OPTIMIZATION OF LENNARD-JONES POTENTIAL PARAMETERS AND BENCHMARK COMPARISON BETWEEN ION MOBILITY CALCULATORS IN FREE MOLECULAR REGIME

A Thesis

Submitted to the Faculty

of

Purdue University

by

Vaibhav R. Shrivastav

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Mechanical Engineering

August 2017

Purdue University

Indianapolis, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Carlos Larriba-Andaluz, Chair

Department of Mechanical Engineering

Dr. Whitney Yu

Department of Mechanical Engineering

Dr. Andres Tovar

Department of Mechanical Engineering

Approved by:

Dr. Sohel Anwar

Chair of the Graduate Program

To my Mom, Dad, and Dada.

ACKNOWLEDGMENTS

I would like to express my gratitude towards my thesis advisor Dr. Carlos Larriba-Andaluz for his unparalleled guidance, immense knowledge, and support towards the completion of my thesis. I want to thank him for having the trust and belief towards me and helping me out at each step. My sincere thanks goes to MD Minal Nahin, Joseph Derrick, and Tianyang Wu for their precious support and help throughout my research work and leisure time, and making this journey memorable. I would also like to acknowledge my friend, Shruti Rawool, who helped me throughout the documentation and review of my thesis. Without her passionate participation and invaluable inputs, the validation could not have been successfully conducted. I am gratefully indebted towards my parents for inspiring me to aim higher and try harder throughout my graduate studies and my thesis research. This accomplishment would not have been possible without their encouragement. Thank you.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

ABSTRACT

Shrivastav, Vaibhav R. M.S.M.E, Purdue University, August 2017. Optimization Of Lennard-Jones Potential Parameters And Benchmark Comparison Between Ion Mobility Calculators In Free Molecular Regime. Major Professor: Carlos Larriba-Andaluz.

Ion Mobility Spectrometry (IMS) is a widely used technique to differentiate charged particles in the gas phase. Although there has been a significant computational development over the past few decades for calculating Ion Mobility and Collision Cross Section (CCS), still there is a need to develop it in terms of its efficiency and performance, to better understand the dynamics of the collision. The work presented here demonstrates the efficiency and performance of newly developed mobility calculator: IMoS. The results were compared to MOBCAL and were found to be in a good agreement for He and N_2 for the same input parameter. IMoS, which has an ability to be parallelized, gave similar values for CCS (within 1% of error) with a speed of two order of magnitude, which is higher than that of MOBCAL. Various options of approximations such as Diffused Trajectory Methods (DHSS, TDHSS) with and without partial charges and Projected Area approximation were considered in this work which lead us to reduce the total computational time required for the calculations. A careful computational study was carried out for 47 organic molecules and few large biomolecules ($>$ 10000 atoms) to demonstrate the similarity and differences in two widely used mobility calculator IMoS and MOBCAL. As the calculations were made faster using IMoS, it was a necessary step to develop an optimization algorithm in order to optimize the Lennard-Jones potential parameters for gas phase calculations used in the Trajectory Method. The process of optimization follows a multiple iterative path, wherein the parameters are completely optimized for all the given elements. A surface plot was generated using tens and thousands of data points for C, H, N,

O, and F to study the relationship between epsilon (ϵ) and sigma (σ) for each element in the N_2 buffer gas. The function (F) used here is a function of experimental CCS and IMoS generated CCS, which was minimized in the process of optimization. These optimized values can be used in the mobility calculator for calculating accurate Collision Cross Sectional values.

1. INTRODUCTION

1.1 Ion Mobility Spectrometry (IMS)

Ion Mobility Spectrometry (IMS) sometimes referred to as plasma chromatography / ion chromatography is becoming a ubiquitous technique and has exponentially grown over the past few decades in the fields of Aerosol Science, Analytical Chemistry, Medical Science, Drug Delivery, Material science, Petroleomics [78,79] and Food Industry [1-9] This technique is being used for various purposes from detection of chemical warfare agents to studying the characteristics of the particle for polymers, macromolecules, nanoparticles, biomolecules, aerosol particles, etc., based on the separation ability and ion transmission efficiency of the instrument. The IMS experimental analysis is utilized to elucidate the theoretical study of ion structures. The behavior of the charged entities in the gas phase through different bath gases (monoatomic or diatomic) yield true values for mobility and Collision Cross Section (CCS) that are necessary to evaluate and establish the re-emission rules for ion-gas molecule collision.[4] This helps in numerical development of the kinetic theory of gases. Ion mobility instruments are cable of differentiating the ion based on their shape and size, which is responsible for defining the mobility of the ion structure in the bath gas. IMS can also be used as a standalone instrument due to its portability and versatility.[22-30]A Mass Spectrometer (MS), which is a highly sensitive and important instrument having high mass resolution, is taken into consideration for differentiating the ion structure based on their mass and charge. [10-21] People over the years have used the combination of IMS-MS to study the structures. It gives multidimensional analysis, which is cost-effective and gives high resolution to study the characteristic of molecules in gas phase.[8] The buffer/carrier gas in the instrument opposes the movement of the ions based on their shape and size. If the

size of the structure is spherical, it will experience less resistance compared to the structure with a straight chain or planar geometry.Hence the average electrical mobility (K) of a structure is separated based on the drift given by the following relation:

$$\mathrm{K}\vec{E} = \vec{V}_{drift} \qquad (1)$$

Based on the study of Michael A. Ewing et. Al [12], the IMS instrument can be subcategorized into two classes: the first class is a nested technique to get a complete spectra in a single separation. This class involves the instrument such as dispersive DTIMS, Traveling Wave Ion Mobility Spectrometry (TWIMS), and Trapped Ion Mobility Spectrometry (TIMS). The second class of separation instrument is the one which is capable of separating the ions for selective mobility and obtains the complete spectrum of the ion. This class involves selective DTIMS, Differential Mobility Analysis (DMA), Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS) also called Differential Mobility Spectrometry (DMS), Overtone Mobility Spectrometry (OMS), circular ion mobility spectrometers, and Transversal Modulation Ion Mobility Spectrometry (TM-IMS). All the above instruments help to increase the separation ability of ions based on their shape and size and to detect the geometry of the structure. Resolving power is the parameter which defines the suitability of the IMS instrument. It can be defined as the ratio of drift time to the full-width at half maxima of the peaks for the structures. Generally, for low-resolution instruments, the resolution lies between 5-60, whereas for high-resolution instruments the resolution lies between 200-1000. [12]

1.2 Collision Cross Section (CCS)

In general, the mobility measurements are related to the CCS. If we have a compact globular structure, it will move fast and have higher mobilities in the buffer gas compared to elongated and open structures. Thus it is important to calculate the collision cross sectional area.

A more useful expression given by Mason-Schamp, for mobility in the free molecular regime is:

$$Z = \left(\frac{\pi m_{red}}{8kT}\right)^{\frac{1}{2}} \frac{3Ze}{4\rho_{gas}\Omega} \qquad (2)$$

Where k is the Boltzmanns constant, T is the background gas temperature, m_{red} is the reduced mass of the particle, ρ_{gas} is the gas mass density, and Ω is the Collision Cross Section. Numerically, there are three ways to calculate the Collision Cross Section and are explained as follows in the order of increasing complexity.

1.2.1 The Projected Area Method

An in-depth knowledge about the cross section of the structure is helpful to deduce the characteristic of the molecule. The first method is known as Projected Area or the PA method. In simple terms, it is the area of the shadow, hence called projection, which is casted by the molecule on a plane. So, the diffusion of the class A molecule in the gas molecule of class B not only depends on the velocity, but also on the size and shape of the molecule of class A. This is responsible for the magnitude of perturbation through buffer gas molecules of class B. In general, it will conclude and confirm the shape and size of the molecule. In order to do so, computational methods are developed, which will calculate the orientationally averaged projection area. The two methods that are generally employed are Monte Carlo simulation and Traveller Salesman Approximation (TSA). This calculation is achieved by first establishing the hard sphere radii to both the structure of interest and the gas molecules that are suppose to collide. This establishes the domain, which encompasses the structure. The gas molecules are randomly thrown at the structure and the number of hits within the prescribed shadow are counted. Next, the total number of gas molecules thrown toward the domain are counted and the ratio of the number of hits inside the shadow to the total number of counts is multiplied by the area of the established domain, which gives the approximate area of the structure. The TSA method includes

capturing the circumference or the boundary of the structure in a different orientation and then averages it accordingly. Note that the below model is an all-atom model and may represent any kind of shape and size. The figure 1 represent only three orientations. In the case of a non-spherical surface, n number of orientation can be made and its average is calculated to find orientationally averaged Collision Cross Section of the structure. The idea is depicted in the figure 1.1:

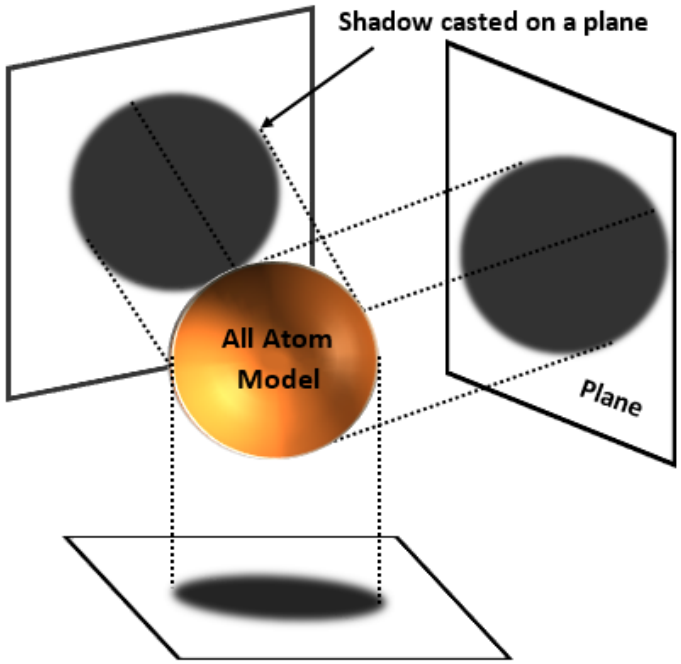

Fig. 1.1. Depiction of orientationally averaged Collision Cross Section using the PA method.

In prior research, orientationally average collision cross section was considered to be the first approximation of collision cross section. This would be the case under the following three conditions explained by Carlos Larriba et. al. as [39]:

- The surface of the structure must be completely convex to avoid multiple scattering effects.

- The rotational and vibrational energies should be negligible compared to that of translational energy for atoms of the structure and gas molecules.
- Long-range potential should not be considered for momentum transfer during the collision.

Keeping the above points under consideration, Carlos Larriba et. al. proposed the Collision Cross Section as:

$$\Omega = L\xi PA \qquad (3)$$

Where PA is orientationally average collision cross-sectional area, "L " is a dimensionless factor accounting for long range potentials / ion-induced dipole potential, and ξ is the dimensionless factor for the momentum transfer, which defines the collision and re-emission of the impinged gas molecules. Ideally, $\xi = 1.36$ [3] is obtained from Millikan data set from experiments. This additional parameter gives more accurate PA, and is true in real situations, where both long range potential as well as the momentum transfer will affect the CCS area because, ideally, no structure is completely convex and multiple scattering is commonly observed. This development certainly leads us to understand accurate physical dimensions of the molecule. Another important advantage is that this method is computationally inexpensive and is reliable for the first hand approximation. We can simulate for more number of orientations in order to get better accuracy in terms of the Collision Cross Section. There are other existing methods such as Projection Superposition Area (PSA) [51] that calculate the projected area but are not widely used.

1.2.2 The Hard Sphere Scattering Method

As the name implies, this method assumes the infinite hard wall potential between the colliding bodies taking into account the scattering effect. It means the impinging gas molecules have to re-emit as soon as it collides with the structure and cannot penetrate the molecule. Two types of hard sphere methods are considered, namely

Elastic Hard Sphere Scattering (EHSS) and Diffused Hard Sphere Scattering (DHSS). Both the methods account for multiple scattering, which is an obvious phenomenon for most of the molecules, especially for the ones with a concave surface structure.

- **Elastic Hard Sphere Scattering (EHSS):** The collision, in this case, happens to be specular and elastic.[41] It means the angle of reflection is the same as that of the angle of incident. The momentum transfer takes place between the frozen structure and a moving gas molecule. This case is true for a monoatomic gas molecule such as Helium but is not true in the case of a diatomic gas molecule such as Nitrogen or Air [38]. There is a slight enlargement in terms of CCS over the Projected Area method. This is caused only by "scattering" (multiple collision events).[4] Though this model is widely used and is simple in calculations, the phenomenon of elastic and specular events are not true for all ions in nature. Hence it results into development of more realistic methods for CCS measurement known as DHSS, which talks about energy exchange and its extent.

- **Diffused Hard Sphere Scattering (DHSS):** Unlike the EHSS method, the DHSS method takes into account the vibrational and rotational energy between the gas molecule and the structure of interest. This method takes into account the diffusive nature, (the incident angle may not be the same as that of the reflecting angle) and inelastic nature of the gas molecule-structure collision. This type of collision increases the value for CCS over the PA method by scattering effect as well as by its diffusing and inelastic nature, which is in good agreement with that of the experimental results. Both the methods are explained in detail in this section taking into consideration the algorithm used to calculate CCS. The following figure depicts the idea of EHSS and DHSS method:

The figure 1.2 demonstrates the probability of reflection for monoatomic and diatomic gas molecule. In the case of a monatomic gas molecule (He), the collision is specular

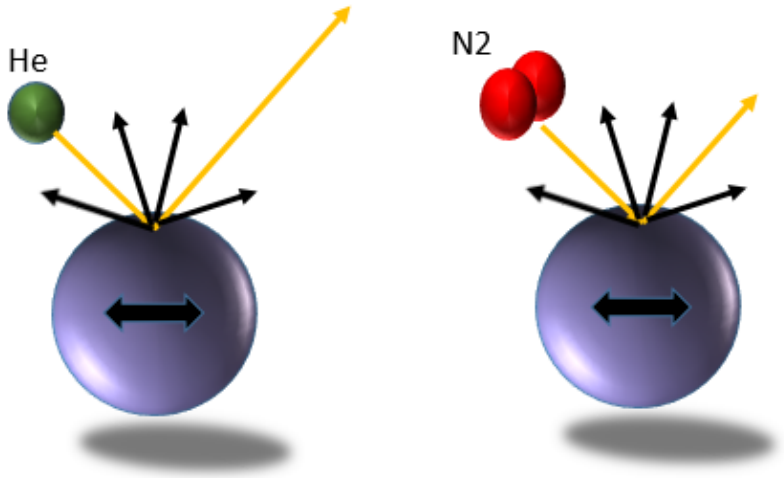

Fig. 1.2. All atom model representing the EHSS and DHSS concept for monoatomic (He) and diatomic gas molecules (N_2).

showing a long arrow for equal reflection and incident angles whereas in a diatomic gas molecule (N_2) the collision is majorly diffused.

1.2.3 The Trajectory Method:

The trajectory method is a rigorous approach for calculating CCS. It takes into account the effects of both long range and short range potential parameters. The combination of these potential parameters can be related as follows: [40, 71]

$$\phi(x,y,z) = 4\epsilon \sum_{i=1}^{n} \left[\left(\frac{\sigma}{r_i} \right)^{12} - \left(\frac{\sigma}{r_i} \right)^{6} \right] - \frac{\alpha}{2} \left(\frac{Ze}{n} \right)^{2} \left[\sum_{i=1}^{n} \left(\frac{x_i}{r_i^3} \right) + \sum_{i=1}^{n} \left(\frac{y_i}{r_i^3} \right) + \sum_{i=1}^{n} \left(\frac{z_i}{r_i^3} \right) \right] \quad (4)$$

Where the first term is the Lennard-Jones 6-12 potential parameter and the second term is ion-induced dipole potential interaction. In equation 4, ϵ and σ are the Lennard-Jones potential parameter. ϵ represents the well depth of the potential curve in meV and defines the depth of the curve responsible for the attraction of two colliding bodies. σ represents the distance in Angstrom unit where the potential value reaches positive. Once the distance R (distance between centers of two bodies)

becomes less than the value of σ, the strong repulsive force is exerted on a close approach of colliding bodies. α represents gas polarizability, x_i,y_i,z_i,and r_i are the relative distances between the atoms and the gas molecule. Though hard sphere scattering is a good approximation, the Trajectory method leads to more exact results for concave surfaces. The CCS calculated from the Trajectory method provides accurate structural assignment of the structure over the Hard Sphere Scattering method when the structure is heavier and more polarizable. The work presented here is about optimizing the Lennard-Jones potential parameter (ϵ and σ) to better understand the effect of these parameters in CCS and mobility calculations.

1.3 Necessity of Numerical Development

There have been a lot of analytical attempts to calculate CCS [43,46-50,52], but due to inherent difficulties and obsequious effects, numerical methods have found its root in approximating such kind of problems. Despite new advancements in the field of experimentation, including a high range of transmission and higher resolution [26, 31-34], there have not been a parallel improvements in the field of computational calculation. Three major reasons for such impedances are as follows:

- Longer time is required for calculations and is difficult to get accurate results that would match the experimental values
- Lack of well established parameters such as Lennard-Jones potential parameter required for gas phase calculation, which can be established by experimental studies
- Lack of reliable methods to perform gas phase molecular dynamics

Presently, many of the existing forcefields that are optimized are used in solvent and are not accurate for gas phase calculations. In our work, we tackled the first two problems. Efficiency and performance were considered and compared to other computational methods which in turn leads us to the conclusion of better time efficiency,

reduced computational cost, and prominent reliable methods used to calculate Collision Cross Section. The in-house developed algorithm IMoS, [35-39] and MOBCAL [40,41] are the two widely used programs for calculating the mobility and CCS in the free molecular regime. In spite of yielding similar CCS result, the work presented here shows that the efficiency and performance of IMoS are few order of magnitude higher than that of MOBCAL. Secondly, we tried to optimize the Lennard-Jones potential parameter, using a MATLAB algorithm. The function used, which is the function of CCS and Lennard-Jones potential parameter, is pushed to a minimum in order to get the optimized value. Tens and thousands of calculations are performed for plotting a surface containing the minimum point for the given function.

2. EXPERIMENTS ON IMS-MS

2.1 Mass Spectrometry Working Principle

Mass Spectrometer is a complex instrument that combines ionization chamber, a mass analyzer, and a detector to identify the mass over charge ratio of the structure of interest. The flow chart in figure 2.1 explains the process involved in Mass Spectrometry.

2.1.1 Quadrupole

This device uses the oscillating electric field to separate the ions based on the m/z using the stability of their trajectories in the electric field. This quadrupole is made of four perfectly parallel rods shown in the figure 2.2, leaving a space in between four rods for ions to travel. A set of two symmetrically opposite rods are of the same charge, so one pair is positively charged and the other is negatively charged. The polarity keeps on oscillating so as to center the ions throughout its journey and also segregating based on the m/z ratio. If the particle is positively charged, it will try to run towards the negative electrode and before it gets discharged to the electrode we change the polarity. This helps the ion to center and also separate at the same time. People use multiple quadrupoles for better transmission efficiency. The function of the quadrupole is varied accordingly. The collision gas is introduced at some point in the second quadrupole that is responsible for further fragmentation of the ions. A third quadrupole is designed for selecting the specific ions which are supposed to be analyzed. Quadrupole can be arranged in n multiple ways for various analysis and scan.

