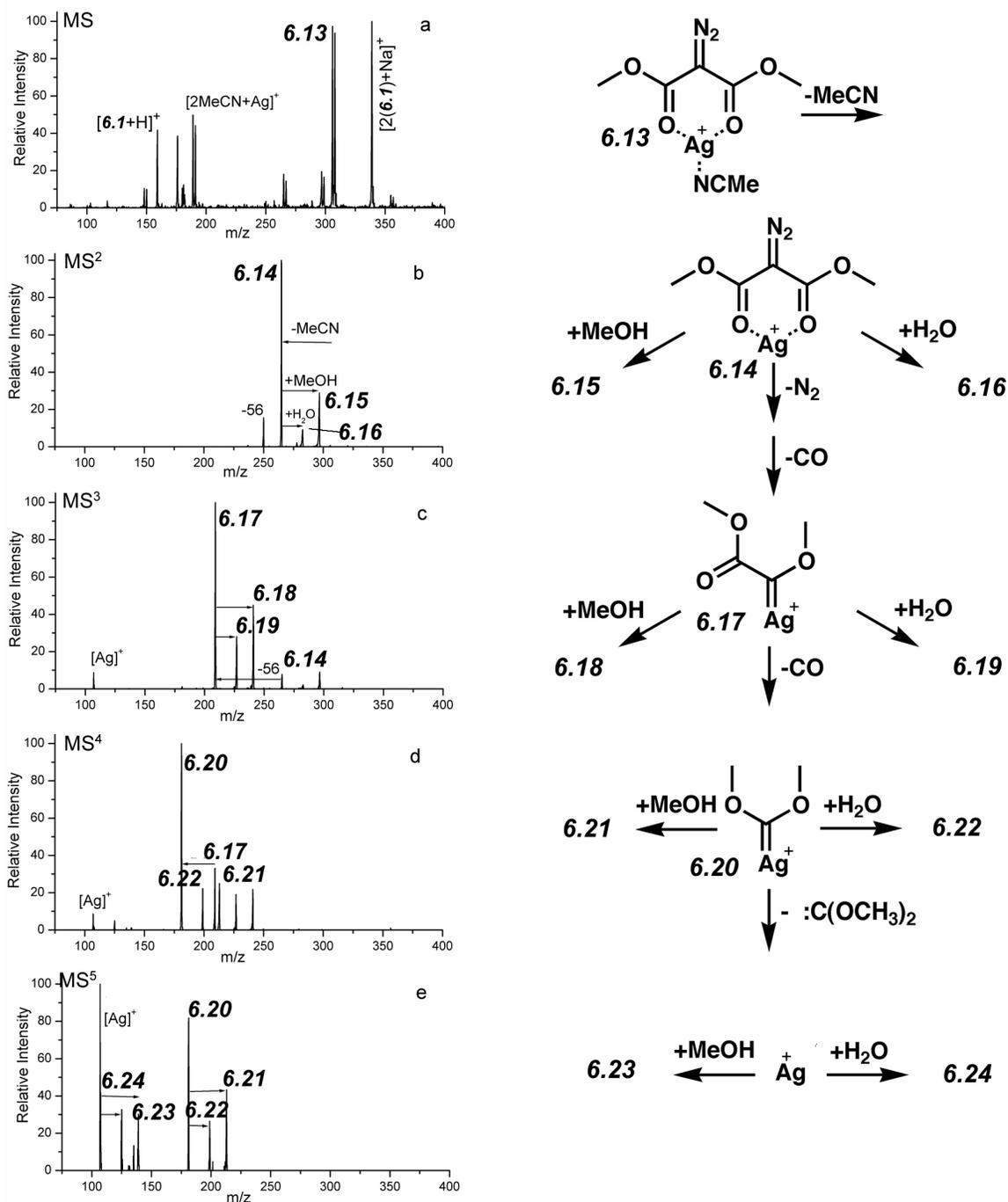
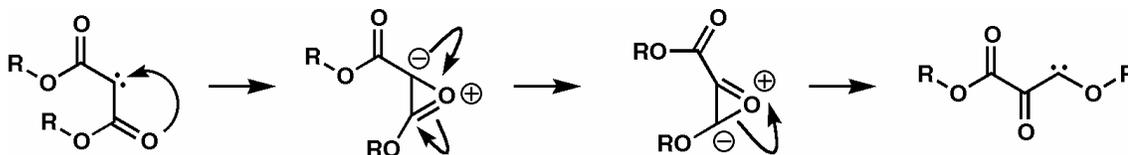


Figure 6.3 Spectra acquired from a mixture of silver (I) and **6.1**. The structures are given to the right of the appropriate spectrum. (a) MS spectrum showing complexation with Ag^+ . (b) MS² spectrum of **6.13**, where the loss of MeCN generates the most abundant product ion peak. (c) MS³ spectrum of **6.14** showing the facile loss of N_2 and

CO. (d) MS⁴ spectrum of **6.17** showing loss of second CO and generation of final silver Fischer carbenes. (e) MS⁵ spectrum of **6.20**, further CAD results in the cleavage of the metal/carbon bond.

Labeling Experiments. Isotopic labeling experiments were performed to confirm the proposed reaction pathways. The results for the two labeled compounds, **6.25** and **6.26**, are given in Table 6.1. Structure **6.25** is labeled with three deuteriums on each of the terminal methyl groups. Comparison of the data in Table 6.1 with that shown in Figure 6.1 reveals that the six deuteriums are retained throughout the entire experiment. This indicates that only interior carbons are lost as CO and that there is no detectable scrambling with MeOH in the ion trap. In order to confirm which carbons are lost and to determine the extent to which rearrangement of the carbene occurs, structure **6.26** was synthesized with ¹³C at the 1,3 positions. Experiments with **6.26** confirm that the two interior ¹³C labeled carbons are lost, as shown in Table 6.1. There is no apparent rearrangement of the carbene via an oxirene intermediate as depicted in Scheme 6.2.¹⁰ These two experiments serve to confirm that N₂ is lost first, as is expected, and that carbon in the 1 and 3 positions specifically, are lost as CO.

Scheme 6.2 Potential carbene scrambling via oxirene (not observed).



Divalent Metals, Protons, and Fixed Charges. The experimental MS data for several related experiments on **6.1** are summarized in Table 6.1. For both copper (II) and nickel (II), the metal is complexed with three molecules of **6.1** or $[3(\mathbf{6.1})+\mathbf{M}]^{2+}$ (where M = Ni or Cu), with the ligands coordinating the metal in a pseudo-octahedral fashion. CAD of these complexes does not lead to the loss of N₂, but instead yields only the loss of an entire coordinating molecule. A small amount of $[\mathbf{6.1}+\text{Cu}+2\text{MeCN}]^{2+}$ is also formed from the copper solution. However, as the data shows in Table 6.1, this complex simply loses one of the MeCN ligands upon CAD. These results illustrate that copper (II) and nickel (II) are not efficient at mediating the formation of Fischer carbenes or subsequent Wolff rearrangements. Protonated **6.1** does not lose N₂ upon CAD either, but instead yields the assortment of fragments shown in Table 6.1.

The results presented thus far suggest that the coordinating charge can dictate the resulting chemistry upon collisional activation in diazo compounds. Structure **6.27** was designed to investigate the energetics of N₂ loss and Wolff rearrangements in the absence of a coordinating charge. In **6.27**, the charge is provided by a fixed quaternary nitrogen from the pyridinium group. As seen in Table 6.1, this compound loses N₂ to yield the most abundant product by CAD in the MS² spectrum. Further collisional activation of the resulting molecule demonstrates that Wolff rearrangement and the loss of multiple CO molecules does occur. However, this process is accompanied by other losses not related to Wolff rearrangement. In fact, the loss of pyridine yields the base peak in the MS³ spectrum. This suggests that the energetics associated with Wolff rearrangement and loss of CO for this reaction are similar to those for the loss of pyridine. Quantitative CAD experiments on related quaternary pyridinium molecules suggest that the loss of pyridine is not likely to require more than ~60 kcal/mol activation energy but may require

substantially less.²⁴ This can be regarded as the upper limit for the activation barrier to Wolff rearrangement and dissociation of CO for these experiments.

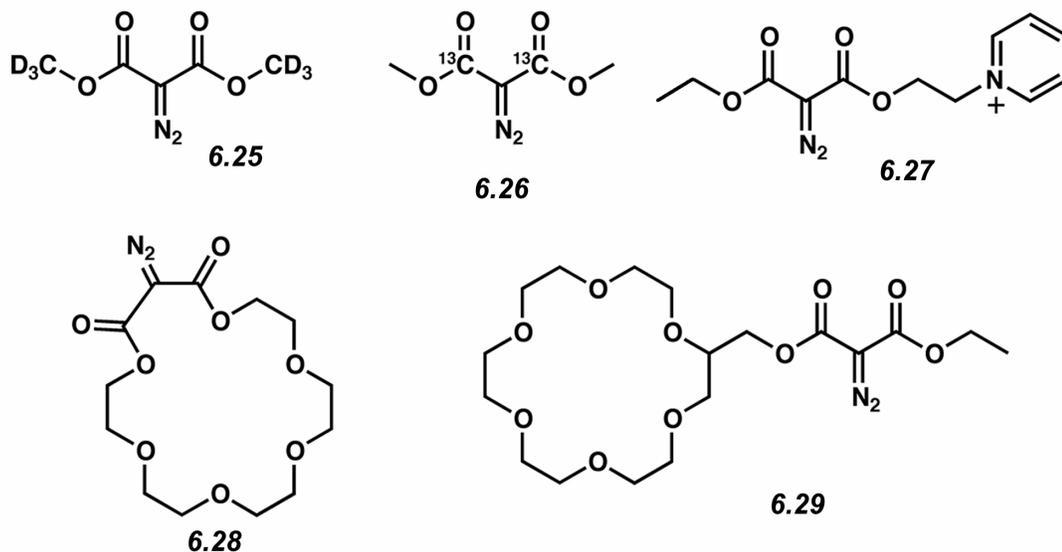
Table 6.1. Summary of MS results.

Molecule	Cation	Initial Ligands	Observed Peaks [†]
6.25	Cu ⁺	MeCN	MS ² on 268 → 193(12), 212(100), 240(15), 268(21) MS ³ on 212 → 161(11), 184(100), 212(80) MS ⁴ on 184 → 143(5), 161(43)*, 175(8)*, 184(100)
6.26	Cu ⁺	MeCN	MS ² on 264 → 207(100), 236(11), 264(2) MS ³ on 207 → 178(100), 207(98)
6.1	Cu ²⁺	2(1)	MS ² on 268.4 → 189.5(100), 268.4(<1)
6.1	Cu ²⁺	2MeCN	MS ² on 151 → 140(100)*, 147(50)*
6.1	Ni ²⁺	2(1)	MS ² on 266 → 187(30), 196(100)*, 203(72)*, 266(<1)
6.1	H ⁺	--	MS ² on 159 → 55(80), 69(6), 87(100), 101(8), 127(93), 145(28), 159(<1)
6.2	Cu ⁺	MeCN	MS ² on 414 → 345(8), 373(100), 414(10) MS ³ on 373 → 289(55), 317(37), 345(42), 373(100)
6.2	Ag ⁺	--	MS ² on 417 → 289(12), 317(17), 333(39), 345(25), 361(96), 389(100), 417(16) MS ³ on 389 → 317(10), 333(26), 361(69), 389(100)
6.27	See structure	--	MS ² on 264 → 113(8), 157(13), 236(100), 264(22) MS ³ on 236 → 157(100), 180(42), 192(40), 208(61), 236(57)

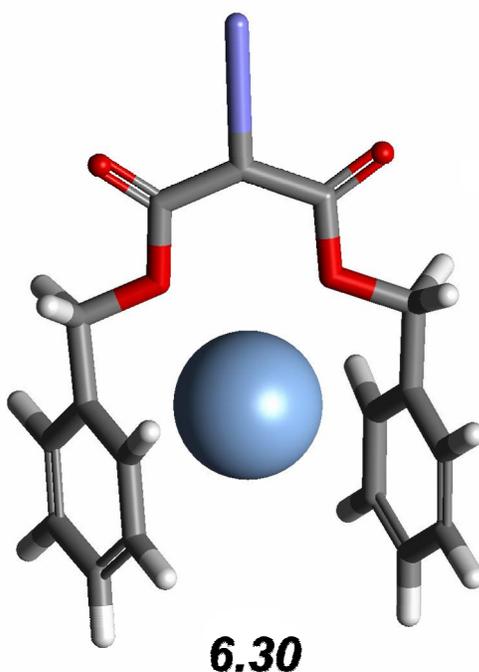
[†]Peaks are given with accompanying relative intensities in parentheses. Some low intensity peaks (under 15% relative intensity) that are not assigned structures or discussed in the text are omitted for clarity.

*Solvent molecule pickups

The pattern of losses for molecule **6.27** contrasts sharply with the results obtained for all of the copper (I) and silver (I) adducts. Metal adducts of **6.1** lose N_2 and the first CO in one step without the addition of further excitation energy. The loss of the second CO occurs upon further activation, but no other competitive products are produced in significant abundance at any stage of the experiment. In the case of **6.27**, N_2 loss occurs without any accompanying loss of CO. Upon further excitation of $[\mathbf{6.27-N}_2]^+$, Wolff rearrangement followed by the loss of the first and second CO molecules occurs spontaneously with the appearance of several other competitive products. In other words, Wolff rearrangement products are no longer favored in the absence of a metal ligand. These results suggest that copper (I) and silver (I) facilitate the Wolff rearrangement and loss of CO shown in Figures 6.1 and 6.2.

Chart 6.1

Effect of Benzyl Groups. In structure **6.2** the methyl groups are replaced by benzyl groups, and with the addition of copper (I) the ESI-MS contains a prominent $[\mathbf{6.2}+\text{Cu}+\text{MeCN}]^+$ peak. As shown in Table 6.1, CAD dissociation of this peak leads primarily to the loss of the MeCN ligand. No appreciable pickup of water or methanol is observed following the loss of MeCN. This suggests that the benzyl groups are able to coordinate the metal ion sufficiently so that the lost ligand is not replaced. There is precedence for this type of coordination from crystal structure studies.²⁵ Interestingly, further CAD of $[\mathbf{6.2}+\text{Cu}]^+$ leads to the sequential loss of N_2 , CO, and CO with each intermediate being observed. No further activation is necessary to observe the loss of all three molecules. Similar experiments with silver (I) and **6.2** yield very similar results (Table 6.1), but in this case the $[\mathbf{6.2}+\text{Ag}]^+$ peak is formed directly from solution in high abundance. Because the metal ion is coordinated to **6.2** by the benzyl groups, the relative conformation of the diazo functionality to the esters is different for copper (I) and silver (I) adducts of **6.1** and **6.2**. DFT calculations suggest a structure similar to the silver (I) adduct shown in **6.30**. The fact that Wolff rearrangement proceeds more easily for **6.30** than for **6.3** suggests that the conformational energy minimum is similar to the preferred conformational orientation that leads to Wolff rearrangement for **6.30**. Conformational effects have been observed to influence Wolff rearrangements previously.¹⁰



Macrocycles. Crown ethers and other related macrocycles are well suited to coordinate a variety of charge bearing groups²⁶ and usually exclude the presence of other solvent adducts once in the gas phase. Therefore, structures **6.28** and **6.29** are well suited to study the effects of various cations and different coordination motifs on the Wolff rearrangement in the absence of third party ligands. The diazo group for **6.28** is included in the ring, bringing the charge in close proximity to the diazo group. However, it should be noted that the coordination geometry between the charge and the diazo is significantly different for **6.1** and **6.28** due to the conformational restraints of the macrocycle. Structure **6.29** is more dynamic, allowing for either coordination or separation of the metal and the diazo group, depending primarily on the temperature of the molecule as explained below.

The CAD results for **6.28** coordinated to copper (I) and sodium are presented in Table 6.2. In both cases, the loss of N₂, CO, and CO occurs predominantly one step at a time, requiring the acquisition of an MS⁴ scan to induce the loss of the second CO. This

suggests that the more conformationally restrained macrocycle requires additional activation energy to induce Wolff rearrangement, but does not necessarily exclude participation of the cation in the reaction. The resonance excitation RF voltages applied at each step of the experiment are (for sodium) 0.67V, 0.83V, 0.99V and (for copper) 0.69V, 0.82V, 0.71V respectively. The excitation voltages are very similar for the first two steps of the experiment for both cations, but copper (I) clearly catalyzes the loss of the second CO more efficiently than sodium ($\Delta V \sim 0.3V$). This is additionally confirmed by the absence of competitive fragments in the copper (I) experiment, whereas in the case of sodium the base peak in the MS⁴ spectrum is not the product of Wolff rearrangement.

The copper (II) adduct of **6.28** can also be prepared in small abundance and the results for the CAD are found in Table 6.2. Copper (II) catalyzes the loss of 44, which is presumably the loss of ethylene oxide from the crown ether portion of the molecule.²⁷ The loss of N₂ is not detected, again suggesting that copper (II) does not catalyze the formation of Fischer carbenes or the subsequent Wolff rearrangement observed with other cations. CAD of the ammonium adduct of **6.28** is also presented in Table 6.2. In this case, the loss of NH₃ followed by subsequent cleavage of protonated **6.28** is the only process observed. This suggests that close proximity of a labile proton can prevent the loss of N₂ and subsequent reactions from occurring. However, it should be noted that this problem is easily corrected by complexation with a more basic amine such as 1-hexylamine. CAD of the hexylamine adduct leads primarily to the loss of N₂ (data not shown). Alternatively, weaker binding by ammonium versus 1-hexylamine may explain the dissociation in the case of ammonium. These experiments confirm the notion that the cation can greatly influence the subsequent chemistry upon excitation of these molecules.

Table 6.2 MS data for crown compounds

Molecule	Cation	Initial Ligand	Observed Peaks [†]
6.28	Na ⁺	--	MS ² on 355 → 327(100), 355(6) MS ³ on 327 → 255(13), 271(6), 299(100), 327(5) MS ⁴ on 299 → 225(100), 255(40), 271(73), 299(18)
6.28	Cu ⁺	--	MS ² on 395 → 339(26), 367(100), 395(18) MS ³ on 367 → 339(94), 367(100) MS ⁴ on 339 → 265(16), 311(100), 339(38)
6.28	NH ₄ ⁺	--	MS ² on 350 → 173(11), 217(12), 261(100), 333(33), 350(72)
6.28	Cu ²⁺	--	MS ² on 197.5 → 175.5(100), 184.3(6)*, 191.2(21)*, 197.5(5)
6.29	NH ₄ ⁺	--	MS ² on 452 → 380(12), 424(100), 452(61) MS ³ on 424 → 175(26), 203(21), 247(19), 291(19), 351(44), 424(100)
6.29	Na ⁺	--	MS ² on 457 → 385(3), 429(100), 457(9) MS ³ on 429 → 299(10), 357(29), 373(100), 385(10), 401(20), 429(30)

[†]Peaks are given with accompanying relative intensities in parentheses. Some low intensity peaks (under 15% relative intensity) that are not assigned structures or discussed in the text are omitted for clarity.

*Solvent molecule pickups

Charge proximity effects. In the copper (I) and silver (I) adducts of **6.2**, the metal ion is coordinated to the benzyl groups and is not in close proximity to the diazo group. It is not surprising that the resulting CAD spectra are very similar for both metals. In general, charges that are not in close proximity to the diazo group have less influence on the resulting CAD patterns. This is well illustrated by comparing **6.28** and **6.29**. Molecular dynamic simulations suggest that the collisional heating of **6.29** leads to extended structures where the lariat side arm extends away from the charge. This leads to a large separation between the charge and the diazo group. For structure **6.28**, heating of the quasimolecular complex does not lead to significant structural changes, leaving the charge in close proximity to the diazo group. The net effect is that the diazo group is $\sim 2\text{\AA}$ further away from the charge in **6.29** when compared to **6.28**. CAD of $[\mathbf{6.28}+\text{NH}_4]^+$ does not lead to the loss of N_2 or subsequent Wolff rearrangement products. By contrast when **6.29** is complexed with ammonium, CAD leads to the loss of N_2 and formation of a carbene (Table 6.2). This implies that the ammonium ion does not interact with the diazo group in **6.29**, but this does not rule out subsequent interactions between the newly generated carbene and the cation. Interestingly, subsequent CAD of $[\mathbf{6.29}-\text{N}_2+\text{NH}_4]^+$ does not lead to any Wolff rearrangement products. This is attributed to the high proton affinity of the carbene,²⁸ which reacts with the ammonium and prevents Wolff rearrangement. CAD of **6.29** complexed with sodium leads to the expected loss of N_2 and the formation of both Wolff rearrangement products in the MS^3 spectrum. Alternatively, weaker binding of ammonium by **6.28** relative to **6.29** may explain the difference in experimental results.

Conformational effects. The degree of conformational restraint present in a molecule can favor or disfavor subsequent Wolff rearrangement. It was established in the previous

paragraph that copper (I) and silver (I) adducts of **6.2** behave very similarly with the simultaneous loss of all products, but it should be noted that the CAD spectra of these two complexes are not similar to any others obtained in the course of this study. We attribute this to a conformational effect where the metal restrains the molecule in a conformation that favors Wolff rearrangement. The opposite effect is observed in $[\mathbf{6.28}+\text{Na}]^+$, which is also conformationally restrained. In this case Wolff rearrangement is not favored, and additional activation energy must be added at each step. Interestingly, the barrier becomes much larger when the diameter of the macrocycle is reduced. This supports the proposition that conformational constraints lead to the larger activation barriers required for the CAD of **6.28** bound to various ions. Unconstrained systems with a remote charge that does not react with the carbene such as **6.27** and $[\mathbf{6.29}+\text{Na}]^+$ all demonstrate similar behavior upon CAD.

Intermolecular reactions. Given the spectator MeCN ligand present in **6.3**, the possibility for intermolecular reactions might exist given the addition of the proper ligand. The results for attaching three different ligands are summarized in Table 6.3. It is easily seen in Table 6.3 that in each of these experiments the ligand is lost (either competitively or exclusively) at some point in the experiment. This suggests that the binding energy of the ligand to the copper (I) cation is not greatly increased by the presence of the additional functional groups. In particular, 5-hexynenitrile is not strongly bound to copper (I). Therefore, if the loss of the weakly bound noncovalent ligand is not observed, but a covalent bond cleavage occurs which interrupts the established sequential rearrangement processes (i.e., loss of N_2 , CO, CO), then it is assumed that an intermolecular reaction has occurred. The reactions that lead to subsequent cleavages are complicated and it is not our intention to fully describe them here. However, it is shown

from the results that alkynes are much more reactive than analogous alkenes. This is not an unusual result given the greater reactivity of alkynes in general. Similarly, the observed reactivity of the phenyl group was unexpected. In the case of **6.2**, the benzyl groups ultimately displaced the ligand and prevented any reactions from occurring.

It is also interesting to note that CAD of [**6.1**+hexynenitrile+Cu]⁺ yielded the loss of a neutral methyl radical in the MS³ spectrum. This unusual loss was confirmed by experiments with **6.26** in which the loss of CD₃ was observed. The loss of methyl radical is accompanied by the pickup of water or methanol. It is unlikely that the methyl group was a metal ligand prior to dissociation, so an alternative explanation is required to explain the pickup of water and methanol. One possible explanation for this involves the oxidation of the copper (I) to copper (II) accompanied by the reduction of the newly formed terminal CO₂. The oxidation to copper (II) would create a new vacant ligand site and lead to the observed pickups. The exact role of the alkyne in this process remains unclear, but it may serve to stabilize the higher oxidation state of the copper.

Table 6.3 Intermolecular reactions with Fischer carbenes.

Molecule	Cation	Initial Ligand	Partial Loss of Ligand	Covalent Bond Cleavage [†]
6.1	Cu ⁺	5-hexynenitrile	yes	MS ³
6.2	Cu ⁺	5-hexynenitrile	yes	no
6.1	Cu ⁺	5-hexenenitrile	yes	no
6.1	Cu ⁺	4-phenylbutyronitrile	yes	MS ³

[†]If covalent bond cleavage occurs either competitively with loss of the ligand or exclusively, the step where this occurs is indicated.

6.4 Conclusions

In summary, the synthesis of copper (I) and silver (I) Fischer carbenes from various diazo malonates in the gas phase is demonstrated for the first time. The carbenes are generated by the facile loss of N_2 . The carbenes prepared by this process can undergo multiple Wolff rearrangements in the gas phase, which subsequently lead to the loss of multiple carbon monoxides. Surprisingly, up to six different carbenes or metallo-carbenoids can be produced in a single experiment. A series of control experiments that elucidate the effects of conformation and metal mediation on Wolff rearrangements are detailed. Although the data has been gathered from gas phase experiments, the results reveal general trends that should be applicable for either enhancing or deterring Wolff rearrangements in solution. Coordinated metal ions profoundly mediate the energetics of these reactions. Silver (I) is most efficient at initiating Wolff rearrangements followed by copper (I), but no mediation is required and molecules labeled with non-participating charges undergo similar chemistry at higher energies. Conformational effects are also found to be important in determining the requisite energy for Wolff rearrangement. Divalent metal ions and protons interfere with the loss of N_2 and do not promote the formation of carbenes.

6.5 References

- ¹ (a) Stevens, A. E.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1978**, *100*, 2584-2585. (b) Stevens, A. E.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1979**, *101*, 6449-6450. (c) Halle, L. F.; Armentrout, P. B.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1981**, *103*, 962-963.
- ² Adlhart, C.; Hinderling, C.; Baumann, H.; Chen, P. *J. Am. Chem. Soc.* **2000**, *122*, 8204-8214.
- ³ Bertani, R.; Michelin, R. A.; Mozzon, M.; Traldi, P.; Seraglia, R.; Busetto, L.; Cassani, M. C.; Tagleatesta, P.; D'Arcangelo, G. *Organometallics* **1997**, *16*, 3229-3233.
- ⁴ (a) Doyle, M. P.; McKervey, M. A.; Ye, T. *Modern Catalytic Methods for Organic Synthesis with Diazo Compounds*; Wiley-Interscience: New York, 1998. (b) Moody, C. J.; Whitham, G. H. *Reactive Intermediates*; Oxford University Press: New York, 1992; 26-50. (c) Straub, B. F.; Hofmann, P. *Angew. Chem. Int. Ed.* **2001**, *40*, 1288-1290.
- ⁵ Julian, R. R.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L. *Angew. Chem. Int. Ed.* **2003**, *42*(9), 1012-1015.
- ⁶ (a) Marzluff, E. M.; Beauchamp, J. L. In *Large Ions: Their Vaporization, Detection, and Structural Analysis*; Baer, T., Ng, C. Y., Powis, I.; Eds.; John Wiley & Sons Ltd.: New York, 1996; 115-143. (b) McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599-614. (c) Hayes, R. N.; Gross, M. L. *Methods Enzymol.* **1990**, *193*, 237-263.
- ⁷ Richardson, D. C.; Hendrick, M. E.; Jones, M. *J. Am. Chem. Soc.* **1971**, *93*, 3790-3791.
- ⁸ Marfisi, C.; Verlaque, P.; Davidovics, G.; Pourcin, J.; Pizzala, L.; Aycard, J.-P.; Bodot, H. *J. Org Chem.* **1983**, *48*, 533-537.
- ⁹ Wolff, L. *Justus Liebigs Ann. Chem.* **1902**, 325, 129.

-
- ¹⁰ (a) Sudrik, S. G.; Chavan, S. P.; Chandrakumar, K. R. S.; Pal, S.; Date, S. K.; Chavan, S. P.; Sonawane, H. R. *J. Org. Chem.* **2002**, *67*(5), 1574-1579 and references therein. (b) McMahon, R. J.; Chapman, O. L.; Hayes, R. A.; Hess, T. C.; Krimmer, H. P. *J. Am. Chem. Soc.* **1985**, *107*, 7597-7606. (c) Fenwick, J.; Frater, G.; Ogi, K.; Strausz, O. P. McMahon, R. J.; Chapman, O. L.; Hayes, R. A.; Hess, T. C.; Krimmer, H. P. *J. Am. Chem. Soc.* **1973**, *95*, 124-132. (d) Pomerantz, M.; Levanon, M. *Tetrahedron Lett.* **1991**, *32*, 995-998. Also see ref. 4a Chapter 9.
- ¹¹ Arduengo, A. J.; Dias, H. V. R.; Calabrese, J. C.; Davidson, F. *Organometallics* **1993**, *12*, 3405.
- ¹² Julian, R.R.; Beauchamp, J. L. *Int. J. Mass. Spectrom.* **2001**, *210*, 613-623.
- ¹³ Falorni, M.; Dettori, G.; Giacomelli, G. *Tetrahedron: Asymmetry* **1998**, *9*, 1419
- ¹⁴ Kametani, T.; Yukawa, H.; Honda, T. *J. Chem Soc. Perkin Trans. 1* **1990**, *3*, 571-577.
- ¹⁵ Lamb, J. D.; Izatt, R. M.; Swain, C. S.; Bradshaw, J. S.; Christensen, J. J. *J. Am. Chem. Soc.* **1980**, *102*, 479-482.
- ¹⁶ Dias, H. V. R.; Polach, S. A. *Inorg. Chem.* **2000**, *39*, 4676-4677.
- ¹⁷ Wentrup, C. *Reactive Molecules: The Neutral Reactive Intermediates in Organic Chemistry*; John Wiley & Sons: New York, 1984; 177-264.
- ¹⁸ Chang, S.-C.; Kafafi, Z. H.; Hauge, R. H.; Billups, W. E.; Margrave, J. L. *J. Am. Chem. Soc.* **1987**, *109*, 4508-4513.
- ¹⁹ (a) Vitale, G.; Valina, A. B.; Huang, H.; Amunugama, R.; Rodgers, M. T. *J. Phys. Chem. A* **2001**, *105*, 11351-11364. (b) Chu, Y.; Yang, Z.; Rodgers, M. T. *J. Am. Mass Spectrom.* **2002**, *13*, 453-468.

²⁰ (a) Barluenga, J.; Lopez, L. A.; Lober, O.; Tomas, M.; Garcia-Granda, S.; Alvarez-Rua, C.; Borge, J. *Angew. Chem. Int. Ed.* **2001**, *40*, 3392-3394. (b) Tulloch, A. A. D.; Danopoulos, A. A.; Kleinhenz, S.; Light, M. E.; Hursthouse, M. B.; Eastham, G. *Organometallics* **2001**, *20*, 2027-2031.

²¹ (a) Likhovorik, I.; Zhendong, Z.; Tae, E. L.; Tippmann, E.; Hill, B. T.; Platz, M. S. *J. Am. Chem. Soc.* **2001**, *123*, 6061-6068. (b) Scott, A. P.; Platz, M. S.; Radom, L. *J. Am. Chem. Soc.* **2001**, *123*, 6069-6076.

²² Similarly, it is impossible to distinguish experimentally whether structures **5** and **17** have already undergone Wolff rearrangement prior to experimental observation.

²³ Shoeib, T.; Aribi, H. E.; Siu, K. W. M.; Hopkinson, A. C. *J. Phys. Chem. A* **2001**, *105*, 710-719.

²⁴ Katritsky, A.R.; Watson, C. H.; Dega-Szafran, Z.; Eyler, J. R. *J. Am. Chem. Soc.* **1990**, *112*, 2471-2478.

²⁵ (a) Dattelbaum, A. M.; Martin, J. D. *Inorg. Chem.* **1999**, *38*, 6200-6205. (b) Wen, M.; Munakata, M.; Suenaga, Y.; Kuroda-Sowa, T.; Maekawa, M. *Inorg. Chim. Acta* **2002**, *332*, 18-24.

²⁶ Bradshaw, J. S.; Izatt, R. M.; Borkunov, A. V.; Zhu, C. Y.; Hathaway, J. K. *Comprehensive Supramolecular Chemistry*, Vol. 1; G. W. Gokel, Pergamon/Elsevier: Oxford, 1996; 35-95.

²⁷ Maleknia, S.; Brodbelt, J. *J. Am. Chem. Soc.* **1993**, *115*, 2837-2843.

²⁸ (a) Vogt, J.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1975**, *97*, 6682-6685. (b) Pliego J. R.; DeAlmeida, W. B. *J. Chem. Soc. Faraday Trans.* **1997**, *93*, 1881-1883.

Chapter 7

Site Specific Sequestering and Stabilization of Charge in Peptides by Supramolecular Adduct Formation with 18-Crown-6 Ether via Electrospray Ionization

Portions published previously in: Julian R. R.; Beauchamp J. L. *Int. J. Mass Spectrom.* **2001**, *210*, 613-623.

7.1 Introduction

The discovery of crown ethers has led to many solution phase applications.¹ However, until recently, the applications to gas phase chemistry have been limited in comparison.^{2,3} Crown ethers are capable of forming supramolecular assemblies in the gas phase, and with the introduction of electrospray ionization (ESI), the means for easily characterizing these species is now available. 18-Crown-6 ether (18C6) binds cationic species. These adducts have been studied in both solution^{4,5} and the gas phase.⁶⁻⁹ The binding affinity of 18C6 is high for metal cations (Table 7.1).⁶ 18C6 also forms stable gas phase adducts with amines.⁷ The selectivity of crown ethers is determined by the cavity

size, which allows 18C6 to bind primary amines better than other candidate crown ethers.^{8,9}

Table 7.1 Experimental^a and Computational^b Bond dissociation energy for 18C6 complexes

Molecule	Bond Dissociation Energy (kJ/mol)
[(Methylamine+H)CE] ⁺	182 ^b
[(Lys+H)CE] ⁺ ^c	150 ^b
[(Arg+H)CE] ⁺ ^c	133 ^b
[(His+H)CE] ⁺ ^c	127 ^b
[Na·CE] ⁺	300 ^a
[K·CE] ⁺	235 ^a
[Rb·CE] ⁺	192 ^a
[Cs·CE] ⁺	170 ^a

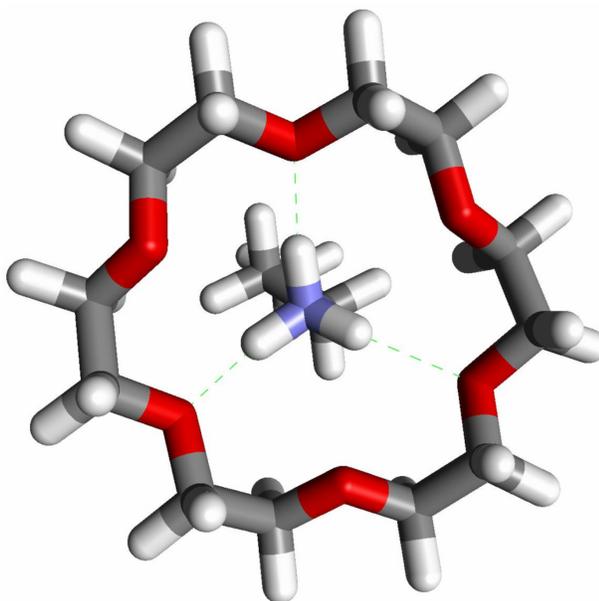
^a taken from reference 6

^b calculated with PM3 relative to separated monomers

^c in each case the CE was attached to the side chain

Lysine is a basic amino acid that plays an important role in the solvation, structure, and activity of proteins.¹⁰ The side chain of lysine terminates in a primary amine. It is possible to form a strong, hydrogen bond complex between 18C6 and protonated lysine. As a model for this, structure 7.1 depicts the noncovalent association between protonated butylamine and 18C6. Alternating oxygens of 18C6 are coplanar and serve as hydrogen bond acceptors for the protonated primary amine. The location and orientation of these three oxygens is optimal for primary amines (average N—O distance 3.1Å) and lends specificity to the binding. The bond dissociation for this interaction is 182 kJ/mol as indicated by PM3 calculations (Table 7.1). The interaction depicted in 7.1 has been applied successfully in our laboratory to study the mechanism of gas phase H/D exchange in small protonated peptides, specifically a series of glycine oligomers.¹¹ These

experiments derived advantage from the ability of 18C6 to sequester protons attached to the protonated N-terminus and inhibit their interaction with deuterated reagents that normally would have effected rapid H/D exchange.



7.1

In the present work, the interaction between lysine and 18C6 is studied with various lysine derivatives to establish the preferred binding site. The ability of 18C6 to form highly charged multiple adducts is investigated through polylysine peptides. An original objective of this study was to use the molecular recognition capabilities of 18C6 to count lysine residues in peptides of unknown sequence. It is shown, however, that complications result from competitive complexation by the protonated side chains of histidine, arginine, and particularly the n-terminus of peptides that do not contain lysine. Nevertheless, 18C6 is shown to be a sensitive chemical probe of molecular structure, and experiments with cytochrome-c and bovine pancreatic trypsin inhibitor (BPTI) suggest that 18C6 can provide information relating to surface functional groups of folded proteins. The binding of 18C6 to peptides and proteins dramatically alters both the ion

abundance and the observed charge state distribution. This charge stabilization is useful because it allows for the facile detection of species that are normally difficult to detect with ESI. The interaction of 18C6 with molecular ions sheds further light on the mechanism of ESI, which remains poorly understood.¹² The results are best rationalized in accordance with the ion evaporation mechanism originally proposed by Iribarne and Thomson.¹³

7.2 Experimental Methodology

All spectra were obtained using a Finnigan LCQ ion trap quadrupole mass spectrometer without modification. The critical instrument settings that yield adduct formation include capillary voltage 14.12V, capillary temperature 200°C, and tube lens offset -39V. Higher capillary temperatures dissociate the 18C6 complexes. The tube lens offset controls the acceleration of ions as they leave the capillary region. The tube lens voltage is minimized to avoid collisions with the He buffer gas. Soft sampling is crucial for the detection of these noncovalent complexes.

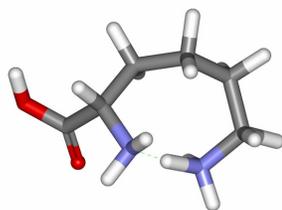
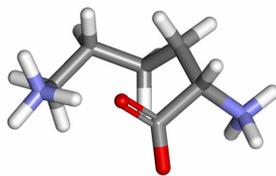
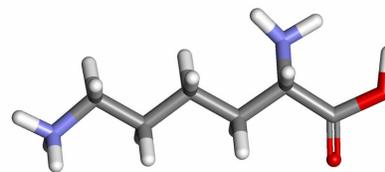
Sample concentrations were typically kept in the ~10 to 100 μM range for all species of interest, unless otherwise noted. All samples were electrosprayed in a mixture of 80:20 methanol/water. 18C6 was added to the sample and electrosprayed with the analyte in order to observe adducts. Samples were electrosprayed with a flow of 3-5 $\mu\text{L}/\text{min}$ from a 500 μL Hamilton syringe for optimal signal. Silica tubing with an inner diameter of .005 in was used as the electrospray tip. No acid was added to any of the samples, unless otherwise noted. Acid has a tendency to decompose 18C6 under electrospray conditions. All chemicals unless otherwise noted were purchased from

Sigma or Aldrich and used without further purification. The cytochrome-c used in this study was taken from horse heart. KK-methyl ester was synthesized by dissolving KK in methanol with 1% H₂SO₄ and heating to 60°C for one hour. The yield was ~85 percent as determined by mass spectrometry.

All semi-empirical calculations were performed with the HyperChem 5.1 Professional Suite. Candidate structures were identified with molecular mechanics using Cerius² version 4.0 by Molecular Simulations Inc. with the Dreiding 2.21¹⁴ force field. Molecular volumes are based on a smoothed Van der Waals radius and were calculated with Cerius². Jaguar 4.0, Schrodinger, Inc., Portland, Oregon, 2000 was used to perform all density functional theory (DFT) calculations at the B3LYP level with the 6-31G** basis set.

7.3 Results and Discussion

Lysine and Lysine Derivatives. A mixture of lysine and lysine methyl ester (MeLys) was electrosprayed with the addition of 18C6. The observed mass spectrum is shown in Figure 7.1a. The complex [Lys+18C6+H]⁺ is formed and is singly charged. The spectrum shows that only a single 18C6 adds to lysine. This is not the case for MeLys, where there is a significant doubly charged double adduct peak corresponding to [MeLys+2(18C6)+2H]²⁺. The C-terminus of MeLys is methylated, precluding deprotonation. The presence of an adduct with two 18C6 ethers attached in the case of MeLys contrasts with the observation that lysine attaches only a single 18C6. The difference in behavior between these species can be attributed to a stable molecular conformation that is only available to lysine.

**7.2****7.3****7.4**

DFT calculations were performed on several possible conformations of protonated lysine (shown as structures **7.2-7.4**). The structures were fully minimized at the B3LYP/6-31G** level. The high proton affinity of lysine (238 kcal/mol) relative to butylamine (220.2 kcal/mol) can be attributed to the favorable intramolecular hydrogen bond¹⁵ formed in structure **7.2**, which must be broken to attach 18C6. Just 3 kcal/mol in energy above **7.2** is the salt bridge structure **7.3**, with the carboxylate between the two protonated amine sites. Structure **7.3** can only bind 18C6 at the n-terminus without significantly disrupting the favorable electrostatic interaction in the salt bridge. 18C6 can attach to the side chain in structure **7.4**, which is 16.1 kcal/mol higher in energy than **7.2**. Thus it is likely that 18C6 will attach preferentially to the n-terminus of lysine to preserve the 13 kcal/mol stabilization afforded by the salt bridge structure. Doubly protonated MeLys can attach two 18C6 ethers in a conformation similar to that shown in **7.4** because it is incapable of forming a salt bridge structure similar to structure **7.3**.

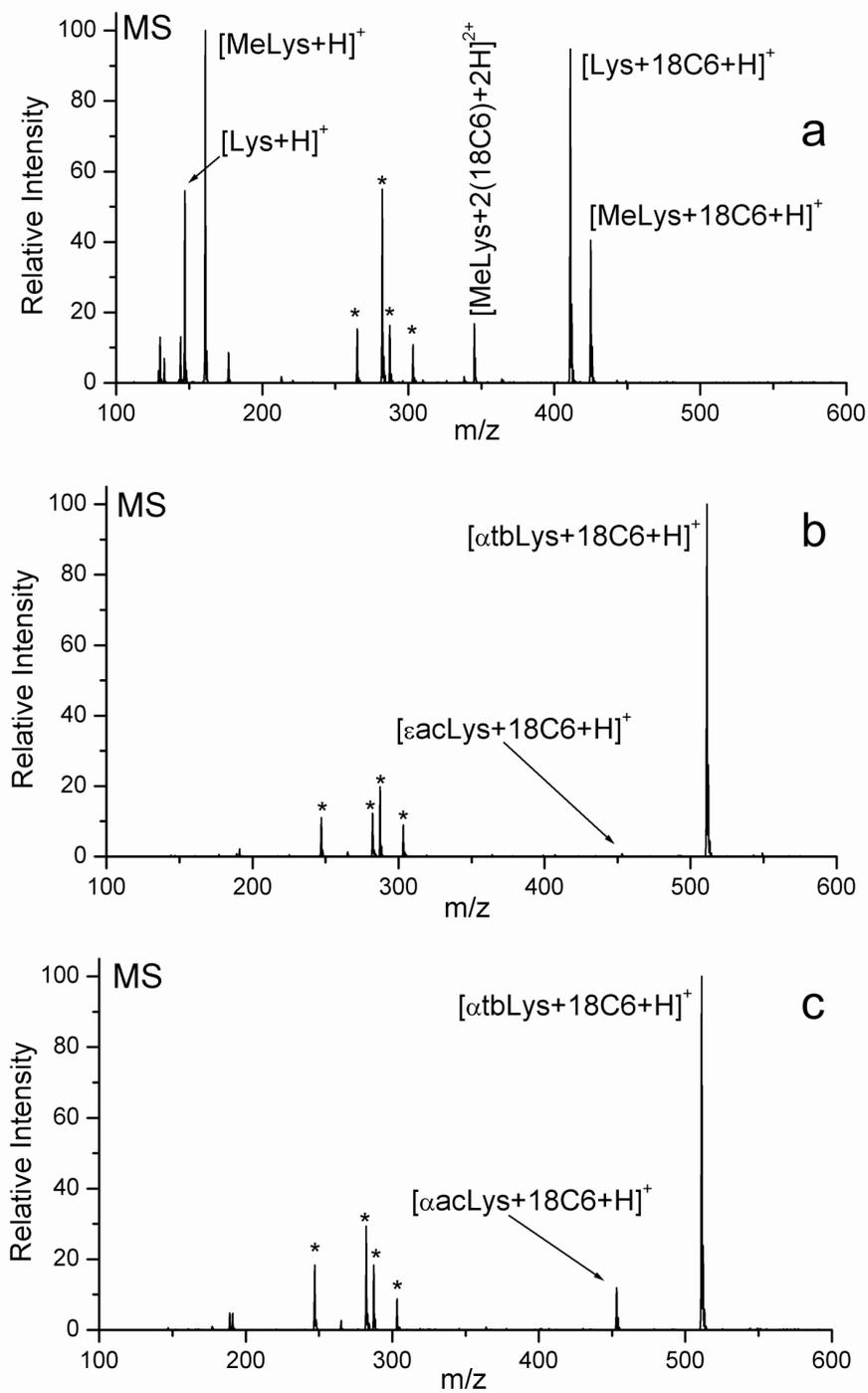
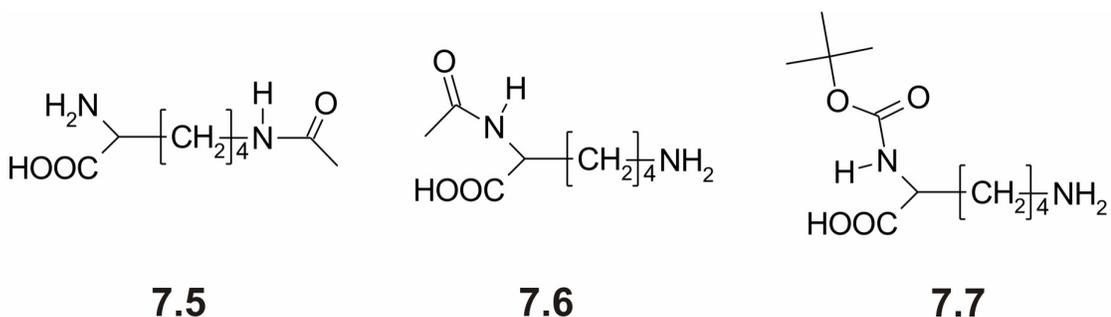


Figure 7.1 (a) Mass spectrum of equimolar mixture of Lys and MeLys with 18C6 showing the difference in binding between these two similar species. The preference for binding to the side chain of lysine is shown by comparison of the competitive binding

between (b) 18C6 with equimolar amounts of ϵ -n-acetyl lysine and α -n-tertbutoxycarbonyl lysine, and (c) equimolar α -n-acetyl lysine and α -n-tertbutoxycarbonyl lysine with 18C6. * indicates peaks corresponding to other 18C6 peaks.

The competitive binding of 18C6 to the n-terminus and the alkyl-ammonium side chain of lysine were probed with lysine derivatives **7.5-7.7**. The results are shown in Figure 7.1b and Figure 7.1c. The effect of blocking the side chain in **7.5** was compared to blocking the N-terminus in **7.6** and **7.7**. The direct comparison of N-acetylated species is not possible due to their identical mass, and therefore the binding affinities were compared relative to **7.7**. In each case, equimolar mixtures of the lysine derivatives were electrosprayed with 18C6. The intensity of the 18C6 complex with **7.5** is 15 times smaller than that for 18C6 complexed with **7.6**. This suggests that 18C6 has a strong preference for binding to the side chain of lysine when there is no possibility of salt bridge formation.



Arginine and Histidine. The side chains of lysine, arginine, and histidine all contain basic sites that are protonated at biological pH and can therefore interact with 18C6. To

examine the binding affinity of 18C6 for each of these side chains a solution of 50 μM each of imidazole, guanidine, and n-butyl amine was electrosprayed with 10 μM 18C6. The spectrum is shown in Figure 7.2a. It is observed that $[\text{nbutylamine}+18\text{C6}+\text{H}]^+$ completely dominates the spectrum, with 3.5% relative intensity for the guanidinium adduct and 1% for the imidazole ring adduct. This result demonstrates the preference of 18C6 for binding to a protonated primary amine, especially considering that n-butyl amine has the lowest proton affinity¹⁶ of these three molecules (Table 7.2). In a similar experiment 50 μM n-butyl amine and alanine were electrosprayed with 10 μM 18C6. Again the results (Figure 7.2b) show that 18C6 has a definite preference for the terminal alkyl amine over the n-terminus.

Table 7.2 Molecular Volume and Proton Affinity of Various Species

Species	Molecular Volume (\AA^3)	Proton Affinity (kJ/mol) ^a
18C6	263.32	967 ^b
NH_4^+	25.84	--
KK	--	1012 ^c
$[\text{KK}+\text{H}]^+$	--	842 ^c
$[\text{KK}+2\text{H}]^{2+}$	--	591 ^c
KKKK	--	1079 ^c
$[\text{KKKK}+\text{H}]^+$	--	967 ^c
$[\text{KKKK}+2\text{H}]^{2+}$	--	760 ^c
$[\text{KKKK}+3\text{H}]^{3+}$	--	680 ^c
Lys	--	996 ^{b,c}
$[\text{Lys}+\text{H}]^+$	--	693 ^c
MeLys	--	1004 ^c
$[\text{MeLys}+\text{H}]^+$	--	696 ^c
Methanol	--	754.3 ^b
Guanidine	--	986.3 ^b
n-butyl amine	--	921.5 ^b
Imidazole	--	942.8 ^b
Glycine	--	886.5 ^b
Alanine	--	901.6 ^b

^a calculated values adjusted to match experimental PA of lysine.

^b taken from reference 23.

^c calculated with PM3 using the experimental heat of formation for H^+ .

Proline, Glycine, and Alanine. Protonated proline does not form an observable adduct with 18C6 (spectrum not shown). Thus only a peptide with a primary amine n-terminus can bind 18C6 at that position. In the absence of basic residues, there is an immediate benefit that results from the ability of 18C6 to bind to a protonated n-terminus. Glycine and Alanine are two amino acids that are difficult to electrospray, particularly in the absence of acid. This is most likely due to the low proton affinities of these amino acids. With the addition of 18C6, glycine and alanine are detected easily by ESI mass spectrometry in the absence of acid (as seen in Figure 7.2c). The isolated charged species can be obtained via mild CID. As discussed further below this observed charge stabilization in transferring ions from solution to the gas phase appears consistent with the ion evaporation model¹³ for ESI.

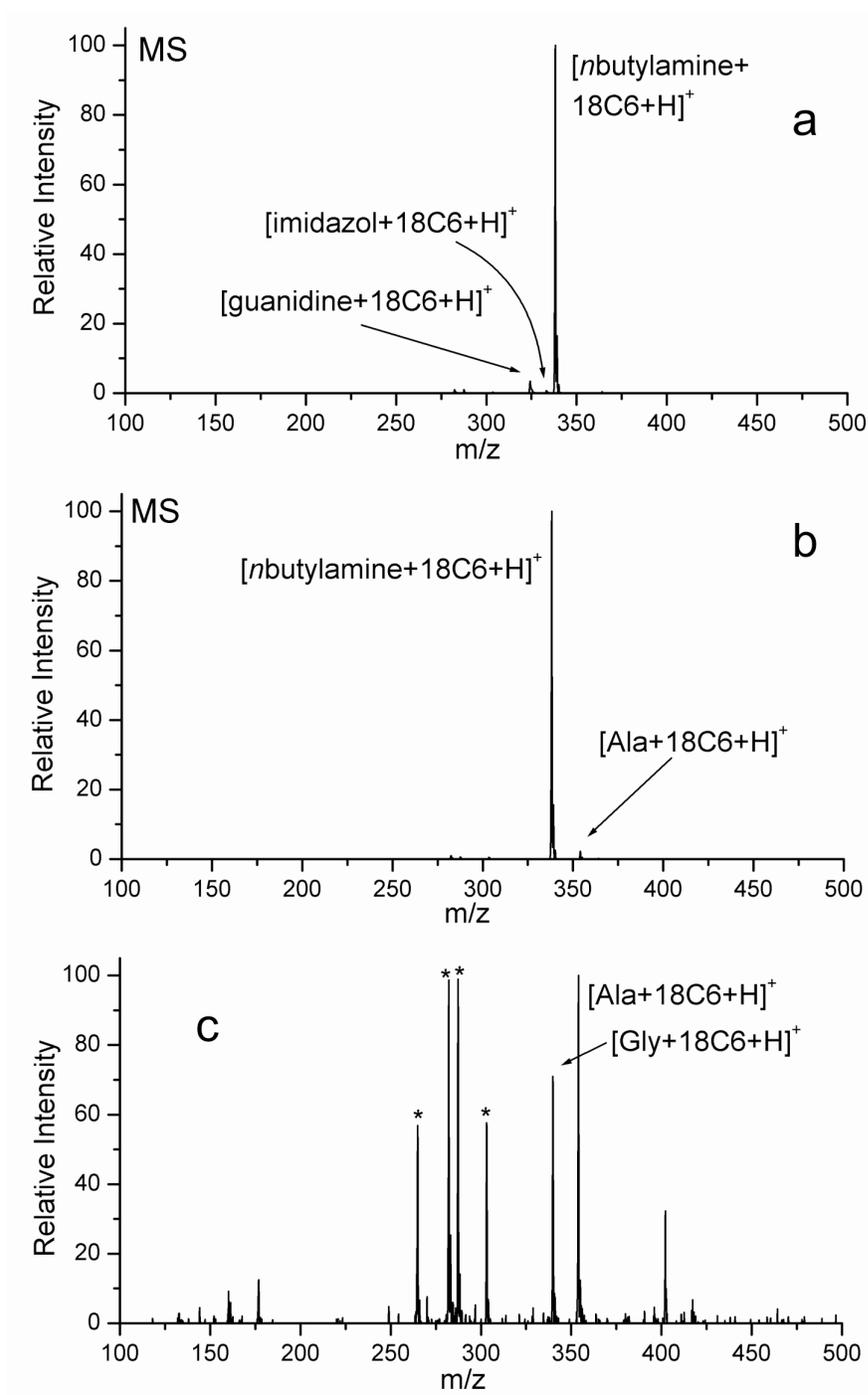


Figure 7.2 (a) Equimolar mixture of n-butyl amine, imidazole, and guanidine with 18C6. These functional groups represent the side chains of lysine, histidine, and arginine respectively and effectively approximate the competitive binding between these amino acids without interference from the N-terminus. (b) Spectra of equimolar mixture of n-

butyl amine and alanine in which the N-terminus is shown to be much less favored than a terminal alkyl amine. (c) In the absence of alkyl amines, glycine and alanine are easily detected by complexation with 18C6 in the absence of acid.

Polylysines. The di-peptide KK was electrosprayed with and without 18C6. The results are shown in Figures 7.3a and 7.3b, respectively. KK is singly charged when electrosprayed under normal conditions, even with the addition of 0.1% acetic acid. With the addition of 18C6, the primary peak is $[\text{KK}+2(18\text{C}6)+2\text{H}]^{2+}$. There are additional minor peaks corresponding to one 18C6 adduct (singly and doubly charged). The +2 charge state that is absent in the spectrum of KK alone becomes the primary peak with the addition of 18C6. In a similar result, methyl-ester dilysine (Me-KK) attaches two 18C6 ethers to a doubly charged cation (spectra not shown). This supports the evidence given above suggesting that the 18C6 attaches to the side chain of lysine. KK is unable to form the salt bridge structure found in lysine, therefore the spectrum of Me-KK with 18C6 present yields results essentially identical to those for KK. This again points to the fact that lysine itself is a special case.

The tetra-peptide KKKK was electrosprayed with and without 18C6 (Figure 7.3c and Figure 7.3d). The peptide is primarily doubly charged under normal electrospray conditions. In the presence of 18C6, $[\text{K}4\text{K}4+4(18\text{C}6)+4\text{H}]^{4+}$ is formed. The complex is quadruply charged as seen in the inset high resolution scan in Figure 7.3d. The side chains are flexible enough to accommodate multiple crown ethers complexed to adjacent lysines. This is illustrated by the PM5 structure **7.8** for the tetra-adduct of KKKK.

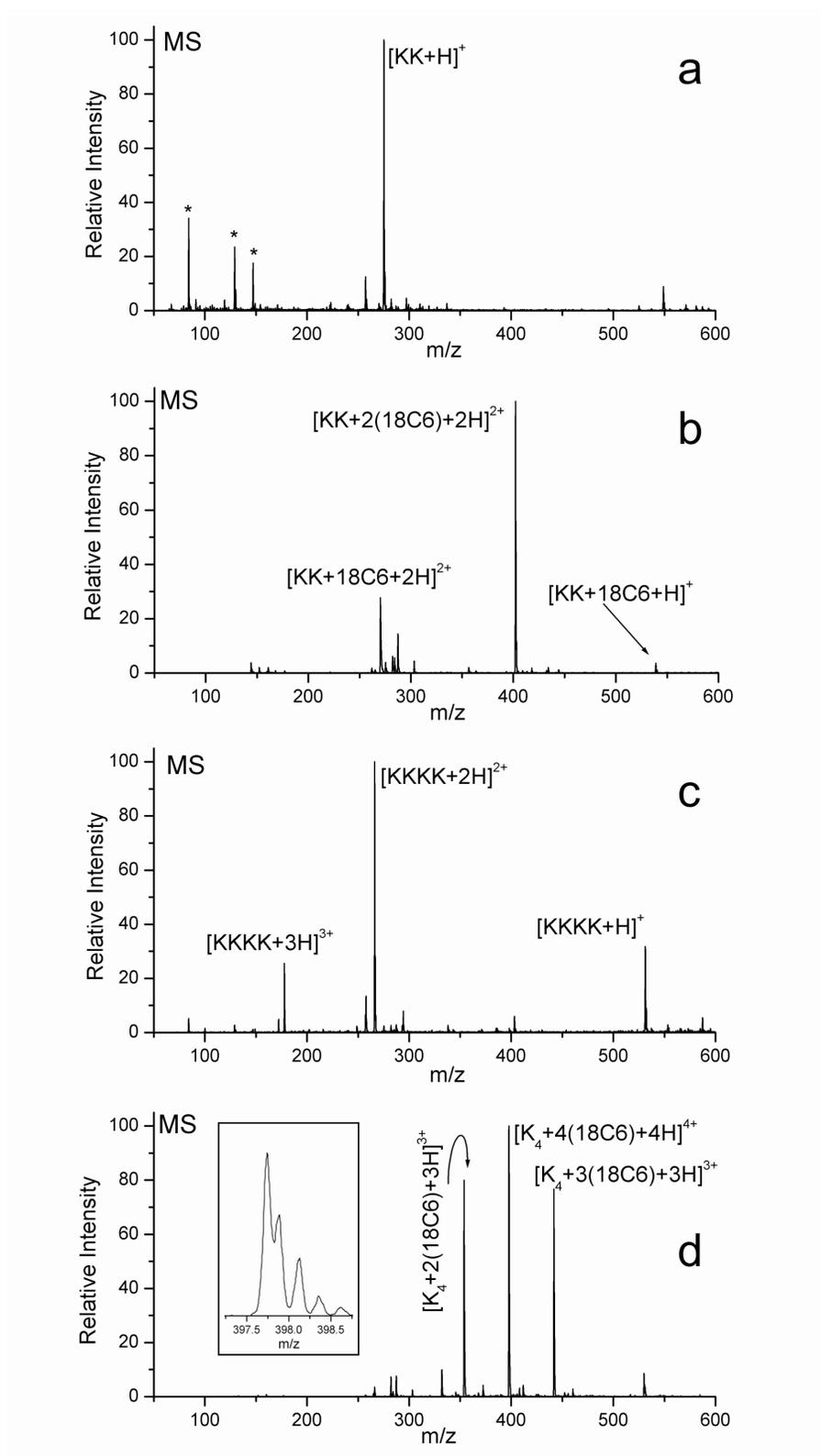
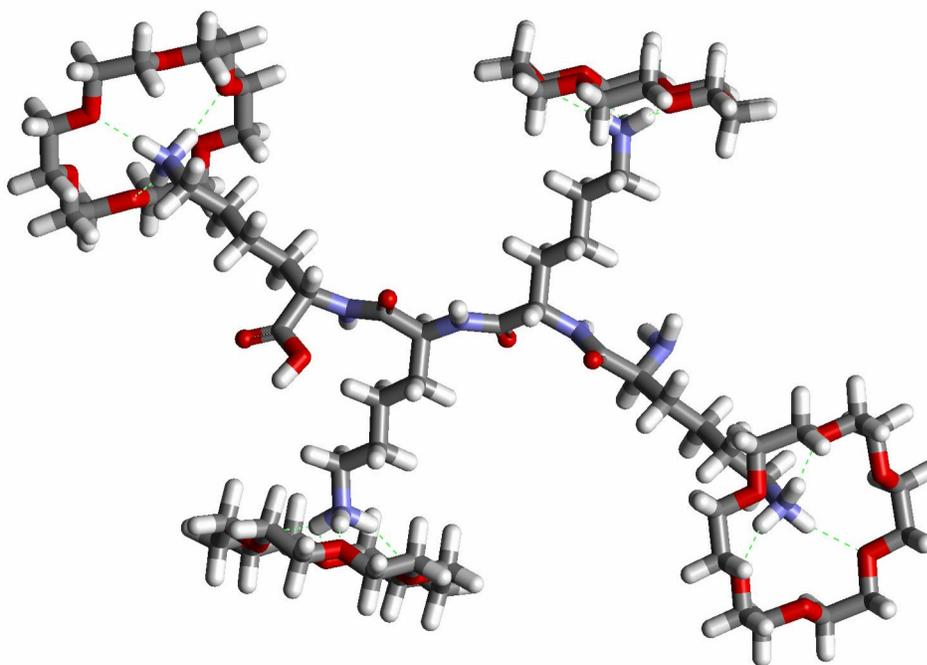


Figure 7.3 Caption on following page.

Figure 7.3 Polylysine spectrum illustrating the charge stabilization between (a), the ESI spectrum of KK with acid under normal ESI conditions, and (b) ESI of KK with 18C6 in the absence of acid. This result is more dramatically repeated in comparison of (c) typical ESI spectrum of KKKK with (d), the ESI spectrum of KKKK with 18C6 (inlaid portion is zoomscan of peak near 398). Notice the increase in charge state for both molecules and the 1:1 Lys to 18C6 complexation. * indicate lysine fragments.