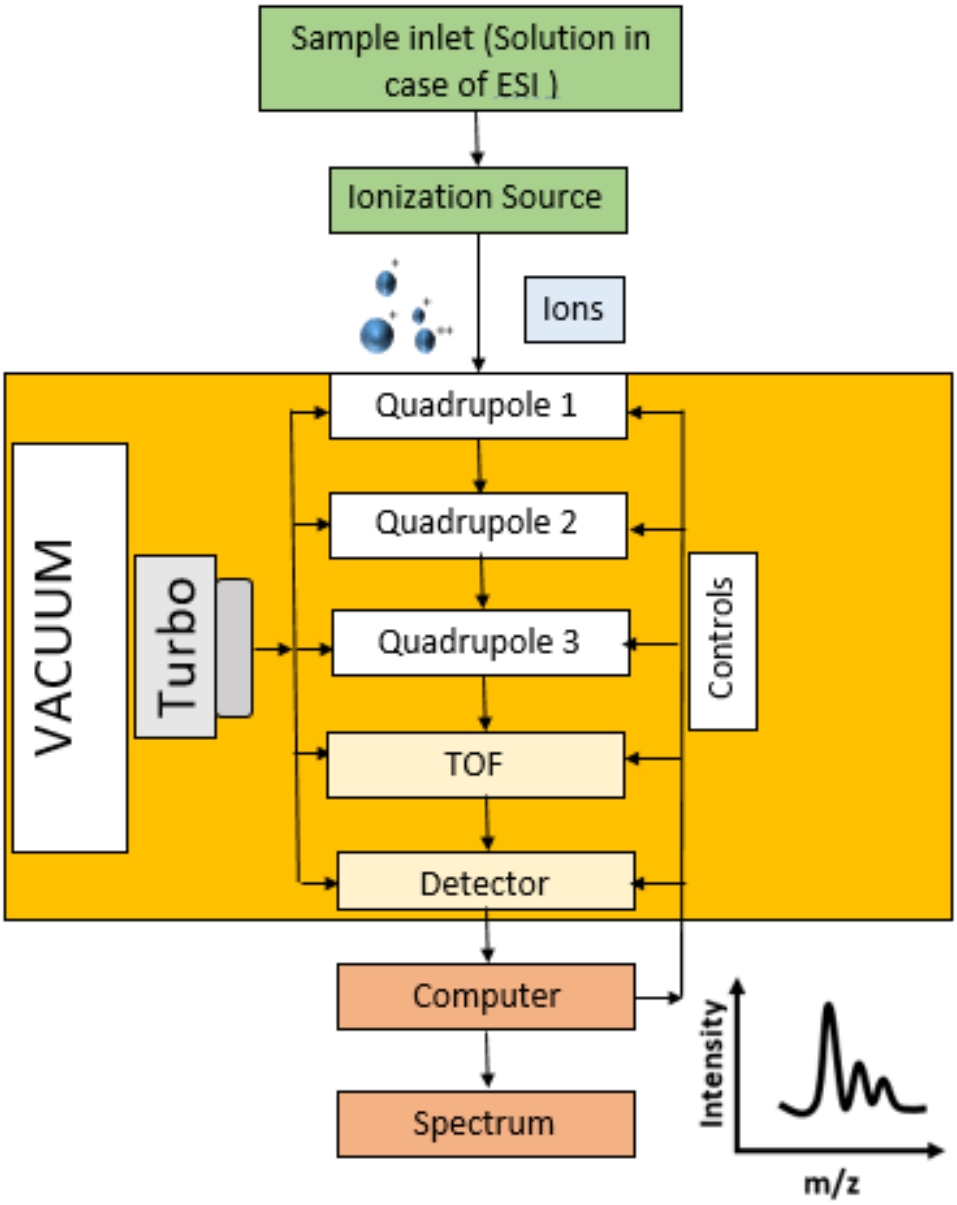

Fig. 2.1. Flow chart of working principle for a Mass Spectrometer system.

2.1.2 Time Of Flight (TOF)

The time of flight chamber is responsible for segregating ions based on their mass to charge ratio. The charged ions once inside the TOF are pushed at the same time and at the same voltage. The ions with less mass to charge ratio will travel

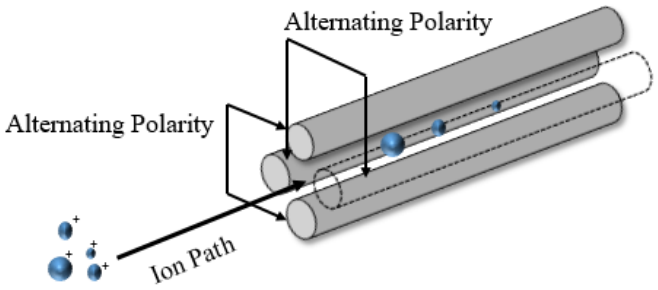

Fig. 2.2. Quadrupole setup

quicker than the ions with higher mass to charge ratio. So the time it takes to travel throughout the TOF chamber based on their respective masses is taken into account. This principle provides an accurate measurement of the mass of different time scale in a well-calibrated instrument. Another important phenomenon is the separation of ions based on their isotopes, for example, in the case of carbon.

2.2 Electrospray Ionization (ESI)

Ionizationization is an important part for studying the structures through IMS and Mass Spectrometry instruments. There are various methods to ionize the particle such as Electron Ionization, Chemical Ionization, Field Ionization, Matrix-Assisted Laser Desorption Ionization, Thermospray, Atmospheric Pressure Ionization, Electrospray Ionization, Atmospheric Pressure Chemical Ionization, Atmospheric Pressure Photoionization etc. In our experiments, we use the ESI technique for ionization of chemical structures. This technique developed by Fenn et. al. [82,83], is used to produce multiple charged structures for the given solution. This technique provides high sensitivity and flexibility to couple with IMS-MS instruments. ESI can be produced by applying a strong electrical field to the liquid (structure of interest + buffer) under atmospheric pressure. The electric field is obtained by applying high

potential of about 4-6 kV between the capillary carrying the liquid and the opposite electrode. The distance between the capillary end and the plate (for DMA acting as an electrode) is about 1 cm. The capillary is sharpened and narrowed up to 70-80 micrometers at one end from where the liquid is sprayed out towards the MS system to lower the onset voltage and electrospray easily. So, the field developed helps to accumulate the charge over the surface of the liquid and breaks carrying multiple charges over the droplet. Later, the solvent portion is also eliminated by counter flow leaving the structure with the charge to enter the system. To understand the phenomenon of formation of charged structure in the gas phase, it is important to understand the working principle of charges and surface tension of the liquid. Charge get accumulated at the tip of the capillary over the liquid surface and creates pressure over the liquid surface. At low voltage, the nature of the liquid droplet will be spherical. As the voltage is increased, the pressure due to charges over the liquid increases and make the spherical droplet to elongate. At onset voltage, which varies according to the solvent, the droplet overcomes the surface tension and forms a cone known as Taylor Cone. This initiates the spray of the smaller droplet with a charge on it. ESI is capable of forming multiple charges over the molecule. These ions formed by the electrochemical process are then supplied to Mass Spectrometer to measure the mass by charge (m/z) ratio. The figure 2.3, explains the production of ESI droplet. The produced charged droplet will reduce further in size due to solvent evaporation giving only charged structure into the system. The concentration of charge on the droplet will depend upon the size of the ions/structure which can be studied by Rayleigh equation, $q^2 = 8\pi^2\epsilon_0\gamma D^3$, where q is the charge, ϵ_0 is permittivity of the environment, γ is the surface tension, and D is the diameter of the droplet. One can also opt to study the negative ions as the electrospray is capable of producing both positively and negatively charged particles.

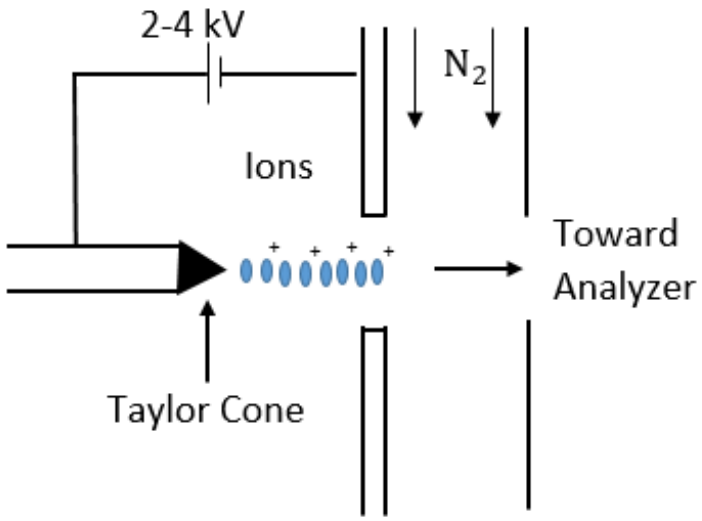

Fig. 2.3. Formation of Taylor Cone through the Electrospray Ionization technique.

2.3 Secondary Electrospray Ionization

Secondary Electrospray Ionization (SEI) is the technique used to generate the ions from separately spraying the charger and evaporating the substance from another chamber. The evaporation process can be carried out by supplying the gas which will help the substance to evaporate and is supplied to another chamber where it is charged by the charger such as electrospray of Acetic Acid, Ammonium Acetate etc. Heating the chamber containing the substance can be another possibility to completely evaporate the substance to the gas phase. This process holds a good application, where the particles can be collected directly from the environmental air with the help of a vent and supplied to the MS system for analysis. The greatest strength of this method is that it does not dissociate the compounds as easily as GC-MS since it does not require an increase in temperature. The figure 2.4 shows the set up for Secondary Electrospray Ionization.

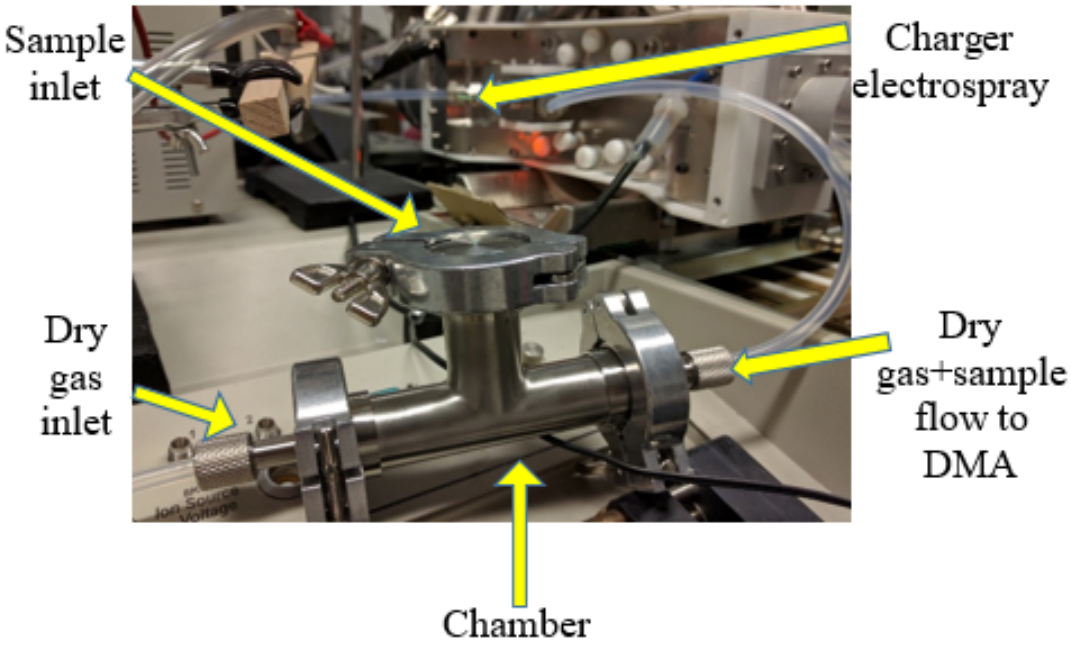

Fig. 2.4. Experimental setup for Secondary Electrospray Ionization.

2.4 Differential Mobility Analyzer (DMA)

A parallel plate Differential Mobility Analyzer is the Ion mobility Spectrometry instrument, which helps to segregate ions based on their mobility when passing through the buffer gas. Having its main application in the field of Aerosol science, DMA is also extensively used in molecular science in the different biomolecule analysis. The working principle of the DMA includes a chamber with sheath flow perpendicular to two parallel electrodes. Charged particles enter the through a slit in the below electrode and are driven by an electrical field E to the opposite electrode while being transported downstream by the sheath flow. A slit present in the upper electrode a distance L downstream of the first slit allows particles of a given mobility to be directed towards the mass spectrometer. This particular mobility can be changed by changing the difference in potential between the two plates and thus, in general, a ramping occurs so as to be able to capture all mobilities and form a complete spectra.

A schematic sketch in figure 2.5 represents a DMA system. Here two kinds of forces

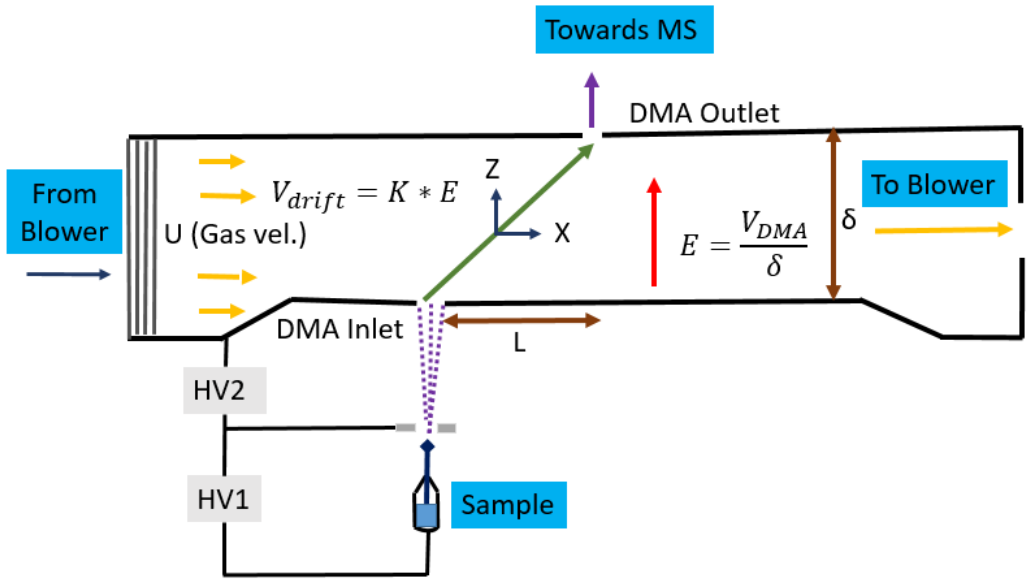

Fig. 2.5. Schematic diagram of working principle for the DMA system

act upon the charged particle: vertical electrical force and horizontal gas flow force, which can be equated through their passing time as follows:

$$t_H = \frac{L}{U} \qquad (5)$$

$$t_V = \frac{\delta}{\vec{V}_{drift}} = \frac{\delta}{Z_P * E} \qquad (6)$$

Equating equation (5) and (6)

$$\frac{L}{U} = \frac{\delta}{KE} = \delta * \frac{\delta}{Z_P * V_{DMA}}$$

The mobility of the ion transmitted from DMA inlet to outlet is given by the equation:

$$Z_P = \frac{U\delta^2}{LV_{DMA}} \qquad (7)$$

Due to DMAs high resolving power, better ion transmission ability, and flexibility in coupling with MS system, it has many promising applications in pharmaceutical, chemical, medical, bioprocess, clinical, and environmental sector.

2.5 Experimental Methodology

2.5.1 Sample Preparation

This is an important step towards producing the charged structure through the electrospray method. The different substance has different solubility and hence should be selected wisely. Nonpolar, semi-polar, and polar substance are widely used for Mass Spectrometry analysis. Hexane and toluene are generally used as the solvent for nonpolar substances. Electron Ionization and Chemical Ionization are generally preferred for the mass range of 10 to >1200 Daltons. Dichloromethane, chloroform, methanol, acetonitrile are the solvent used for semi-polar substances. EI, CI, MALDI and ESI are the techniques used for semi-polar molecules. Water and acid/base are the solvents that go easily with polar substances. ESI is the best techniques to use for the polar molecule as they can be easily electrosprayed.

2.5.2 Electrospray Setup

The prepared solution is electrosprayed from Polymicro silica capillary. This capillary has an inner diameter of 40 m and outer diammeter of 360 m. The solution can be stored in 1.5-2 ml polypropylene vial and is charged with acid if necessary. The vial is pressurized with the help of a syringe so as to allow the flow of solution into the capillary. A thin platinium wire is used for conducting the current through the soultion kept at arround 2-3 kV (this can be varied according to the requirement) above the front plate of DMA.

2.5.3 Method for Urine Analysis

We performed urine sample analysis in MD Sciex QSTAR triple quadrupole TOF Mass Spectrometer with DMA installed for mobility differentiation. The urine sample was analyzed with Acetic Acid and Ammonium Acetate as the charger for two differ-

ent experiments. We used both direct ESI and secondary ESI for analysis purposes and the working environment in mentioned below.

2.6 Direct Electrospray Ionization (ESI)

Urine in Ammonium Acetate

20 μL of the urine sample with 20 L of Ammonium Acetate (20 mM solution Ammonium Acetate). This was then added in 50/50 methanol and water to make a final 1 ml of solution. The figure 2.6 shows the m/z versus Intensity plot for the direct ESI with Ammonium Acetate as the charger. A bunch of data separated in their mobilities can be observed in this graph. The peaks demonstrate the signal obtained from different structures based on their mobility.

Urine in Acetic Acid Acetate

A solution of 20 μL of the urine sample with 30 L of Acetic Acid (30 mM solution of Acetic Acid) was prepared. This solution was then added in 50/50 methanol and water to make a final 1 mL of solution. The highest peak observed here is a trace for Tetrahepthyl Ammonium Bromide, which was used peviously for other experiements. There is a possibility to find the traces of the structures from the previous experiments. It can be easily cleared from the analysis. The figure 2.7 shows the intensity graph for urine sample in Acetic Acid.

2.7 Secondary Electrospray Ionization (SEI)

A drop of urine (as provided) was placed in a closed chamber where a sheath flow evaporates the urine sample which is transported towards a buffer (20mM of Acetic Acid and Ammonium Acetate in 50/50 Methanol/Water). It charges the urine in the gas phase which then goes into the DMA-MS system. The flow rate for the capillary was acquired at the pressure of 150-200 mbar and about 2.5-3 kV applied to the capillary providing 60-90 nA current to the liquid. The results of the urine analysis are studied here.

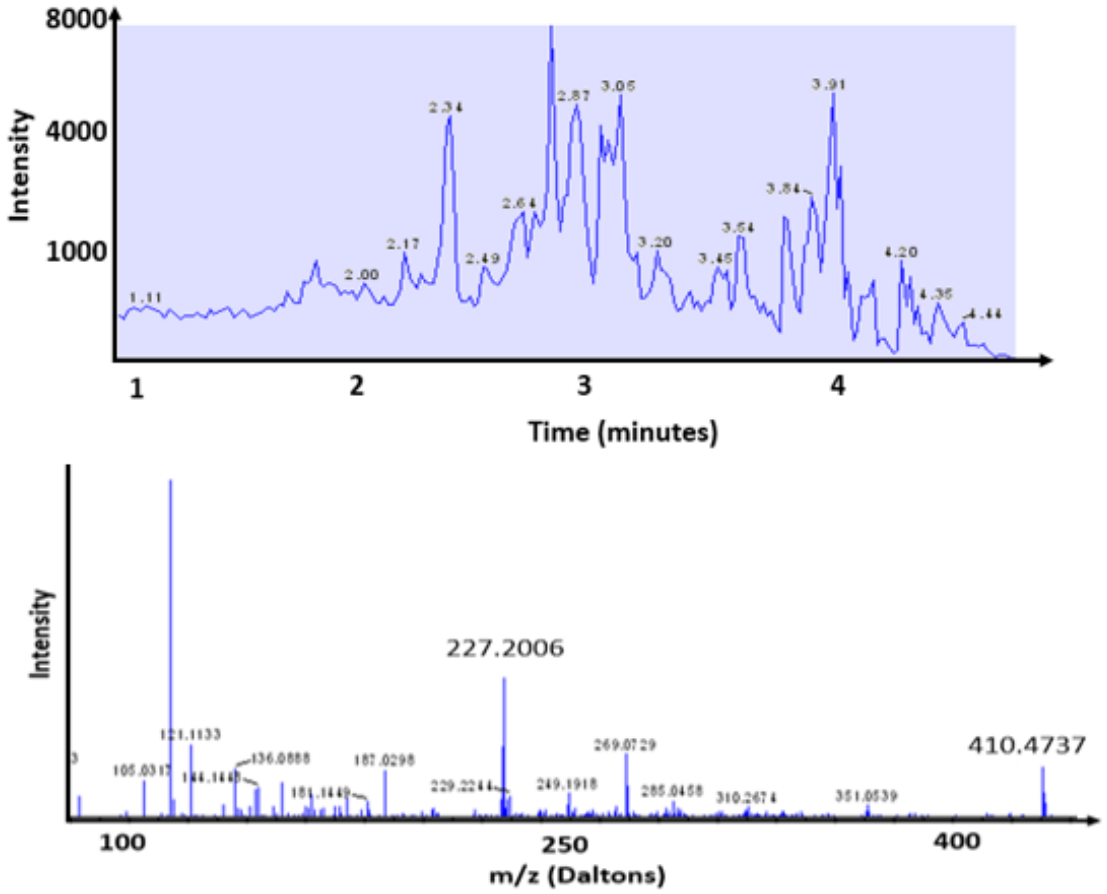

Fig. 2.6. Mass spectrum of urine sample in Ammonium Acetate for ESI

Urine in Ammonium Acetate

The figure 2.8 shows the spectra for the SEI method. The spectrum consist of various peak showing the performance of MS system.

Urine in Acetic Acid Acetate

The figure 2.9 shows the spectrum in 20 mM Acetic Acid ESI. The intensity obtained was low. The contour plot in figure 2.10 above shows the intensity for the peaks having voltage rise on X-axis and m/z ratio on Y-axis. The intensity shown here is for the peak of urea and has a resolution of about 37. The contour demostrates the efficency of the system together with DMA installed. The obtained peak of urea was then compared to the online isotope distribution calculator validates the results.

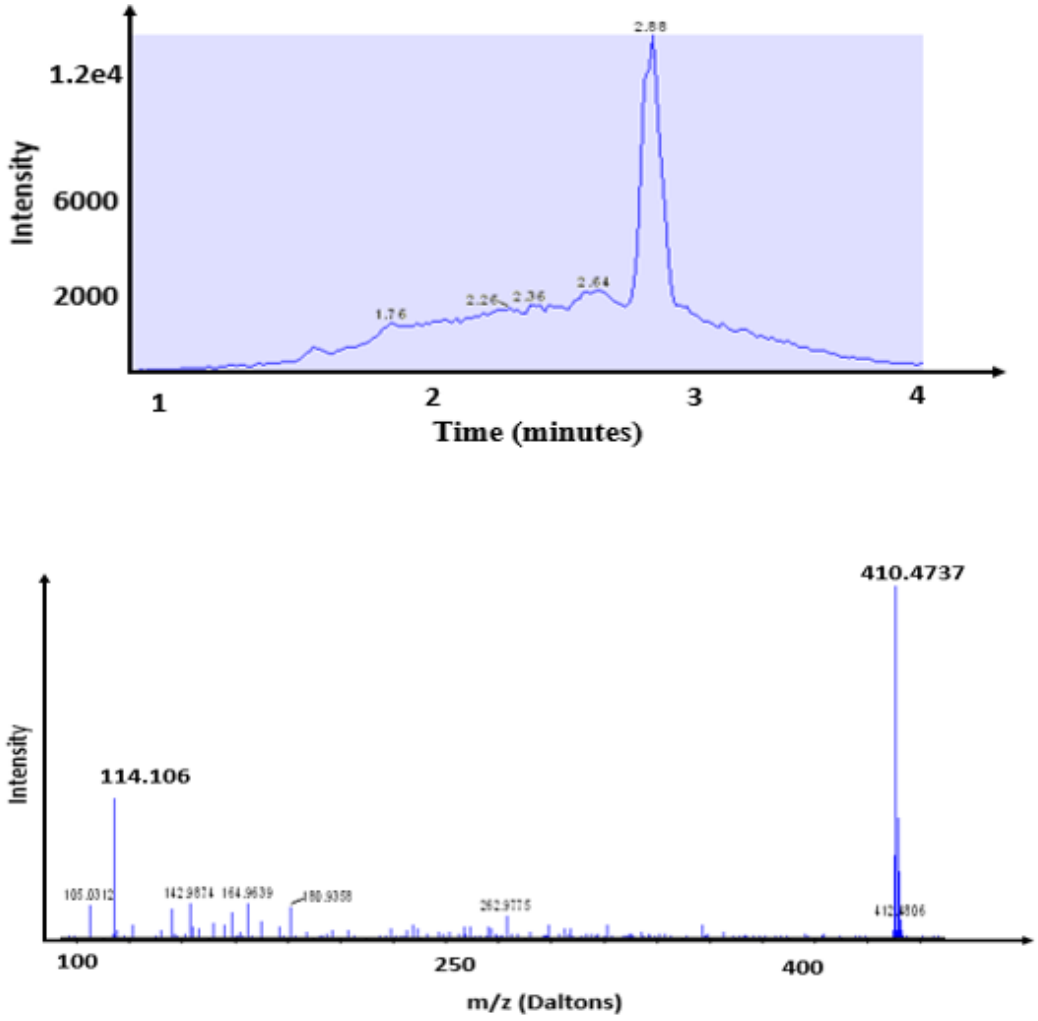

Fig. 2.7. Mass spectrum of urine sample in Acetic Acid for ESI

The peaks from the experiment and isotope distribution calculator were found to be similar in nature validating the experimental analysis and is demonstrated in figure 2.11.

2.8 Discussions

- The most abundant peak was found for Creatinine (114 m/z), Malic acid (135 m/z), and Urea (61 m/z) throughout the analysis.

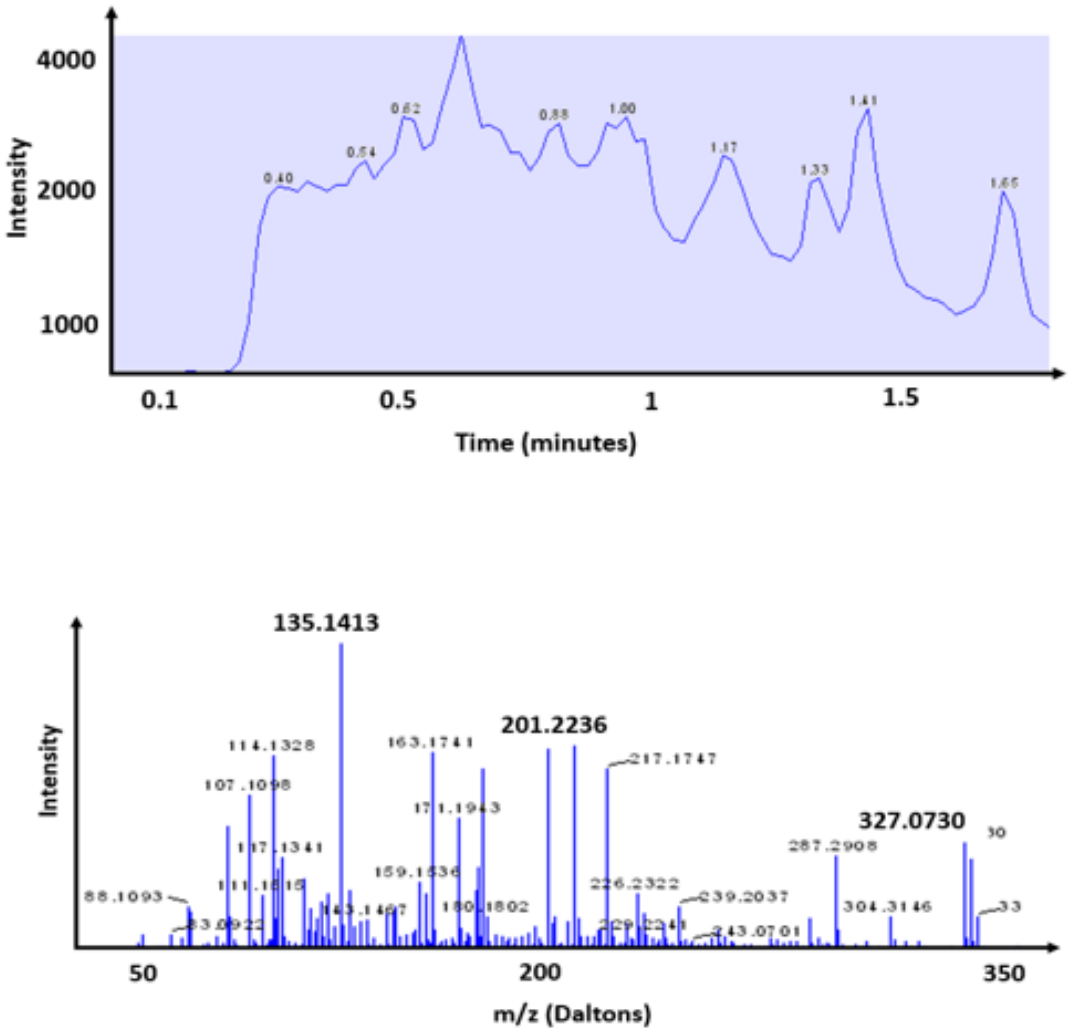

Fig. 2.8. Mass spectrum of urine sample in Ammonium Acetate for SEI

- 410 peak is seen due to an earlier experiment performed with Tetrahepthylammonium, which was then cleaned and cannot be seen in later experiments.
- The intensity of peaks was high for Ammonium Acetate charger.
- Peaks with higher m/z are yet to be determined. The higher m/z in the spectra may be for a sugar related structure such as Sucrose.
- Others peak apart from one mentioned above were also present (according to excel sheet of urine analysis), but with very low intensity.

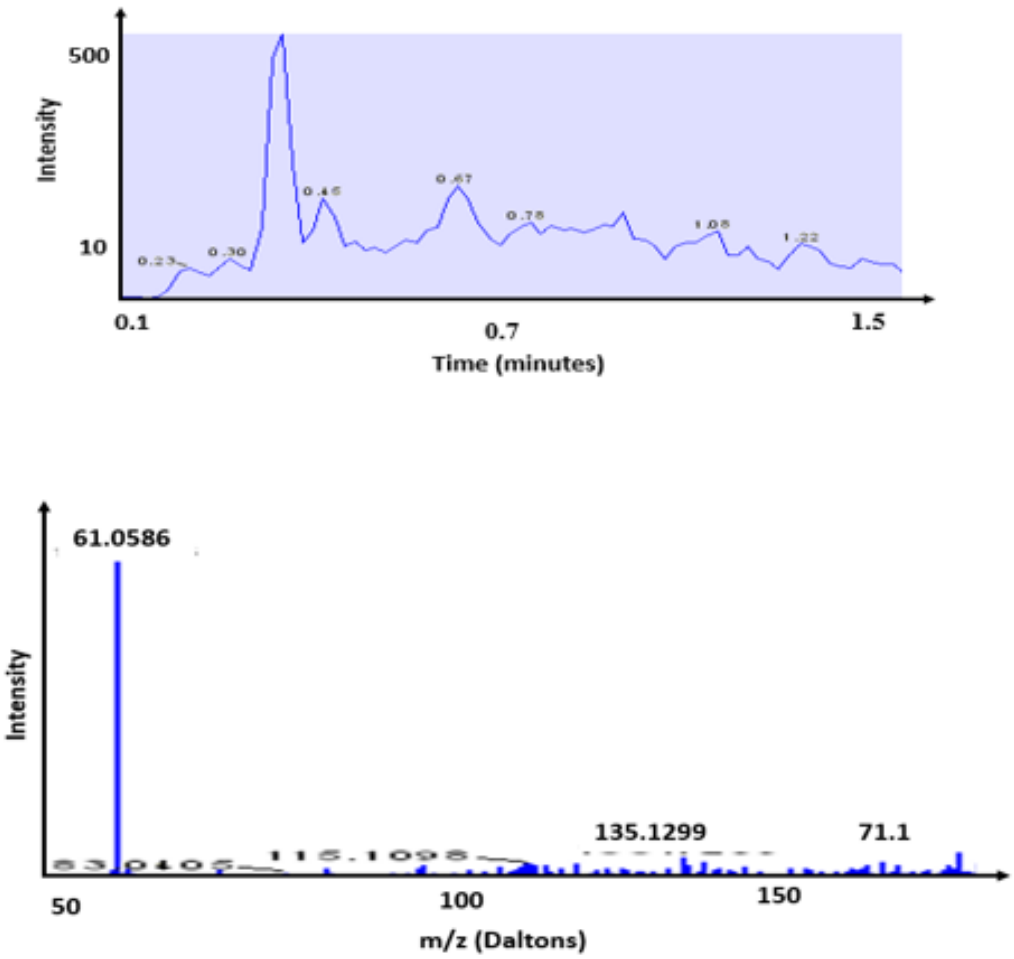

Fig. 2.9. Mass spectrum of urine sample in Acetic Acid for SEI

- We also found peak for caffeine (195 m/z) which was from the environment (not from urine sample). This happened because the chamber was exposed to the environment on different occasions, which can be the source of caffeine.
- As it did not appear in further experiments or appeared with very low intensity, it can be said that that, Caffeine was from the environment.

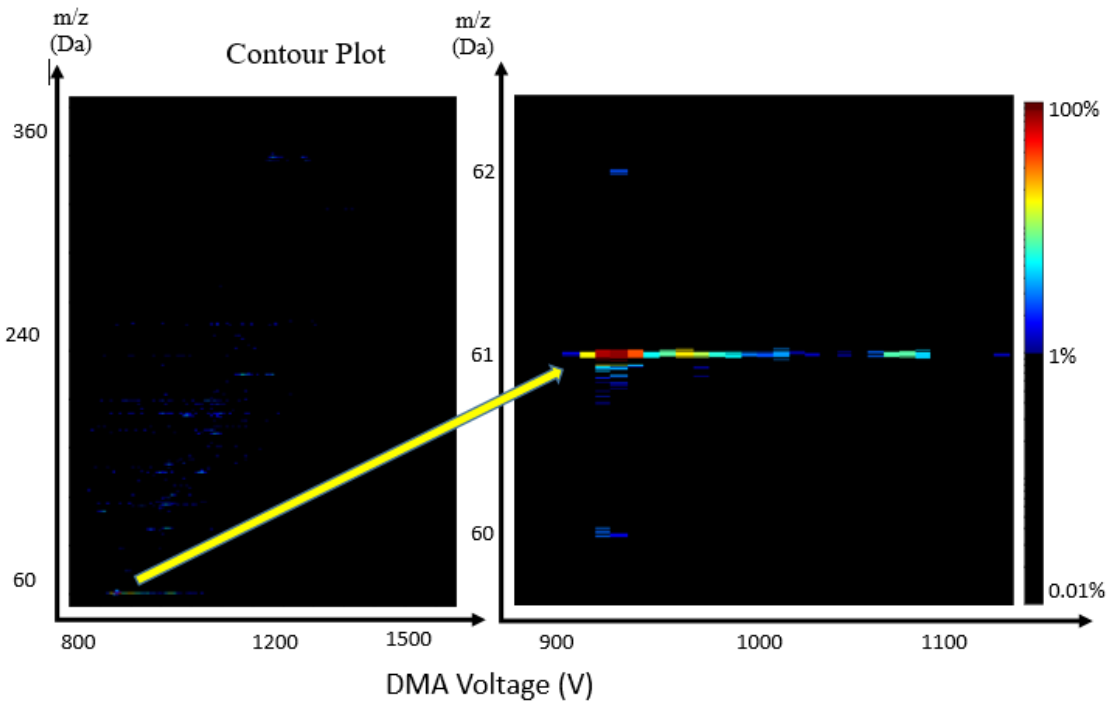

Fig. 2.10. Contour plot of Urea

- It shows that instrument/ secondary ESI method is highly sensitive to give peaks for substance from the environment, which serves our purpose of using Secondary ESI.