7.8

Small Model Peptides. Two peptides containing both Arg and Lys were examined. The dipeptide RK is both singly (59%) and doubly (41%) charged under normal electrospray conditions. The spectrum of RK mixed with 18C6 is shown in Figure 7.4a. The primary peak is a doubly charged species corresponding to $[RK+18C6+2H]^{2+}$. A typical electrospray spectrum for Tuftsin (TKPR) shows that this peptide prefers to be

singly (84%) charged. As shown in Figure 7.4b, it becomes primarily doubly charged with the addition of 18C6. The most abundant crown ether adduct for TKPR with 18C6 is $[\text{TKPR}+18\text{C6}+2\text{H}]^{2+}$.

The peptide PHPFHFVTK (referred to as Turn) contains one lysine, two histidines, and proline at the n-terminal position. The spectrum of Turn in Figure 7.4c was taken under normal electrospray conditions and demonstrates a preference for the doubly charged state. When 18C6 is added (Figure 7.4d), the charge state increases to +3 and there is primarily one crown ether adduct, i.e., $[\text{Turn}+18\text{C6}+3\text{H}]^{3+}$. The singly charged peptide does not have a corresponding 18C6 adduct peak, suggesting that the labile proton preferentially resides on one of the histidines. The doubly charged peptide attaches only one 18C6, suggesting that the charges reside on opposite ends of the molecule with one proton on the c-terminal lysine and the other on the histidine adjacent to the n-terminus. The addition of a third proton likely occurs on the remaining histidine, and can be accompanied by attachment of a second 18C6, leading to the peak corresponding to a double 18C6 adduct.

The results suggest that for biologically relevant multifunctional peptides, 18C6 has a strong but not exclusive preference for binding to the side chain of lysine. As an alternative binding site, the peptide n-terminus is significantly less basic and more sterically constrained than the side chain of lysine.¹⁷ Arginine and histidine are more basic than lysine but lack the stronger specific binding of 18C6 to protonated primary amines. In small peptides 18C6 can bind to adjacent lysines, and there is a one to one correlation between the number of attached 18C6 ethers and lysine residues in what is usually the most intense peak in the spectrum.

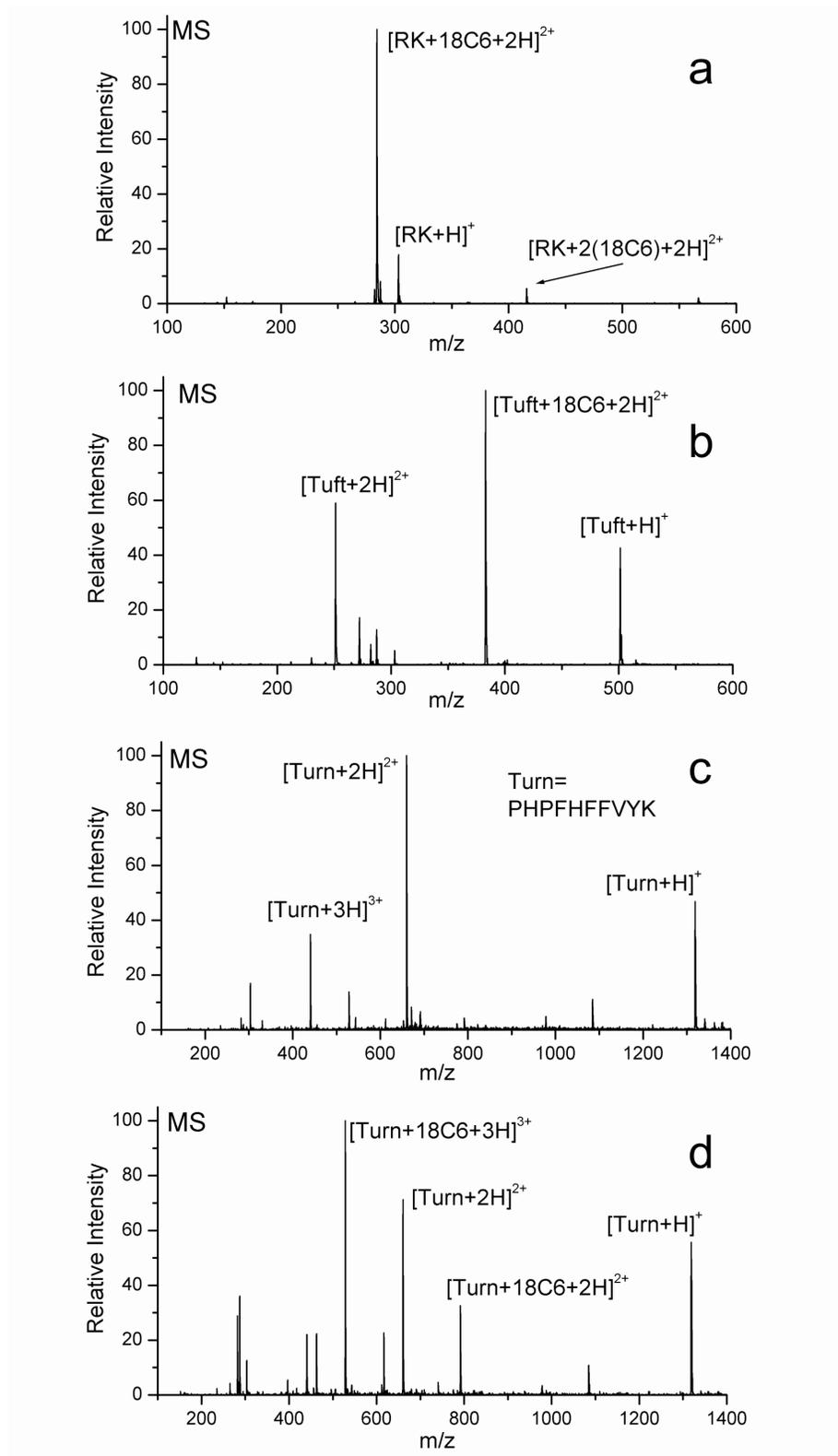


Figure 7.4 Caption on following page.

Figure 7.4 (a) Mass spectrum of peptide RK with 18C6. (b) Mass spectrum of tuftsin with 18C6. (c) Typical ESI spectrum of Turn, which with the additions of 18C6 (d) mass again shows an increase in charge state. Some competitive binding by histidine is present in the doubly charged cation.

Large Peptides and Proteins. It is tempting to infer that the number of lysines in a peptide can be quantified by counting the number of 18C6 ethers attached to the most abundant adduct. This approach will often work. However, there are several caveats that can make this determination very difficult for a peptide of unknown sequence. For example, molecular conformation may prevent 18C6 from attaching to buried or sterically hindered lysine residues. Cytochrome-c has 19 lysine residues, but attaches a maximum of four 18C6 ethers at room temperature (Table 7.3). This is less than the eleven surface exposed lysines of the folded cytochrome-c. Interestingly, however, there are only four exposed lysines (the maximum number of 18C6 ethers attached to any charge state of cytochrome-c) which are not involved in salt bridges. When cytochrome-c is denatured in solution by heating to 80°C in the presence of 18C6 and then cooled to room temperature, up to eleven crown ethers attach (Table 7.4). 18C6 appears to prevent the refolding of cytochrome-c since nearly identical spectra were obtained when the same solution was electrosprayed two days later.

Table 7.3 18C6 Adduct Distribution with Horse Heart Cytochrome-C at 20 °C

Charge State	Maximum # of 18C6 Adducts	# of 18C6 in Most Abundant Peak
19	4	4
18	4	3
17	3	3
16	4	1
15	4	1
14	4	1
13	3	1
12	4	1
11	4	1
10	3	1
9	2	1
8	-	-
7	-	-

Table 7.4 18C6 Adduct Distribution with Horse Heart Cytochrome-C heated to 80 °C for One Minute and Cooled to 20 °C.

Charge State	Maximum # of 18C6 Adducts	# of 18C6 in Most Abundant Peak
19	9	9
18	10	10
17	9	9
16	11	9
15	9	8
14	11	10
13	7	7
12	7	6
11	6	5
10	5	3
9	-	-
8	-	-
7	-	-

As noted above, 18C6 can dramatically alter the charge state and abundance of ions observed by ESI. In proteins, this effect is most dramatic where arginine is the primary charge carrier. For example, in Figure 7.5a, BPTI (Bovine pancreatic trypsin inhibitor) shows a dramatic shift in charge state distribution when compared to cytochrome-c (Figure 7.5b). BPTI has 6 arginines and 4 lysines, whereas cytochrome-c has only 2 arginines with 19 lysines. For BPTI, the addition of 18C6 enables the lysine residues to retain available protons in the presence of the more basic arginine residues. In the case of cytochrome-c, most of the exposed lysine residues are already charged and the addition of 18C6 does not greatly affect the charge state distribution.

For proteins, the task of lysine quantification is probably impossible. However the preliminary results presented here are sufficiently intriguing to warrant more detailed studies of a larger number of proteins by varying conditions to effect protein unfolding in the presence of 18C6.

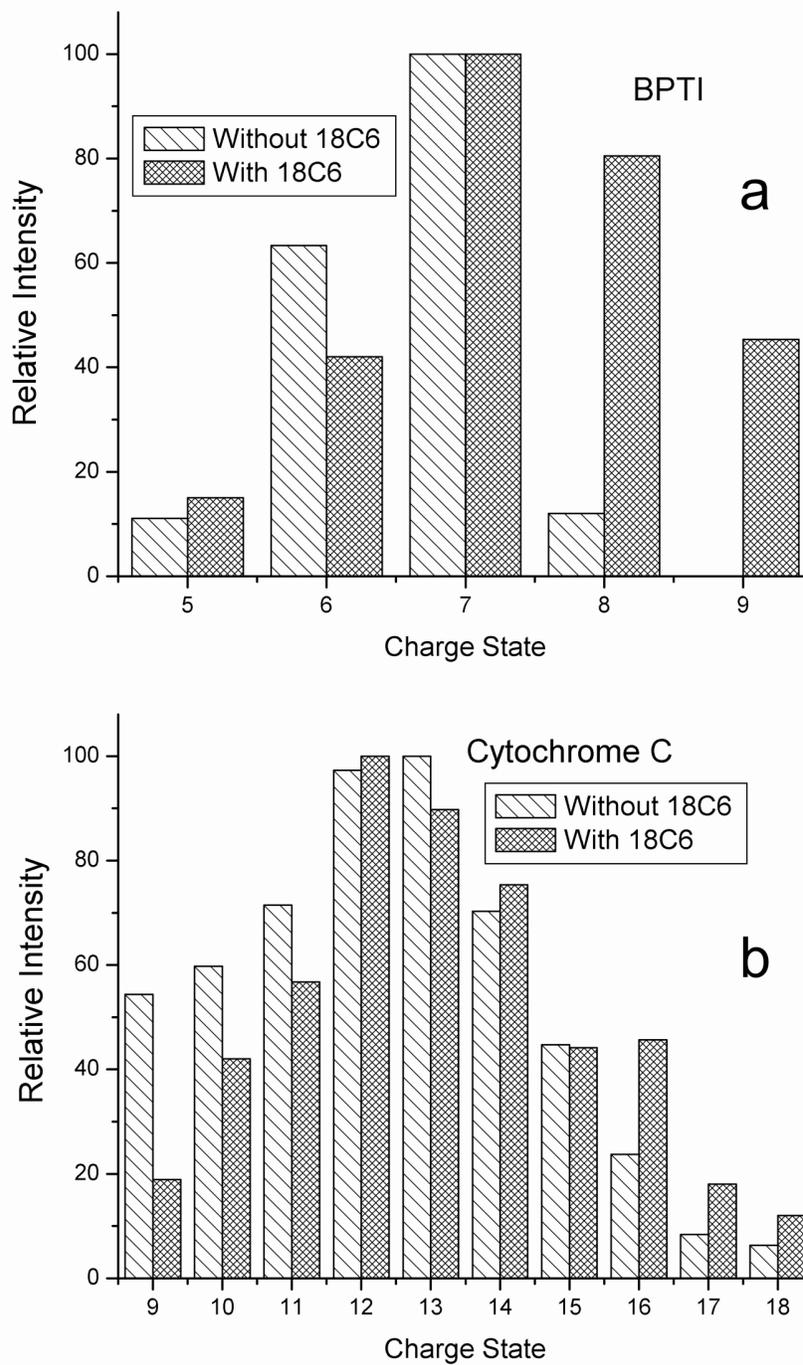


Figure 7.5 (a) Charge state distribution of BPTI showing the effect of adding 18C6. (b) Cytochrome c taken from horse heart shows less charge stabilization, probably due to fewer arginine residues and the inaccessibility of lysine residues in surface salt bridges.

Charge Stabilization and Increased Ion Abundance. The maximum charge state for a peptide produced by ESI is determined by a variety of factors including pH, functional group proton affinities, and solvent. The addition of 18C6 leads to an increase in the charge state for most molecular species. The calculated and experimental proton affinities for the various charge states of several molecules are shown in Table 7.2. Based on these proton affinities, $[\text{KKKK}+3\text{H}]^{3+}$ should be the highest charge state produced for KKKK, because the proton affinity of methanol is higher than that for $[\text{KKKK}+3\text{H}]^{3+}$. This prediction is in agreement with the experimental results found in Figure 7.3c and other theoretical approaches predicting the maximum charge state produced with ESI.^{18,19} With the addition of 18C6, the +4 charge state is not only observed, but it represents the most abundant ion in the spectrum.

There are essentially two competing processes that control ion evaporation in a highly charged droplet. Coulombic repulsion (or the electric field due to excess charge) drives the charges near the surface of the droplet, while the free energy of solvation retains the charges within the solvent. As a droplet shrinks due to normal evaporation the coulombic repulsion between molecules of like charge increases. Rayleigh fissioning will lead to smaller droplets.²⁰ Recently experiments in our labs confirmed that when droplets undergo Rayleigh fissioning, a methanol droplet releases 15-20% of the charge with little accompanying loss of solvent.²¹ The resulting offspring droplets are very highly charged, and ion evaporation is postulated to occur in these very small droplets when coulombic repulsion overcomes the free energy of solvation for a particular ion. The exact nature of this event is unclear.

It has been shown previously by studies of nonpolar analytes that relative ion abundance can be increased by increasing the surface activity of an analyte.^{22,23} 18C6 adducts also lead to enhanced ion yields in electrospray ionization. 18C6 excludes solvent from the ion and is often used as a phase transfer catalyst suggesting that the resulting complexes are neither hydrophilic nor hydrophobic. The surface activity of an ion complexed with 18C6 is increased by reducing both the polar nature of the ion and the magnitude of the free energy of solvation. The Born equation serves as a first approximation for the free energy of solvation.²⁴ In a semi-quantitative approximation, the calculated molecular volumes listed in Table 7.2 are taken to be spherical. The corresponding radius is used in the Born equation. Using this approach, the addition of 18C6 reduces the free energy of solvation for NH_4^+ by 204 kJ/mol. Based on this result, it is expected that $[\text{K}^+\text{K}^+\text{K}^+\text{K}^+ + 4(18\text{C}6) + 4\text{H}^+]^{4+}$ would be very easily desolvated in comparison with $[\text{K}^+\text{K}^+\text{K}^+\text{K}^+ + 4\text{H}^+]^{4+}$. The coulombic repulsion within the droplet and low magnitude of the free energy of solvation would increase the surface activity and concomitantly the ion abundance of the 18C6 adduct.

The preferential charging and desolvation of molecular species with 18C6 attached has an interesting consequence. In an experiment with KKKK, the concentration of 18C6 was varied from slightly below the stoichiometric ratio for the number of available binding sites (i.e. the number of lysines) to one 18C6 for four lysines. As seen in Figure 7.6, the spectrum changes very little despite the lowering of the 18C6/KKKK ratio well below the stoichiometric equivalency. The fact that the $[\text{K}^+\text{K}^+\text{K}^+\text{K}^+ + 4(18\text{C}6) + 4\text{H}^+]^{4+}$ peak dominates the spectrum shown in Figure 7.6b suggests that solution phase concentrations are clearly not represented in the gas phase ion abundance of these species.

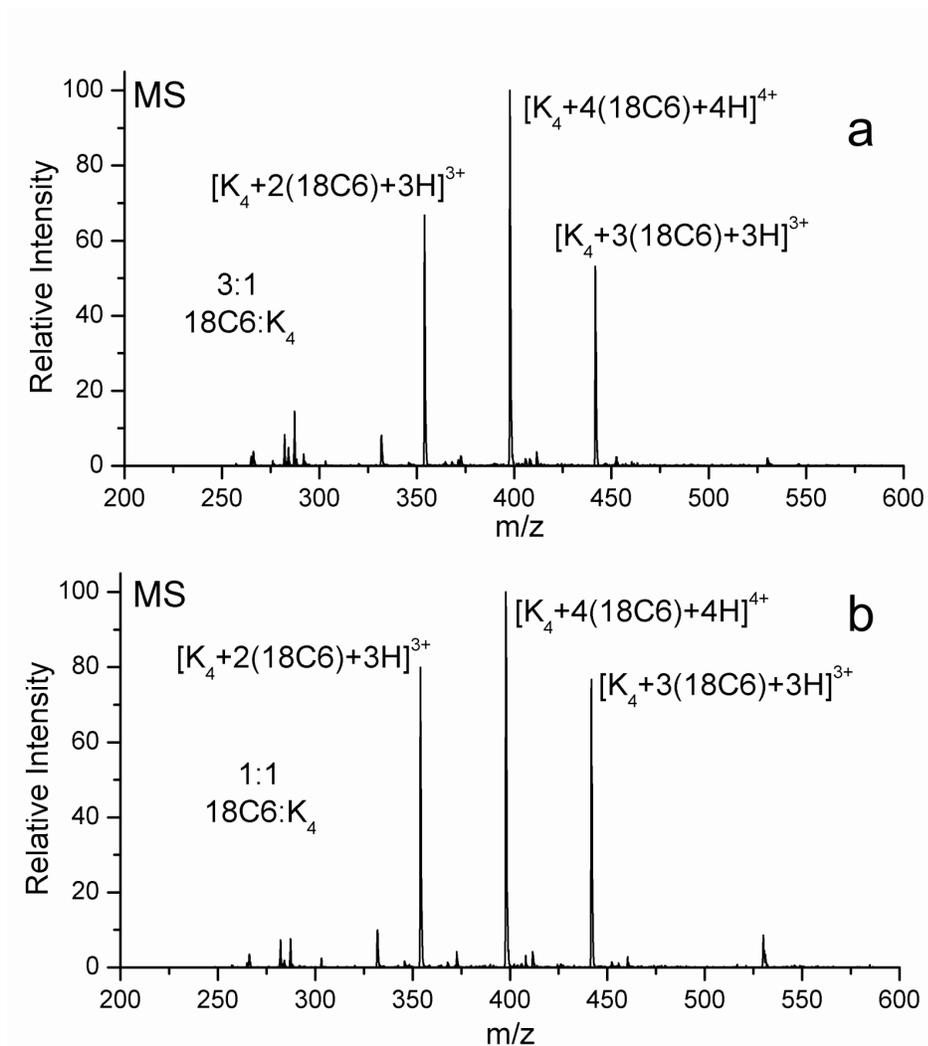


Figure 7.6 Effect of concentration on 18C6 binding. (a) 3:1, (b) 1:1 (18C6 to KKKK) concentrations respectively. The solution phase concentrations are not accurately reflected in the gas phase distribution.

Only a very small fraction of the analyte introduced via ESI ever reaches the detector. In fact, for a typical ESI methanol droplet with a diameter of 15 microns there are approximately 1.5×10^6 charges when the droplet is at the Rayleigh limit.²¹ Using these numbers, there are approximately 35 molecules/charge in a droplet that was formed from a 50 μ M solution (this should be regarded as a lower limit for the analyte/charge ratio because it neglects evaporation that may have already concentrated the solution). Under these conditions, there are far more molecules of analyte in an ESI droplet than there are charges to allow for their detection. Taking this into account, if the rate at which an ion is expelled from the droplet is increased, then the ion abundance will be greatly increased. The lack of sufficient charge introduces a time dependence on the ion abundance where those ions that are “evaporated” first will be the most abundant.

7.4 Conclusion

Lysine forms a stable salt-bridge complex with 18C6 complexed to the N-terminus. In peptides, however, the side chain of lysine is found to be the preferred binding site. 18C6 can even sequester charge on several adjacent lysines as demonstrated by the abundance of $[\text{KKKK}+4(18\text{C}6)+4\text{H}]^{4+}$ (structure 7.8, Figure 7.3d). Lysine residues can usually be quantified in small peptides by counting the number of 18C6 ethers attached to the peptide in the most abundant peak of the spectrum. In larger peptides and proteins, 18C6 is sensitive to and may be used to detect changes in conformation. It is clear that 18C6 can prevent the refolding of a denatured protein for an extended period of time. Competitive complexation by the side chains of histidine and arginine is low. The binding of 18C6 is found to dramatically alter the ion abundance and the charge state

distribution of most species to which it attaches. This charge stabilization is useful because it allows for the facile detection of species that are normally difficult to detect with ESI, but it also complicates the quantification of lysine through molecular recognition.

18C6 will form a stable adduct to primary amines in both solution and the gas phase. This complexation reduces the magnitude of the free energy of solvation for the ions and facilitates their removal to the gas phase. The result is that the charge state distribution and ion abundance of the peptide are affected. With further research in this area, it may be possible to expand the capabilities of ESI by tailoring molecular interactions that allow for the detection of others species that are difficult to electrospray. In the case demonstrated here, the charge state and ion abundance are increased primarily for lysine rich molecules; however, there is no limitation that precludes this principle from operating in other systems. Further experiments may lead to fine-tuned reagents which are highly selective for binding to specific residues.

7.5 References

-
- ¹ Hiraoka, M. *Crown Compounds: their characteristics and applications* Elsevier Scientific Publishing Company: New York, 1982.
- ² Schalley, C. A. *Int. J. Mass Spec.* **2000**, *194*, 11-39.
- ³ Pope, R. M.; Shen, N.; Hofstadler, S. A.; Dearden, D. V. *Int. J. Mass Spec. Ion Proc.* **1998**, *175*, 179-186. Cunniff, J. B.; Vouros, P. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1175-1182.
- ⁴ Rudiger, V.; Schneider, H. J.; Solov'ev, V. P.; Kazachenko, V. P.; Raevsky, O. A. *Eur. J. Org. Chem.* **1999**, 1847-1856.
- ⁵ *Comprehensive Supramolecular Chemistry*, vol. 1 (Ed.:G. W. Gokel), Pergamon/Elsevier: Oxford, 1996.
- ⁶ More, M. B.; Ray, D.; Armentrout, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 417-423.
- ⁷ Williamson, B. L.; Creaser, C. S. *Int. J. Mass Spec.* **1999**, *188*, 53-61.
- ⁸ Maleknia, S.; Brodbelt, J. *J. Am. Chem. Soc.* **1993**, *115*, 2837-2843.
- ⁹ Meot-ner, M. *J. Am. Chem. Soc.* **1983**, *105*, 4912-4915.
- ¹⁰ (a) He, Z. M.; Zhang, Z. D. *J. Protein Chem.* **1999**, *18*, 557-564. (b) Vila, J. A.; Ripoll, D. R.; Villegas, M. E.; Vorobjev, Y. N.; Scheraga, H. A. *BioPhys. J.* **1998**, *75*, 2637-2646.
- ¹¹ Lee, S. -W.; Lee, H. -N.; Kim, H. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1998**, *120*, 5800-5805.
- ¹² Fenn, J.; Rosell, J.; Nohmi, T.; Shen, S.; Banks, F. *Biochemical and Biotechnological Applications of Electrospray Ionization Mass Spectrometry* **1996**, *619*, 60-80.

-
- ¹³ (a) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287. (b) Thomson, B. A.; Iribarne, J. V. *J. Chem. Phys.* **1979**, *71*, 4451.
- ¹⁴ Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, *94*, 8897-8909.
- ¹⁵ Campbell, S.; Beauchamp, J. L.; Rempe M.; Lichtenberger D. L. *Int. J. Mass Spec.* **1992**, *117*, 1-3.
- ¹⁶ E. P. Hunter and S. G. Lias, "Proton Affinity Evaluation" in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. W.G. Mallard and P.J. Linstrom, February 2000, National Institute of Standards and Technology, Gaithersburg MD, 20899 (<http://webbook.nist.gov>)
- ¹⁷ Kulikov, O. V.; Krestov, G. A. *Pure & Appl. Chem.* **1995**, *67*, 1103-1108.
- ¹⁸ Wong, S. F.; Meng, C. K.; Fenn, J. B. *J. Phys. Chem.* **1988**, *92*, 546-550.
- ¹⁹ (a) Schnier, P. D.; Gross, D. S.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1086-1097. (b) Schnier, P.D.; Price, W. D.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 972-976.
- ²⁰ Lord Rayleigh *Proc. Roy. Soc.* **1882**, *14*, 184-186.
- ²¹ Smith, J. N.; Flagan, R. C.; Beauchamp, J. L. *J. Phys. Chem. A* **2002**, *106*, 9957-9967.
- ²² Cech, N. B.; Enke, C. G. *Anal. Chem.* **2000**, *72(13)*, 2717-2723.
- ²³ Tang, L.; Kebarle, P. *Anal. Chem.* **1993**, *65*, 3654-3668.
- ²⁴ Born, M. *Phys. Z.* **1920**, *1*, 45.

Chapter 8

Molecular Recognition of Arginine in Small Peptides by Supramolecular Complexation with Dibenzo-30- Crown-10 Ether

Previously published in: Julian, R. R.; Akin, M.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L. *Int. J. Mass Spectrom.* **2002**, *220*, 87-96.

8.1 Introduction

The field of supramolecular chemistry is devoted in part to the study of noncovalent molecular interactions.¹ Among the myriad applications of supramolecular chemistry, the focus here will be on the discovery and development of specific receptors designed to recognize biologically relevant species. Recent efforts have produced substantial progress in this endeavor utilizing a variety of hosts including crown ethers², lariat crown ethers³, cyclodextrins⁴, calixarenes⁵, cyclophanes⁶, cyclofructans⁷, sulfonate dyes⁸, and monosaccharides⁹ among others. Although the majority of these studies have been performed in solution, the number of gas phase studies of supramolecular chemistry has been increasing¹⁰ and several of the above examples refer to gas phase work.

For the molecular recognition of a specific side chain in a peptide or protein, the charged amino acids arginine, histidine, lysine, glutamic acid, and aspartic acid offer the best targets of opportunity. In the gas phase these remain distinguished as the most basic

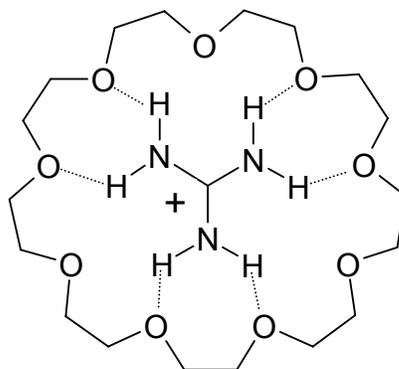
and most acidic of the twenty amino acids. The acidic side chains are difficult to distinguish due to the fact that both terminate with a carboxylic acid and differ only in the number of methylenes joining the acidic functionality to the backbone. The different chemical functionalities of the basic side chains enhance the possibility of achieving specificity. Lysine and arginine have received the most attention with regards to molecular recognition. We have recently demonstrated the excellent molecular recognition capabilities of 18-crown-6 (18C6) as a specific host for the side chain of lysine in the gas phase.² In these experiments, 18C6 is mixed into a solution with a lysine containing peptide which is then electrosprayed into the gas phase. The 18C6 forms a specific complex with the alkyl ammonium group on the protonated side chain of lysine, stabilized by three hydrogen bonds. This complex forms in solution and is transferred intact to the gas phase using electrospray ionization. The quantification of lysine residues in small peptides is possible with this technique through the formation of multiple adducts.²

In related gas phase work, Freiss and Zenobi have utilized a series of sulfonates for the molecular recognition of arginine.¹¹ The molecular recognition is primarily due to electrostatic attraction between the basic guanidinium group of arginine and an acidic disulfonate. The complexes are detected utilizing matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS). The N-terminus is additionally recognized by this technique, yielding one additional peak beyond the number of arginines in the peptide.

In addition, several cleverly designed molecules have been developed for the molecular recognition of arginine in solution. Dougherty and coworkers synthesized a

cyclophane based host which utilizes a combination of hydrophobic, cation- π , and ion-pairing interactions to form a stable complex with arginine.⁶ Schrader and coworkers have developed a series of bisphosphonate receptors that utilize a combination of hydrogen bonding, electrostatic interactions, and cation- π effects to recognize alkyl guanidinium groups.¹² Bell and coworkers are responsible for the “arginine cork”, which demonstrates arginine selectivity through the formation of several specific hydrogen bonds, accompanied by electrostatic attraction.¹³

In contrast, Lehn and coworkers first suggested that large macrocycles might serve as neutral receptors for arginine, based on the recognition of the guanidinium group by large crown ethers.¹⁴ This idea is supported by several related experiments. In synthetic chemistry, guanidinium has been successfully used as a template for the synthesis of large macrocycles such as 27-crown-9 (27C9) as shown in **8.1**.¹⁵ X-ray crystallography has shown that both 27C9 and 30-crown-10 (30C10) crown ethers and their derivatives are good receptors for guanidinium, forming inclusion complexes in the solid phase with a 1:1 stoichiometry.¹⁶ In these complexes the molecules are coplanar and stabilized primarily through a network of hydrogen bonds as shown in **8.1**. For guanidinium itself, the 27-crown ethers were found to be the ideally sized hosts. These results suggest that 27C9 or 30C10 could be utilized for the molecular recognition of arginine in gas phase experiments in a manner analogous to that for 18C6 described above.²



8.1

In the present work, experiments are conducted with various macrocycles to determine which is ideally suited to host the side chain of arginine. Competitive collision induced dissociation (CID) experiments are utilized to determine the relative bond dissociation energetics for the various crown/peptide interactions. The molecular recognition capabilities of 18C6 are utilized in combination with larger macrocycles, yielding additional information about the amino acid composition of peptides.

8.2 Experimental Methodology

All data were obtained using a Finnigan LCQ ion trap quadrupole mass spectrometer without modification. Soft sampling is crucial for the detection of these noncovalent complexes. The critical instrument settings that yield adduct formation include capillary voltage 14.12V, capillary temperature 200°C, and tube lens offset -39V. Higher capillary temperatures dissociate the crown ether/peptide complexes. The tube lens offset controls the acceleration of ions as they leave the capillary region. The tube lens voltage is minimized to reduce the number of energetic collisions with the He buffer gas.

Sample concentrations were typically kept in the ~10 to 100 μM range for all species of interest, unless otherwise noted. All samples were electrosprayed in a mixture of 80:20 methanol/water. Crown ethers were added in a 2-3-fold excess to the sample and electrosprayed with the analyte in order to observe adducts. Samples were electrosprayed with a flow of 3-5 $\mu\text{L}/\text{min}$ from a 500 μL Hamilton syringe for optimal signal. Silica tubing with an inner diameter of .005 in. was used as the electrospray tip. No acid was added to any of the samples, unless otherwise noted. All chemicals unless otherwise noted were purchased from Sigma or Aldrich and used without further purification. 27-crown-9 was synthesized according to well established techniques.¹⁷

All calculations were performed with the HyperChem 5.1 Professional Suite. Candidate structures were identified with molecular mechanics and were fully optimized at the PM3 semi-empirical level.