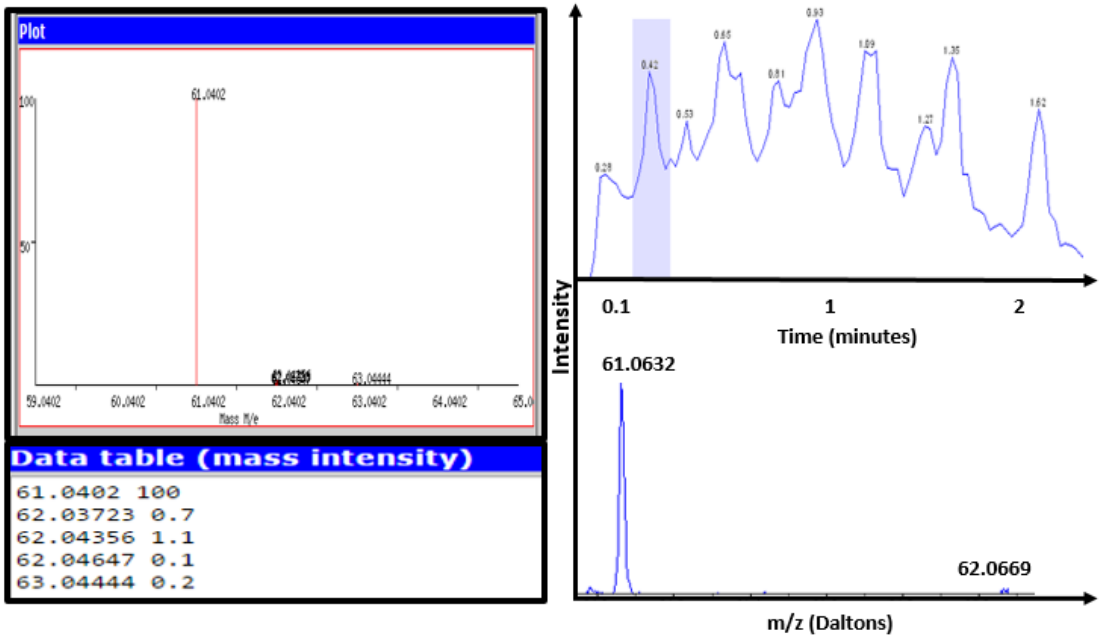

Fig. 2.11. Mass spectrum comparison for the peak of Urea

3. ION MOBILITY CALCULATOR FOR FREE MOLECULAR REGIME

3.1 MOBCAL

MOBCAL or MOBility CALculator is one of the important algorithms developed at MFJ research group, Indiana University around the year 1996. This algorithm uses momentum transfer theory and kinetic theory of gases loosely based on Onsagers[53,54] reciprocal relation. MOBCALs collision cross section defines momentum transfer during collision of the gas molecules and the structure, and then it averages over all the possible orientations and velocities[40,42] given by:

$$\Omega_{avg}^{1,1} = \frac{1}{8\pi}\int_0^{2\pi} d\theta \int_0^{\pi} sin\phi d\phi \int_0^{2\pi} d\gamma \frac{\pi}{8}\left(\frac{\mu}{2k_BT}\right)^3 \int_0^{\infty} g^5 e^{-\frac{\mu g^2}{2K_BT}} dg \int_0^{\infty} 2b(1-cos\chi)db \qquad (8)$$

In the preceding equation, θ,ϕ,and γ represent the orientation angles, g is the relative velocity, b is the impact parameter, μ is the reduced mass, and χ is the deflection/scattering angle. The above equation can be reduced for a collision between two hard spheres as follows: [41]

$$\Omega_{avg}^{1,1} = \frac{1}{8\pi}\int_0^{2\pi} d\theta \int_0^{\pi} sin\phi d\phi \int_0^{2\pi} d\gamma \pi b_{min}^2 \qquad (9)$$

Here,b_{min} is the hard sphere contact distance. This equation definitely ignores the long-range potential interaction between colliding gas molecules and the charged structure. This is quite useful for the structure with convex geometry and is equal to the orientationally averaged projection area. In equation 5, the deflection angle χ is calculated numerically which depend upon the orientation angles, the impact parameter, the gas molecule velocity, and the potential parameter, which is given by [56, 57]:

$$\chi\left(\theta,\phi,\gamma,g,b\right)=\pi-2b\int_{r_m}^{\infty}\frac{dr}{r^2\sqrt{1-\frac{b^2}{r^2}-\left(\frac{\frac{\phi(r)}{m_{red}g^2}}{2}\right)}} \qquad (10)$$

Where r_m in the integral is the distance of the closest approach, $\phi(r)$ is the potential interaction, and r is the position of the gas molecule which is to be considered in the program. The position r(x, y, z) will depend on the starting point and initial velocity of the gas molecules. The figure 3.1 depicts the idea of movement and momentum transfer of a gas molecule around the structure. For numerical calculation

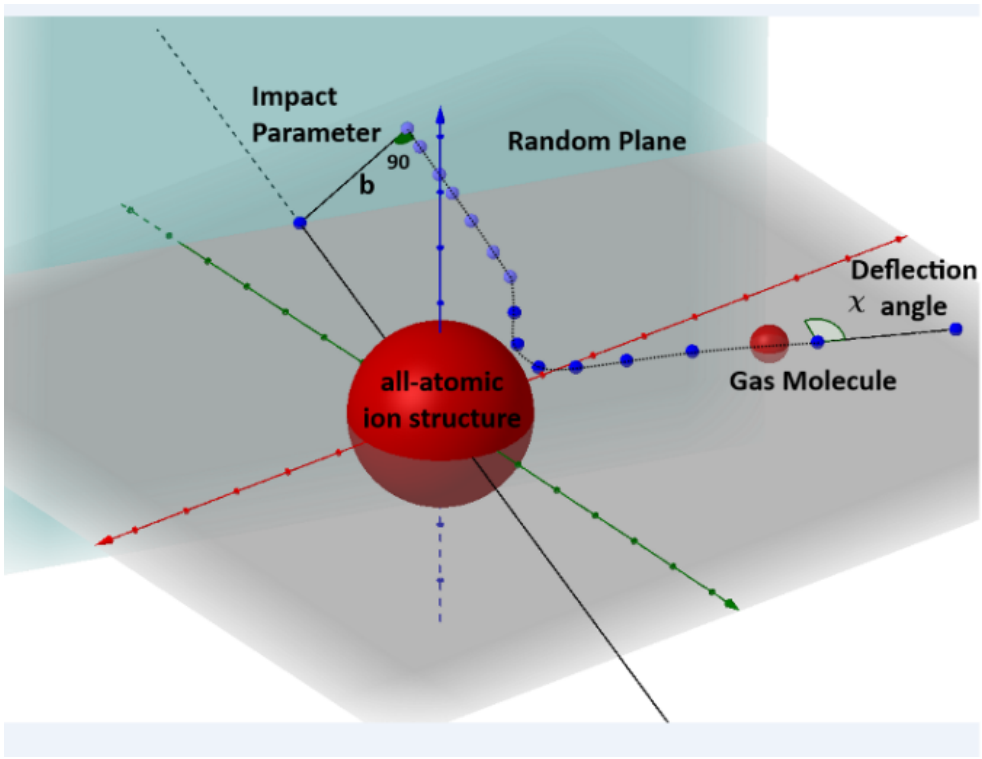

Fig. 3.1. Depiction of the process that MOBCAL undertakes to calculate the deflection angle of a gas molecule from a random plane orientation

in MOBCAL, a random plane is chosen at first. The gas molecule with a different impact factor b and gas velocities are emitted perpendicular to the plane towards the structure. Depending upon the potential interaction, the trajectories are followed throughout the calculation. If no potential is assumed, the trajectory is supposed to

be linear, and if they are considered then the path of the gas molecule will tend to curve using the time step algorithm (4th order Runge-Kutta). This curved trajectory will depend on the placement of charges on the structure and multiple potential interactions between the colliding bodies. A couple of assumptions are laid down in order to simplify the calculation for small values of E/N, where N is the gas number density:

- The first assumption is that the velocity of the gas molecule is larger than that of ion.
- The second assumption is that energy will be completely conserved during the process of collision, so that the deflection angle by itself will be capable of defining the momentum transfer.

The gas molecule trajectory is monitored until the gas molecule exists in the specified domain and the deflection angle is recorded. The used Lennard-Jones potential parameter must be optimized using the experimental CCS value. [40,41,58-71]

3.1.1 MOBCAL Program Description

MOBCAL is a Fortran code mobcal.f, compiled using Fortran 77. It has another file known as mobcal.run, which is its input file. It has three specified lines containing the command to read the input: the first line describes the name of the file, which contains the information of the coordinates of the atoms, the second line is the name of the output file, and the third line specifies the random seed given to the computer. The file which contains the coordinate has the extension .mfj. This has name of file title in the first row. The second line specifies the number of coordinate sets to be averaged. The third line, ang specifies the unit Angstrom for the coordinates. The fourth line contains the scaling factor generally specified as 1.0000. Starting from line 5 of the code, for each atom x, y, z, coordinates are specified followed by the mass in an integer for every atom. In the MOBCAL code, it is important to define the

number of trajectories on line 627 (imp = 25). A detailed description can be found on the MFJ Research Group from Indiana University.[81]

3.2 IMoS (Ion Mobility Spectrometry)

IMoS is a newly developed in-house algorithm by Carlos Larriba-Andaluz for mobility calculation in the free molecular regime. This algorithm provides an efficient and fast way compared to other mobility calculator to calculate the Collision Cross Section and mobility of the given charged structures. A number of studies are carried out regarding numerical calculations using IMoS providing evidence of its efficiency and reduced computational time. Benchmark comparison for IMoS and MOBCAL is explained here in detail [80]. IMoS provides numerical development in the calculation of Collision Cross Section, especially for diatomic gases, which have stalled in recent years. It takes into consideration the energy accommodation factor during the collision process giving results that are in good agreement with experimental results, long-range potentials, and calculations for complex structures in different background gases. Here, we explain the working principle of the IMoS algorithm, a detailed description about files required for IMoS, and its usage.

3.2.1 The IMoS Working Principle

The IMoS algorithm suite is based on the general free molecular momentum transfer approach during colliding bodies. Real gas conditions are put forth in calculations based on full Maxwell-Boltzmann distribution approximation, wherein gas molecules are allowed to collide with the structure of interest from all possible directions. This allows the simulation to follow real gas conditions, even taking into account gas molecule-gas molecule collision. This phenomenon occurs at finite Kn number as the structure and gas molecule size are similar. The collision of gas molecules from all directions lead to the formation of symmetric drag tensor by considering only three perpendicular directions. Hence, three orientations of the given structure are enough

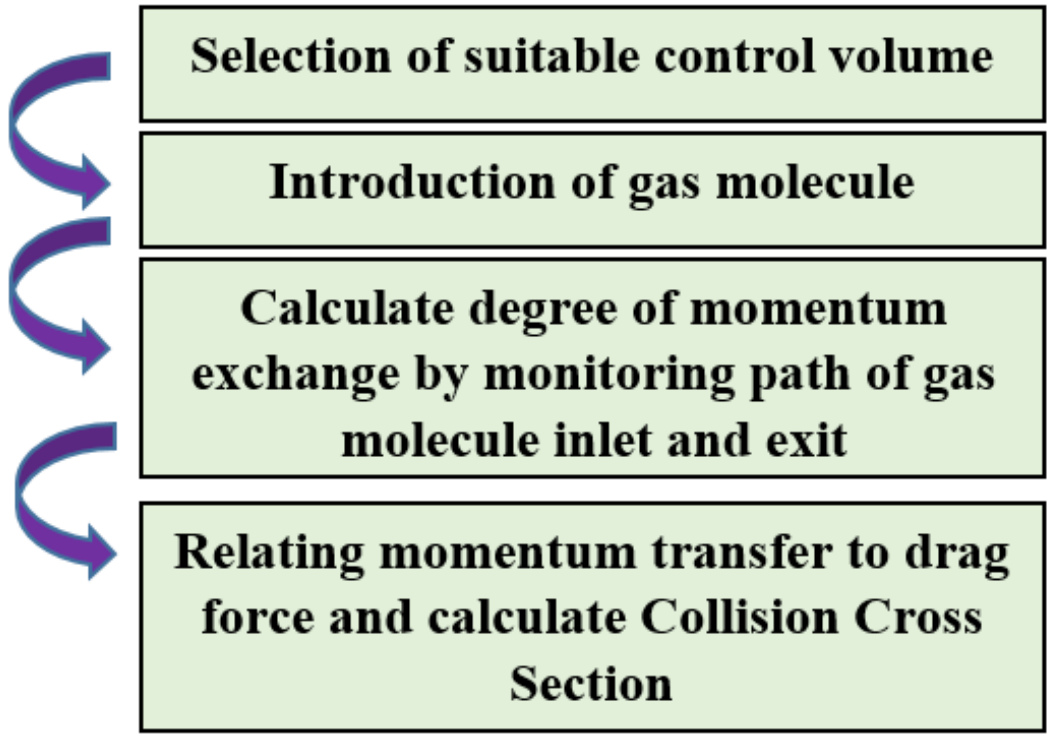

Fig. 3.2. Flow chart for the process of calculation in IMoS

to calculate the Collision Cross Section. The general calculation steps are incorporated in the block diagram in figure 3.2.

The introduction of a gas molecule in the control volume is an initial step for the calculation. The gas molecule velocity vector distribution function is denoted as ρ^*, where '*' expresses the dimensionless case $\left(\frac{\rho^{'}}{\rho^*} = h^2\right)$, which is normalized using the most probable gas molecule speed $h = \left(\frac{2KT}{m_{gas}}\right)^{\frac{1}{2}}$ and is given by:

$$\rho^* = \left(\frac{1}{\pi}\right)^{\frac{3}{2}} \exp\left(-(||\vec{c}^* - \vec{V}^*||^2)\right) \qquad (11)$$

Where c is the velocity of the gas molecule and V is ion velocity. By solving the above equation assuming a low Mach number (ratio of the speed of bulk velocity

to that of the mean thermal speed is comparatively small), nonlinear effects can be removed using Chapman-Enskog linearization [55] and using series expansion of $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ yields

$$\rho^* = \left(\frac{1}{\pi}\right)^{\frac{3}{2}} \exp\left(-(\vec{c}^{*2} + \vec{V}^{*2} - 2\vec{c}^*\vec{V}^*)\right) \qquad (12)$$

Here $\vec{V}^{*2}$ is very small compared to the gas molecule velocity $\vec{c}^{*2}$ and hence can be neglected.

$$\rho^* = \left(\frac{1}{\pi}\right)^{\frac{3}{2}} \exp\left(-(\vec{c}^{*2}).exp(2\vec{c}^*\vec{V}^*\right) \qquad (13)$$

$$\rho^* = \left(\frac{1}{\pi}\right)^{\frac{3}{2}} \exp\left(-\vec{c}^*\right).(1 + 2\vec{c}^*\vec{V}^*)$$

Neglecting higher terms in the above expansion for exp(x) for the same reason.

$$\rho^* = \left(\frac{1}{\pi}\right)^{\frac{3}{2}} \exp\left(-\vec{c}^*\right) + \left(\frac{1}{\pi}\right)^{\frac{3}{2}} \exp\left(-\vec{c}^*\right).2\vec{c}^*\vec{V}^* \qquad (14)$$

$$\rho^* = \rho_0^* + (2\vec{c}^*\vec{V}^*)\rho_0^* = \rho_0^* + \rho_1^* \qquad (15)$$

Where ρ_0 is classical Maxwell-Boltzmann velocity distribution (absence of bulk velocity)

$$\rho_0 = \left(\frac{m_{gas}}{2\pi kT}\right)^{\frac{3}{2}} exp\left(-\frac{m_{gas}c^2}{2kT}\right) \qquad (16)$$

As the first term in equation 11 does not have a bulk velocity vector its contribution in momentum transfer must be zero and can be neglected. The gas molecule entering the control volume follows ρ_1^* distribution, which helps in increasing the efficiency of the computational time.

The next step is to choose the control volume for the calculation. A control volume is selected so as to define the working domain and boundary conditions for the collisions. The net momentum transfer can be easily inferred from the gas molecule inlet and outlet conditions. A polyatomic structure can be enclosed in the cuboidal control volume if the structure of interest is skewed (somewhat rectangular or stretched in the shape) or spherical control volume if the structure is globular in nature. For this

to happen, the size of the control volume is determined for both cases and the domain with minimum volume is selected by IMoS for further calculation. The control volume is depicted in the figure 3.3.

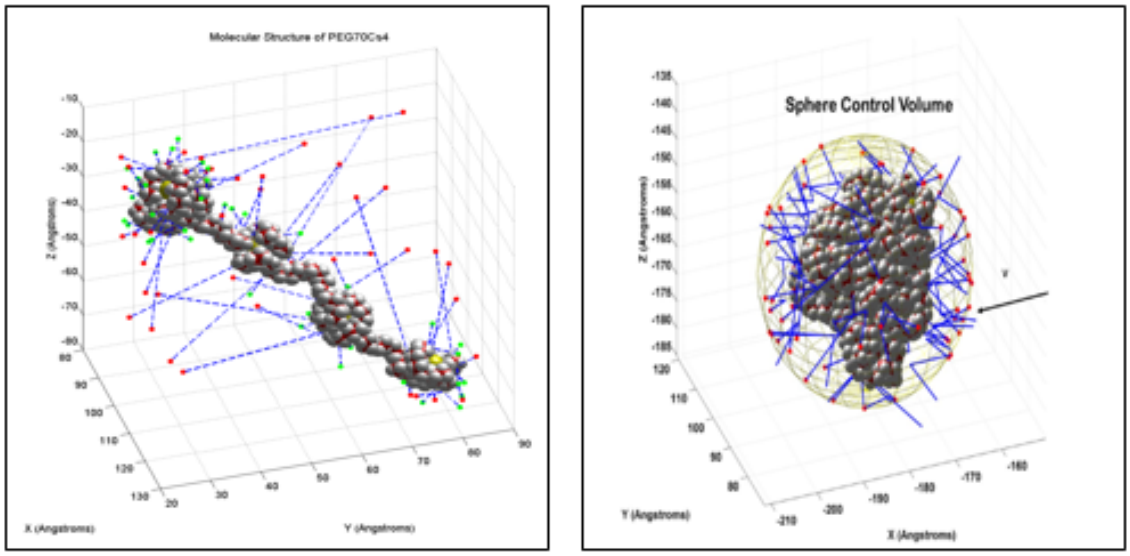

Fig. 3.3. Domain selection for two different shapes of the structures [38]

The flow rate entering the control volume is given by. [49]

$$Q = \oint \iiint_{c.n<0} \rho(\vec{C}_{gas}.\vec{n}) d^3 c_{gas} dA \qquad (17)$$

Where n is the outward normal to the surface.

In the above equation, c.n<0 accounts only for the gas molecules entering the domain, which will be considered for the calculation purposes, and is not the limit for integration. Because ρ_0 is neglected, there is possibilty of numerically negative gas flow accounting for negative momentum transfer, which should be taken under consideration during the calculation of net momentum transfer during collision. The volumetric flow rate should be calculated through all the walls of the selected domain. Considering the gas velocity distribution from equation 11, the volumetric flow rate can be simplified as follows:

$$Q = \oint \int_{c.n<0} \rho_1(\vec{C}_{gas}n)d^3c_{gas}dA \qquad (18)$$

$$Q = \oint \int_{c.n<0} 2(2\vec{c}^*\vec{V}^*)\rho_0(\vec{C}_{gas}n)d^3c_{gas}dA \qquad (19)$$

The above equations help us to set the volumetric flow rate through the control volume. Now, the program establishes appropriate emission angles and velocities for the given control volume. For instance, if a cuboid is chosen for calculation, the emission angle and velocities must obey the following probability of distribution for the walls parallel to the bulk flow:[39]

$$\frac{k_1}{\pi^{\frac{3}{2}}} \int_0^{\frac{\pi}{2}} \int_0^{2\pi} \int_0^{\infty} 2V^*c^{*2}cos^2e^{-(c^{*2})}c^{*2}sin^2(\theta)cos(\phi)d\theta d\phi dc^* = 1 \qquad (20)$$

The above equation is separable and the angle can be sampled accordingly and k_1 is the normalizing constant. The calculation of the volumetric flow rate is followed by momentum transfer calculation. Once the initial position and the velocity vector is established, a Verlet algorithm is used to track the gas molecule trajectories from the inlet through the gas molecule exit considering all the remission laws, potential interaction for different atoms, and its respective charges. The IMoS algorithm efficiently implies various re-emission laws to calculate Collision Cross Section which is in good agreement with the experimental results. Re-emission trajectories are calculated once the collision point is inferred. Several re-emission laws are available in the past literature. Specular re-emission rules, various diffused re-emission rules with random re-emission angles, and thermal accommodation factor are taken into account.The momentum transfer is then calculated by the difference in the initial and final velocities and the mass of the gas. The average drag force, momentum transfer is related to the mobility through as follows:

$$Z(\Omega) = \frac{\vec{V}_{drift}}{<\vec{F}_D>} \qquad (21)$$

The mobility then can be related to the Collision Cross Section (Ω) as follows:

$$Z = \left(\frac{\pi m_{red}}{8kT}\right)^{\frac{1}{2}} \frac{3Ze}{4\rho_{gas}\Omega} \qquad (22)$$

Later, the Happel and Brenners method [72, 73] can be utilized to differentiate between true and average mobility.

This algorithm reckons various re-emission rules based on Epstein [46] observations. When the analytical and experimental Collision Cross Section for 1.5-10 nm spherical particles [52,39] are compared, it yields $\Omega = 1.36PA$. With that fact, specular and elastic re-emission rules, which are adduced in the MOBCAL algorithm, are unable to produce a good match with the experimental results for diatomic gases. Hence, there was a need to introduce the DHSS and the DTM method which accounts for diffused and inelastic behaviour of collision. Two commonly used diffused re-emission model put forth by Epstein can be used to more closely match the experimental values. The first one is to be considered for complete thermal accomodation, where the re-emisson is a mixuture of specular as well as diffused collision, with the re-emission speed sampled through Maxwell-Boltzmann distribution of the surface temperature. This gives rise to accomodation coefficient , which can be used in order to decide what fraction should be diffused and what fraction should be specular to match the experimental values. This rule has instant energy exchange (vibrational and rotational) for gas molecules and the structure, which yields $\Omega = (1 + \frac{\pi}{8})PA \approx 1.39PA$. Hence, to match the experimental result, α was taken as 0.91. It means that 91% of collision is diffused while rest are specular. The second diffusive rule takes into account the random angle of re-emission with retained translational energy. For this case, $\Omega = (1 + \frac{9\pi}{64})PA \approx 1.44PA$ is observed. For this case, α should be taken as 0.81 to match the experimental results. IMoS provides an option to include potential interaction for the charged structure and gas molecule interaction same as that of MOBCAL. Hard sphere potential gives rectilinear trajectories with instant re-emission of the gas molecule. The phenomenon of scattering can be observed, meaning that multiple collision events occur for a single gas molecule throughout its journey for the selected domain. Specular and elastic collision are referred to as Exact/Elastic Hard Sphere Scattering or EHSS [41] which is incorporated in MOBCALs official version, while IMoS can also perform diffused and inelastic collision, [4,38,39] which are generally

observed for heavier gases such as N_2 and Ar termed as Diffuse Hard Sphere Scattering or DHSS. Other potentials are also considered (apart from hard sphere potentials) for the Trajectory Method. The most commonly used potential is the Lennard-Jones potential parameters in combination with the ion-nduced dipole potential given by equation 4. For diatomic Nitrogen gas molecules, which has linear structure, two other important parameters were considerd. The first is the ion-quadropole potential and the second is the orientation of the nitrogen, which was also modified for MOBCAL. For this to be considered, one negative charge of 0.4825e was placed on each nitrogen and one positive that is 2q or 0.965e was placed at the center of the nitrogen molecule. Hence the ion-quadrupole potential can be expressed as: [74,75] IMoS provides an option to include potential interaction for the charged structure and gas molecule interaction same as that of MOBCAL. Hard sphere potential gives rectilinear trajectories with instant re-emission of the gas molecule. The phenomenon of scattering can be observed, meaning that multiple collision events occur for a single gas molecule throughout its journey for the selected domain. Specular and elastic collision are referred to as Exact/Elastic Hard Sphere Scattering or EHSS [41] which is incorporated in MOBCALs official version, while IMoS can also perform diffused and inelastic collision, [4,38,39] which are generally observed for heavier gases such as N_2 and Ar termed as Diffuse Hard Sphere Scattering or DHSS. Other potentials are also considered (apart from hard sphere potentials) for the Trajectory Method. The most commonly used potential is the Lennard-Jones potential parameters in combination with the ion-nduced dipole potential given by equation 4. For diatomic Nitrogen gas molecules, which has linear structure, two other important parameters were considerd. The first is the ion-quadropole potential and the second is the orientation of the nitrogen, which was also modified for MOBCAL. For this to be considered, one negative charge of 0.4825e was placed on each nitrogen and one positive that is 2q or 0.965e was placed at the center of the nitrogen molecule. Hence the ion-quadrupole potential can be expressed as: [74,75]

$$\phi_{IQ}(x,y,z) = \sum_{j=1}^{3}\sum_{i=1}^{n}\left(\frac{Z_i Z_j e^2}{r_{ij}^3}\right) \qquad (23)$$

Where index $j = 1$ and 3 denote the two negative charges on the nitrogen atom and $j = 2$ denotes the center charge. The index i denotes the charge on the ion. Now the orientation for N_2 molecule is considered according to the weighted impact parameter and all possible orintation are made. IMoS has its own way to take the oreinataion for diatomic gasses [35], and the ion-quadrupole potential and weighted gas molecule orientation is already incorporated in the IMoS algorithm.

3.2.2 Other Collision Cross Section Methods:

IMoS gives an option to select a combination among the potential interaction and the Collision Cross Section calculation methods described in the previous section. The following is a list for the combination of calculations that can be made using IMoS.[4] The way to input this is explained later in the IMoS.cla section:

- TEHSS (trajectory-elastic hard sphere scattering): The gas molecule and the structure are modeled as hard sphere. Re-emission is elastic and of specular type with ion-induced dipole potential considered.

- TDHSS (trajectory-diffuse hard sphere scattering): The gas molecule and the structures are modeled as hard sphere. Re-emission is diffused and of inelastic type with only ion-induced dipole potential considered.

- TMLJHe (trajectory method-Lennard-Jones-Helium): With the assumption of the specular-elastic collision, Lennard-Jones potentials (4-6-12) is considered in this method. It does not involve the optimized value for the parameter.

- TMLJN_2 (trajectory method-Lennard-Jones-Nitrogen) : This is similar to that of TMLJHe, having an optimized parameter for the N_2 provided by Campuazano et al.[71]

3.2.3 IMoS Suite Explanation

Explanation of the IMoS.cla file

The IMoS.cla file is an input file having all the relevant information for IMoS to operate. It contains a set of instructions to be followed and executed. Below is the explanation of each line in the IMoS.cla file: Header section:

excelfile Savefile Gas

$C60positions.xlsx\backslash savefolder\backslash C60positions.txtHe$

These are the first two lines in IMoS.cla. The first line contains the title for the next line such as Excel file used, the saved file name, and the gas used in the calculation, and is not read in the program. The second line contains the name of the Excel or a pdb file which has to be read by the IMoS program. It has the path for the save folder with the user defined name for the output file, where the output file is stored with the respective results for the structures contained in the Excel file. This has to be separated by a space and not by a tab for the program to read it correctly. Followed by the output file name is the gas which can be used for calculation. IMoS has N_2, He, Ar, CO_2, SF_6, and Air as the option for gas to be used for calculation purposes. The notation for the gas is case sensitive. The gas input here will read the LJtable.xlsx sheet for the given gas. This Excel sheet contains the Lennard-Jones potential parameter (ϵ and σ) for the number of elements for the respective bath/buffer gas.

interface 1 0

This above parameter is used if the user needs a GUI. Otherwise, for the GUI-free interface, the value is set to interface 0 0. This will allow the command prompt to pop up and show the desire calculation on a command window on a Windows Operating System. Both the numbers are important and should be used.

Body section:

fromvalue 1

tovalue 10

The above parameter sets the range for the calculation for a given Excel file. A single Excel file may contain a number of Excel sheets, each containing coordinates and charge value for different molecule/structure. The user can define the number of structure to be calculated by selecting fromvalue to tovalue in an Excel file. For the case of a pdb file, it will calculate all the structures given that it has more than one model and will always start the calculation from the first model. If there is only one model, we have to make sure the fromvalue and tovalue is the same to avoid the repetition of calculations. There is a certain way of writing a pdb file if more than one model is included in a pdb file, which are mentioned as follows:

- You must have a TER line to terminate a chain in your model
- You must have an ENDMDL line to terminate your model
- You must have a MODEL line that starts your new model
- The following example for Humanin is demonstrated: ATOM 390 HB1 ALA A 24 1.791 10.249 -16.257 1.00 0.00 H
  ATOM 391 HB2 ALA A 24 0.371 9.436 -16.916 1.00 0.00 H
  ATOM 392 HB3 ALA A 24 1.609 8.504 -16.072 1.00 0.00 H
  TER 393 ALA A 24
  ENDMDL
  MODEL 13
  ATOM 1 N MET A 1 8.093 -8.495 15.332 1.00 0.00 N
  ATOM 2 CA MET A 1 6.757 -8.803 15.906 1.00 0.00 C
  ATOM 3 C MET A 1 5.730 -7.748 15.511 1.00 0.00 C

partialc 0

Charge 1

The above parameters give the charge situation for the selected structure. The parameter partialc 0 is set to 0 or 1 to define whether to include partial charges or not on any of the atoms of a structure. The second parameter Charge 1 is to set the

charge on the structure and is read from the IMoS.cla file under the following two conditions:

1. If there is no total charge assigned on the Excel sheet regardless of how many charges are put on the individual atom.

2. For any pdb file regardless of any charges assigned to individual atom.

One important thing to be noted here is, if we specify the value of charge from the IMoS.cla file, the charge (positive or negative) will be placed in the geometric center of the structure. This placement is good for a structure with a straight chain or globular geometry but is not a good idea for hollow structure/ring structure (for example, Kekulene) to assign the charge in the center as there is no atom in the center and will affect the calculation where potential parameters are taken into consideration. For such cases, we can assign charges on individual atoms in the Excel file. This will help us to get rid of any computational error to miss-assign the structure.

Mgas 28

This is a mass of gas used for the calculation in Da. Specify this value accordingly for the buffer gas you would like the calculation to be performed.

radgas 1.5

Radius for buffer gas is specified in this line. For N_2 and He the values is 1.5 and 1.2 respectively in Angstroms. This is helpful for the PA and the Hard Sphere method.

Polarizability 1.7

This is the value for polarizability of the gas used in A^3. The value for N_2 and He is 1.73 and 0.2073 respectively. Polarizability (α) is used in equation 4.

Pressure 101325

The value of pressure is specified in Pascals. This pressure mimics the real gas molecule collision environment.

Mweight 104

It defines molecular weight of the molecule/structure in Da. This line is read by the program if the molecular weight is not specified in the Excel file.

Temperature 310

This value assigns the temperature in Kelvin used in the speed of the gas and mobility calculations. The structure, however, remains unaffected by any specified value.

redCoef 1.1

The factor reduction coefficient helps to reduce the size of the structure by the percentage applied to the coefficient. In this case, the dimension (distance from the geometric center and radius) is reduced by 10%. Generally, it is advised to use the original structure keeping redCoef 1.0.

NrotationsPA 2000

This parameter defines the number of random orientation required for the PA algorithm to calculate the Collision Cross Section. This is active only if the PA method is selected for the calculation purposes.

NrotationsEHSS 3

NrotationsTM 3

The above parameters establish the number of rotations of molecules required for the calculation. This value is set to 3 to calculate the drag in three perpendicular directions. For the above condition to execute, the value for MOBCAL is to be set to 0. Otherwise, if MOBCAL is set to 1, the program will behave in a similar fashion as that of MOBCAL. In the MOBCAL=1 case, NrotationsEHSS and NrotationsTM are set to 30, which is a good amount of rotation for calculation.

Acommodation 0.0

The accommodation factor as explained in an earlier section decides what percentage of the collision is specular and elastic in nature. Its value ranges from 0 to 1. If 0 is chosen, for instance, the collision in the algorithm will be 100% specular and elastic (EHSS). This factor is only used in calculating the EHSS/DHSS method. It is also used in the TM and DTM method only if the Lennard-Jones potential parameter is neglected and only hard sphere potential is implied.

Timestep 100

The Timestep parameter is used to set the time steps in the Trajectory Method as

this method uses the Velocity verlet algorithm. The higher the value are, the smaller will be the time step in the calculation. This value is not changed usually, but if heavier charged structure is selected, the value might be increased.

Boxdomain 16

This value is supposed to remain constant for every calculation.

Diffuse? 0

This value ranges from 0 to 1, depending upon the accommodation value selected. If the value is set to 0, it means the collision is specular, and if it is set to 1 the collision is diffused in nature.

reemvel 6

There are eight types of re-emission setups depending on their distribution pattern:

- Elastic (mean): Gas molecule's re-emit velocity is taken from the mean of incoming gas molecule from a skewed Maxwellian distribution
- Elastic: Gas molecules are re-emitted as the same velocity as that of impingement
- Maxwell
- 92% Maxwell
- Maxwell Distribution
- 92% Maxwell Distribution
- Other
- Other Distribution

Mean values of distribution (Maxwell, 92% Maxwell, and other) are generally recommended for the EHSS/DHSS calculation. The user can select the distribution of the gas molecule re-emission accordingly.

Other 0

Other is activated if reemvel is chosen to be 7 or 8. In that case, the value for other is set as 1 and Maxwellian Distribution in employed.

Simplify 0

This option helps to simplify the structure by eliminating internal atoms that will not affect Hard Sphere collision calculations. We can set this to 1 if we are using the EHSS/DHSS method or can be set to 0 if we use TM method. For the TM method, consideration of whole structure is necessary because of the effect of Lennard-Jones potentials.

PA 1, EHSS/DHSS 1, TM 1, DTM 0

The above-mentioned parameters are the methods employed for calculation. It has either 0 or 1. IMoS provides the flexibility in calculating more than one methods at a time.

PATSA 0

The above parameter is used to decide, what kind of Projection Approximation you want to use. If 0 is selected, it will follow the Monte Carlo approach or if 0 is selected then Traveler Salesman (TSA) is used for calculation, though the TSA method is about 30 times faster and is recommended. Although the structure is hollow, the Monte Carlo simulation is preferred.

Cutoff 0

For large structures (>10000 atoms) the cutoff value is set to 1 only if we're using Lennard-Jones potential to establish a cut-off value. Generally, this value is set to 0.

Mobcal 0

This parameter is used to set the principle direction. If [Mobcal] = 0, then [NrotationsEHSS] and [NrotationsTM] must be set to 3 and pral. Directions will be calculated. If [Mobcal]=1, then a MOBCAL approach where random directions are used and averaged.

LennardJones 1

This parameter is used only for TM calculations. If the value is set to 0 then the TM

method will use 4-∞ potential, otherwise, if the value is set as 1, 4-6-12 potential is used.

qpol 0

This parameter tells us whether (1) or not (0) to include the ion-quadrupole potential in case of diatomic gases.

TDHSS 0

It tells us whether the calculation to be made includes trajectory-diffused hard sphere scattering.

SimplifiedTM 1

This parameter tells us whether a simplification of the collision integral is used or not. If we set the value to 0, then 10 to 100 times more gas molecule has to be selected is order to get the accuracy. Hence, it is recommended to set the value to 1.

seed 17

This parameter gives random numbers to be considered in the calculations.

Numthreads 7

This parameter tells about the number of processor we are using for the calculations.

chargelines 55 62

This parameters tell the program the line location of the charge on the pdb file.

Explanation of the Excel File:

Excel file is the input file for IMoS, which contains the information of the structures for calculation. The file should be given a name and the same name should be specified in the IMoS.cla file, which will then read the data from the Excel sheet, a number of Excel sheets can be used representing different structures in a single excel file.

3.3 Discussion

Both numerical programs were tested on the same Intel i7 3.6 GHz chip with 4 physical cores. MOBCAL is not parallelized and it was running one instance at a

Table 3.1.
Excel sheet specification used in IMoS

Columns	**Specification for each column**
1	Specifies the type of atoms/elements and is string
2-4	Specifies the x, y, and z coordinates for the center of the atom
5	Specifies the radii of the atom in Angstrom unit
6	Specifies the charge on the atom. It can be integer or non-integer (i.e. partial charges). No charge will be assigned if left blank.
7	Specifies the total charge on the molecule used in calculation
8	Contains L-J parameters and color of the ion based on its amu.

time. IMoS can be parallelized (OMP) and was ran once instance at a time with different number of threads (1, 5 and 7 threads). While a linear increase in speed is not expected when using hyper threaded parallelization, the increase in speed between 5 and 7 threads is quite notable. We also note here that MOBCAL could be easily parallelized and a linear increase in speed would also be expected. When studying large biomolecules (>1000 atoms), MOBCAL was too inefficient to be used and only the parallelized version of IMoS was used in a Linux environment using a 2.8 GHz chip with 16 threads. Finally, the structures studied herein (listed in the supplementary information) have been taken from literature and not all of them are optimized for the gas phase. As such, the CCSs that appear in this manuscript will not necessarily agree with experimentally measured CCSs; the structures simply serve to facilitate comparison between the two methods. Similar results obtained in comparison were random structures chosen to validate the methods. Among the structures studied, supplementary information shows CCSs calculations for proteins (kindly provided by Iain Campuzano [71]), tetraalkylammonium salts, fullerenes, ionic liquids, polyethylene glycol homopolymer molecules, and proteins (kindly provided by Morgan Lawrentz and Iain Campuzano). Here, we tested both the program in terms of its efficiency

and performance considering all the parameters involved are equivalent so as to compare on the same platform and potential of each program can be seen and studied. Simulations were run with the same number of gas molecules (or trajectories in the case of MOBCAL) for both programs; $9 * 10^5$ molecules in He and $4 * 10^5$ molecules in N_2. The performance (CCS values and time needed for calculation) of MOBCAL was compared to IMoS 1, 5, and 7 threads execution. Different potentials were used in IMoS and were compared to MOBCAL for a large number of atoms (> 1000).