8.3 Results and Discussion

DB30C10. The results for the complexation of DB30C10 with three peptides are shown in Figure 8.1. In each case, the peptide is electrosprayed from a solution with DB30C10 present in 2-3-fold excess. The peptide KPPGFSPFR (a bradykinin analog that we will designate Kbk) forms a single DB30C10 adduct $[\text{Kbk}+\text{DB30C10}+2\text{H}]^{2+}$, as shown in Figure 8.1a. Kbk contains only one C-terminal arginine, and the results in Figure 8.1a are consistent with the molecular recognition of arginine by the DB30C10. Figure 8.1b shows the results for YGGFMRGL (Enk). In this case, a single DB30C10 adduct corresponding to $[\text{Enk}+\text{DB30C10}+2\text{H}]^{2+}$ is observed, which is also consistent with the single internal arginine present in the peptide. These promising results suggested that the extension of this method to the recognition of multiple arginines would be possible. However, as seen in Figure 8.1c, bradykinin (Bk, sequence RPPGFSPFR) does not form the expected double DB30C10 adduct. In fact, the single adduct $[\text{Bk}+\text{DB30C10}+2\text{H}]^{2+}$ is only observed in low abundance.

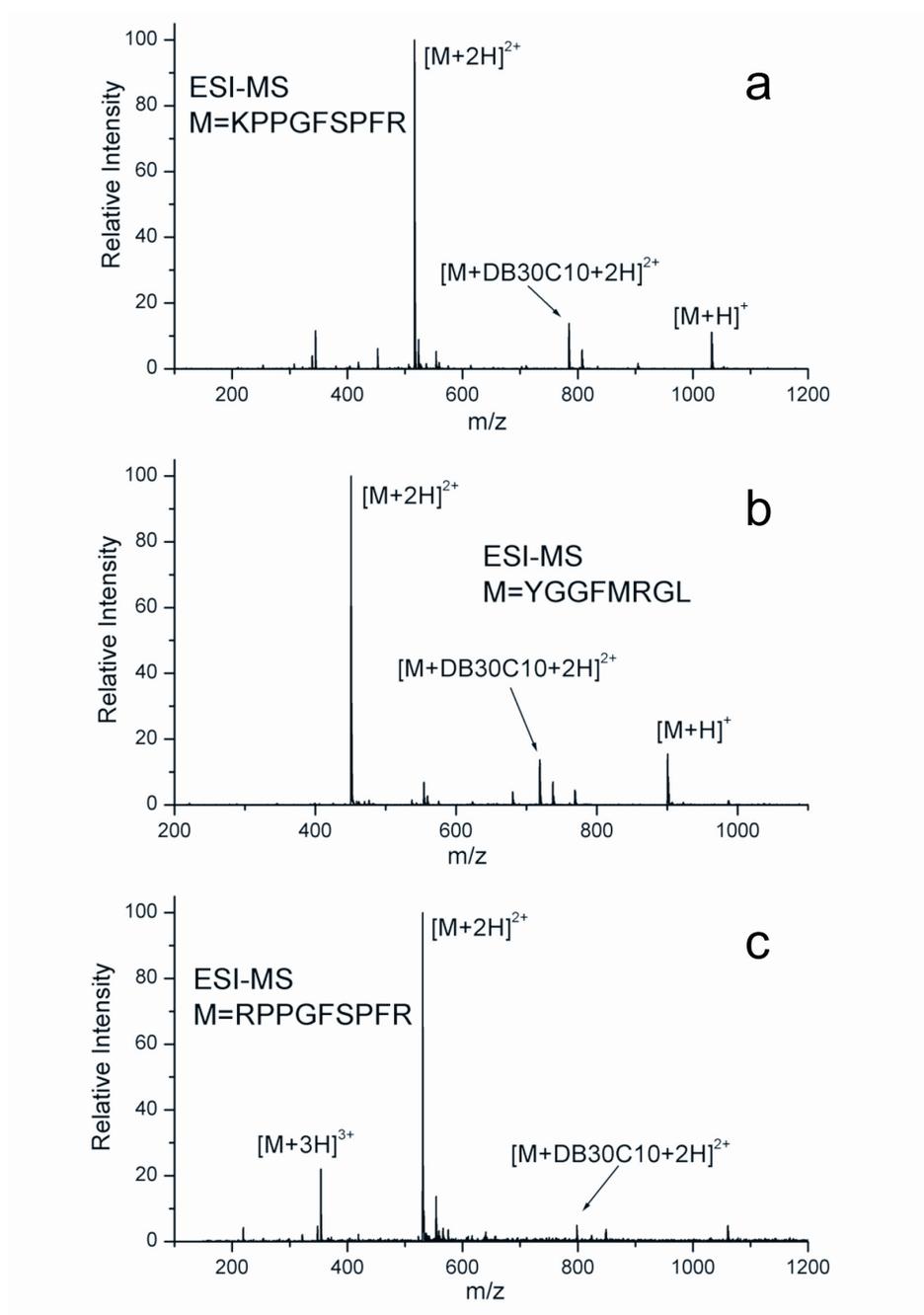


Figure 8.1 (a) ESI-MS of KPPGFSPFR (Kbk) with DB30C10. This peptide forms one adduct with DB30C10, corresponding to the single arginine present. (b) ESI-MS of YGGFMRGL (Enk) with DB30C10. The single expected adduct is observed. (c) The ESI-MS of RPPGFSPFR (Bk) with DB30C10 yields only a single, low intensity adduct where two are expected.

These results contrast with those obtained by the supramolecular complexation of lysine by 18C6. Figure 8.2 shows the results of electrospraying a mixture of Kbk with 18C6 present in solution. As we have recently demonstrated in our labs,² 18C6 forms stable gas phase adducts preferentially with lysine. This interaction leads to the base peak in Figure 8.2, which corresponds to the $[\text{Kbk}+18\text{C6}+2\text{H}]^{2+}$ adduct peak. In Figure 8.2, the adduct peaks are much more intense than the uncomplexed peptide peaks. Furthermore, multiple adduct peaks are detected, despite the fact that 18C6 attaches only weakly to the n-terminus and the side chain of arginine.

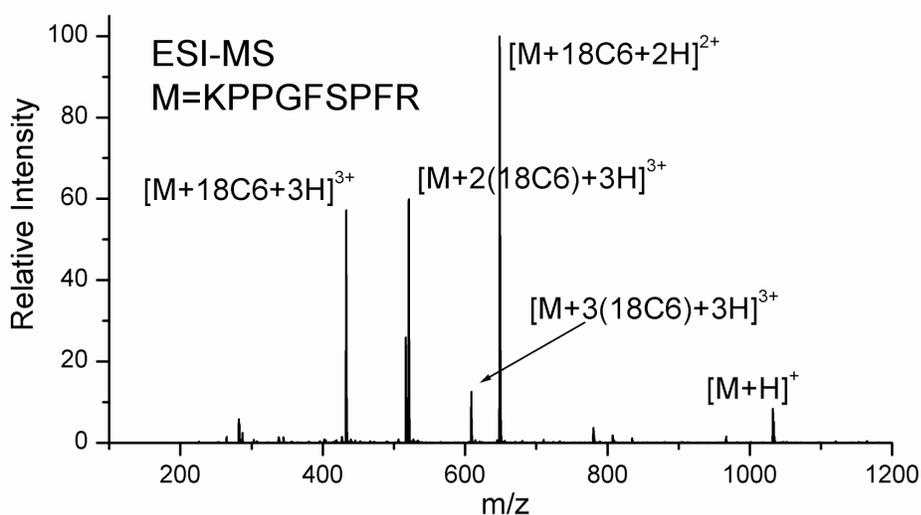
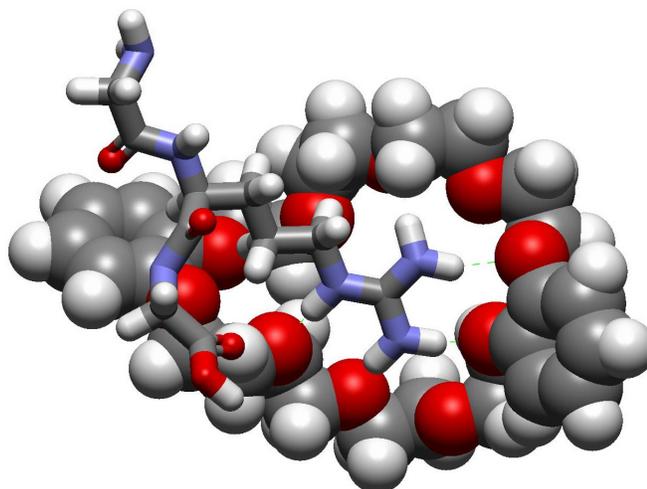


Figure 8.2 ESI-MS resulting from the addition of 18C6 to a solution of KPPGFSPFR. In this case, the adducts are the most intense peaks present, and multiple 18C6 adducts are easily observed.

One interpretation of this data would suggest that the low abundance of the DB30C10 peaks in Figure 8.1 is the result of very weak gas phase binding between DB30C10 and the side chain of arginine, but further experimentation detailed below shows this not to be the case. Recent related work in our labs has shown that the solution phase binding constant for crown/ion complexes is related to the abundance of the adduct peak observed in the gas phase.¹⁸ Solution phase data on related crown ethers suggests that a low solution phase binding constant for arginine by DB30C10 may be a contributing factor to the low observed abundance of adducts in the gas phase. This may account for the low adduct abundance observed for bradykinin. Furthermore, molecular modeling demonstrates that the size of DB30C10 may lead to unfavorable steric interactions when two or more DB30C10 molecules are complexed with a peptide. The relative sizes are shown with the example of DB30C10 bound to GRG in **8.2**. Despite the low abundance of the gas phase adducts, DB30C10 forms strongly bound complexes with arginine once in the gas phase as will be detailed below.

**8.2**

As discussed above, Boer and coworkers have suggested that the 27-crown ethers are the optimally sized hosts for guanidinium, because the crystal structure data indicates that the 30-crown ethers are slightly too large.¹⁶ In order to ascertain the optimal host for arginine in the gas phase, we performed a competitive CID experiment in which both 27C9 and DB30C10 were attached to the dipeptide RR.¹⁹ Figure 8.3 shows the results for the CID of the $[\text{RR}+\text{DB30C10}+\text{27C9}+\text{2H}]^{2+}$ complex. Two fragments are observed which both correspond to the loss of neutral crown ether, but the DB30C10 is retained by RR more often than 27C9 by a factor of 4. This suggests that arginine, with the larger alkyl-guanidinium side chain, prefers the larger DB30C10 host. Molecular modeling and crystal structures¹⁶ indicate that for 27C9 to interact with arginine, the guanidinium group must pivot out of the plane of the crown ether, weakening the hydrogen bonds that stabilize the complex. Thus for arginine, the 30-crown ethers appear to be optimally sized.

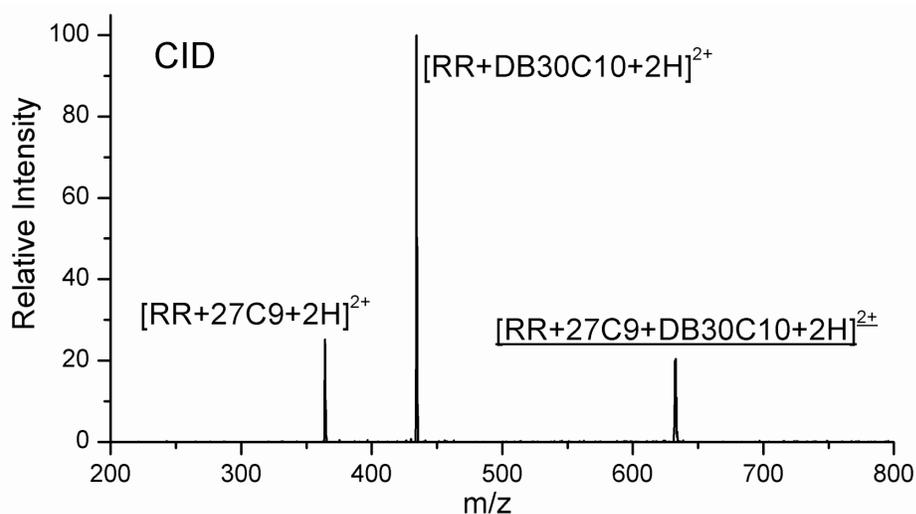


Figure 8.3 CID spectrum of diarginine with 27C9 and DB30C10 both attached. The DB30C10 is retained preferentially, demonstrating a higher bond dissociation energy for the larger crown. The peak with an underlined label is being dissociated.

DB30C10 and 18C6. The formation of DB30C10/peptide complex does not preclude the addition of 18C6 adducts if the peptide contains lysine residues in addition to arginine. In other words, it is possible to form a complex between a peptide and both crown ethers simultaneously. Figure 8.4 shows the results for such an experiment with the peptide TKPR. This peptide contains one lysine and one arginine, which should yield at least three different complexes, two corresponding to each crown attached individually and one complex with both crowns attached. As expected, Figure 8.4 shows that $[\text{TKPR}+\text{DB30C10}+3\text{H}]^{3+}$, $[\text{TKPR}+18\text{C6}+2\text{H}]^{2+}$, and $[\text{TKPR}+\text{DB30C10}+18\text{C6}+3\text{H}]^{3+}$ are all observed. Therefore the crown ether based arginine and lysine molecular recognition techniques appear to be mutually compatible. In fact, for reasons which remain unclear, the presence of both crowns enhances the abundance of both adducts. Furthermore, a mixed peak such as $[\text{TKPR}+\text{DB30C10}+18\text{C6}+3\text{H}]^{3+}$ raises the issue of which crown ether is more strongly bound.

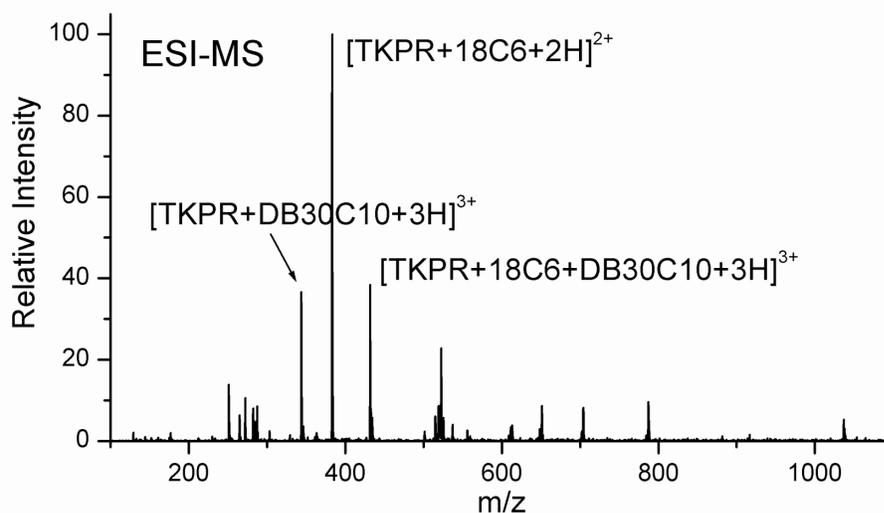


Figure 8.4 ESI-MS of TKPR with both 18C6 and DB30C10. 18C6 attaches to the lysine and DB30C10 attaches to the arginine, demonstrating the compatibility of the two experiments.

Figure 8.5 shows the results of a series of experiments conducted to determine the relative binding strengths of 18C6 and DB30C10 to lysine and arginine, respectively. In each experiment, a peptide containing arginine and lysine is electrosprayed with both 18C6 and DB30C10. The peak corresponding to [peptide+18C6+DB30C10] is then isolated and subjected to mild CID, leading to the loss of the crown ether with the lowest gas phase binding energy. The peptides RK, TKPR, and Kbk were chosen for their range in size and relative positions of lysine and arginine. As seen in Figure 8.5a-c, the DB30C10 remained attached to the peptide preferentially over 18C6 by a substantial margin in every case. Figure 8.5d shows the results for the same experiment with tetralysine. In this case, the peptide contains no arginine residue. Although peptides lacking arginine will not bind DB30C10 to a significant extent, we found that it was possible to form, in low abundance, non-specific complexes with multiple, sequential lysines present. The results are reversed in Figure 8.5d, with the 18C6 remaining attached to tetralysine preferentially over DB30C10.

These results further support the notion of selectivity for complexation of 18C6 and DB30C10 with lysine and arginine, respectively. If adduct formation were the result of non-specific interactions with the peptides, then the relative binding energetics would be difficult to predict. As clearly illustrated by the data in Figure 8.5, this is not the case. For arginine containing peptides, DB30C10 has a higher binding energy. In the absence of arginine, lysine containing peptides bind 18C6 with a higher affinity. Each crown binds specifically to a particular amino acid. The bond strength of the [18C6+NH₄]⁺ complex is 71±3 kcal/mol.²⁰ The experimental bond dissociation energy for 18C6 attached to a protonated primary amine in the gas phase is not known, but is estimated to

be >50 kcal/mol.²¹ The bond dissociation energy for DB30C10 attached to a guanidinium ion should prove to be even higher in the gas phase. The higher binding energy likely results from the more extensive hydrogen bonding along with favorable ion-dipole interactions (see structure **8.1**).

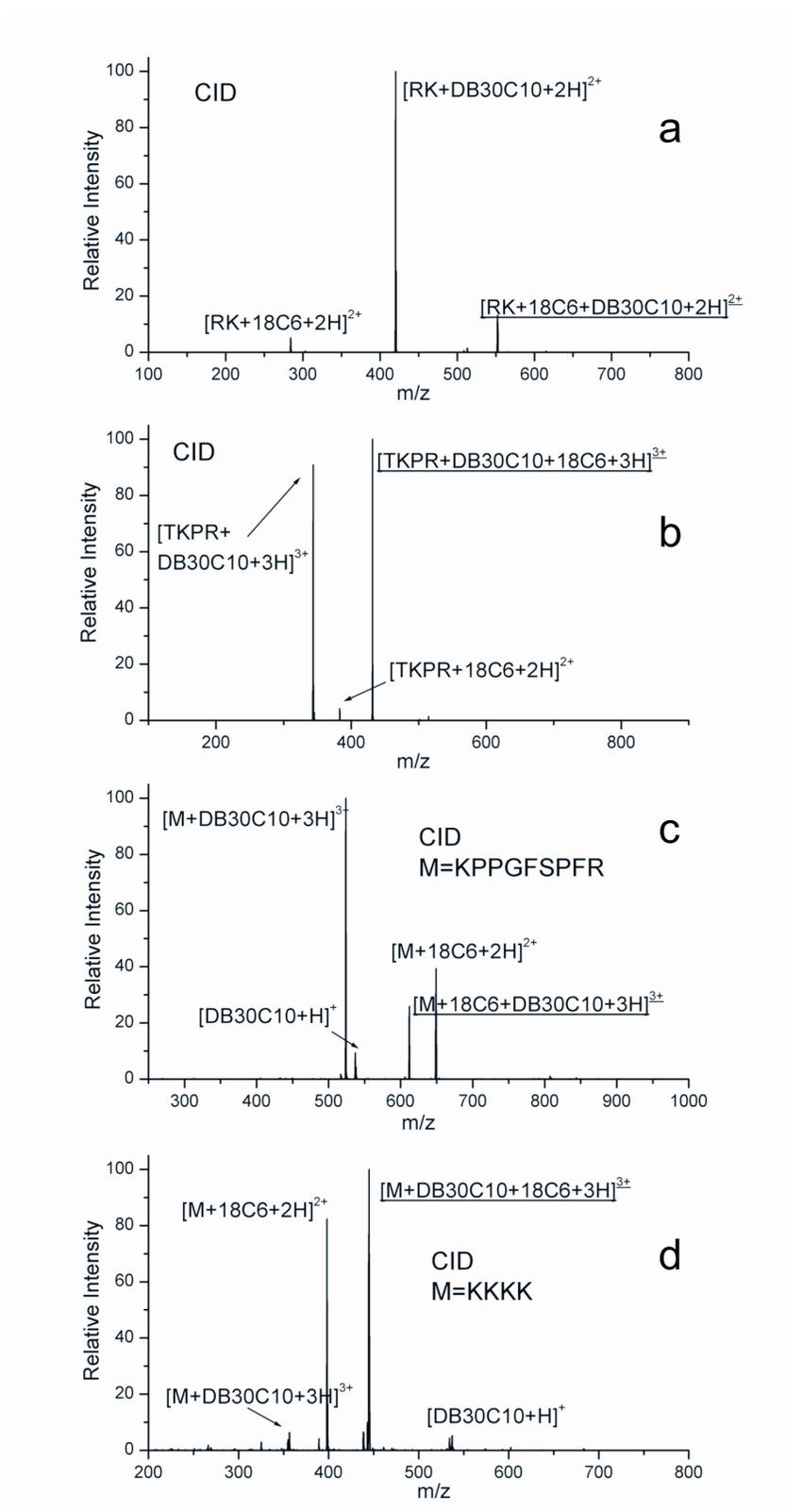


Figure 8.5 (Caption on following page)

Figure 8.5 CID spectra for the mixed [Peptide+18C6+DB30C10] species for several peptides. (a) CID of the RK mixed cluster shows retention of the DB30C10, suggesting a higher binding energy for arginine recognition. (b) The same result is obtained with TKPR, with the arginine now on the C-terminus. (c) The larger peptide KPPGFSPFR also yields a higher binding affinity for arginine recognition by DB30C10. (d) Control experiment in which tetralysine (which has no arginine residues) is used. In this case, the 18C6 is retained, demonstrating a higher affinity for lysine than DB30C10. This suggests that 18C6 is specific for lysine recognition and similarly that DB30C10 specifically binds to arginine. In each case, the peak with an underlined label is being dissociated.

The combination of the 18C6 and DB30C10 experiments yields more information about the composition of a peptide than can be obtained from each experiment individually. If a peptide contains only lysine, then experiments such as that shown in Figure 8.5d will demonstrate that there are no arginines present. If the peptide contains both arginine and lysine, then the number of lysines can be quantified and the presence of at least one arginine can be confirmed by experiments such as those shown in Figures 8.5a-c. Figure 8.6 shows the results for the competitive dissociation of the corresponding [peptide+18C6+DB30C10] peak for two peptides that contain only arginine. In the absence of lysine, 18C6 will form a weakly bound complex with the N-terminus of the peptide. Upon CID only the loss of the 18C6 is observed. The DB30C10 is retained without competition from 18C6 adducts for arginine containing peptides that do not contain lysine (Figure 8.6). This allows for the discrimination between arginine and arginine and lysine containing peptides because only peptides that contain both amino acids will yield CID peaks in which both 18C6 and DB30C10 are retained. This would allow the lysine and arginine containing fragments of a tryptic digest to be identified without any prior knowledge about the sequence of a protein.

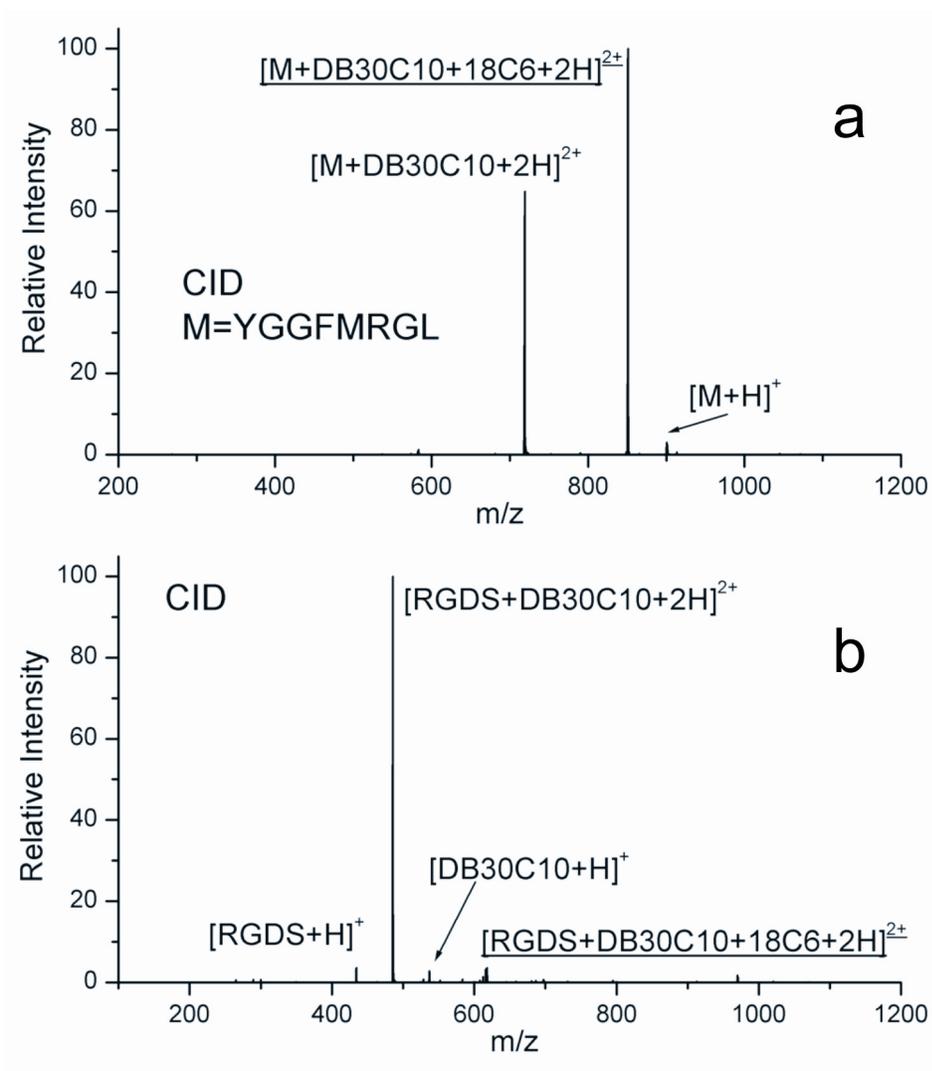


Figure 8.6 In the absence of lysine, CID on the [Peptide+18C6+DB30C10] peak yields no competitive dissociation, with all of the DB30C10 being retained over the 18C6. Similar results are obtained for both (a) YGGMRGL and (b) RGDS. This allows for the distinction of lysine versus arginine containing peptides without prior knowledge of the sequence. In each case, the peak with an underlined label is being dissociated.

8.4 Conclusions

DB30C10 forms a specific, strongly bound complex with the alkyl guanidinium group in the side chain of arginine in the gas phase. This interaction can be used to identify the presence of arginine in a peptide without any prior knowledge of its sequence. DB30C10 is a better host for arginine than 27C9 because the larger crown ether affords the extra space needed to accommodate the alkyl portion of the side chain. Comparison with lysine recognition by 18C6 demonstrates that the relative bond dissociation energy for the DB30C10 system is greater. Notwithstanding, the bond dissociation energy for 18C6 is greater than that for DB30C10 in lysine containing peptides that do not contain arginine, demonstrating the specificity of both interactions. This is the first example of the molecular recognition of arginine by a neutral host in the gas phase.

Further development in the field of molecular recognition of biological molecules will lead to advances in proteomics and our understanding of how biological systems employ molecular recognition themselves. The results presented in this paper allow for the partial sequencing of peptides in the gas phase, but the ability to use DB30C10 as a delivery agent for a wider range of site specific chemistry in the gas phase is even more promising. This can be accomplished through the development of lariat crown ethers, which have a variable side chain covalently attached. Similar work utilizing 18C6 is currently in progress in our laboratory, but the high binding energy of the DB30C10 offers an excellent opportunity to develop chemistry that can be induced on the peptide by the lariat prior to dissociation of the crown ether. Through such reagents, the selective cleavage of peptides in the gas phase, or the determination of gas phase structures could theoretically be accomplished.

8.5 References

-
- ¹ For a good general review see: *Supramolecular Chemistry* Beer, P. D.; Gale, P. A.; Smith, D. K. Oxford University Press, New York, 1999.
- ² Julian, R. R.; Beauchamp, J. L. *Int J. Mass Spectrom.* **2001**, *210*, 613.
- ³ Galan, A.; Andreu, D.; Echavarren, A. M.; Prados, P.; de Mendoza, J. *J. Am. Chem. Soc.* **1992**, *114*, 1511.
- ⁴ (a) Hauser S. L.; Johanson E. W.; Green H. P.; Smith P. J. *Org. Lett.* **2000**, *2(23)*, 3575.
(b) Ramirez, J.; He, F.; Lebrilla, C. B. *J. Am. Chem. Soc.* **1998**, *120*, 7387.
- ⁵ Ludwig, R.; Fresen. *J. Anal. Chem.* **2000**, *367*, 103.
- ⁶ Ngola S. M.; Kearney P.C.; Mecozzi S.; Russell K.; Dougherty D. A. *J. Am. Chem. Soc.* **1999**, *121*, 1192.
- ⁷ Sawada, M.; Shizuma, M.; Takai, Y.; Adachi, H.; Takeda, T.; Uchiyama, T. *Chem. Commun.* **1998**, 1453.
- ⁸ Salih, B.; Zenobi, R. *Anal. Chem.* **1998**, *70*, 1536.
- ⁹ Krishna, P.; Prabhakar, S.; Manoharan, M.; Jemmis, E. D.; Vairamani, M. *Chem. Commun.* **1999**, 1215.
- ¹⁰ Schalley, C. A. *Int. J. Mass Spec.* **2000**, *194*, 11.
- ¹¹ Friess S. D.; Zenobi R. *J. Am. Soc. Mass Spectrom.* **2001**, *12(7)*, 810.
- ¹² (a) Rensing, S.; Arendt, A.; Springer, A.; Grawe, T.; Schrader, T. *J. Org. Chem.* **2001**, *66*, 5814. (b) Schrader, T. H. *Tetr. Lett.* **1998**, *39*, 517.
- ¹³ Bell, T. W.; Khasanov, A. B.; Drew, M. G. B.; Filikov, A.; James, T. L. *Angew. Chem. Int. Ed.* **1999**, *38*, 2543.
- ¹⁴ Lehn, J. M.; Vierling, P.; Hayward, R. C. *J. Chem. Soc., Chem. Commun.* **1979**, 296.

-
- ¹⁵ (a) Madan, K.; Cram, D. J. *J. Chem. Soc., Chem. Commun.* **1975**, 427. (b) Kyba, E. P.; Helgeson, R. C.; Madan, K.; Gokel, G. W.; Tarnowski, T. L.; Moore, S. S.; Cram, D. J. *J. Am. Chem. Soc.* **1977**, *99*, 2564.
- ¹⁶ Boer, J. A. A.; Uiterwijk, J. W. H. M.; Geever, J.; Harkema, S.; Reinhoudt, D. N. *J. Org. Chem.* **1983**, *48*, 4821.
- ¹⁷ Chenevert, R.; D'Astous, L. *J. Heterocyclic Chem.* **1986**, *23*, 1785.
- ¹⁸ Julian, R. R.; Beauchamp, J. L. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 493-498.
- ¹⁹ Although these two arginines are not identical, it is anticipated from molecular modeling that the bond dissociation energies will be largely determined by the interaction of the crown ether with the side chain. Secondary interactions such as hydrogen bonds with the backbone should be largely equivalent for both of these crowns.
- ²⁰ Meot-Ner, M.; Sieck, L.W.; Leibman, J.F.; Scheiner, S. *J. Phys. Chem.* **1996**, *100*, 6445.
- ²¹ Colorado, A.; Brodbelt, J. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 1116.

Chapter 9

The Unusually High Proton Affinity of Aza-18-Crown-6 Ether: Implications for the Molecular Recognition of Lysine in Peptides by Lariat Crown Ethers

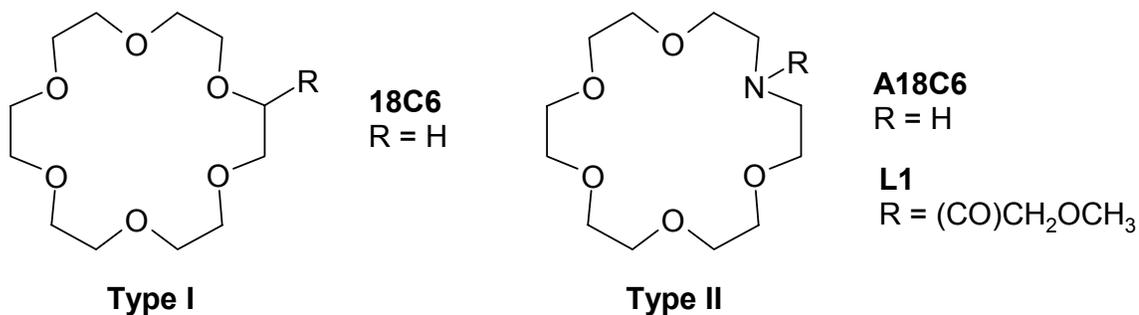
Published previously in: Julian, R. R.; Beauchamp, J. L. *J. Am. Soc. Mass Spectrom.* **2002**, *13*:5, 493-498.

9.1 Introduction

The study of supramolecular chemistry by mass spectrometry in the gas phase is rapidly expanding.¹ Self complementary molecular assemblies have been used to encapsulate a variety of ionic guests.² The macrobicyclic cryptate effect has been demonstrated for cryptands complexed with various metal ion guests in the gas phase.³ Other studies have shown that chirality can affect the noncovalent supramolecular association of amino acids into clusters.⁴ Crown ethers have been studied extensively due to their ability to bind metal cations.⁵ The ability of crown ethers, particularly 18-crown-6 (18C6), to complex with ammonium⁶ or alkylammonium⁷ ions in the gas phase has also been demonstrated. We have recently applied this property of 18C6 to the

molecular recognition of the biologically important amino acid lysine in small peptides and proteins.⁸

It would be useful to combine the molecular recognition of 18C6 with the added chemical functionality available in lariat crown ethers (crown ethers with a side chain covalently attached).⁹ Such lariat crown ethers would have increased chemical tunability and could be used to initiate or mediate chemical reactions with biological species in a site selective manner by utilizing the molecular recognition of lysine. For example, it might be possible to initiate the gas phase cleavage of peptide bonds in the vicinity of lysine residues with an appropriate lariat crown ether. The most common motifs for making lariat crown ethers are depicted as **Type I** and **Type II** below. **Type II** lariat crown ethers are by far the most common due to the extra bond afforded by the nitrogen heteroatom. **Type II** lariat crown ethers are flexible, allowing the side chain to interact with either face of the crown cavity.¹⁰ **Type II** lariat crown ethers also avoid the generation of the chiral center which complicates the synthesis of **Type I** lariat crowns.¹¹ This suggests that the interaction of aza-18-crown-6 (A18C6) and **Type II** lariat crown ethers with lysine should be investigated to determine if the molecular recognition abilities of 18C6 can be coupled with the added functionality of lariat crown ethers.



In the present work, the molecular recognition capabilities of A18C6 and L1 are compared with those of 18C6. A18C6 does not demonstrate the molecular recognition of

lysine that is observed for 18C6. This is attributed to the high proton affinity of A18C6, which serves to preferentially remove a proton from any alkyl ammonium ion upon introduction to the gas phase. The proton affinity of A18C6 is determined by the kinetic method to be 250 ± 1 kcal/mol.¹² Semi-empirical calculations suggest a highly symmetrical structure for protonated A18C6 in which the three most distant oxygens are able to fold back and hydrogen bond with the protonated nitrogen. L1, by contrast, has a lower proton affinity than A18C6 and forms noncovalent adducts with the side chain of lysine. However, the relative abundance of the noncovalent adducts of L1 with lysine is small in comparison with 18C6. The relative affinities of A18C6, L1, and 18C6 for the ammonium ion are determined by the kinetic method.¹² The order of the relative ammonium ion affinities is $18C6 \gg L1 > A18C6$, which is representative of the ability of each crown to form noncovalent adducts with lysine. The implications of these results for the design of lariat crown ethers capable of interacting with lysine are discussed.

9.2 Experimental Methodology

All data were obtained using a Finnigan LCQ ion trap quadrupole mass spectrometer without modification. Soft sampling is crucial for the detection of these noncovalent complexes. The critical instrument settings that yield adduct formation include capillary voltage 14.12V, capillary temperature 200°C, and tube lens offset -39V. Higher capillary temperatures dissociate the crown ether/peptide complexes. The tube lens offset controls the acceleration of ions as they leave the capillary region. The tube lens voltage is minimized to avoid collisions with the He buffer gas. For the proton affinity experiments, the capillary temperature was reduced to 140°C and up to 1% (v/v) acetic

acid was added to the mixture. The standard kinetic method was applied to measure the proton affinity of A18C6.¹³ 1,1,3,3-tetramethylguanidine (TMG), 1,5-Diazabicyclo[4.3.0]non-5-ene (DBN), 1,8-Diazabicyclo[5.4.0]undec-7-ene (DBU), 1,3,4,6,7,8-Hexahydro-1-methyl-2H-pyrimido[1,2-a]pyrimidine (HMPP), and 1,3,4,6,7,8-Hexahydro-2H-pyrimido[1,2-a]pyrimidine (HPP) were used as the reference bases.

Sample concentrations were typically kept in the ~10 to 100 μM range for all species of interest, unless otherwise noted. All samples were electrosprayed in a mixture of 80:20 methanol/water. Crown ethers were added to the sample and electrosprayed with the analyte in order to observe adducts. Samples were electrosprayed with a flow of 3-5 $\mu\text{L}/\text{min}$ from a 500 μL Hamilton syringe for optimal signal. Silica tubing with an inner diameter of .005 in was used as the electrospray tip. No acid was added to any of the samples, unless otherwise noted. Unless otherwise noted, all chemicals were purchased from Sigma or Aldrich and used without further purification. L1 was synthesized by the dropwise addition of methoxy acetyl chloride (CAS 38870-89-2) to A18C6 in dichloromethane with one equivalent of triethylamine. The mixture was refluxed for four hours and the solvent was removed under a flow of dry nitrogen. The ratio of product to starting material was approximately 5:1, from analysis of the mass spectrum of the crude reaction mixture, indicating a reasonable yield.

All semi-empirical calculations were performed with the HyperChem 5.1 Professional Suite. Candidate structures were identified with molecular mechanics using Cerius² version 4.0 by Molecular Simulations Inc. with the Dreiding 2.21¹⁴ force field. Structures were fully optimized at the PM3 semi-empirical level.

9.3 Results and Discussion

To determine if **Type II** lariat crown ethers can specifically bind to lysine, we choose A18C6 as the simplest model compound and examined a mixture of A18C6 with tetralysine. The results are compared with those for 18C6. Figure 9.1a shows the electrospray mass spectrum for tetralysine under typical conditions with the addition of 0.1% (v/v) acetic acid. The peptide is primarily doubly charged, although there are significant contributions from the triply and singly charged peaks as well. When 18C6 is added to the solution as shown in Figure 9.1b, the multiply charged, multiple adduct $[K_4+4(18C6)+4H]^{4+}$ is the base peak. This spectrum demonstrates the excellent molecular recognition and charge stabilization capabilities of 18C6, which have been examined in greater detail previously.⁸ In contrast, the addition of A18C6 to a solution of tetralysine does not result in the formation of a noncovalently bound crown/peptide adduct peak (Figure 9.1c). Furthermore, the doubly and triply charged tetralysine peaks present in Figure 9.1a are not detected in Figure 9.1c, and the singly charged $[A18C6+H]^+$ peak dominates the spectrum. Through the replacement of one oxygen by nitrogen, A18C6 has effectively lost the ability of 18C6 to form stable supramolecular complexes with protonated lysine residues in peptides. The two primary factors that account for this observation are discussed below.

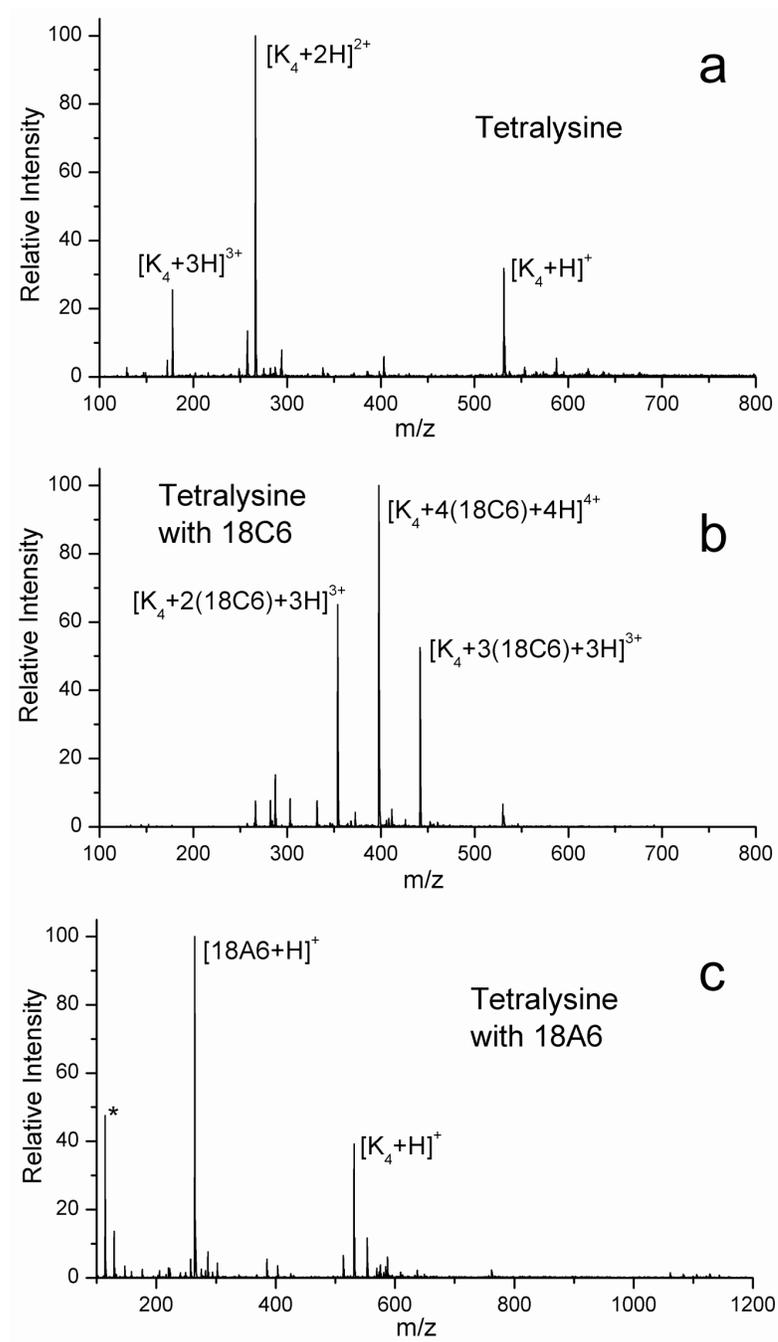


Figure 9.1 (a) Electrospray mass spectrum of tetralysine under typical conditions. (b) ESI-MS of a mixture of tetralysine and 18C6, demonstrating abundant noncovalent molecular cluster formation. (c) ESI-MS of tetralysine with A18C6. No adduct formation is observed. * corresponds to an A18C6 fragment.

Scheme 9.1 Reference bases.

High proton affinity of A18C6. Clearly, the nitrogen present in A18C6 will raise the proton affinity of this compound higher than that of 18C6. The standard kinetic method¹² was employed to determine the proton affinity of A18C6. The structures and abbreviations for the reference bases are given in Scheme 9.1. The collision-induced dissociation (CID) of the proton bound dimer $[A18C6+H+DBU]^+$ shown in Figure 9.2 is representative of the data collected for this experiment. No other dissociation pathways were observed for any of the dimers, and changing the collision energy did not significantly alter the product distribution. The plot of the $\ln([A18C6+H]^+/[Ref+H]^+)$ versus the proton affinity of the reference compounds is shown in Figure 9.3. From the data in Figure 9.3, the proton affinity for A18C6 is determined to be 250 ± 1 kcal/mol. It has been pointed out that caution should be utilized when employing the kinetic method with multidentate ligands because entropic effects for different dissociation pathways are not accounted for.¹⁵ However, the basicity of the nitrogen heteroatom will localize the extra proton in the case of A18C6, limiting any entropic effects which would require special consideration.

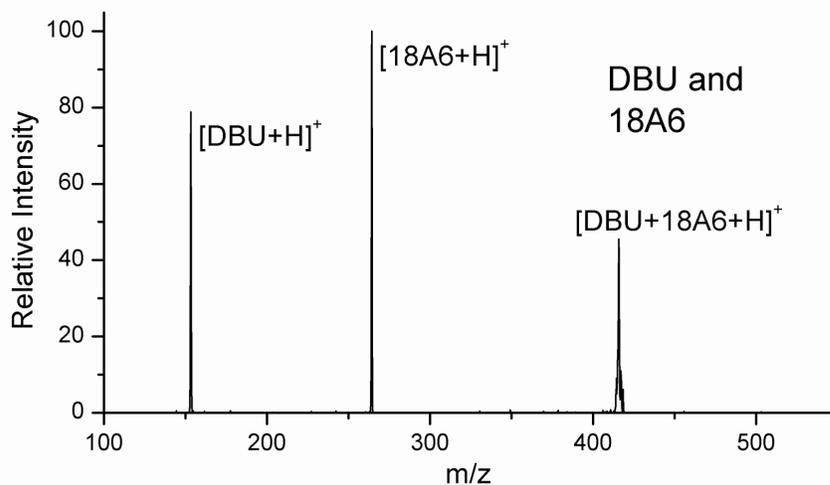


Figure 9.2 Example of data used to determine the proton affinity of A18C6 by the kinetic method. The dissociation of $[\text{DBU}+\text{A18C6}+\text{H}]^+$ is shown.

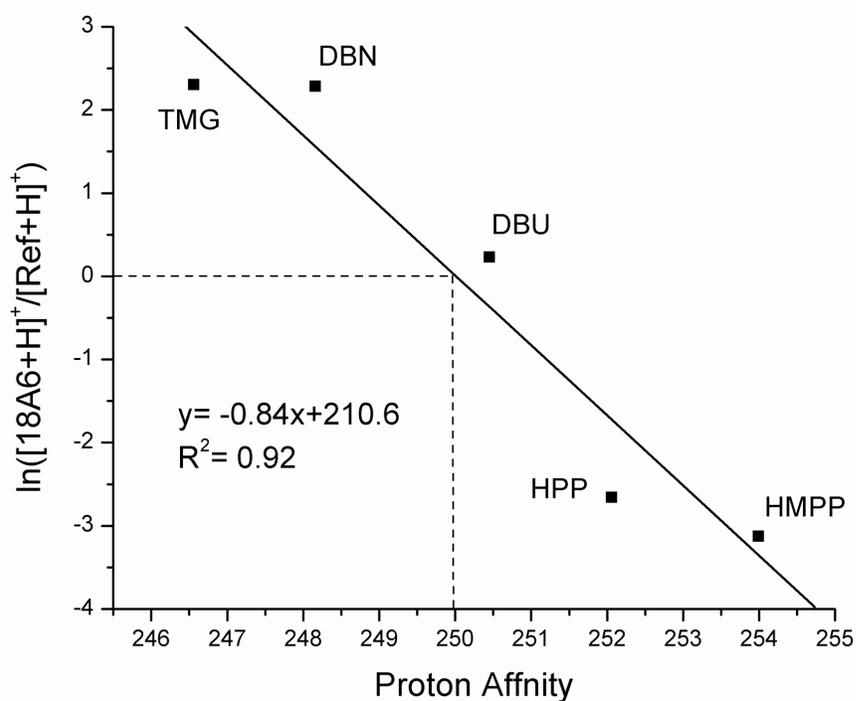
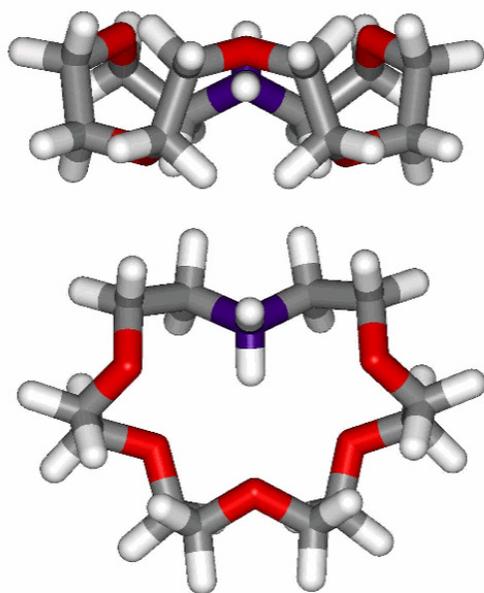
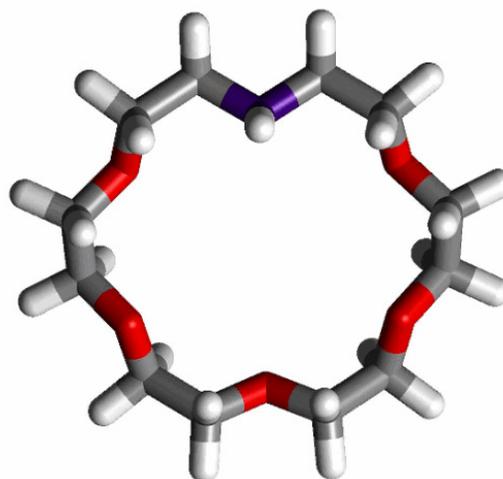


Figure 9.3 Kinetic method analysis of data acquired using the five reference bases given in Scheme 9.1. The proton affinity of A18C6 is determined to be 250 ± 1 kcal/mol. The proton affinities of the reference bases are 246.56, 248.16, 250.45, 252.06, and 253.99 kcal/mol, respectively.¹⁵

The proton affinity of A18C6 is nearly 20 kcal/mol higher than the proton affinity of 18C6. The proton affinity of diethyl amine is 227.6 kcal/mol, and the addition of two oxygens in diethanolamine (228 kcal/mol) does not significantly raise the proton affinity of that compound. There are no other single nitrogen containing secondary amines within at least 10 kcal/mol of this value in the NIST database.¹⁶ PM3 semi-empirical calculations reveal that $[A18C6+H]^+$ has a highly symmetrical structure (**9.1**). The three most distant oxygens are able to fold back and hydrogen bond with the proton, which is localized on the nitrogen because it is by far the most basic site in the crown. The structure is more compact than that normally associated with A18C6 (**9.2**), and the distance between coplanar heteroatoms (between the nitrogen and oxygens in the 7 and 13 positions) shrinks from 5.2 Å to 3.7 Å. The gas phase protonation of A18C6 causes a contraction of the guest cavity, which will lead to the loss of any interaction with a proximate primary amine. This contrasts sharply with the behavior of 18C6.



9.1 Two Views



9.2

The proton affinity of A18C6 is much greater than that of lysine (236 kcal/mol) or n-butyl amine (220.2 kcal/mol). This suggests that any complex formed in solution between A18C6 and the side chain of lysine will lead to a proton transfer to the A18C6 upon entry to the gas phase. As discussed above, the proton transfer will cause the contraction of the A18C6 guest cavity and dissociation of the complex. Thus the molecular recognition of lysine by A18C6 will result in the A18C6 stripping a proton off of the side chain of lysine. Solution phase data on benzyl ammonium ions indicate that the association of A18C6 with ammonium ions is weak in polar solvents such as water and methanol.¹⁷ The combination of the high proton affinity of A18C6 and its low solution phase association with ammonium ions explain the inability of A18C6 to form any peptide/crown ether adducts in the gas phase. Furthermore, the high proton affinity accounts for the reduced charge state distribution of tetralysine in Figure 9.1c, because any complexes that are formed will lead to the net loss of charge for lysine containing peptides.

L1 forms adducts with lysine. The results presented thus far suggest that if A18C6 is to be used as the base molecule for a lariat crown ether that will form noncovalent adducts with lysine, then the proton affinity of the nitrogen heteroatom must be lower in the lariat crown ether than it is in A18C6. The standard method of attaching side chains results in a tertiary amine at the point of attachment. In general, tertiary amines have higher proton affinities than secondary amines, which would suggest that there would be no adduct formation with such a lariat crown ether. This is indeed the case. Fortunately, amide nitrogens such as that found in n-methyl-acetamide¹⁶ (212.4 kcal/mol) have much

lower proton affinities than secondary amines. L1 is a lariat crown ether where the side chain is attached by amidation of the nitrogen heteroatom.

L1 forms noncovalent complexes with the side chain of lysine that can be observed in the gas phase. This is shown in the ESI-MS spectrum for dilysine and L1 in Figure 9.4. The adduct peaks are not nearly as intense as those observed for 18C6, but there is sufficient signal to isolate and observe the CID spectra for each adduct (Figures 9.4b and 9.4c). Molecular modeling at the PM3 semi-empirical level suggests that the methyl-ether oxygen in the side chain of L1 does not participate in the binding of an alkylammonium ion to the lariat crown ether. Therefore, we attribute the enhanced molecular recognition abilities of L1 relative to A18C6 to the lower proton affinity of the lariat crown.

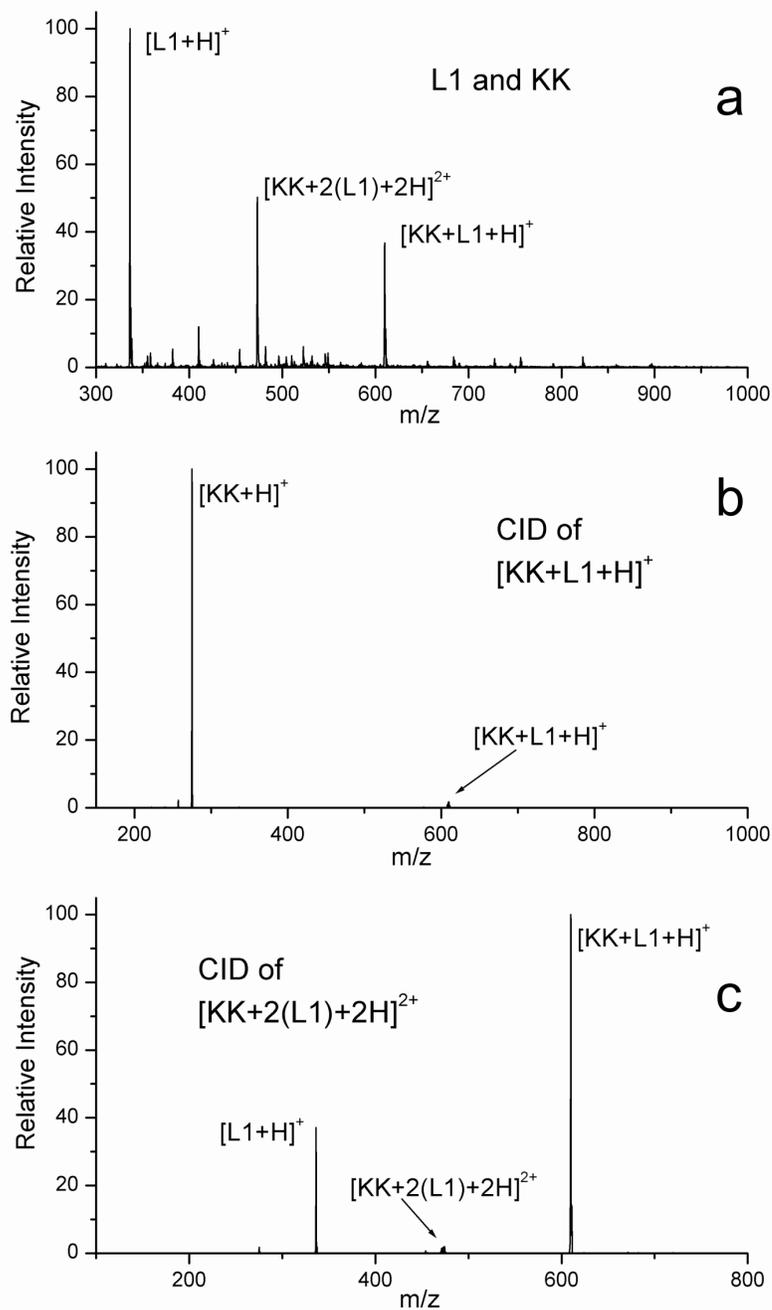


Figure 9.4 (a) ESI-MS of dilysine and L1. Adduct formation with L1 is observed. (b) CID spectrum of $[KK+L1+H]^+$, resulting in the loss of L1. (c) CID spectrum of $[KK+2(L1)+2H]^{2+}$, resulting in the loss of protonated L1, yielding the complex $[KK+L1+H]^+$.

Ammonium ion affinities. The kinetic method can be used to determine affinities for ions as well as protons. The relative ammonium ion affinities of various crown ethers should be correlated with their abilities to form noncovalent adducts with lysine. The $[18C6+NH_3+A18C6+H]^+$ complex is very difficult to isolate which suggests that it is weakly bound, and leads to increased baseline noise (Figure 9.5a). Dissociation of the complex during the process of isolation leads to a very prominent $[18C6+NH_3+H]^+$ peak in the absence of any excitation. There is no dissociation leading to the $[A18C6+NH_3+H]^+$ ion. As shown in Figure 9.5b, upon collision induced dissociation the ammonium ion is retained entirely by the 18C6. The $[A18C6+H]^+$ peak is absent from the CID spectrum as well.

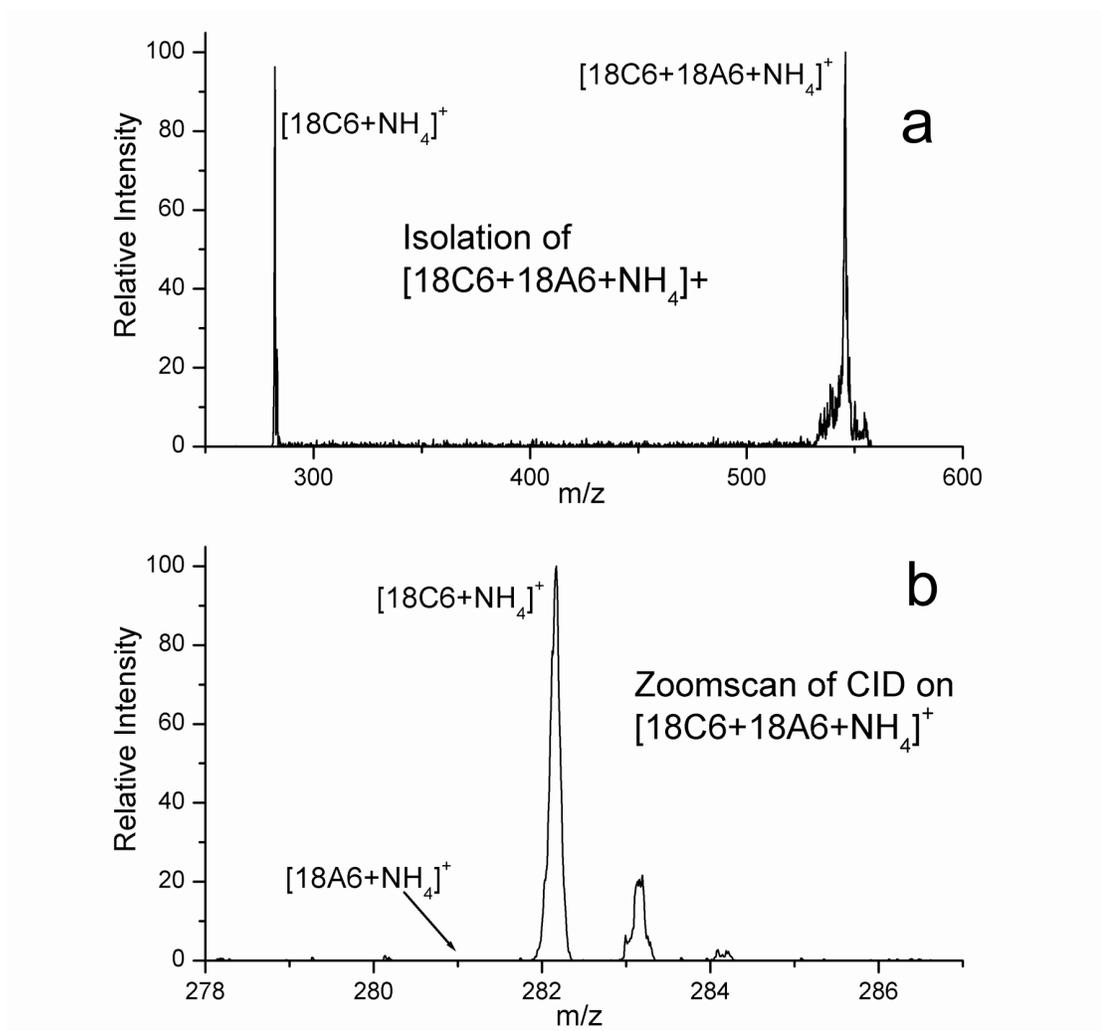


Figure 9.5 (a) Isolation spectrum for $[18C6+NH_3+A18C6+H]^+$ showing dissociation of the adduct peak to yield the 18C6 ammonium adduct exclusively. (b) Zoomscan of the CID spectrum for the $[18C6+NH_3+A18C6+H]^+$ peak, again yielding only $[18C6+NH_3+H]^+$. No protonated A18C6 was observed.

Figure 9.6a shows the isolation of $[L1 + NH_3 + 18C6 + H]^+$. In this case, dissociation of the isolated peak is very minimal. Figure 9.6b shows that the ammonium ion is retained by 18C6 and suggests that 18C6 has a much higher ammonium ion affinity than L1. In addition, comparison with the data in Figure 9.5 suggests that L1 has a greater ammonium ion affinity than A18C6. Direct comparison of the ammonium ion affinities of L1 and A18C6 was not possible because we were unable to prepare the appropriate ammonium ion bound dimer. This is attributed to the high proton affinity of A18C6, which leads to protonation of the A18C6 followed by dissociation prior to detection of the complex. These combined results suggest that the relative ammonium ion affinities for A18C6, L1, and 18C6 as determined qualitatively by the kinetic method are: $18C6 \gg L1 > 18AA6$. Not surprisingly, this indicates that the ammonium ion affinity for a crown ether is correlated with the capacity of that molecule to form noncovalent adducts with lysine.