3.3.1 Comparison of Collision Cross Sections

Figure 3.4 shows the CCS comparison calculated in MOBCAL and IMoS in He as a bath gas. Three methods used for calculations are PA, EHSS, and the Trajectory Method. About 2000 random orientations were employed to calculate CCS for the Projected Area method. For calculation purposes, 47 organic ions including bioactive molecules, tetra-alkyl ammonium salts, fullerenes, ionic liquids, polyethylene glycol homopolymer molecules, and proteins were used. Van der Waals radii were used to sum the radius of the atom and the gas molecule, which provides a good platform for comparing the PA and EHSS methods. For both the programs the temperature was kept constant at 301k and the polarizability (α) for ion-induced dipole potential was taken as 0.2073 A^3. List of molecules and the values of Collsion Cross Sections in He can be found in Appendix A. Vdw radii used in calculations for the MOBCAL program (adding Helium gas molecule) is listed in table 3.2. The Lennard-Jones potential parameter used in calculations for both the program are the same and listed in table 3.3. In the figure 3.4, the difference in the CCS for PA and the EHSS method is well below 1% and can be observed normally below 0.1%. In the case of the Trajectory Method, the difference is slightly higher but still within the 1% difference, except for the case of Fullerenes, which are about 1.5% on average. The reason for such a difference could not be pinpointed exactly, but it likely arises because of the inherent differences in the algorithms. IMoS considers collisions from all directions

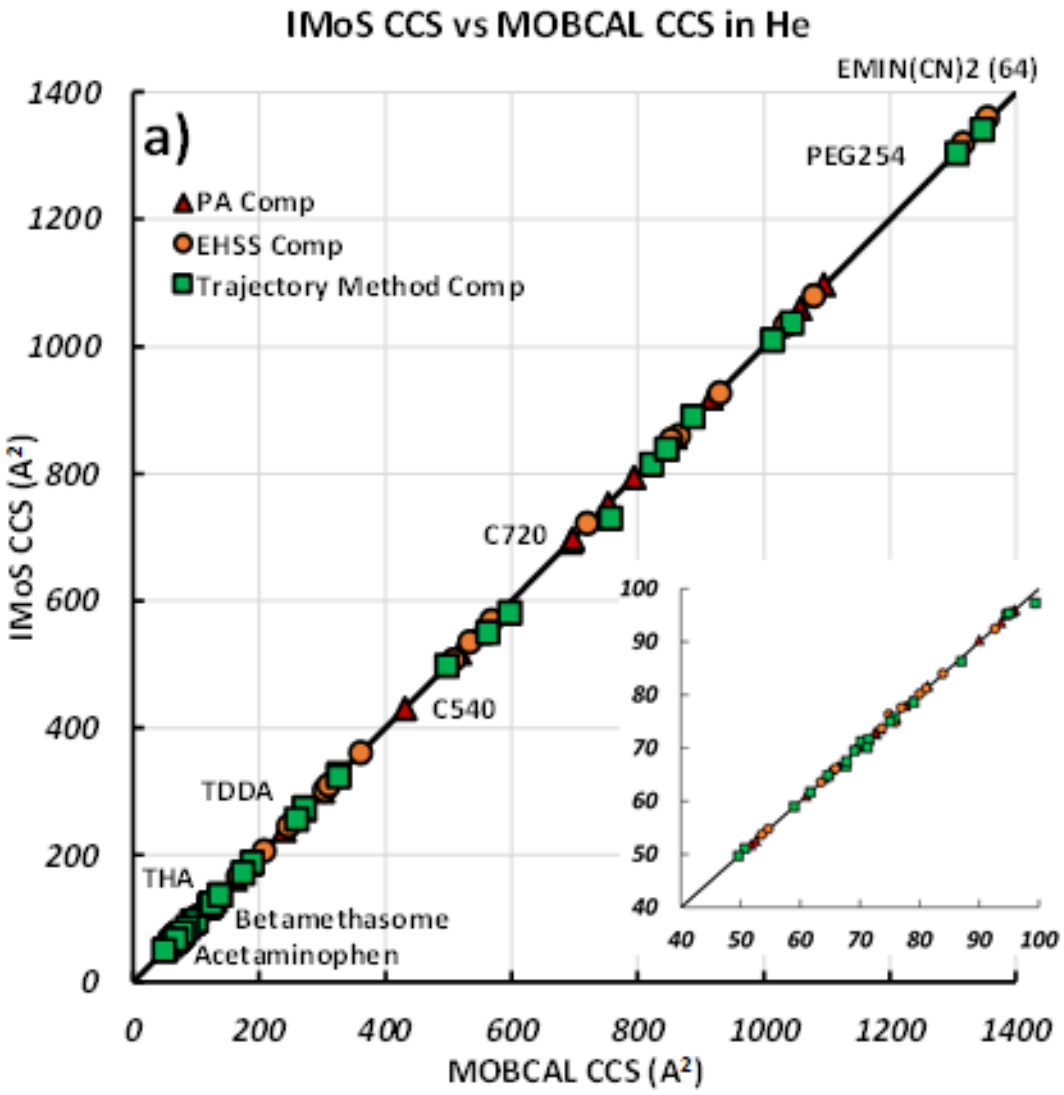

Fig. 3.4. CCS calculation comparison between MOBCAL and IMoS for PA, EHSS, and TM in He with Lennard Jones and ion-induced dipole potentials.

at every orientation with correctly distributed number of gas molecules entering the front and the back of the domain (more on the front). The relative velocities of the gas molecules from the front and the back are also slightly different (to account for the ion moving) which makes the trajectories different (See discussion in [39]). Finally, IMoS uses an adaptive time step (vs. MOBCALs constant time step) that minimizes

Table 3.2.
vdw radii (adding He gas molecule) values used in calculations for MOBCAL program

Element	**Radii Values (A)**	**Element**	**Radii Values (A)**
H	2.2	Na	2.85
C	2.2	Si	2.85
O	2.2	S	2.85
N	2.2	Fe	2.85
F	2.2	P	2.85

Table 3.3.
Lennard-Jones potential parameters

He		N_2	
ϵ	σ	ϵ	σ
(H)0.0989235	2.261	0.251829132	1.898616579
(C)0.21252132	3.0126	0.572561771	3.225486966
(O)0.1717344	2.4344	0.432705251	3.074994711
(N)0.2361348	3.3473	0.527096623	3.571906174
(F)0.1717344	2.4344	0.395004044	3.014650405
(Cs)0.28336176	4.01676	0.5814	4.20081

errors by keeping the acceleration term from being too large. This becomes especially important for large clusters of atoms and charges where the force term can become very large. In all, IMoS and MOBCAL treat trajectories differently and should yield slightly different results when using the same Lennard-Jones pairs. However, a small variation of the Lennard-Jones parameters (a factor of 1.015) in IMoS will yield values that are within 0.5% of those provided by MOBCAL for fullerenes. Calculations

were also made for the Nitrogen environment and the comparison is demonstrated in figure 3.5. The results presented are for CCS in N_2 for 16 different molecules. The

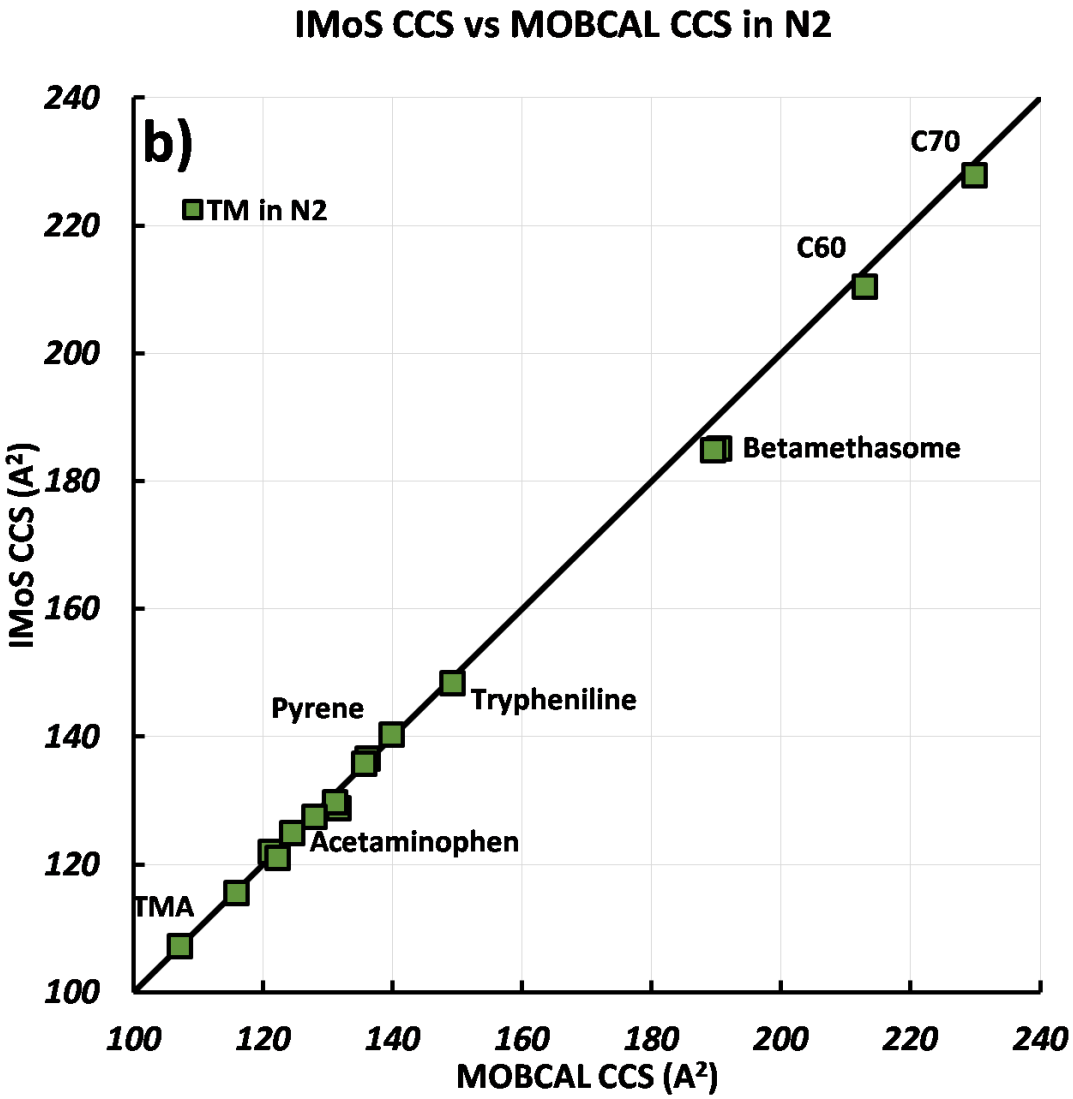

Fig. 3.5. CCS calculation comparison between MOBCAL and IMoS for TM in N2 considering ion quadrupole potential

temperature was maintained at 301K and the polarizability of 1.7 A^2 was used for N_2. In this figure, we showed only the Trajectory Method and other two methods (PA and EHSS) are not presented as they are intrinsically same as that of the He counterpart. The ion-quadrupole potential explained in equation 19 is incorporated in

the calculations, which tends to increase the computational time and hence precludes the study of the molecules containing large number of atoms (>100) for MOBCAL. Again the error difference for both IMoS and MOBCAL was less than 1% for majority molecules and slightly higher than 2% for Betamethasone and Dexamethasone. The calculations can also be performed using 4-6-12 potential with optimized Lennard-Jones parameters instead of the ion-quadrupole potential. Optimization of the L-J parameter includes the effect of the ion- quadrupole potential, which then will be computationally expensive as in He. With the above graphs for CCS, it can be seen that both the algorithms quantitatively agree well and now performance (time required for calculations) can be demonstrated and the efficiency of the programs can be tested accordingly. CCS values for the molecules are attached in table B2 appendix B.

3.3.2 Performance of IMoS and MOBCAL

Figure 3.6 displays the total time required to complete the calculation with respect to a number of atoms in He for IMoS and MOBCAL. MOBCAL calculates all three methods (PA, EHSS, and TM) simultaneously. We do not attempt to modify the program to calculate CCS for an individual method and instead allows it to compute the total computational time of the program. This was reasonable so as to not modify the code as the TM method itself takes 90% of the computational time for the calculations studied in all cases here. On the other hand, though IMoS has the option to take into account for individual method calculation, we choose to calculate all the three methods simultaneously for comparison purposes. From the figure 3.6, it can be clearly seen that the time for both methods increases approximately linearly with the number of atoms with a minimum base cost which can be observed as the number of atoms decreases. As seen in this work, IMoS was allowed to run for 1, 5, and 7 core processors. When IMoS and MOBCAL are compared on one to one core ratio, IMoS seems to be one order of magnitude faster and/or efficient than MOB-

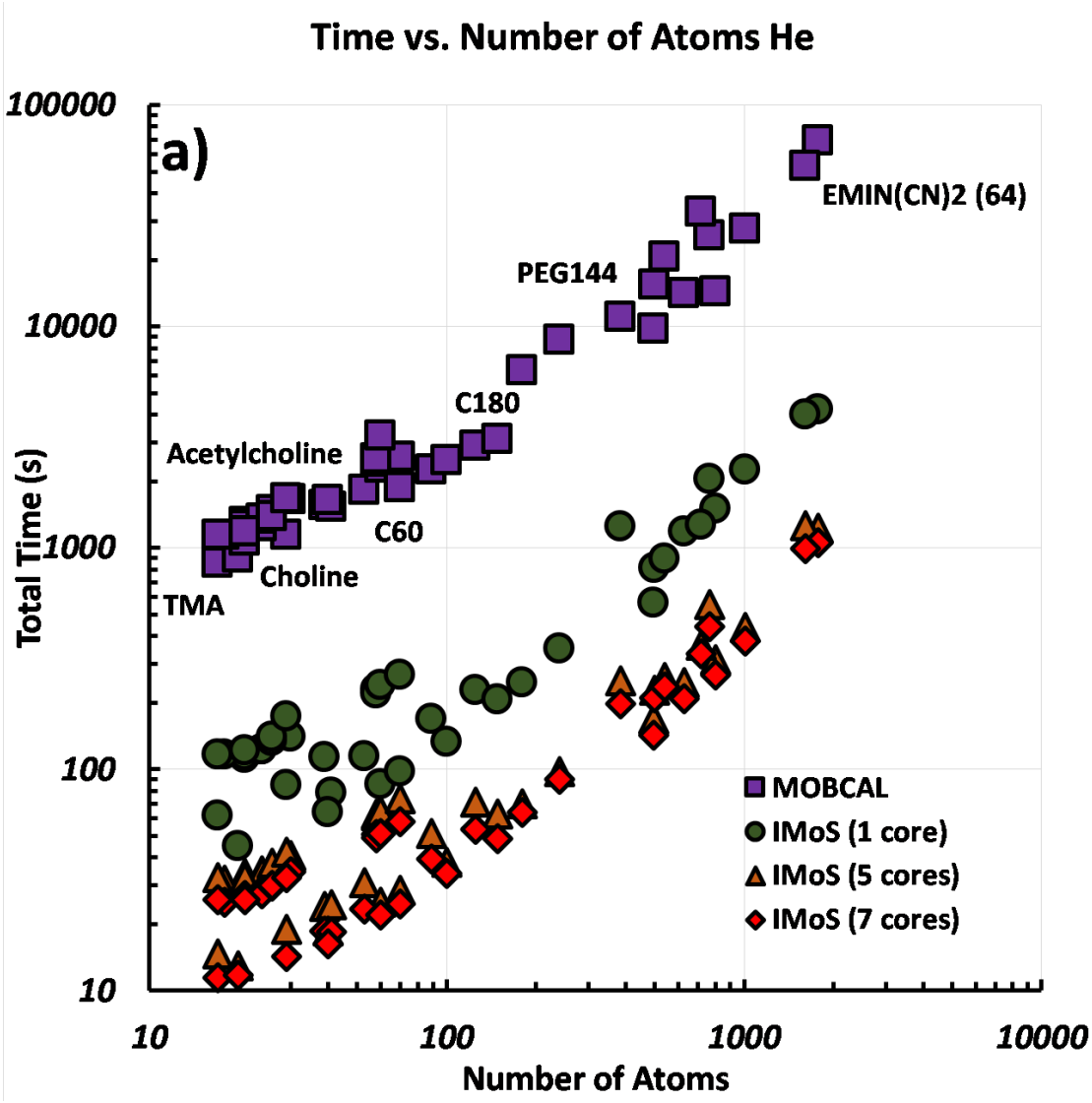

Fig. 3.6. Total time to perform PA, EHSS, and TM methods as a function of the number of atoms for MOBCAL and IMoS (for 1, 5, and/or 7 cores) in He

CAL. The difference can be increased and IMoS can perform more efficiently using 7 core (4 physical plus 4 virtual machines), which increases the performance by two order of magnitude. The increase in speed based on a number of core is not linear but significantly close. These calculations are performed in the free molecular regime and the calculation is performed individually for each gas molecule leading to linear

increase. Again, this calculation can be easily parallelized by IMoS. With that said, one can see the difference to increase several folds when 16, 24, or higher number of cores are used. Supercomputing cluster provides an excellent tool to calculate for large molecules (> 1000 atoms per molecule). Appendix B shows the time calculation in He. Figure 3.7 displays the time calculation for IMoS (calculated for 7 cores) and

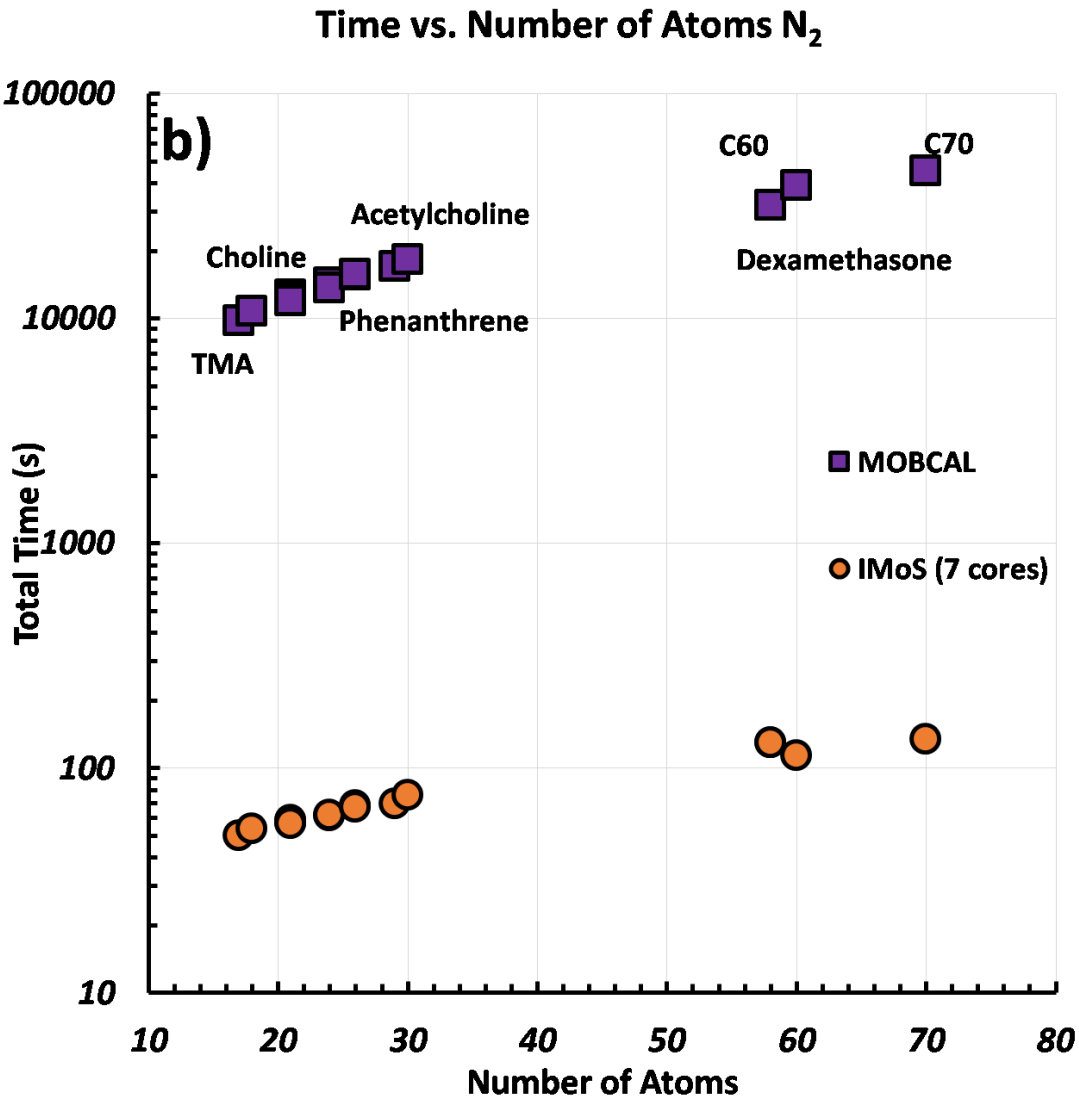

Fig. 3.7. Total time to perform PA, EHSS, and TM methods as a function of the number of atoms for MOBCAL and IMoS (for 1, 5 and/or 7 cores) in N2.

MOBCAL in diatomic N_2 as the function of number of atoms. From the above graph, it can be observed that the increase of time in seconds with respect to the number of atoms is linear. As stated earlier, in case of N_2, the ion-quadrupole potential was established and used in calculation, which dramatically increased the computational time and becomes burdensome to use MOBCAL. While IMoS still requires approximately 100 seconds for the 70 atom structure, the computational time for MOBCAL is already in the tens of thousands of seconds for tens of atoms precluding a reasonable comparison of both calculations for structures that are larger than a hundred atoms. Time calculation for the structure are attached in Appendix B.

3.3.3 Simplification of Methods for Faster Calculations

As discussed earlier, the Trajectory Method requires more time for calculating CCS. Although IMoS has an advantage of parallelization, performing a full Trajectory Method is an arduous task for large biomolecules (having tens of thousands of atoms in its structure) and will require huge amount of time for calculations in the N_2 bath gas. A couple of assumptions can be put forth in order to simplify the calculation process. These simplifications generally depend upon the behavior of the structure (atom and/or charge)-gas molecule interaction. If we take the case of a larger molecule, the interaction between the gas molecule and all the atoms becomes less significant. This is due to the fact that the atoms located at or near the center are surrounded by other atoms, which do not allow it to contribute its effect in the collision with gas molecules. Now, with this known fact, we can substitute a reasonable assumption for such conditions. The assumption also alters according to the use of gas molecule in the calculations and some simplifications can be made leading to a general correction. This process can either be done by adding different contributions or by substituting these contributions for a more general parameter. Such type of simplifications mentioned below are implemented in IMoS and will be tested and compared.

EHSS can be a good substitution for a monoatomic gas such as He, which is in good agreement with the Trajectory Method. In the case of Helium, the attractive component of the Lennard-Jones potential is small. It means that the depth of the well is less and decays at the rate of $\frac{1}{r^6}$ as the gas molecule moves away from the atom/charge. The repulsive component is still active and is large when a gas molecule collides with the atom. These two effects can easily be substituted using a slightly larger atom and considering hard sphere potentials.

PA: The Projected Area method can also be a fair substitution for smaller molecules in He as seen from figure 3.8, which seems to work well and is in good agreement with TM. Shown in figure 3.8 is the ratio between TM and EHSS and TM and PA for several structures in He ranging in tens of thousands of atoms. Note that at large sizes, the PA by itself cannot yield good estimates while EHSS still remains an acceptable solution even at tens of thousands of atoms. Although seen from the figure 15, for larger molecules, PA does not yield a good estimate with respect to TM as the ratio of PA and TM extends to 1.35, EHSS provides a good estimation even for tens of thousand of atoms. The reason behind this observation lies in the scattering effect that occurs in a larger molecule which is counted by the EHSS method and not by the PA method. This observation can lead us to the conclusion that the EHSS method can be a better simplification for TM having the ability to work faster and produce precise CCS measurement as that of TM for the larger molecule in a Helium environment. An error incorporated in approximation is just as significant as the error in experiments. The sound agreement between EHSS and TM in He is shown here is well known, and empirical and semi-empirical laws have been put in place to relate the two calculations [4, 35, 76]. As described earlier, EHSS offers a good alternative to TM in He, but fails in the case of heavier diatomic gas molecules such as N_2 or Ar even for small atoms (although it seems to asymptotically converge to the TM CCS for clusters larger than 30000 atoms).[4] Noted from the figure 3.9, for ions upto 70 atoms, the TM/EHSS ratio is higher upto 1.8 even for the smaller molecule. This later drops down to 1.25 and 1.5 for larger molecules. The large difference observed

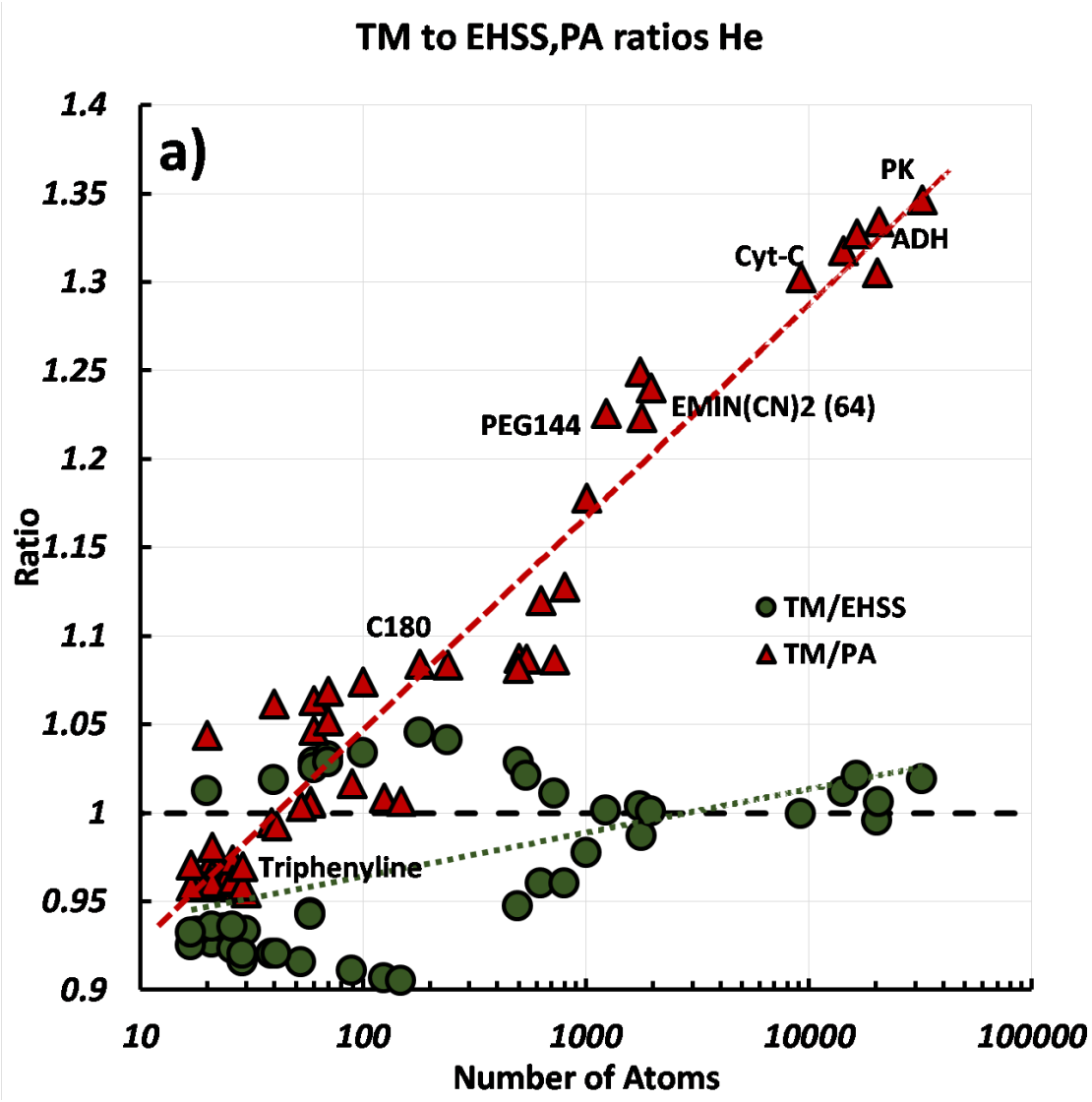

Fig. 3.8. Ratio of TM/EHSS (green circles) and TM/PA (red triangles) CCS vs. the number of atoms in He

for the smallest ions (up to 30 atoms) can be attributed to the ion-induced dipole potential. This potential rapidly decays and should only account to approximately 1% for the 70 atom molecule (C70 fullerene), [39] where there remains an observable difference. This observed discrepancy in heavier gases can arguably be related to the remission of the impinging gas molecules being diffused instead of specular gas molecules (when considering frozen ion structures), as noted by Epstein in 1924,[46]

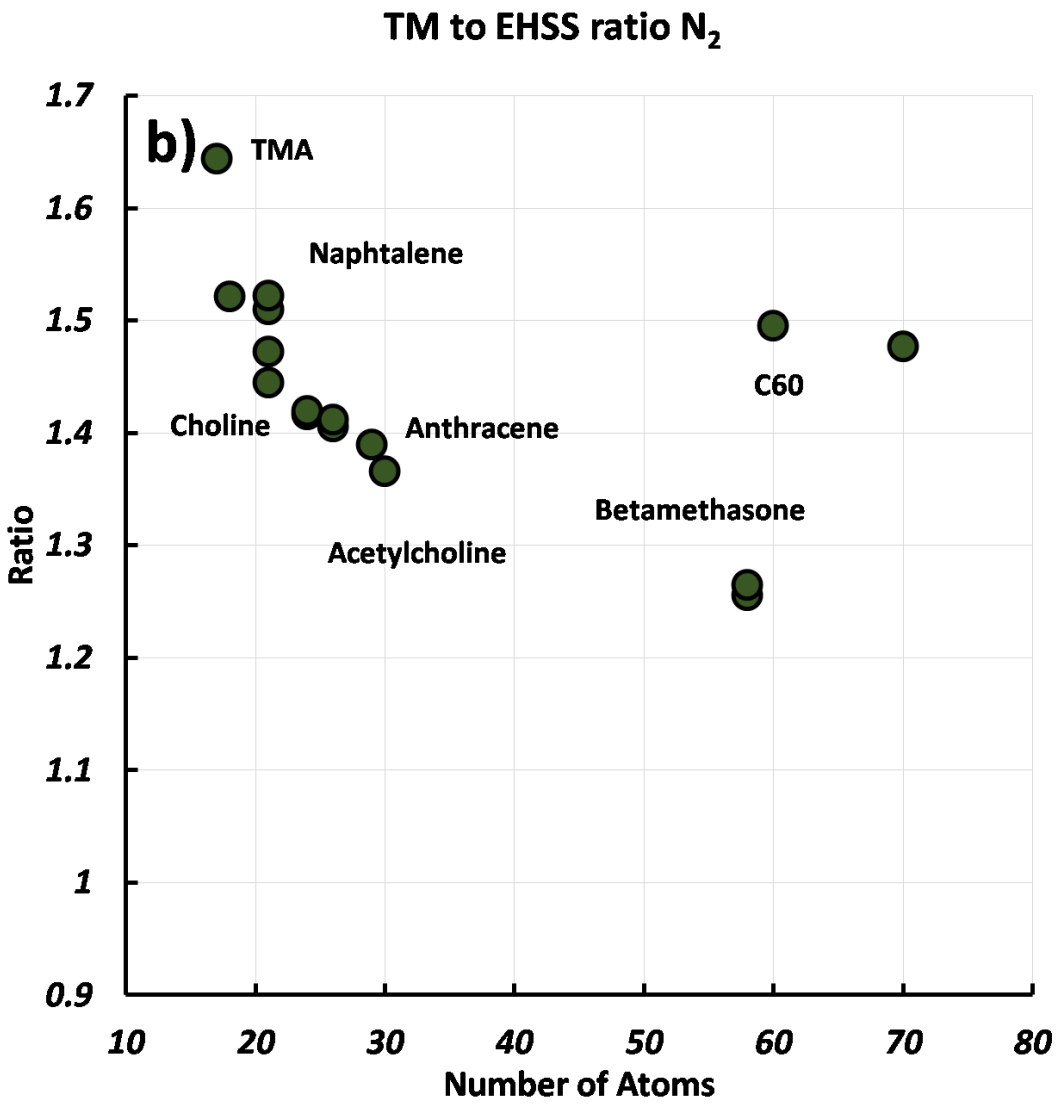

Fig. 3.9. Ratio of TM/EHSS CCS as a function of the number of atoms in N_2.

precluding the possible use of EHSS in heavier gases.[4, 38, 39] The diffusing nature of the remission could either come from the physical action-reaction transfer caused by a much heavier atom (N_2 vs He) on a vibrating atom, which could not directly be explained by the TM method since it does not consider translation/vibrations of atoms, or it could come from the effect of multiple Lennard-Jones potential wells (much stronger than in the case of He) of collision-adjacent atoms that re-orient the gas molecule in a non-specular direction. Whether it is an effect of physical momen-

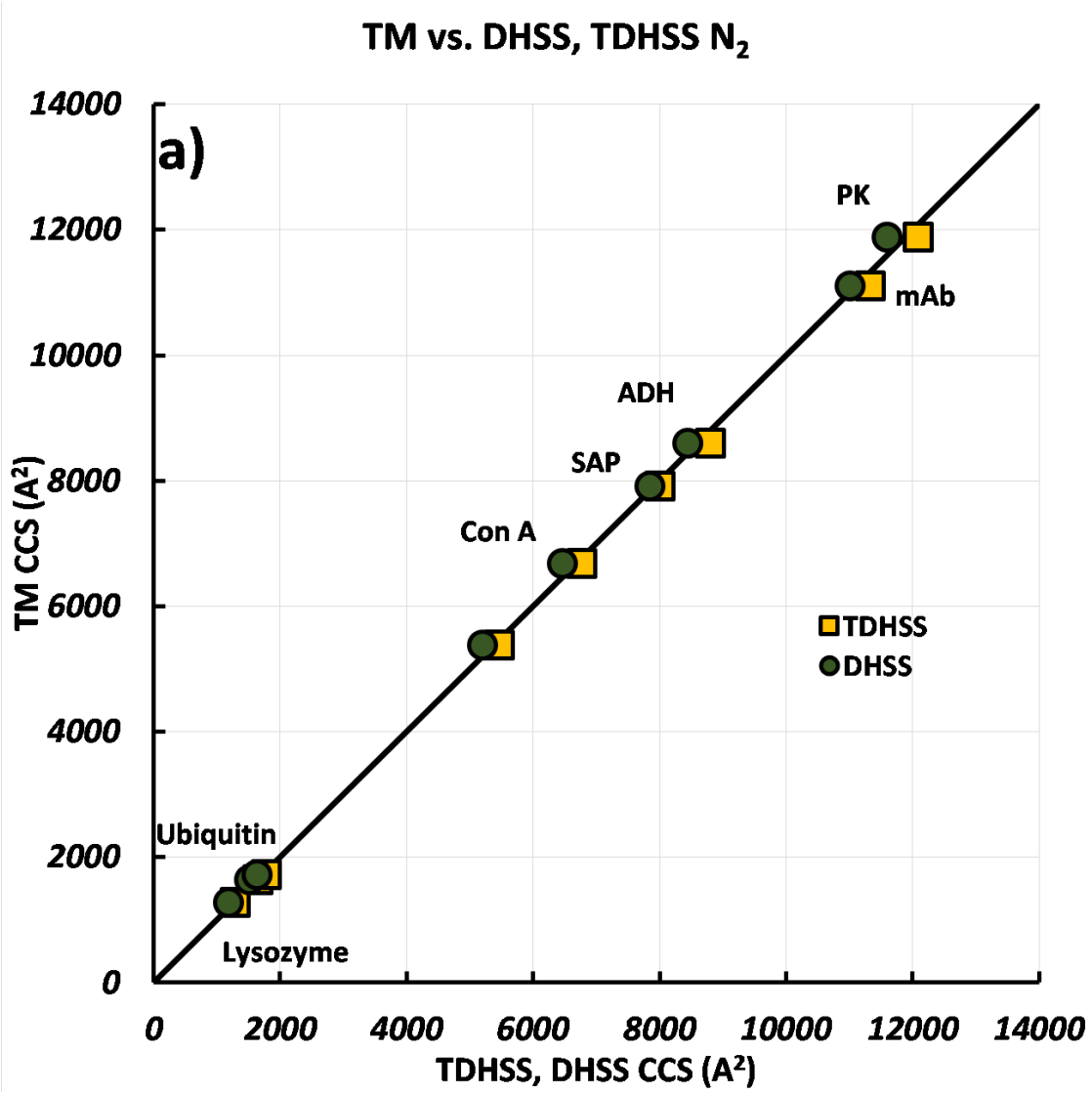

Fig. 3.10. Comparison of calculated CCS between TM and diffuse scattering methods with (TDHSS, yellow squares) and without (DHSS, green circles) ion-induced dipole potential for multiply charged relatively large molecules (>1000 atoms) in N_2.

tum transfer or just a potential interaction or a mixture of both, an approximation could tentatively be proposed instead of the full Trajectory Method for larger ions. The most logical approximation that one can immediately apply is to consider the remission to be completely diffusing (DHSS) using a slightly modified velocity distribution. This modified distribution in IMoS is chosen in this case so as to match the

constant value inferred by Robert Millikan in his oil-drop experiments of 1.36.[52] This new simplification will help to reduce the calculation time and number of operations at every iteration as there will be no need for calculating the Lennard-jones potential parameter between gas molecules and atoms. To demonstrate the performance of such simplifications, an attempt was made to calculate the Collision Cross Section for the above explained simplified methods for larger molecules (> 1000 atoms) and compare them. Figure 3.10 shows the comparison for DHSS, TDHSS (considering ion induced dipole potential), and Trajectory Method in Nitrogen. With the given simplification, the results obtained from considering ion-induced dipole potential is remarkable and within a few percentage of error for large sizes of molecules. Another condition exists for calculations using partial charges on the atoms of the structure of interest. It has been observed that it considers the high computational cost for the calculation of ion-induced dipole potential between every partial charge on an atom and the respective gas molecule. In such cases, interaction potential has a minimal effect on the gas molecule and decays as the gas molecule go away from the center of the atom, provided that all directions are equally probable and the small dipole effect will cancel out. An ingenious simplification can be made in this case for the larger molecule. We can locate the overall charge on the molecule either in the center (again, keeping in mind the structure should not be hollow) or locate the charge on appropriate charge carriers, keeping the rest of the atom in neutral condition. Figure 3.11 shows the difference in CCS between applying partial charges (taken from Amber forcefield) or applying the charges directly at the center for large biomolecules. As expected, the error associated with not applying partial charges to the atoms is negligible ($< 2\%$) for such large molecules. With this huge simplification, it is necessary to see the efficiency (time required) of the calculations.