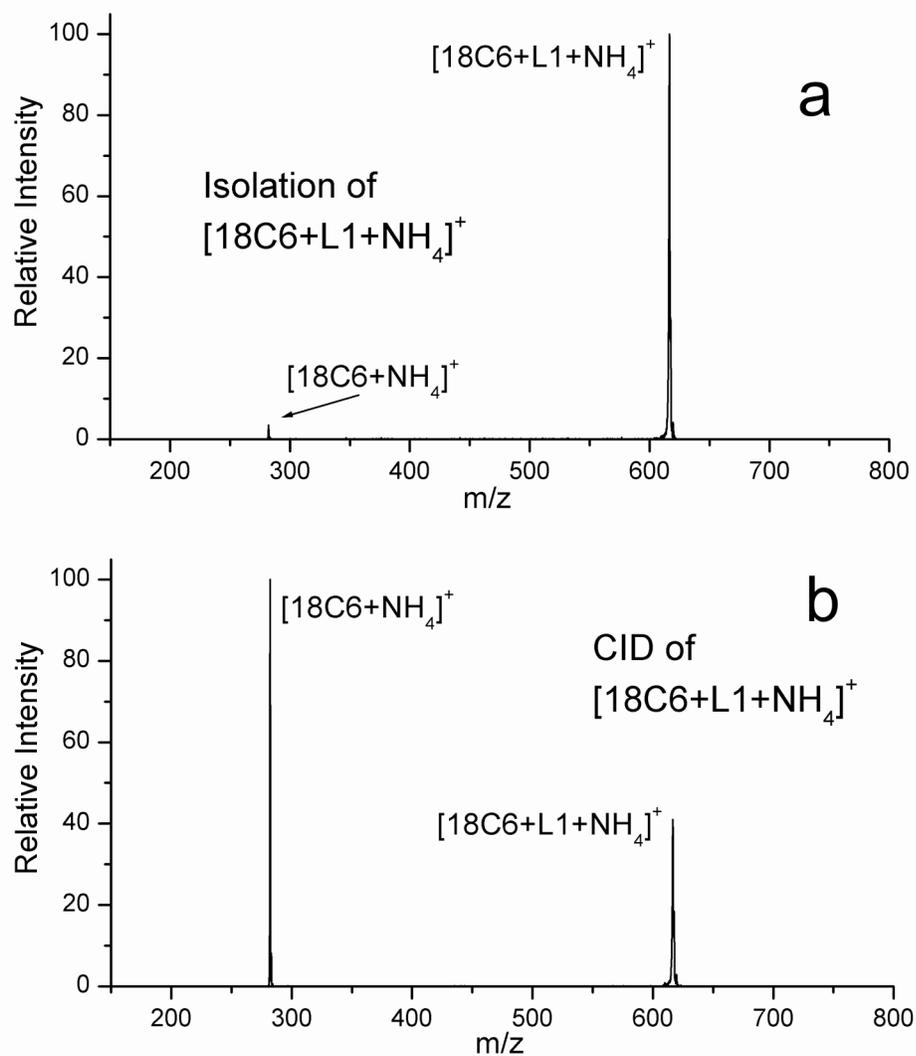


Figure 9.6 (a) Isolation spectrum for $[18C6+NH_3+L1+H]^+$ showing only minimal dissociation of the adduct peak. (b) CID spectrum for the $[18C6+NH_3+L1+H]^+$ peak, again yielding only $[18C6+NH_3+H]^+$.

9.4 Conclusions

The proton affinity of A18C6 is 250 ± 1 kcal/mol as determined by the kinetic method. This proton affinity is unusually high for a secondary amine, and is attributed to a highly symmetrical, hydrogen bond stabilized structure. The high basicity of A18C6 inhibits the formation of noncovalent adducts with lysine containing peptides. A18C6 will deprotonate an alkylammonium ion upon introduction to the gas phase, leading to a net decrease in the charge state of lysine containing peptides, such as observed for tetralysine in Figure 1. It is also likely that the protonation of A18C6 in solution is responsible for the greatly reduced solution phase binding of A18C6 to benzyl ammonium ions relative to 18C6.¹⁷ The ordering of ammonium ion affinities for the crown ethers determined in the present study, $18C6 \gg L1 > A18C6$, parallels their affinity for lysine in small peptides.

The amidation of A18C6 reduces the proton affinity and allows for the complexation of lysine containing peptides. Relative to 18C6, L1 does not bind to the side chain of lysine in great abundance. This suggests that the intrinsic binding energy to lysine is lower for L1 than it is for 18C6 and mitigates the utility of such derivatives in the design of effective reagents for the selective cleavage of peptides. Therefore, the lariat side arm of any future crown designed to do chemistry in the gas phase must either enhance the binding energy to the peptide and/or be highly reactive. Otherwise, CID of the complex in question will simply lead to dissociation back into the component molecular species as is the case with dilysine and L1. The **Type I** lariat crown ethers are likely to be the better choice for the development of biomimetic reagents for selective peptide cleavage, despite the synthetic and conformational limitations imposed by attaching the side chain to carbon.

9.5 References

-
- ¹ Schalley, C. A. *Int. J. Mass Spec.* **2000**, *194*, 11-39.
- ² Schalley, C. A. Rivera, J. M.; Martin, T.; Santamaria, J.; Siuzdak, G.; Rebek, J. *Eur. J. Org. Chem.* **1999**, 1325-1331.
- ³ Chen, Q.; Cannell, K.; Nicoll, J.; Dearden, D. V. *J. Am. Chem. Soc.* **1996**, *118*, 6335-6344.
- ⁴ (a) Julian, R. R.; Hodyss, R.; Kinnear, B.; Jarrold, M. F.; Beauchamp, J. L. *J. Phys. Chem. B* **2002**, *106*, 1219-1228. (b) Counterman A. E.; Clemmer D. E. *J Phys Chem B* **2001**, *105* (34), 8092-8096. (c) Cooks R. G.; Zhang D.X.; Koch K. J.; Gozzo F. C.; Eberlin M. N. *Anal. Chem.* **2001**, *73* (15), 3646-3655.
- ⁵ (a) More, M. B.; Ray, D.; Armentrout, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 417-423, (b) Wu, H.-F.; and Brodbelt, J. S. *J. Includ. Phenom. Mol.* **1994**, *18*, 37-44
- ⁶ Maleknia, S.; Brodbelt, J. S. *J. Am. Chem. Soc.* **1993**, *115*, 2837-3843.
- ⁷ Lee, S. -W.; Lee, H. -N.; Kim, H. S.; Beauchamp J. L. *J. Am. Chem. Soc.* **1998**, *120*, 5800-5805.
- ⁸ Julian, R. R.; Beauchamp, J. L. *Int J. Mass Spectrom.* **2001**, *210*, 613-623.
- ⁹ Dishong, D. M.; Diamond, C. J.; Cinoman, M. I.; Gokel, G. W. *J. Am. Chem. Soc.* **1983**, *105* (3), 586-593.
- ¹⁰ Schultz R. A., Dishong D. M., Gokel G. W. *Tetrahedron Lett* **1981**, *22* (28), 2623-2626.
- ¹¹ *Comprehensive Supramolecular Chemistry*, vol. 1 (Ed.:G. W. Gokel), Pergamon/Elsevier: Oxford, 1996.
- ¹² Cooks, R. G.; Kruger, T. L. *J. Am. Chem. Soc.* **1977**, *99*, 1279-1281.

-
- ¹³ McLuckey, S. A.; Cameron, D.; Cooks, R. G. *J. Am. Chem. Soc.* **1981**, *103*, 1313.
- ¹⁴ Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, *94*, 8897-8909.
- ¹⁵ Chu, I. -H.; Zhang, H.; Dearden, D. V. *J. Am. Chem. Soc.* **1993**, *115*, 5736-5744.
- ¹⁶ E. P. Hunter and S. G. Lias, "Proton Affinity Evaluation" in NIST Chemistry WebBook, NIST Standard Reference Database Number 69, Eds. W.G. Mallard and P.J. Linstrom, February 2000, National Institute of Standards and Technology, Gaithersburg MD, 20899 (<http://webbook.nist.gov>)
- ¹⁷ Rudiger, V.; Schneider, H. -J.; Solov'ev, V. P.; Kazachenko, V. P.; Raesky, O. A. *Eur. J. Org. Chem.* **1999**, 1847-1856.

Chapter 10

Molecular Mousetraps: Gas Phase Studies of the Covalent Coupling of Noncovalent Complexes Initiated by Reactive Carbenes Formed by Controlled Activation of Diazo Precursors

Published previously in: Julian, R. R.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L.
Angew. Chem. Int. Ed. **2003**, *42*(9), 1012-1015.

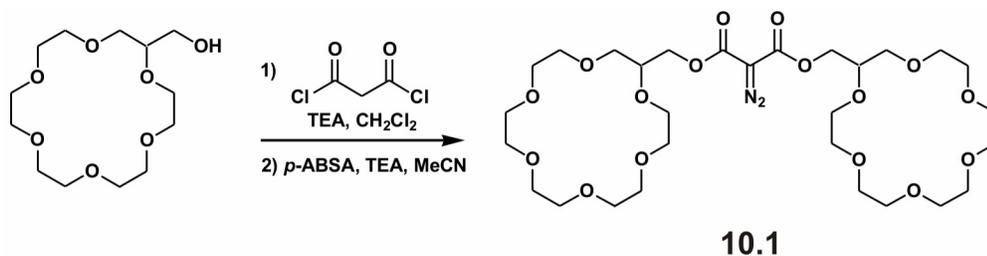
10.1 Introduction

Molecular recognition is a powerful technique that can be used to generate noncovalently bound host/guest complexes for a variety of purposes.^{1,2} These noncovalent complexes are easily transferred to the gas phase by electrospray ionization (ESI).³ Attempts to effect intermolecular reactions between the cluster components are often frustrated by the lability of noncovalent complexes due to the relatively weak interactions that hold them together. In the present work, we have successfully initiated intermolecular reactions in noncovalent clusters. First, a strongly bound host/guest complex is formed in solution and transferred to the gas phase by ESI. Second, a diazo group which has been incorporated into the host is efficiently and easily converted into a highly reactive carbene⁴ by low energy collision activated dissociation (CAD).⁵ This

carbene^{6,7} then reacts in an intermolecular fashion, covalently binding the host/guest complex. These reagents are herein referred to as “molecular mousetraps”.

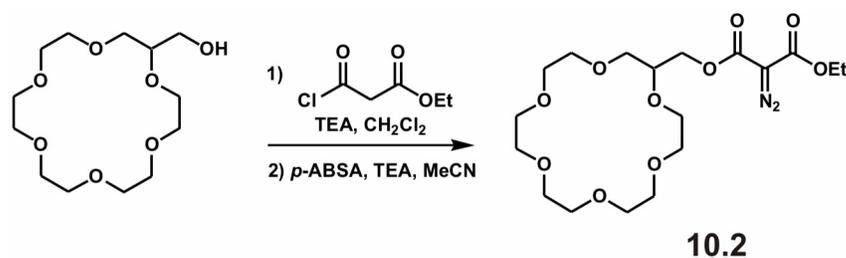
10.2 Experimental Section.

General Information: Due caution should always be used when handling diazo compounds. Reactions were performed in flame-dried glassware under a nitrogen atmosphere using freshly distilled solvents. All other reagents were used as received from commercial sources. Reaction temperatures were controlled by an IKA Mag temperature modulator. ¹H NMR spectra were recorded on a Varian Mercury 300 spectrometer (at 300 MHz) and are internally referenced to the chloroform peak (7.27 ppm) relative to Me₄Si. Data for ¹H NMR spectra are reported as follows: chemical shift (δppm), multiplicity, coupling constant (Hz), and integration. IR spectra were recorded on a Perkin Elmer Paragon 1000 spectrometer and are reported in frequency of absorption (cm⁻¹). Preparatory reversed phase HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25 x 100 mm, 100 μm C18 column equipped with a guard.



Compound 10.1: To a stirred, dry solution of 18-crown-6-methanol (50.0 μl, 0.16 mmol), dichloromethane (1.5 ml), and triethylamine (25 μl, 0.18 mmol) was added malonyl dichloride (9.0 μl, 0.09 mmol). The mixture was heated to reflux for eight hours, cooled, and then evaporated *in vacuo*. The residue was dissolved in acetonitrile (1.2 ml), and treated with triethylamine (220 μl, 1.58 mmol). To this solution was added

p-acetamidobenzenesulfonyl azide (31.9 mg, 0.13 mmol), and the mixture was stirred for ten hours. The solvent was removed *in vacuo*, the residue dissolved in a minimal amount of dichloromethane (500 μ l), and the undesired salts were precipitated out of solution with the addition of ether (5 ml). Filtration through celite and removal of solvent *in vacuo* yielded **10.1** (41.8 mg, 81% yield). A small sample (~15 mg) was chromatographed to analytical purity by HPLC, (0.1% (wt/v) TFA in water, 8.0 ml/min, 0.30% acetonitrile/min, 83-85 min). FTIR (thin film) 3429, 2918, 2143, 1743, 1691, 1595, 1454, 1356, 1251, 1108 cm^{-1} ; ^1H NMR (300 MHz, CDCl_3) δ 4.44 (dd, $J = 3.85$, 11.5 Hz, 1H), 4.26 (dd, $J = 5.49$, 11.0 Hz, 1H), 3.87-3.58 (m, 23H); MS m/z 683.3 (H^+).



Compound 10.2: To a stirred, dry solution of 18-crown-6-methanol (50.0 μ l, 0.16 mmol), dichloromethane (1.5 ml), and triethylamine (33 μ l, 0.24 mmol) was added ethyl malonyl chloride (28 μ l, 0.22 mmol). The mixture was heated to reflux for eight hours, cooled, and then evaporated *in vacuo*. The residue was dissolved in acetonitrile (750 μ l), and treated with triethylamine (30 μ l, 0.22 mmol). To this solution was added *p*-acetamidobenzenesulfonyl azide (53.1 mg, 0.22 mmol), and the mixture was stirred for ten hours. The solvent was removed *in vacuo*, the residue dissolved in a minimal amount of dichloromethane (500 μ l), and the undesired salts were precipitated out of solution with the addition of ether (5 ml). Filtration through celite and removal of solvent *in*

vacuo yielded **10.2** (59.8 mg, 87% yield). A small sample (~15 mg) was chromatographed to analytical purity by HPLC (0.1% (wt/v) TFA in water, 8.0 ml/min, 0.30% acetonitrile/min, 82-85 min). FTIR (thin film) 2879, 2142, 1755, 1689, 1457, 1326, 1102, 762; ^1H NMR (300 MHz, CDCl_3) δ 4.45 (dd, $J = 3.85, 12.1$ Hz, 1H), 4.31 (q, $J = 7.14$ Hz, 2H), 4.27 (m, 1H), 3.85 (t, $J = 4.95$), 3.80 (s, broad, 1H), 3.67 (s, broad, 2H), 1.32 (t, $J = 7.14$ Hz, 3H); MS m/z 435.2 (H^+).

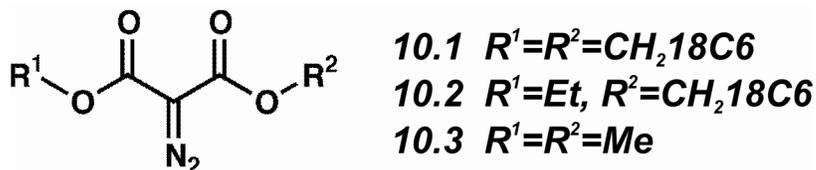
Mass Spectrometry. All spectra were obtained using a Finnigan LCQ ion trap quadrupole mass spectrometer without modification. Sample concentrations were typically kept in the ~10 to 100 μM range for all species of interest. All samples were electrosprayed in a mixture of 80:20 methanol/water. The appropriate host was added to the sample and electrosprayed with the guest in order to observe adducts. Semi-empirical calculations for Figure 10.1 were performed on HyperChem 5.1 Professional Suite using the PM3 parameter set.

Theory. Calculations to determine the singlet/triplet splittings were performed on structures fully optimized at the B3LYP/CCPVTZ(-F) $^+$ level. Comparison of this methodology with previous computational and experimental results for the following carbenes CH_2 , HCCl , HCF , CCl_2 , CF_2 , and HCCHO yielded results within (on average) ± 0.6 kcal/mol of the best experimental or theoretical value.⁸ Zero point energy corrections were not included. Reactions were modeled at the B3LYP/6-31G** level by minimizing structures containing both reactants, with several different starting geometries. Initial geometries included likely starting points for the most probable reaction mechanisms i.e. hydrogen abstraction, concerted insertion, and ylide formation.

The DFT calculations were carried out using Jaguar 4.1 (Schrödinger, Inc., Portland, Oregon).

10.3 Results and Discussion

We have synthesized and examined the chemistry of the prototypical molecular mousetraps **10.1** and **10.2**. 18-crown-6 (18C6) is a well-known host for protonated primary amines, both in solution and in the gas phase.⁹ For example, we have recently shown that 18C6 selectively binds to lysine residues in small peptides.^{3a} Mousetrap **10.1** is designed to bind molecules with either one or, preferentially, two protonated primary amines. Mousetrap **10.2**, with a single 18C6, binds to a single protonated primary amine. Structure **10.3** was used as a model compound in computations.



The interaction between **10.1** and doubly protonated 1,6-diaminohexane (DAH) is shown in Figure 10.1. This complex forms in solution and can be transferred intact to the gas phase by ESI, as seen in Figure 10.2a. The complex can be isolated and subjected to CAD as shown in Figure 10.2b. The sole product results from a neutral loss of 28 Da, which is interpreted to be the loss of N_2 from the diazo group. Significantly, the result shown in Figure 10.2b provides evidence for covalent bond cleavage in preference to dissociation of the complex. The loss of N_2 from the diazo should yield the corresponding carbene (**:10.1**) as a highly reactive, short lived intermediate. This carbene can then react intermolecularly or intramolecularly.

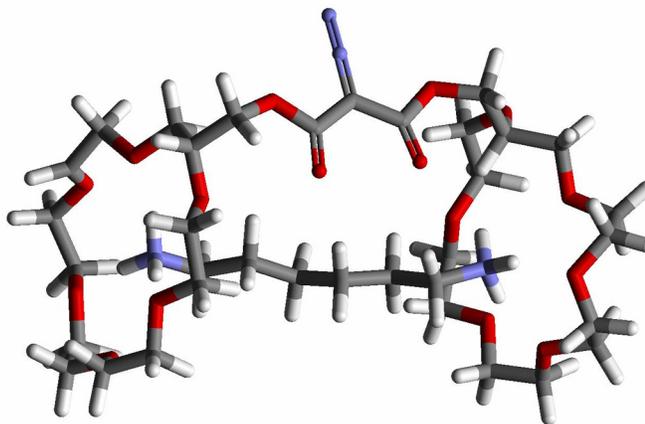


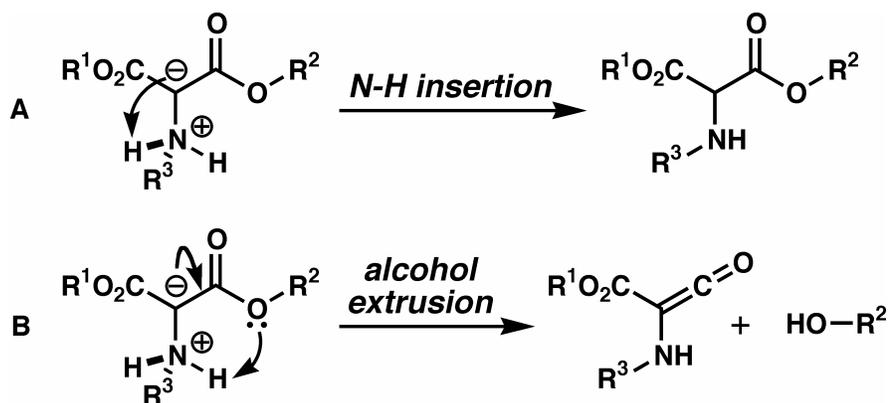
Figure 10.1 One possible structure for the noncovalent adduct of **10.1** and doubly protonated 1,6-diaminohexane (DAH) in the gas phase, as determined by PM3 semi-empirical calculations.

The product from Figure 10.2b is subjected to further collisional activation in Figure 10.2c. The majority of the product ion intensity results from covalent bond cleavage with loss of a crown or part of a crown and retention of DAH. The fragmentation of the host without the accompanying loss of the guest provides evidence that an intermolecular reaction involving covalent coupling of the complex by C-H insertion of the carbene has occurred. For the doubly protonated DAH, the complexation of the protonated primary amines by the crown ethers reduces the likelihood of an N-H insertion reaction by the carbene.¹⁰ It is also observed in Figure 10.2c that some of the DAH simply dissociates from the complex, suggesting an intramolecular process¹¹ is competitive in this case.

Singly charged DAH has a lower binding energy to **10.1** than the doubly charged species, yet Figure 10.2d illustrates that $[:\mathbf{10.1}+\text{DAH}+\text{H}]^+$ is generated with high efficiency from the $[\mathbf{10.1}+\text{DAH}+\text{H}]^+$ complex. The loss of nitrogen is accompanied by an additional loss of 294 Da, which can be accounted for by the loss of hydroxymethyl

18C6. This additional loss is observed for all complexes of both **10.1** and **10.2** in which there is an unprotonated primary amine or alcohol available (in experiments with **10.2**, the loss of ethanol is also observed). DFT calculations at the at the B3LYP/CCPVTZ(-F)⁺ level on **10.3** yield a singlet ground state with a singlet/triplet splitting of 3±1 kcal/mol. This suggests that the singlet state is certainly accessible and perhaps favorable, which is in agreement with experimental results.¹² DFT calculations at the B3LYP/6-31G** level on **10.3** and methylamine lead to the formation of an ammonium ylide without barrier. The ammonium ylide is a local minimum on the potential energy surface, and previous reports have suggested that all carbenes will initially react with amines by the formation of an intermediate ylide.¹³ From this ammonium ylide two reaction pathways with minimal barriers are possible as shown in Scheme 10.1, and it should be pointed out that both lead to covalent attachment of the host/guest complex by intermolecular reactions. One leads to formal N-H insertion, and the other leads to the loss of an alcohol and the generation of a ketene.

Scheme 10.1.



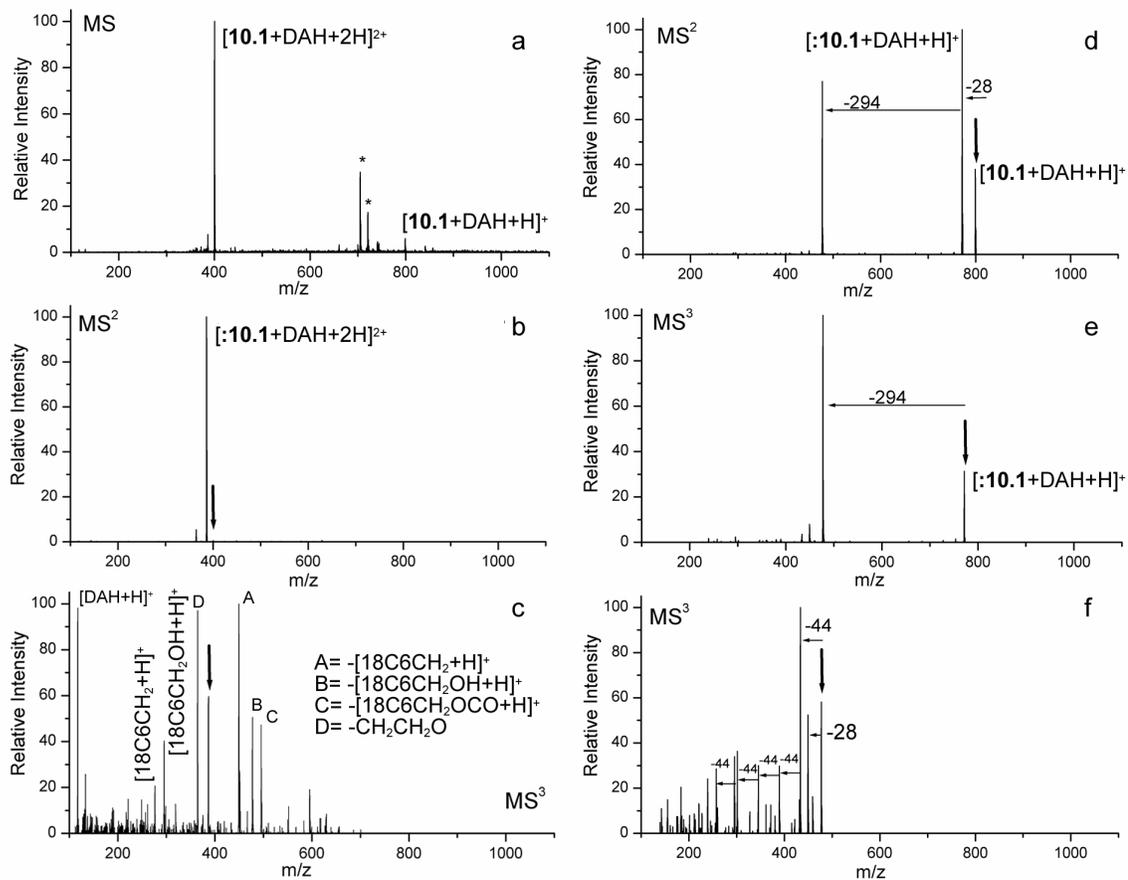


Figure 10.2 (a) Mass spectrum of **10.1** with DAH. (b) CAD spectrum of $[\mathbf{10.1}+\text{DAH}+2\text{H}]^{2+}$ which shows the almost exclusive loss of 28 Da, indicating the generation of carbene **:10.1**. (c) MS^3 on the $[\mathbf{10.1}+\text{DAH}+2\text{H}]^{2+}$ peak. Dissociation is accompanied by covalent bond cleavage, suggesting C-H insertion by **:10.1** and the formation of a new molecule. (d) CAD of $[\mathbf{10.1}+\text{DAH}+\text{H}]^+$, leading entirely to intermolecular reaction products. (e) MS^3 on $[\mathbf{10.1}+\text{DAH}+\text{H}]^+$, resulting only in the loss of 294 Da. The absence of complex dissociation suggests covalent attachment. (f) MS^3 on $[\mathbf{10.1}+\text{DAH}+\text{H}-294]^+$. In the absence of both crowns, DAH is retained, confirming the covalent coupling of the complex. A bold downward arrow indicates the peak being subjected to CAD. * 1+alkali metal adduct peaks.

Further excitation of the isolated $[:\mathbf{10.1}+\text{DAH}+\text{H}]^+$ peak following the loss of nitrogen yields exclusively the loss of 294 Da as shown in Figure 10.2e. The N-H insertion product shown in Scheme 10.1A has a proton on the secondary amine. Transfer of this proton to the ester can lead to the loss observed in Figure 10.2e by alcohol extrusion. It is also possible, though unlikely, that the ammonium ylide formed in Figure 10.2d could be sufficiently long-lived to yield this product directly.

Figure 10.2f offers several critical results. First, the loss of 28 Da is probably the loss of CO from the ketene product shown in Scheme 10.1B. Second, the fragment being subjected to further collisional activation in Figure 10.2f contains only a single remaining crown. The primary losses are multiple $\text{CH}_2\text{CH}_2\text{O}$ fragments from this remaining crown. The data reveals the sequential removal of nearly the entire remaining crown ether without the loss of the guest molecule. In the absence of both crowns, the retention of the guest can only be explained by a newly formed covalent bond.

10.4 Conclusion

These studies demonstrate that reagents which bind to specific functional groups in complex molecules can be derivatized to introduce the means to covalently couple them to the target molecules with appropriate methods of activation. We have combined 18C6, which binds strongly to protonated primary amines,^{3a} with a diazo precursor to a reactive carbene to form a potent “molecular mousetrap” that can be used to target lysines in peptides or proteins. More details of the chemistry and applications of these and related molecular mousetraps are described in Chapter 11.

10.5 References

-
- ¹ Peczuh, M. W.; Hamilton, A. D. *Chem. Rev.* **2000**, *100*, 2479-2494.
- ² Beer, P. D.; Gale, P. A.; Smith, D. K. *Supramolecular Chemistry*; Oxford University Press, New York, 1999.
- ³ (a) Julian, R. R.; Beauchamp, J. L. *Int. J. Mass. Spectrom.* **2001**, *210*, 613-623; (b) Smith, R. D.; Bruce, J. E.; Wu Q. Y.; Lei, Q. P. *Chem. Soc. Rev.* **1997**, *26*, 191-202; (c) Veenstra, T. D. *Biophys. Chem.* **1999**, *79*, 63-79. (d) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.
- ⁴ A neutral, two electron carbene is formed. Other studies have focused on carbene radical cations such as CH_2^+ , see for example: Flammang, R.; Nguyen, M. T.; Bouschoux, G.; Gerbaux, P. *Int. J. Mass Spectrom.* **2002**, *202*, A8-A25.
- ⁵ (a) Marzluff, E. M.; Beauchamp, J. L. in *Large Ions: Their Vaporization, Detection, and Structural Analysis*; Baer, T., Ng, C. Y., Powis, I.; Eds.; John Wiley & Sons Ltd.: New York, 1996; 115-143; (b) McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599-614; (c) Hayes, R. N.; Gross, M. L. *Methods Enzymol.* **1990**, *193*, 237-263.
- ⁶ (a) Rice, F. O.; Glasebrook, A. L. *J. Am. Chem. Soc.* **1934**, *56*, 2381-2383; (b) Herzberg, G. *Proc. Roy. Soc. Lond. Ser. A.* **1961**, *262*, 291-317; (c) Moss, R. A.; Jones, M., Jr. Eds. *Carbenes, Vols. 1 and 2*; Wiley: New York, 1973, 1975; (d) Rynbrandt, J. D.; Rabinovitch, B. S. *J. Phys. Chem.* **1970**, *74*, 4175-4176; (e) Rynbrandt, J. D.; Rabinovitch, B. S. *J. Chem. Phys.* **1971**, *54*, 2275-2276; (f) Poutsma, J. C.; Nash, J. J.; Paulino, J. A.; Squires, R. R. *J. Am. Chem. Soc.* **1997**, *119*, 4686-4697; (g) Paulino, J. A.; Squires, R. R. *J. Am. Chem. Soc.* **1991**, *113*, 5573-5580; (h) Leopold, D. G.; Murray, K. K.; Miller, A. E. S.; Lineberger, W. C. *J. Chem. Phys.* **1985**, *83*, 4849-4865 and

references therein; (i) Bertani, R.; Michelin, R. A.; Mozzon, M.; Traldi, P.; Seraglia, R.; Busetto, L.; Cassani, M. C.; Tagleatesta, P.; D'Arcangelo, G. *Organometallics*, **1997**, *16*, 3229-3233.

⁷ (a) Doyle, M. P.; McKervey, M. A.; Ye, T. *Modern Catalytic Methods for Organic Synthesis with Diazo Compounds*; Wiley-Interscience: New York, 1998; (b) Moody, C. J.; Whitham, G. H. *Reactive Intermediates*; Oxford University Press: New York, 1992; 26-50.

⁸ Scott, A. P.; Platz, M. S.; Radom, L. *J. Am. Chem. Soc.* **2001**, *123*, 6069-6076.

⁹ (a) Bradshaw, J. S.; Izatt, R. M.; Borkunov, A. V.; Zhu, C. Y.; Hathaway, J. K. *Comprehensive Supramolecular Chemistry, Vol. 1*; G. W. Gokel, Pergamon/Elsevier: Oxford, 1996; 35-95; (b) Maleknia, S.; Brodbelt, J. *J. Am. Chem. Soc.* **1993**, *115*, 2837-2843; (c) Dearden, D. V.; Dejsupa, C.; Liang, Y.; Bradshaw, J. S.; Izatt, R. M. *J. Am. Chem. Soc.* **1997**, *119*, 353-359; (d) Schalley, C. A. *Mass Spectrom. Rev.* **2001**, *20*, 253-309; (e) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.

¹⁰ 18C6 locks protons onto primary amines, see ref 3a and Lee, S.-W.; Lee, H.-N.; Kim, H. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1998**, *120*, 5800-5805.

¹¹ The likely competing reaction is Wolff rearrangement of the carbene. See, for example, (a) Likhovorik, I.; Zhendong, Z.; Tae, E. L.; Tippmann, E.; Hill, B. T.; Platz, M. S. *J. Am. Chem. Soc.* **2001**, *123*, 6061-6068; (b) ref 8.

¹² Both results obtained the current work and in: Richardson, D. C.; Hendrick, M. E.; Jones, M. *J. Am. Chem. Soc.* **1971**, *93*, 3790-3791, where Wolff rearrangement is observed, which must proceed through the singlet state.

¹³ Pliego, J. R.; Almeida, W. B. *J. Phys. Chem. A* **1999**, *103*, 3904-3909.

Chapter 11

Biomimetic Approaches to Gas Phase Peptide

Chemistry: Combining Selective Binding Motifs with

Reactive Carbene Precursors to Form

Molecular Mousetraps

Accepted for publication in: Julian, R. R.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L.
Int. J. Mass Spectrom.

11.1 Introduction

In the post genomic world of proteomics,¹ many substantial advances will be made through experiments conducted in the gas phase. Therefore, understanding and (ultimately) controlling gas phase peptide chemistry is of paramount importance. For example, the study of gas phase peptide chemistry has revealed that selective cleavage of the peptide backbone will occur at aspartic acid residues.^{2,3} It has been further demonstrated that this cleavage occurs by a displacement reaction that yields a stable five-membered ring. Understanding this phenomenon allows for the accurate prediction of peptide cleavages in aspartic acid containing peptides. Furthermore, C-terminal peptide sequencing via a similar mechanism where the C-terminal amino acids are sequentially removed has also yielded promising, if limited, results.⁴ Unfortunately, this C-terminal sequencing is restricted to peptides with 8 amino acids or less, severely

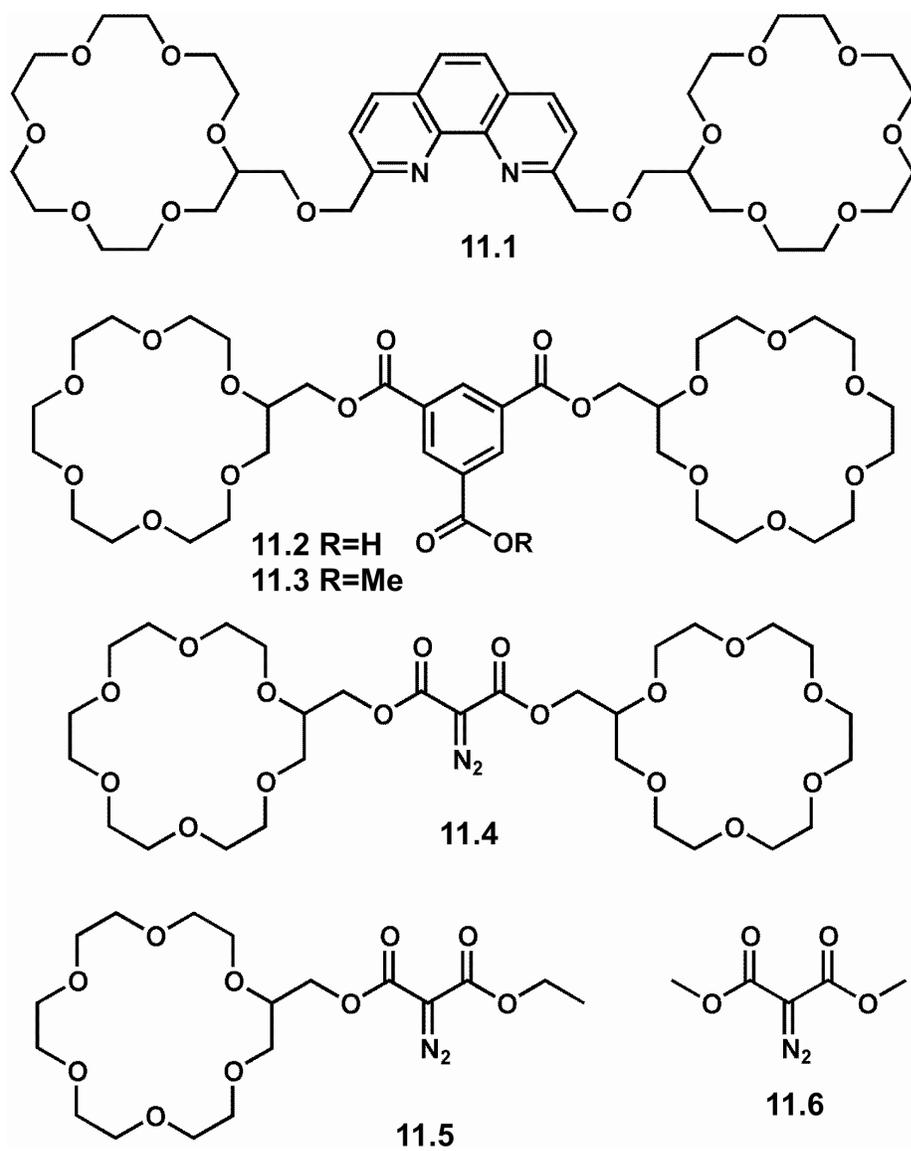
curtailing the ability of this technique to sequence proteins in the gas phase. The addition of transition metals can also mediate peptide chemistry in the gas phase.^{5,6} Preliminary studies have shown that Zn^{2+} , Ni^{2+} , and Co^{2+} will attach to histidine and promote peptide fragmentation at this residue.⁵ These experiments were carried out on a very limited sampling of peptides, but the resulting cleavages were highly specific. Similarly, Fe^{2+} complexes with cysteine containing peptides enhanced the number of cleavages observed at the cysteine residues when the peptide was collisionally activated. These important initial results illustrate that peptide chemistry can be influenced by the addition of appropriate reagents.

In an effort to develop experimental methodologies for the controlled manipulation of peptides in the gas phase, we have undertaken a systematic study to develop *de novo* reagents capable of selectively attaching to and reacting with peptides in the gas phase. It is envisioned that these reagents will be capable of initiating a wide range of chemical reactions, such as peptide backbone cleavage at specific residues, as described above. With appropriate modifications, the type of reagents proposed herein could serve as fluorescent probes, chemical cross-linkers, and sequence specific binding agents. In the present work we consider two reagents (**11.1** and **11.2**) designed to initiate selective cleavage of the peptide backbone near lysine residues. In addition, we present reagents **11.4** and **11.5** designed to covalently attach to lysine containing peptides following appropriate activation which generates a reactive carbene center. We first discuss considerations that led to the design of these reagents.

Reagents **11.1-11.5** all rely on molecular recognition of specific amino acid side chains to form specific noncovalent complexes in the gas phase. Fortunately, a significant amount of work developing reagents that selectively recognize and noncovalently attach

to specific amino acid side chains has already been reported.^{7,8,9} These side chain hosts represent the first and simplest form of a biomimetic reagent, one that is only capable of recognition. It should be noted that the facile formation of noncovalent complexes is critical to the success of this type of experiment, and proper conditions for enhancing noncovalent complex formation have been studied extensively.^{10,11}

Chart 11.1



18-crown-6 ether (18C6) was chosen as the recognition and binding motif because it is both synthetically flexible and amenable to noncovalent complexation. It is well known for its ability to bind both metal cations and protonated primary amines in both solution and in the gas phase.¹² This ability is particularly useful for the recognition of lysine, because the side chain of this amino acid terminates in a primary amine. 18C6 complexes protonated primary amines through a combination of three hydrogen bonds and ion-dipole interactions. Noncovalent complexes with 18C6 bound to protonated primary amines can be transferred into the gas phase by electrospray ionization mass spectrometry (ESI-MS). When added to a solution containing a peptide, the 18C6 complex with the peptide is typically the most abundant peak in the spectrum.⁸ Appropriately modified lariat crown ethers behave similarly, forming noncovalent complexes that can be transferred to the gas phase as shown previously¹³ and in the present work.

Lariat crown ethers **11.1** and **11.2** were synthesized and tested to determine their ability to selectively cleave various peptides. While these reagents did not demonstrate optimal reactivity, yielding limited results with regards to peptide cleavage, they serve to illustrate the important factors in biomimetic reagent design. Greater success was achieved with reagents **11.4** and **11.5**, which are designed to covalently attach to peptides. These two reagents containing diazo groups were complexed with a series of small molecules and peptides. Collisional activation was utilized to generate a carbene from the diazo functionality without dissociation of the complex. The intermolecular reactions were studied with ESI-MS and density functional theory (DFT). Sequential MSⁿ spectra revealed covalent bond attachment between the constituents of the complex subsequent to the generation of the carbene. The data demonstrate that the insertion

reactions are sensitive to the presence or absence of N-H and O-H functional groups. The present work expands upon our previous work reporting initial experiments with this class of reagents that we have termed molecular mousetraps.¹⁴

10.2 Experimental

Mass Spectrometry. All spectra were obtained using a Finnigan LCQ ion trap quadrupole mass spectrometer without modification. The critical instrument settings that yield adduct formation include capillary voltage 5-15V, capillary temperature 200°C, and tube lens offset -30 to -50V. Higher capillary temperatures can dissociate the noncovalent complexes. The tube lens offset controls the acceleration of ions as they leave the capillary region. The tube lens voltage is minimized to avoid collisions with the He buffer gas. Soft sampling is crucial for the detection of these noncovalent complexes.

Sample concentrations were typically kept in the ~10 to 100 μM range for all species of interest. All samples were electrosprayed in a mixture of 80:20 methanol/water. The appropriate host was added to the sample and electrosprayed with the guest in order to observe adducts. Collision activated dissociation (CAD) was performed by isolating and then exciting the isolated peak by colliding it with He buffer gas. Samples were electrosprayed with a flow of 3-5 $\mu\text{L}/\text{min}$ from a 500 μL Hamilton syringe for optimal signal. Silica tubing with an inner diameter of .005 in was used as the electrospray tip.

Calculations. The energetics of the carbene insertion reactions were quantitatively evaluated by carrying out reactions with the model compound **11.6**. The structures of all reactants were fully minimized, and several different reaction mechanisms were tested. Initial structures included likely starting points for hydrogen abstraction, concerted

insertion, and ylide formation. The starting structures for each of these possibilities corresponded respectively to: one hydrogen directed at the carbene, symmetrical presentation of the H-C-H or O-H bonds, and one lone pair directed at the carbene. The DFT calculations were carried out using Jaguar 4.1 (Schrödinger, Inc., Portland, Oregon). PM5 semi-empirical calculations were carried out using CACHE Worksystem Pro 5.04 (Fujitsu, Inc., Beaverton, Oregon).

Experimental Details for Syntheses: Due caution should always be used when handling diazo compounds. Reactions were performed in flame-dried glassware under a nitrogen atmosphere. Solvents were dried and purified using activated alumina columns. Diethylamine was distilled from CaH_2 . 18-crown-6-methanol was dried prior to use by heating (~ 100 °C) under vacuum. All other reagents were used as received from commercial sources. Reaction temperatures were controlled by an IKAmag temperature modulator.

Compound 11.1: To a stirred solution of diethylamine (13 μL , 0.123 mmol) in THF (500 μL) at 0 °C was added nBuLi (60 μl , 2.1 M, 0.126 mmol) dropwise. The mixture was stirred for 10 min and then transferred via syringe to a solution of 18-crown-6-methanol (30 μL , 0.109 mmol) in THF (500 μL) stirred at -78 °C. The solvent was removed under reduced pressure as the reaction warmed to room temperature. Excess diethylamine was removed by two consecutive additions of THF (1 mL) and removal under reduced pressure. The residue was then redissolved in THF (1 mL) and 2,9-bis(bromomethyl)-1,10-phenanthroline¹⁵ (19 mg, 0.052 mmol) in CH_2Cl_2 (4 mL) was added. The resulting solution was stirred for 24 hours, and then ether (10 mL) was added to precipitate the salt byproduct, which was removed by filtration through celite. The

removal of solvent under reduced pressure yielded **11.1** (37.5 mg, 0.047 mmol, 91% yield) in sufficient purity for experimental use.

Compound 11.2: To a stirred solution of 18-crown-6-methanol (47 μl , 0.150 mmol), triethylamine (25 μl , 0.179 mmol), and dichloromethane (4.5 ml) was added 1,3,5-benzenetricarbonyl trichloride (20.4 mg, 0.077 mmol). The mixture was heated to reflux for 12 hours, and then H_2O (1.5 mL) was added and the mixture was again heated to reflux for 1 hour. The solvent was removed under reduced pressure, the residue dissolved in a minimal amount of dichloromethane (500 μl), and the undesired salts were precipitated out of solution with the addition of ether (5 ml). Filtration through celite and removal of solvent under reduced pressure yielded **11.2** (54.2 mg, 0.071 mmol, 95% yield) in sufficient purity for experimental use.

Compound 11.3: An identical procedure as that for the formation of **11.2** was followed with the exception that the reaction was quenched with MeOH (500 μL) instead of H_2O to yield **11.3** (49.1 mg, 0.063 mmol, 82% yield) in sufficient purity for experimental use.

Compounds 11.4 and 11.5 were prepared according to the method described in Chapter 10.

11.3 Results and Discussion

Transition metals have been observed to influence peptide dissociation in previous gas phase experiments.^{5,6} In an attempt to utilize the reactivity of transition metals for the selective cleavage of peptide bonds, reagent **11.1** was developed. **11.1** consists of two 18C6 ethers linked by a phenanthroline moiety, which can bind a variety of transition metals. Figure 11.1a shows that **11.1** forms an abundant noncovalent complex with the

peptide KK and copper (I). Collisional activation of the base peak $[\mathbf{11.1}+\text{KK}+\text{Cu}+\text{H}]^{2+}$ results primarily in dissociation of the complex into $[\mathbf{11.1}+\text{Cu}]^+$ and $[\text{KK}+\text{H}]^+$ with an additional prominent peak corresponding to the loss of 44 Da from $[\text{KK}+\text{H}]^+$. This loss is most likely explained as elimination of CO_2 from the C-terminus. In Figure 11.1c, collisional activation of the much less abundant complex $[\mathbf{11.1}+\text{KK}+\text{Cu}+2\text{H}]^{3+}$ yields the loss of CO_2 directly. In the absence of the copper (I) ion, no loss of 44 Da is observed for either charge state, suggesting that copper (I) effectively initiates this reaction. Unfortunately, this chemistry only occurs with very short peptides that end with KK or RK, and reagent **11.1** did not initiate any other cleavages. A wide variety of peptides and different transition metals including Ag(I), Fe(III), Co(II), Zn(I), Zn(II), Mn(II), Ni(II), Pd(II) and Cu(II) were tested. Many of these experiments failed to produce an abundant noncovalent complex, and when the complex was formed and isolated the result was simple dissociation in every case where the peptide contained an internal KK sequence.

These results can be rationalized by insufficient binding energy of the noncovalent complex in the gas phase. The presence of a cationic transition metal trapped between two positively charged lysine residues results in unfavorable coulombic interactions that effectively reduce the binding energy of the complex. The binding energy is reduced by $\sim 80 \pm 10$ kcal/mol for inserting a singly charged transition metal ion as determined by PM5 calculations. This explains why only minimal complexation (or none) occurs for internal KK sequences, and the reduced binding also leads to the exclusive dissociation of these complexes upon collisional activation. A deprotonated C-terminus effectively mitigates the unfavorable interactions and increases the binding energy by neutralizing the central positive charge. Therefore, reagent **11.1** is suitable for selectively attaching

near the C-terminus of peptides that end in KK or RK/KR, however it did not prove effective at cleaving peptides in the gas phase.

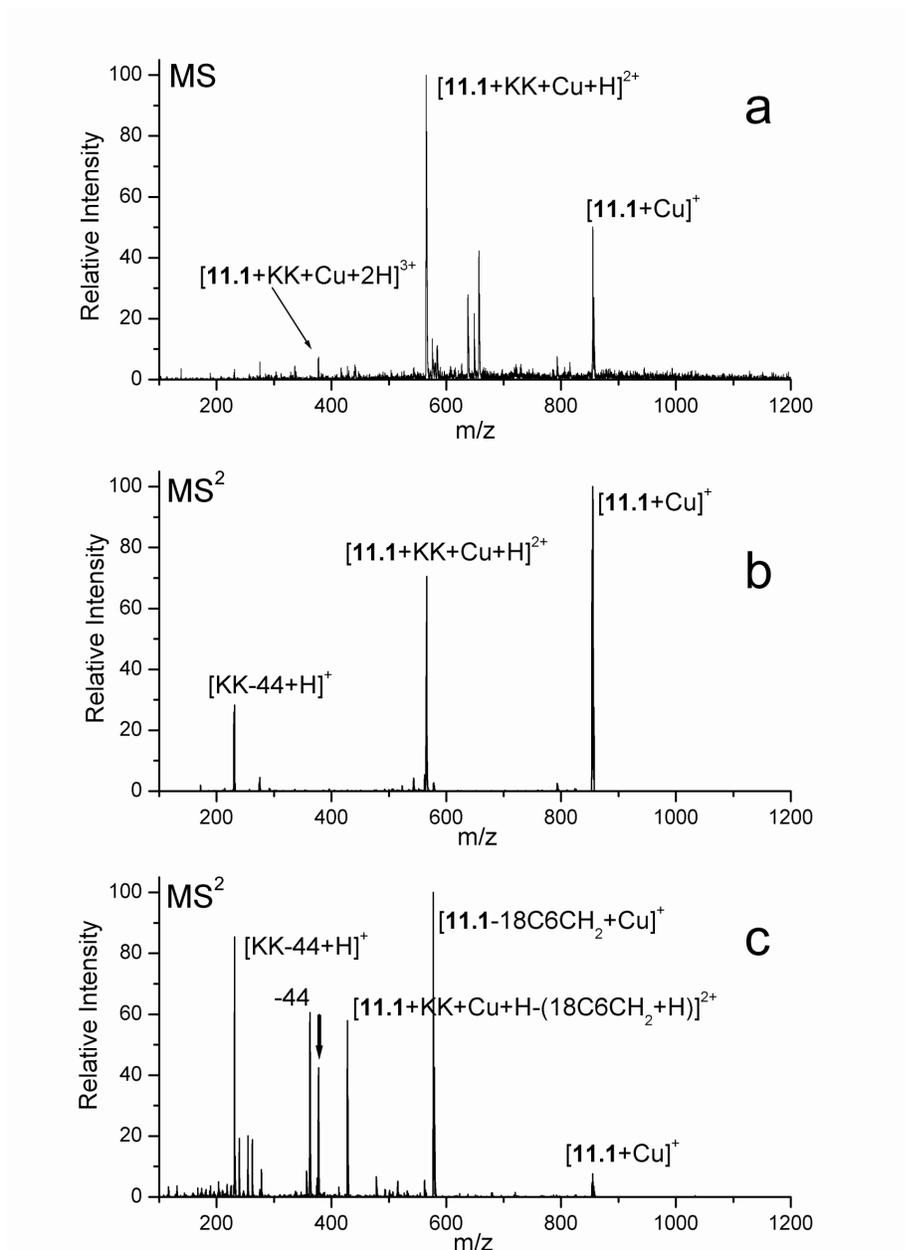


Figure 11.1 (a) ESI-MS of **11.1**, copper (I), and KK. An abundant complex is formed. (b) MS² on the doubly charged complex results in simple dissociation. (c) MS² on the triply charged complex results in the loss of CO₂ from the C-terminus of the peptide. Bold downward arrow indicates the peak being subjected to collisional activation.

As mentioned above, selective cleavage at aspartic acid residues has been observed in the gas phase previously, indicating that acid/base chemistry may provide an alternate route for cleaving peptides in the gas phase.^{2,3} Reagent **11.2** was designed based upon this premise. **11.2** contains two 18C6 ethers linked by benzoic acid. Deprotonation of the acid is assisted by favorable electrostatic interactions upon complexation with two protonated lysine residues. The ESI mass spectrum for a solution of **11.2** and KKKK is shown in Figure 11.2a. The doubly charged adduct [**11.2**+KKKK+2H]²⁺ forms the base peak in the spectrum. Collisional activation of this peak results primarily in dissociation of the complex. However, there are additionally two peaks corresponding to the loss of water and the N-terminal lysine. To verify that this chemistry was initiated by the benzoic acid, an additional experiment was conducted where the acid was converted to a methyl ester (**11.3**). The results are shown in Figure 11.2c and are nearly identical to those shown in Figure 11.2b.

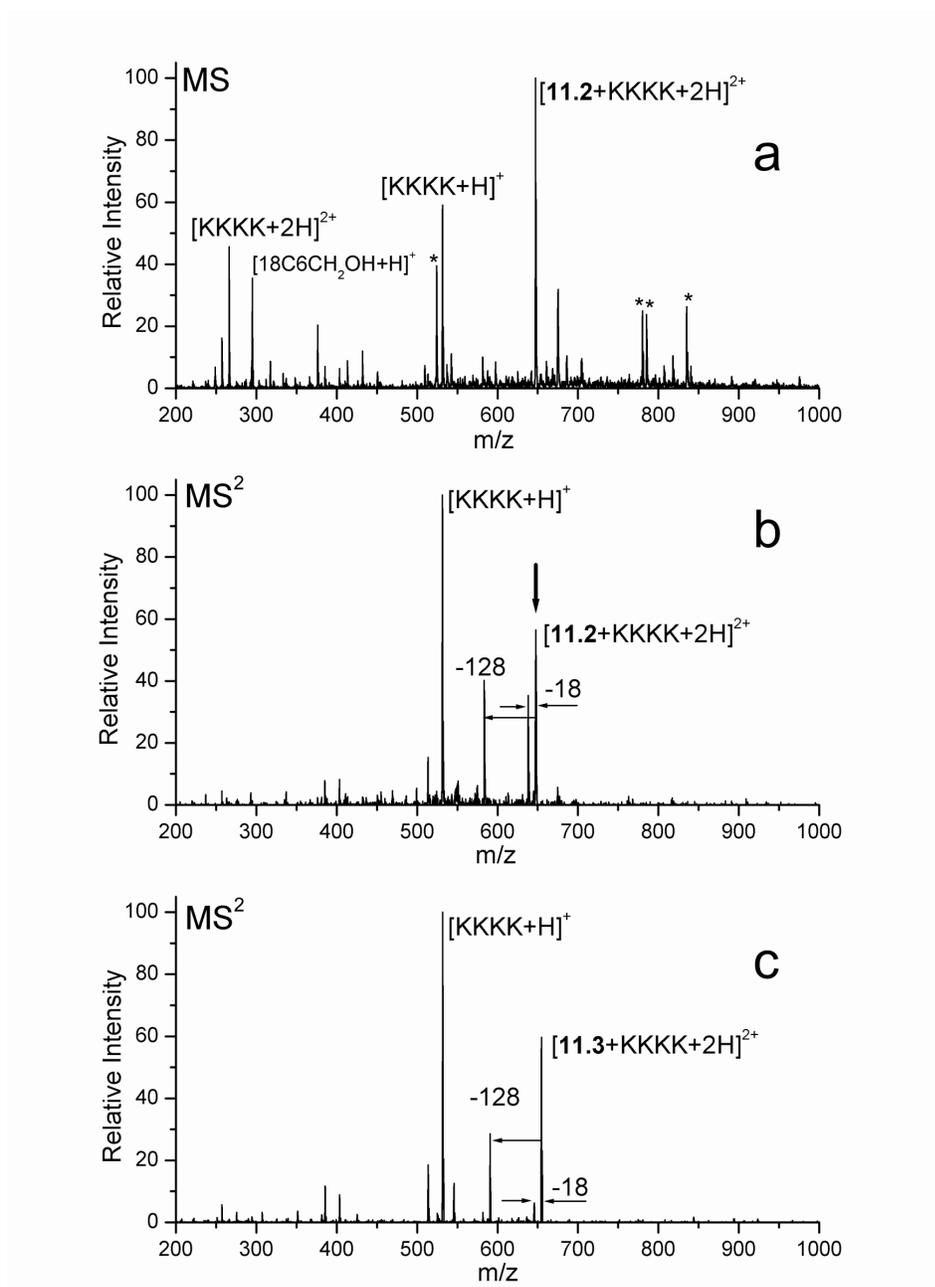
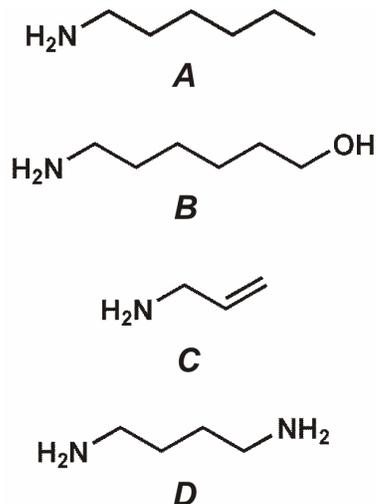


Figure 11.2 (a) ESI-MS of KKKK with **11.2**, demonstrating excellent recognition. (b) MS² on the base peak leads to the loss of the N-terminal lysine. (c) Control experiment with **11.3** yields same results as in (b), suggesting that **11.2** is merely a spectator adduct and does not initiate the cleavage of the N-terminal lysine. Bold downward arrow indicates the peak being subjected to collisional activation.

Therefore it is likely that **11.2** is merely a spectator adduct, which is sufficiently strongly bound to remain attached after a covalent bond cleavage has occurred but does not directly affect the cleavage process. Earlier studies of selective cleavages at aspartic acid residues suggest that this process is favored due to the proximity of the aspartic acid side chain to the peptide backbone, with acidity enhanced by a proximal positive charge.² The observation that the similar reactivity of glutamic acid (with the addition of a single methylene) is greatly reduced in comparison suggests that the reaction has very special geometrical constraints. It may be that the acidic group in **11.2** cannot exploit the same reaction pathway as inferred for aspartic acid cleavages because it is not held in close proximity to the peptide backbone. Nevertheless, the results from **11.2** are important because they demonstrate that biomimetic reagents with multiple crown ethers have sufficient binding energy to mitigate dissociation in favor of peptide cleavage processes.

Although the cleaving of peptide bonds remains an important goal, covalent attachment to peptides is another important reaction that is often used for cross linking peptides and proteins.¹⁶ Molecular mousetraps **11.4** and **11.5** are designed to covalently attach to peptides containing lysine residues or any other molecule which contains a protonated primary amine. Both **11.4** and **11.5** contain a reactive diazo group, which yields a highly reactive carbene upon collisional activation. Experimental and theoretical results for the interactions of **11.4** with 1,6-diaminohexane have been reported previously.¹⁴ In order to understand the underlying chemistry, we have performed several experiments with simple small molecules to further elucidate the reaction pathways.

Chart 11.2



Reactions with Small Molecules. In Figure 11.3a, the ESI spectrum for a solution of 1-aminohexane (*A*) and **11.4** is shown. The complex corresponding to $[\mathbf{11.4}+\mathbf{A}+\mathbf{H}]^+$ clearly forms the base peak in the spectrum, demonstrating the excellent recognition of **11.4** for protonated primary amines. This complex is subjected to collisional activation in Figure 11.3b. The loss of N_2 is the only major product observed, yielding the reactive carbene (denoted by $:\mathbf{11.4}$) in nearly 100% yield. Theoretical results at the B3LYP/6-31G** level with methane and the similar carbene $:\mathbf{11.6}$ suggest that C-H insertion occurs with little or no barrier in a concerted fashion.¹⁷ In Figure 11.3b, the carbene ($:\mathbf{11.4}$) can react with *A* by C-H insertion at various points along the hydrocarbon chain. This is confirmed in Figure 11.3c, where no dissociation of *A* is observed after further collisional activation. Instead several covalent bond cleavages are observed, corresponding to the loss of a $\text{CH}_2\text{CH}_2\text{O}$ link from **18C6** and another corresponding to the loss of an entire crown. This suggests that C-H insertion does in fact occur and leads to the covalent attachment of the host/guest complex.

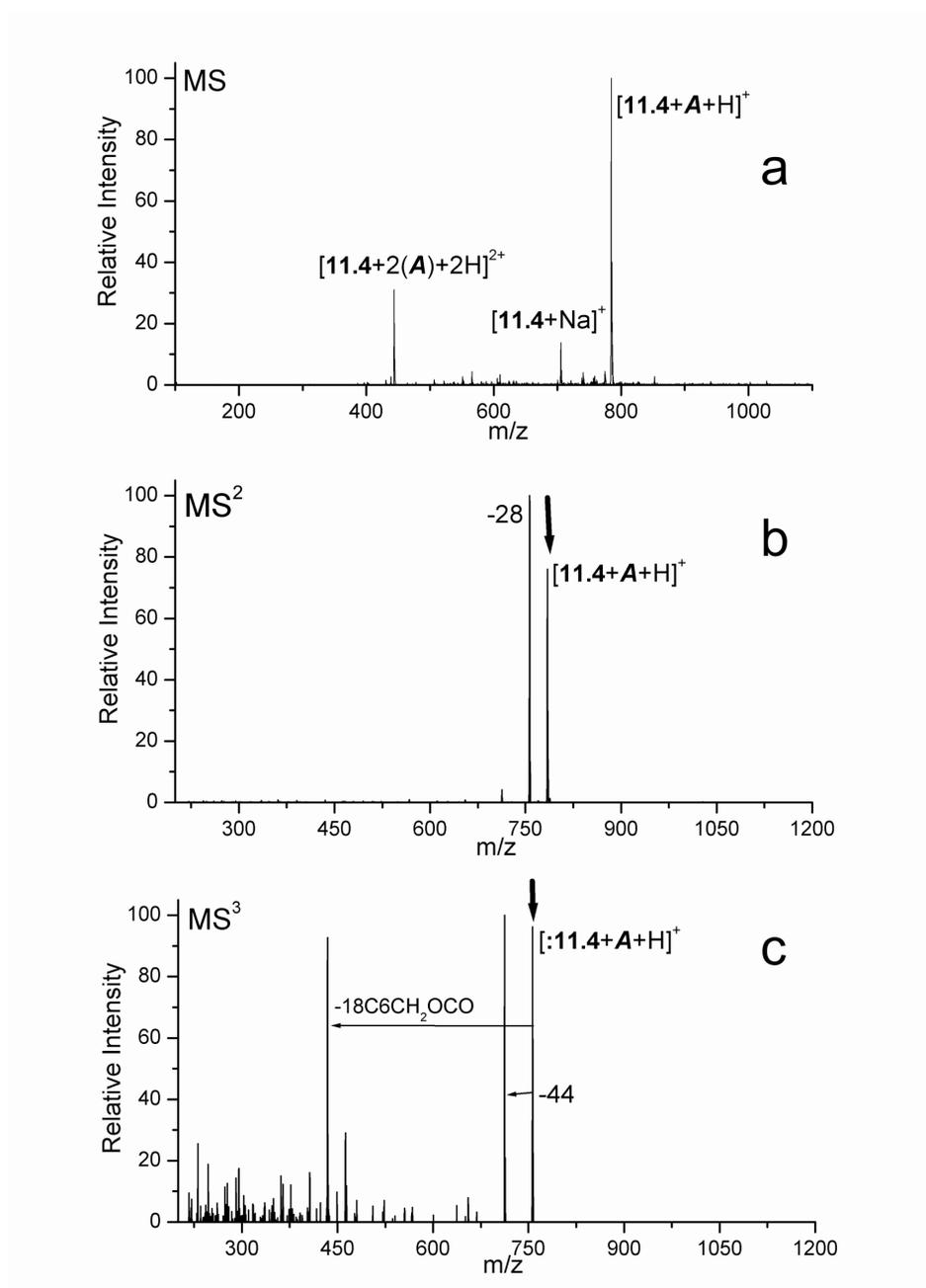


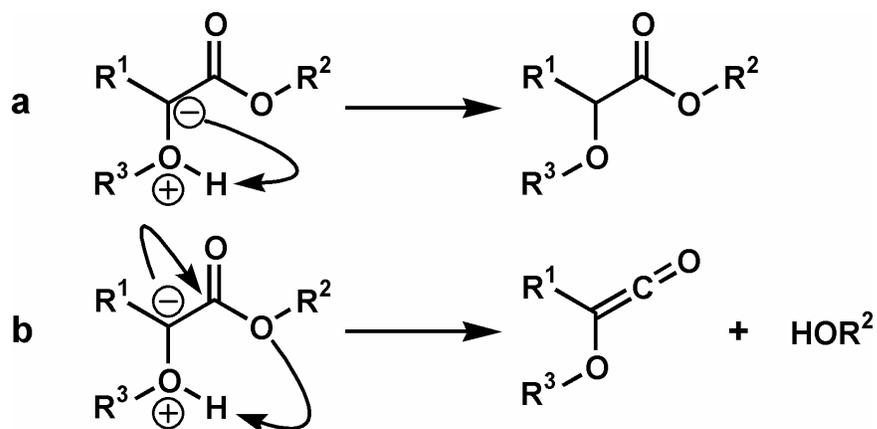
Figure 11.3 (a) ESI-MS of 1-aminohexane (*A*) and **11.4** demonstrating excellent recognition. (b) MS² on the base peak leads primarily to the loss of N₂ and the generation of the corresponding carbene. (c) Further excitation does not result in loss of *A*, suggesting that an intermolecular insertion reaction has occurred. Bold downward arrow indicates the peak being subjected to collisional activation.

Hydroxyl groups are found in three amino acid side chains and can exhibit enhanced reactivity towards carbenes. Figures 11.4a and 11.4b show the results for CAD experiments with **11.4** and 1,6-aminohexanol (**B**) which is used as a model compound. In Figure 11.4a, the CAD of $[\mathbf{11.4+B+H}]^+$ leads to similar results to those obtained previously for 1,6-diaminohexane.¹⁴ The initial loss of N_2 is accompanied by an additional loss of $18\text{C}_6\text{CH}_2\text{OH}$. The MS^3 spectrum is shown in Figure 11.4b for the CAD of $[\mathbf{:11.4+B+H}]^+$. The loss of $\text{CH}_2\text{CH}_2\text{O}$ leads to the base peak in Figure 11.4b while the loss of $18\text{C}_6\text{CH}_2\text{OH}$ is secondary. The loss of $\text{CH}_2\text{CH}_2\text{O}$ is not present in the MS^2 spectrum in Figure 11.4a. This suggests that the loss of $18\text{C}_6\text{CH}_2\text{OH}$ in Figure 11.4a and 11.4b proceed by two different reaction mechanisms and that the two products produced in Figure 11.4a are generated competitively rather than consecutively.

The two proposed reaction pathways are shown in Scheme 11.1 and are similar to those proposed for the comparable 1,6-diaminohexane system.¹⁴ DFT calculations on **:11.6** and H_2O at the B3LYP/6-31G** level support the formation of an intermediate oxonium ylide. The formation of the ylide proceeds without barrier from several different starting geometries. Precedence for this mechanism can be found in previous studies, which have revealed oxonium ylide formation in reactions of various alcohols with carboethoxycarbene, a closely related molecule.¹⁸ All of the experimental and theoretical data support the reaction mechanisms shown in Scheme 11.1 for any system with an alcohol (unprotonated amines react by a very similar pathway as shown previously).¹⁴ In fact, the additional loss of 294 in the MS^2 spectrum is indicative of the presence of alcohols and amines. In Figure 11.4c, further excitation of the complex following the loss of one 18C_6 results primarily in the loss of the other 18C_6 without the accompanying

loss of any **B**. In the absence of both crowns, the retention of the **B** can only be explained by an insertion reaction which has transformed the noncovalent complex into a molecule.

Scheme 11.1



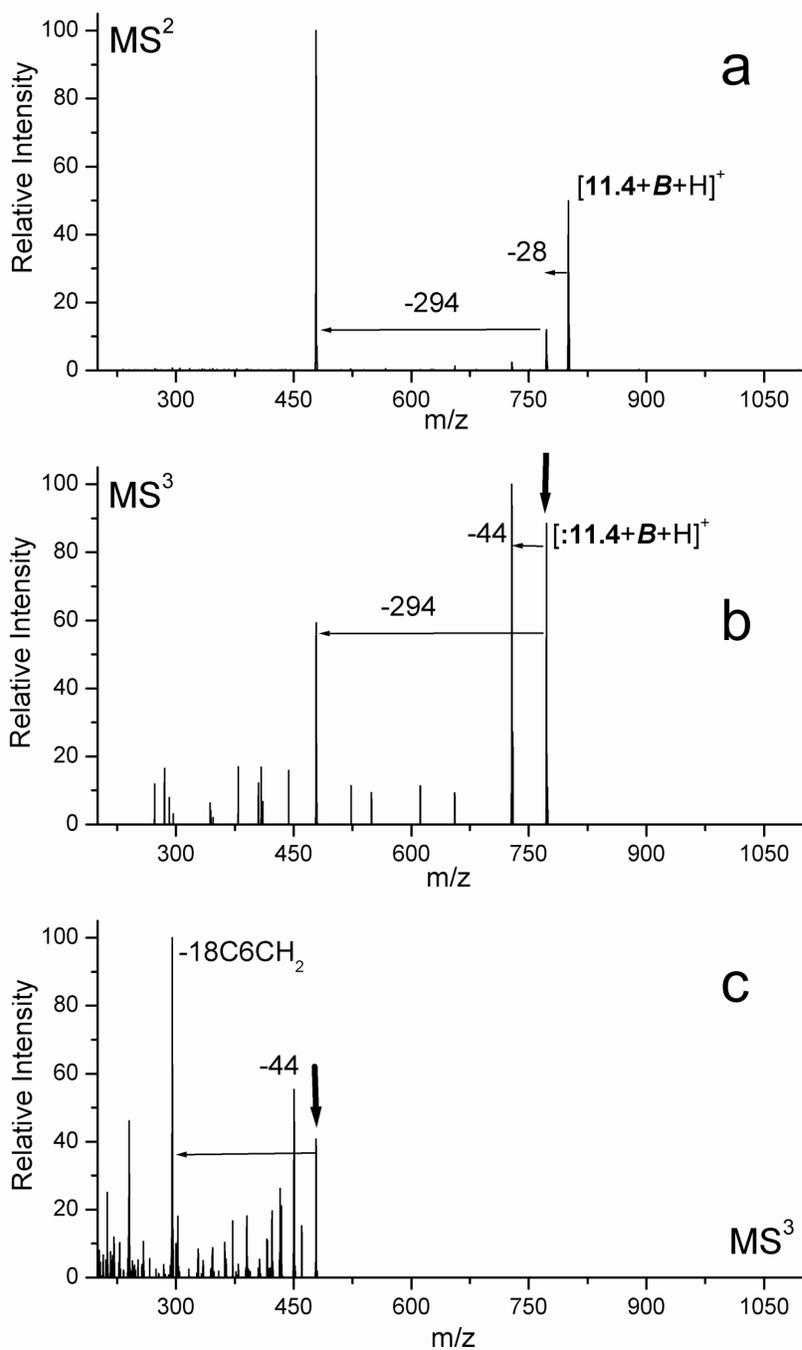


Figure 11.4 a) MS² on [4+B+H]⁺ yields similar results to those for 1,6-diaminohexane. b) MS³ spectrum is notably different, suggesting that peaks produced in (a) occur competitively. c) Further excitation of the complex does not result in any dissociation of the **B**. Bold downward arrows indicate peaks being subjected to CAD.

Reagent **11.5** contains only a single crown ether connected to a diazo functional group, with an ethyl ester connected to the side opposite 18C6. The results for complexing allylamine (**C**) and 1,4-diaminobutane (**D**) with reagent **11.5** are given in Figure 11.5. Collisional activation of the complex $[\mathbf{11.5}+\mathbf{C}+\mathbf{H}]^+$ results primarily in the loss of N_2 (Figure 11.5a). Further excitation of the product peak yields the loss of neutral EtOH and a multitude of other peaks in Figure 11.5b. However, dissociation of **C** is not observed, suggesting that covalent attachment has been achieved. Carbene insertion into double bonds is a well documented phenomena in solution and is the most likely explanation for the results observed here.¹⁹

Experiments with **D** and **11.5** yield results similar to those obtained with host **11.4** and protonated 1,6-diaminohexane except that the loss of EtOH is observed in addition to the loss of $18\text{C}6\text{CH}_2\text{OH}$. In Figure 11.5c it is shown that the loss of EtOH is approximately twice as abundant as the loss of $18\text{C}6\text{CH}_2\text{OH}$. This is consistent with the proposed reaction mechanisms. In Figure 11.5d, a fragment that contains no 18C6 is subjected to CAD. **D** (mass 88 Da) does not dissociate from the complex. Since there is no crown ether present to bind to a primary amine, the data in Figure 11.5d offers compelling evidence that indeed what was once a noncovalent complex is now a molecule.

All of the data obtained by reactions with small molecules suggests that covalent attachment occurs rapidly and almost exclusively when the complex containing **11.4** or **11.5** is subjected to CAD. The corresponding carbenes (**:11.4** and **:11.5**) can undergo insertion reactions with a wide variety of different functional groups. The appropriate next step is to see whether these reagents can covalently attach to peptides themselves.

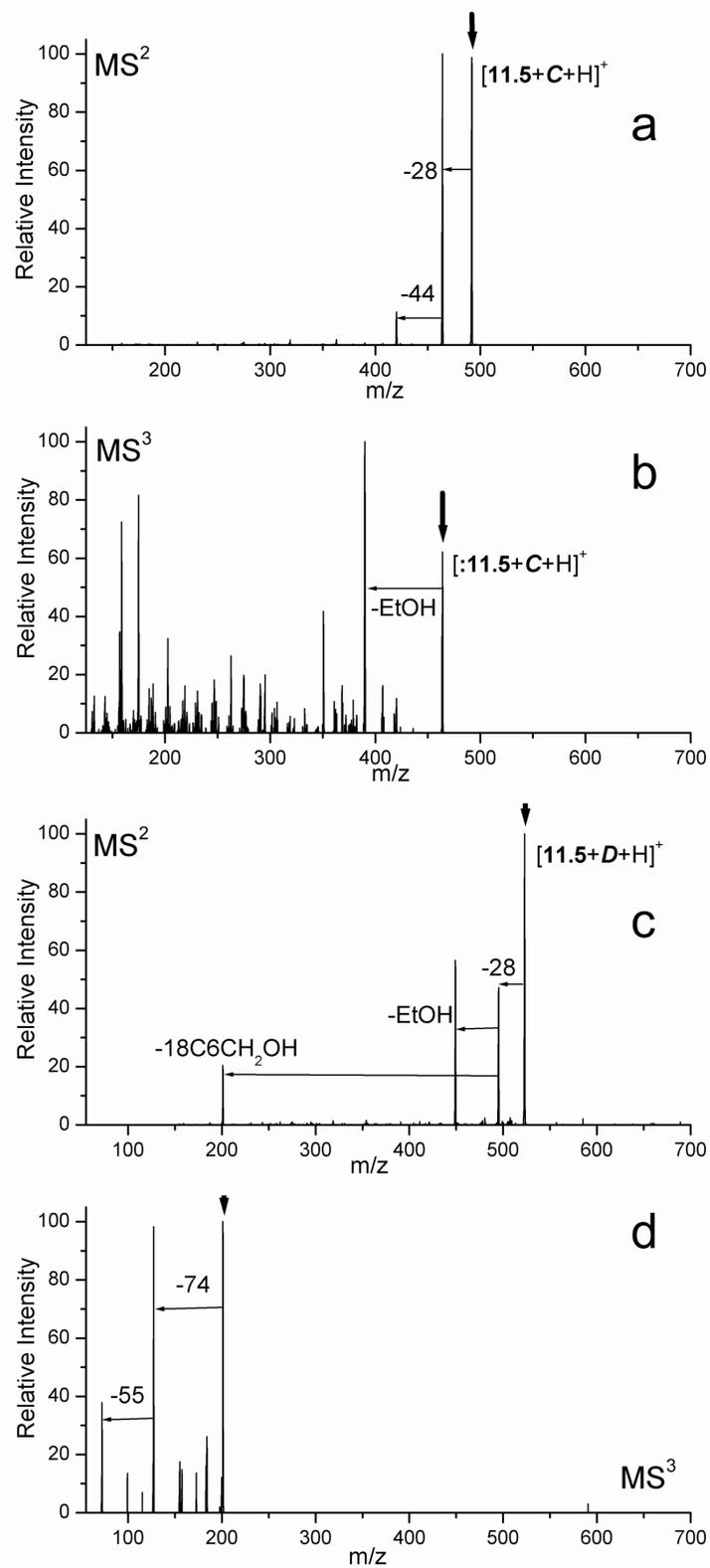


Figure 11.5 Caption on next page.

Figure 11.5 Experimental results for host **2**. a) MS² on complex with allylamine (**C**) results in the loss of N₂. b) MS³ spectrum reveals many fragmentation pathways, none of which lead to dissociation of the guest. c) CAD spectrum of 1,4-diaminobutane (**D**) and **11.2** loses N₂, EtOH, and 18C6CH₂OH. d) MS³ spectrum on peak containing no 18C6 ring fragments rather than dissociating, offering compelling evidence for covalent attachment between the host and guest. Bold downward arrows indicate peaks being subjected to CAD.

Peptides. Reagent **11.4** is designed to bind to peptides containing two lysine residues. The ESI spectrum for **11.4** complexed with the simple peptide KGK is shown in Figure 11.6a. Abundant adduct peaks are observed, indicating excellent recognition. In Figure 11.6b, the [**11.4**+KGK+2H]²⁺ peak is subjected to collisional activation. The loss of N₂ leads to the base peak in the spectrum, with an additional loss of 294 Da being observed as well. No dissociation is observed, suggesting that the appropriate combination of high binding energy and low activation barriers has been achieved for reagent **11.4**. Further collisional activation in Figure 11.6c does not lead to any dissociation of KGK, again confirming that an intermolecular reaction has occurred. Very similar results are obtained for other peptides containing two lysines in close proximity, such as INLKAIAALVKKVL, AAKRCAA, and KK. If the singly charged [**11.4**+KGK+H]⁺ complex is subjected to CAD, then a neutral loss of **11.4** yields the only observed product. This appears to suggest that two crown ethers are necessary to achieve sufficient binding energy for the intermolecular reaction to occur. However, it will be demonstrated below that this is not the case.

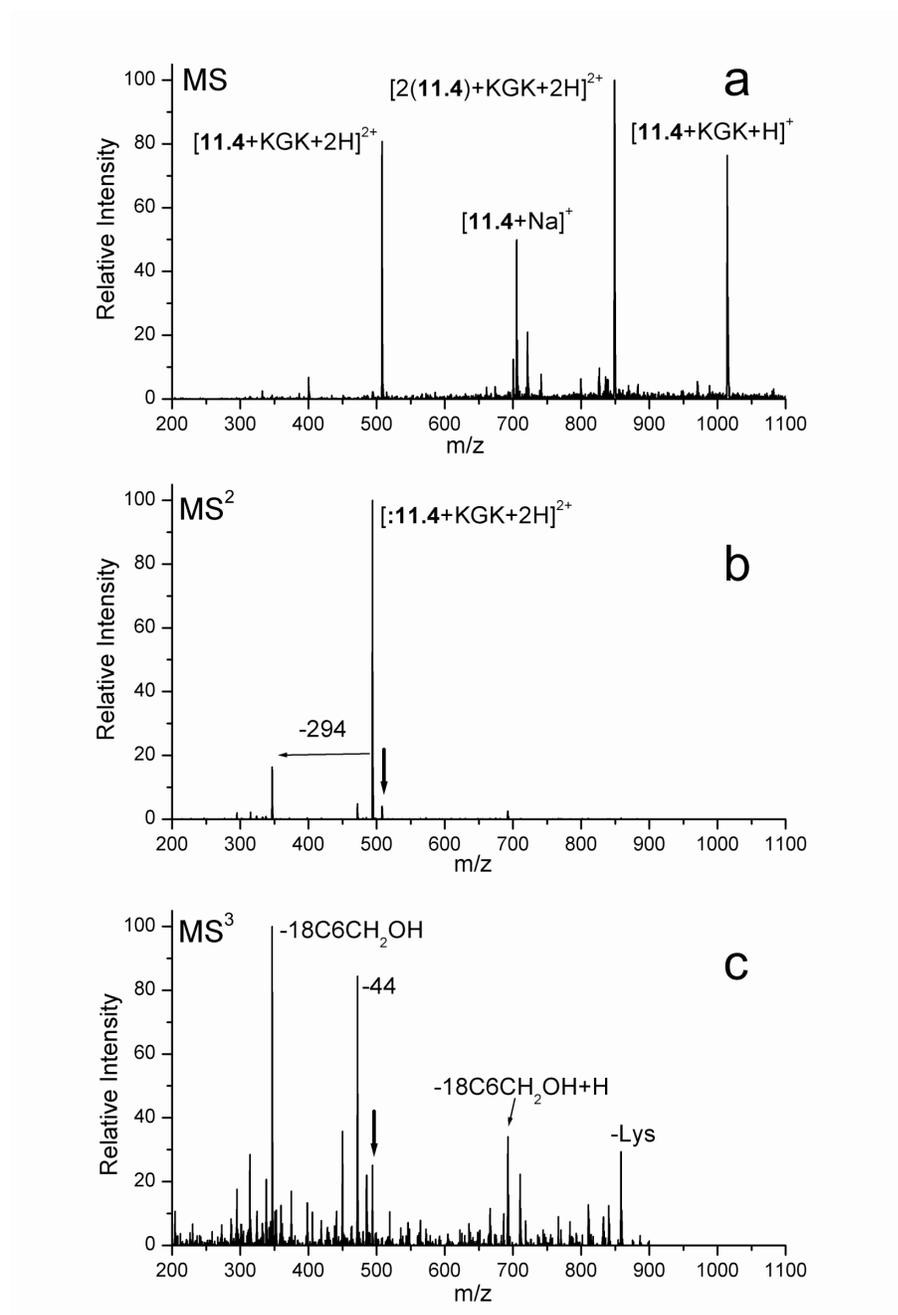


Figure 11.6 (a) ESI-MS of KGK and 11.4 demonstrating abundant noncovalent complex formation. (b) MS² on $[11.4+KGK+2H]^{2+}$ yields the expected loss of N₂. (c) Further CAD reveals that the peptide has been trapped by the molecular mousetrap and the two molecules are now covalently attached. Bold downward arrows indicate peaks being subjected to CAD.

Reagent **11.5** only contains a single 18C6 and will therefore only bind to a single lysine, reducing the overall binding energy relative to **11.4**. The spectrum in Figure 11.7a shows that **11.5** forms abundant noncovalent complexes with KGK. Isolation and collisional activation of the doubly charged complex $[\mathbf{11.5}+\text{KGK}+2\text{H}]^{2+}$ results exclusively in the loss of N_2 , generating the reactive carbene in Figure 11.7b. Reisolation and further activation of this peak in Figure 11.7c yields fragments corresponding to the loss of the C-terminal and N-terminal lysine residues and the loss of 17 Da (presumably NH_3). Simple dissociation is not observed, indicating covalent attachment through an intermolecular reaction occurred. The loss of lysine from both termini of the peptide suggests that either attachment of the crown is not selective for one lysine over another, or that the insertion reaction is not selective, or both.

CAD of the singly charged $[\mathbf{11.5}+\text{KGK}+\text{H}]^+$ complex again results in loss of the neutral mousetrap **11.5** exclusively as seen in Figure 11.7d. The exact cause for this interesting behavior is not known, but the results can be explained by at least two possibilities. Either the binding energy of the complex is enhanced by the addition of a second proton, or the absence of the second proton enables a lower energy dissociation pathway. Regardless of the cause, it is observed in general that complexes with higher charge states tend to favor intermolecular reactions, while lower charge state complexes tend to favor simple dissociation. In very similar reactions to those shown in Figures 11.5b and 11.5c, reagent **11.5** has been covalently attached to many peptides including: INLKAIAALVKKVL, AAKRCAA, KPPGFSPFR, GGK, and GGKAA.

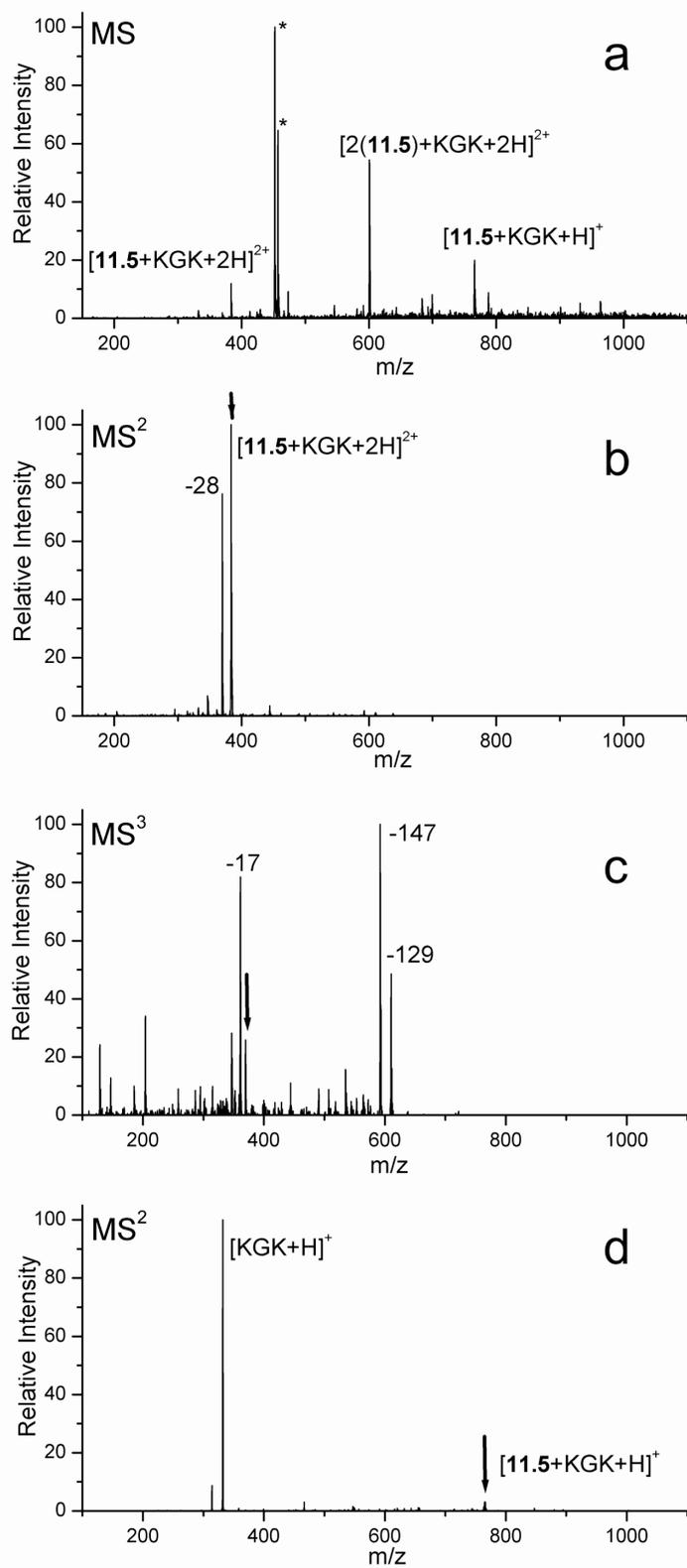


Figure 11.7 Caption on next page.

Figure 11.7 (a) ESI-MS of KGK and **11.5** demonstrating abundant noncovalent complex formation. (b) MS² on [**11.5**+KGK+2H]²⁺ yields the expected loss of N₂. (c) Further CAD reveals that the peptide has been trapped by the molecular mousetrap and the two molecules are now covalently attached. (d) MS² on the singly charged complex results in simple dissociation. Bold arrows indicate peaks being subjected to CAD.

11.4 Conclusion

These experiments demonstrate that development of biomimetic reagents capable of directing peptide chemistry in the gas phase is possible. The first successful examples of such reagents have been given. The search for a reagent that selectively cleaves peptides in the gas phase is still ongoing, but we have shown that with the proper combination of high binding energy and low reaction barriers, it is possible to initiate intermolecular reactions in noncovalent complexes with peptides. Furthermore, it is shown that sufficient binding energy to favor peptide cleavage over complex dissociation can be achieved with two 18C6 ethers attached to two lysines. Molecular mousetraps capable of covalently attaching to any lysine containing peptide are presented herein. This type of molecule represents the first step towards the development of gas phase cross-linking reagents. It is anticipated that the knowledge acquired from these initial results will allow for the development of other reagents capable of initiating controlled peptide chemistry in the gas phase. Although in the present study we have only considered the activation of these adducts to initiate covalent attachment in the gas phase, it should also be possible to effect similar chemistry in solution, with carbene formation initiated by either photochemical or metal catalyzed processes.²⁰

11.5 References

¹ Vihinen, M. *Biomol. Eng.* **2001**, *18*, 241-248.

² (a) Tsaprailis, G.; Arpad, S.; Nikolaev, E. N.; Wysocki, V. H. *Int. J. Mass Spectrom.* **2000**, *195/196*, 467-479. (b) Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. *J. Am. Chem. Soc.* **1999**, *121*, 5142-5154.

³ Lee, S. -W.; Kim, S. K.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1998**, *120*, 3188-3195

⁴ Lin, T.; Glish, G. L. *Anal. Chem.* **1998**, *70*, 5162-5165.

⁵ Hu, P.; Loo, J. A. *J. Am. Chem. Soc.* **1995**, *117*, 11314-11319.

⁶ Nemirovskiy, O. V.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1285-1292.

⁷ Some reagents have been developed specifically for the gas phase: (a) Friess S.D.; Zenobi R. *J. Am. Soc. Mass Spectrom.* **2001**, *12(7)*, 810. (b) Julian, R. R.; Akin, M.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L. *Int. J. Mass Spectrom.* **2002**, *220*, 87-96.

⁸ Julian, R. R.; Beauchamp, J. L. *Int. J. Mass. Spectrom.* **2001**, *210*, 613-623.

⁹ For solution phase reagents, please see: (a) Bell, T. W.; Khasanov, A. B.; Drew, M. G. B.; Filikov, A.; James, T. L. *Angew. Chem. Int. Ed.* **1999**, *38*, 2543. (b) Galan, A.; Andreu, D.; Echavarren, A. M.; Prados, P.; de Mendoza, J. *J. Am. Chem. Soc.* **1992**, *114*, 1511. (c) Ludwig, R. Fresen. *J. Anal. Chem.* **2000**, *367*, 103. (d) Ngola S. M.; Kearney P. C.; Mecozzi S.; Russell K.; Dougherty D. A. *J. Am. Chem. Soc.* **1999**, *121*, 1192. (e) Rensing, S.; Arendt, A.; Springer, A.; Grawe, T.; Schrader, T. *J. Org. Chem.* **2001**, *66*, 5814, (f) Schrader, T. H. *Tetr. Lett.* **1998**, *39*, 517.

¹⁰ (a) Schalley, C. A. *Mass Spectrom. Rev.* **2001**, *20*, 253-309. (b) Smith, R. D.; Bruce, J. E.; Wu Q. Y.; Lei, Q. P. *Chem. Soc. Rev.* **1997**, *26*, 191-202; (c) Veenstra, T. D. *Biophys. Chem.* **1999**, *79*, 63-79. (d) Loo, J. A. *Int. J. Mass Spectrom.* **2000**, *200*, 175-186.

¹¹ (a) Schwartz, B. L.; Light-Wahl, K. J.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 201-204. (b) Nemirovskiy, O. V.; Ramanathan, R.; Gross, M. L. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 809-812. (c) Brodbelt, J. S. *Int. J. Mass Spectrom.* **2000**, *200*, 57-69. (d) Eckart, K.; Spiess, J. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 912-919.

¹² (a) Bradshaw, J. S.; Izatt, R. M.; Borkunov, A. V.; Zhu, C. Y.; Hathaway, J. K. *Comprehensive Supramolecular Chemistry*, Vol. 1; G. W. Gokel, Pergamon/Elsevier: Oxford, 1996; 35-95. (b) Maleknia, S.; Brodbelt, J. *J. Am. Chem. Soc.* **1993**, *115*, 2837-2843.

¹³ Julian, R. R.; Beauchamp, J. L. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 493-498.

¹⁴ Julian, R. R.; May, J. A.; Stoltz, B. M.; Beauchamp, J. L. *Angew. Chem. Int. Ed.* **2003**, *42(9)*, 1012-1015.

¹⁵ a) Chandler, C. J.; Deady, L. W.; Reiss, J. A. *J. Heterocyclic Chem.* **1981**, *18*, 599. b) Weijnen, J. G. J.; Koudijs, A.; Schellekens, G. A.; Engbersen, J. F. J. *J. Chem Soc. Perkin Trans. 2* **1992**, 830.

¹⁶ Steen, H.; Jensen, O. N. *Mass Spectrom. Rev.* **2002**, *21*, 163-182.

¹⁷ The insertion reaction only occurs when the H-C-H bond is presented symmetrically to the carbene, suggesting a minimal barrier may exist. Higher level DFT calculations at the at the B3LYP/CCPVTZ(-F)⁺ level on **:11.6** yield a singlet ground state with a singlet/triplet splitting of 3±1 kcal/mol suggesting that these reactions may proceed through the singlet state.

¹⁸ Toscano, J. P.; Platz, M. S.; Nikolaev, V.; Popic, V. *J. Am. Chem. Soc.* **1994**, *116*, 8146-8151.

¹⁹ Moss, R. A.; Jones, M., Jr. Eds. *Carbenes, Vols. 1 and 2*; Wiley: New York, 1973, 1975.

²⁰ (a) Doyle, M. P.; McKervey, M. A.; Ye, T. *Modern Catalytic Methods for Organic Synthesis with Diazo Compounds*; Wiley-Interscience: New York, 1998. (b) Moody, C. J.; Whitham, G. H. *Reactive Intermediates*; Oxford University Press: New York, 1992; 26-50.