Figure 3.12 demonstrates the comparison for Trajectory Method (not using ion-quadrupole potential) and the TDHSS method for both with and without partial charges. 16 cores were employed for the calculation. As the effect of ion quadrupole potential is minimal for a larger molecule, it adds to the high computational time.

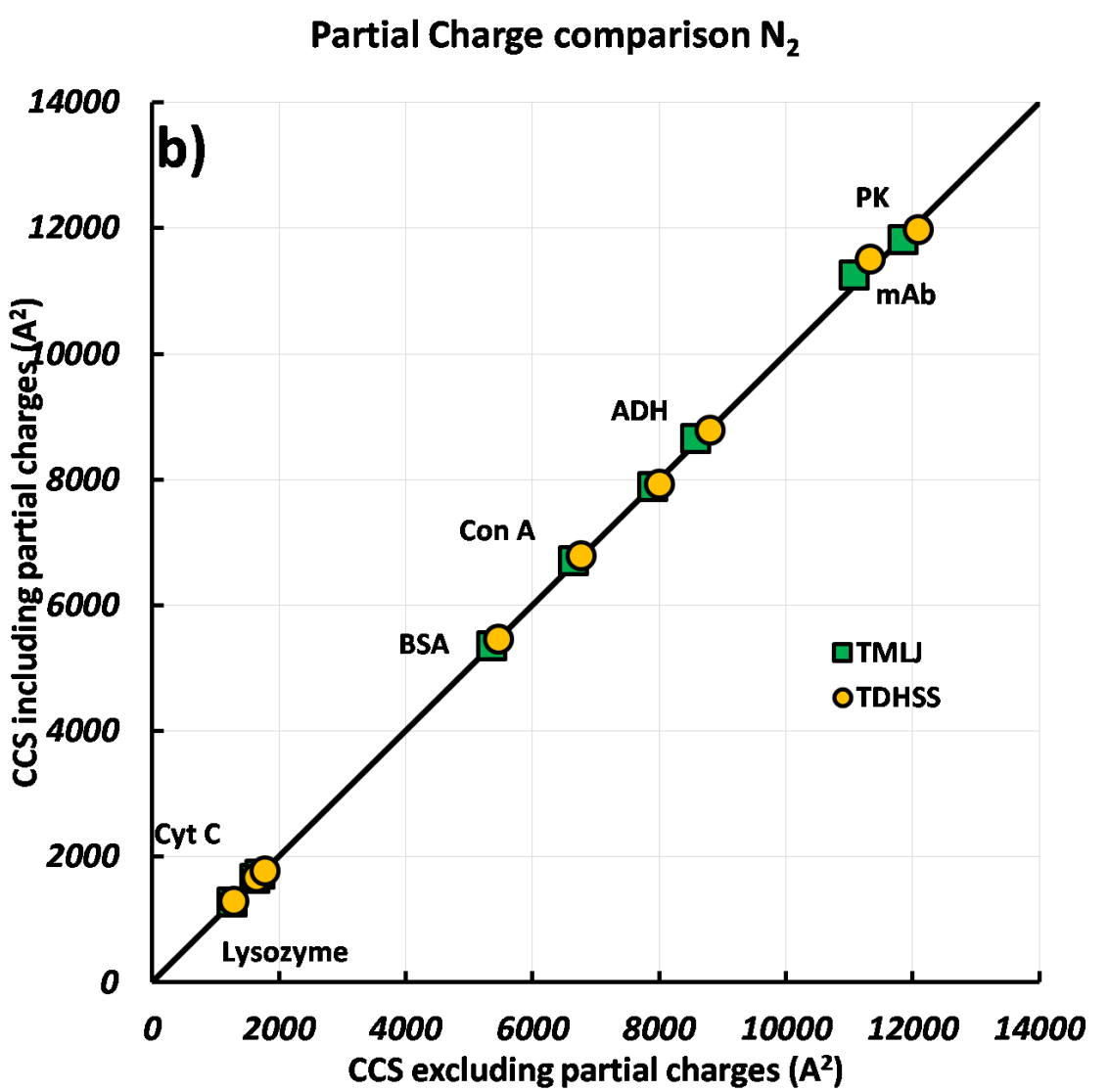

Fig. 3.11. Comparison between assigning partial charges and centering the absolute value of the charge in the center or specific ions. This comparison is performed for the regular TM method (TMLJ) and for the Diffuse Scattering Trajectory Method (TDHSS).

A reduction by a factor of 30 was observed when TDHSS without a partial charge is considered for calculation. This is significant for Protein Kinase (PK protein) where the time is reduced from several hours to around 15 min.

The process of simplification can be made much easier if one knows the contribution of ion-induced dipole potential and enhancement of CCS due to the scattering effect.

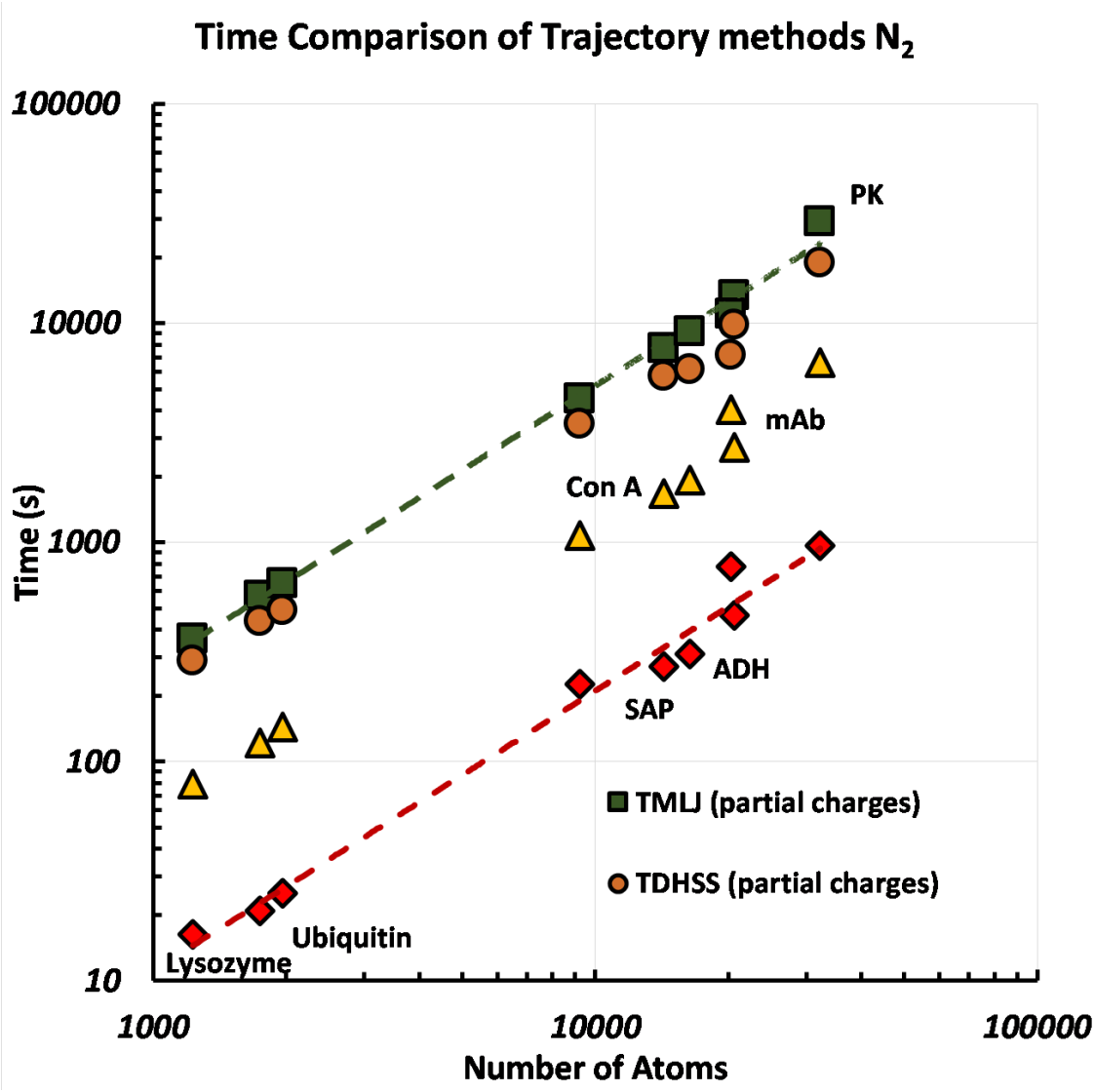

Fig. 3.12. Time comparison between different approximations of Trajectory Methods as a function of the number of atoms.

Recalling equation 3, $\Omega = L\xi PA$, one can just calculate the average Projected Area and multiply by "L" (correction factor to take into account the ion induced dipole potential) and ξ (is the reemission enhancement) to get the true Collision Cross Section. Remission enhancement is also referred to as accommodation coefficient and the value is supposed to be constant around 1.36 in case of Nitrogen [52] but needs

to be derived experimentally for Helium. The "L" can be approximated for a sphere in Nitrogen by the following equation:

$$L \simeq \left[1 + A\varphi_e\left(1 + \frac{1}{\xi}\left(\frac{5}{22} + \frac{5}{7}\varphi\right)\right)\right] \qquad \varphi_e \leq 1 \qquad (24)$$

where A=(1/5,3/5) is a numerically derived parameter selected to be 2/5. Here, is chosen to be 1.36, and $\varphi_e = U_{pol}\frac{\left(\frac{d_p+d_g}{2}\right)}{kT}$ denotes the polarization potential $U_{pol} = \frac{\alpha Z^2 e^2}{8\pi\epsilon\epsilon_0 r_u^4}$ evaluated at the surface $r_i = \big(d_p + d_g\big)/2$ assuming that all charges are in the center. In the case of non-spherical particles,(dp+dg)/2 can be approximated by the value $\left(\frac{PA}{\pi}\right)^{\frac{1}{2}}$ as long as the atom structure does not deviate too much from a globule. This semi-empirical equation is formed by a correction due to direct impingement of molecules plus a correction due to grazing molecules (those inside the factor $\frac{1}{\xi}$) that do not directly impinge but still transfer momentum. Figure 3.13, shows the comparison for the approximation made in equation 3 and Trajectory Method for the larger molecule. This approximation works well for room temperature conditions and holds good even in slight change in temperature. When we analyze this graph, it can be seen that the calculation time has reduced to minimum for about few seconds because now you need to only calculate the average Projected Area for tens of thousands of atoms. Many of these simplifications are applicable for coarse grained molecule, as most of the simplifications applied inherently takes into account for coarse grained structures. [77] TDHSS without partial charges or the simplification used in eq. (12) should yield very similar results whether applied to an all-atomic structure or to a simplification of the same structure using only approximations of major components.

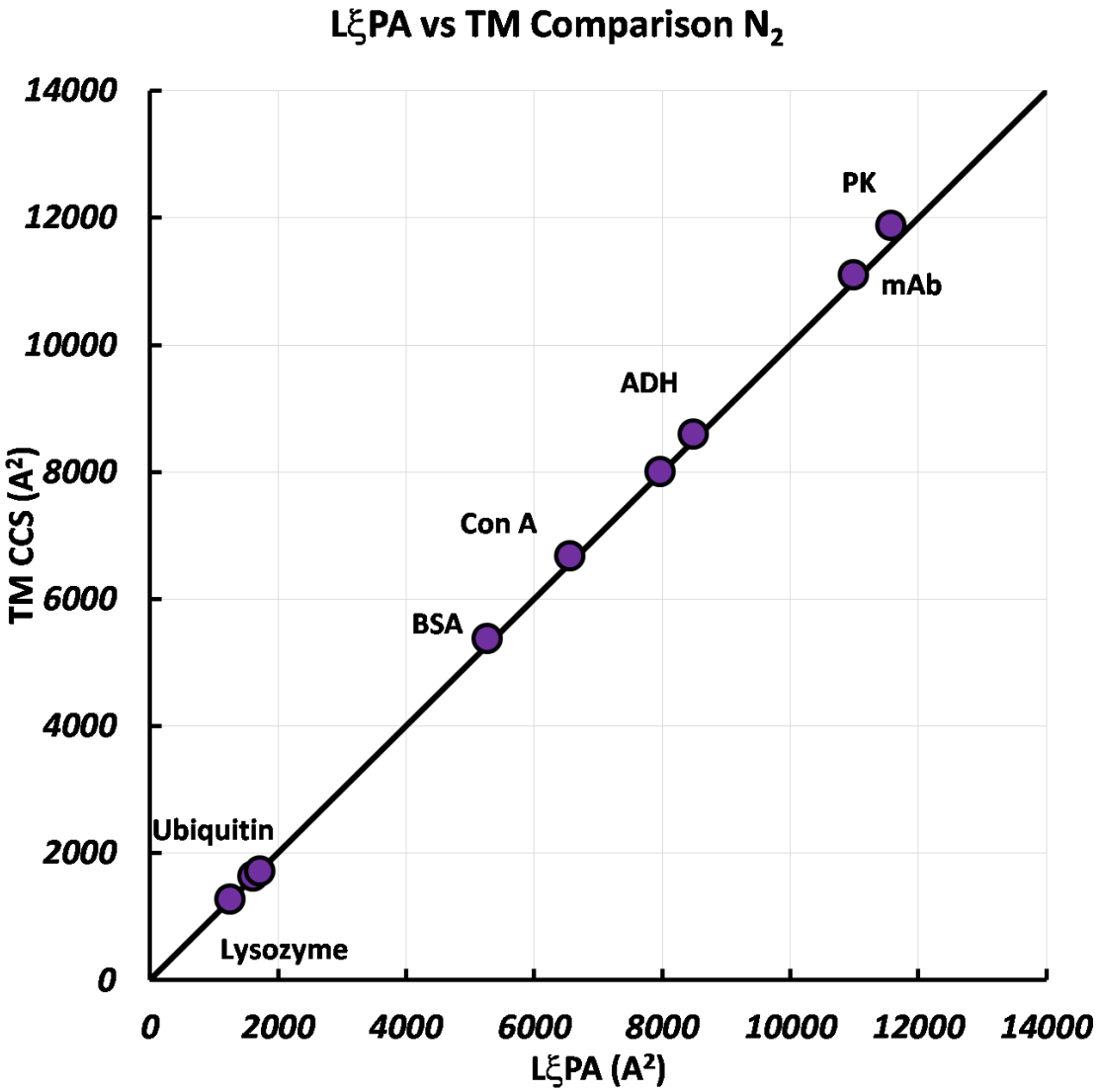

Fig. 3.13. Relation between the theoretical approximation L ξ PA and the TM method using the correction from equation 3

4. OPTIMIZATION OF LENNARD-JONES POTENTIAL PARAMETERS FOR TRAJECTORY METHOD

Optimization of the Lennard-Jones potential parameter (σ: Distance (r) between two interacting bodies; in this case, ion structures and gas molecules, where potential energy becomes positive and ϵ: Depth of potential well) is a necessary step to determine relatively similar Collison Cross Section to that of experimental values. Thus, the Lennard-Jones potential parameter needs to be identified correctly in order to calculate Collision Cross Section using the Trajectory Method. Optimization is the way to determine the close values for these parameters for which the selected function (consisting of the Collision Cross Sections for both experimental and numerical) through the Trajectory Method using the Lennard-Jones potential parameter, approaches to a minimum (close to zero). These values are then considered to be the global minimum for a given function where experimental Collision Cross Section are predetermined. Mobility calculator in free molecular regime uses different sets of the Lennard-Jones potential parameter. Later, they were re-optimized for different bath gases and used for the calculations.[84] Even though these optimized values gave a close prediction for the Collision Cross Section, it has to be considered again for optimization through numerical methods, so as to get a clear picture of how these parameters affect the Collision Cross Sectional area. This study provides a better understanding of the interaction of potential parameters with the numerical measurement of Collision Cross Section. In this section, we will go through the methodology developed for optimization, the algorithm used, outcome, and inferences of obtained results. We will discuss the minimum obtained for the given function and criteria to obtain a set for the Lennard-Jones potential parameter for Carbon (C), Hydrogen (H), Nitrogen (N), Oxygen (O), Fluorine (F).

4.1 Methodology

In order to perform optimization, we require a function that we will minimize. Here, a function consists of a set of the Collision Cross Section for a number of chemical compounds, whose experimental values are known beforehand.

$$F(\epsilon,\sigma) = \left(1 - \frac{\Omega_1(IMoS)}{\Omega_1(EXP)}\right)^2 + \left(1 - \frac{\Omega_2(IMoS)}{\Omega_2(EXP)}\right)^2 + \left(1 - \frac{\Omega_3(IMoS)}{\Omega_3(EXP)}\right)^2 + ... + \left(1 - \frac{\Omega_{n-1}(IMoS)}{\Omega_{n-1}(EXP)}\right)^2 + \left(1 - \frac{\Omega_n(IMoS)}{\Omega_n(EXP)}\right)^2 \quad (25)$$

In the above equation, Ω represents Collision Cross Section values for IMoS in the numerator and experimental values in the denominator. Each term in the equation is squared in order to get rid of any negative number. The above equation satisfies the condition for reaching the minimum. As we can see, when the numerical value (IMoS value) for Collision Cross Section reaches close to or equal to the experimental value, the term will approach zero giving minimum for the function. All the terms are finally added together to give the global minimum value. The terms in the above equation can be considered according to the users specification. It truly depends on a number of experimental results we have, so that we can generate the same number of IMoS Collision Cross Section values and check for the function.

We choose 16 molecules/structures for calculation. N-ethylaniline, acetaminophen, tetramethylammonium, tetraethylammonium, choline, acetylcholine, naphthalene, phenanthrene, anthracene, pyrene, triphenylene, C60, C70, Dexamethasone, Betamethasone, and Paracetamol are the structures for which experimental values are available [71]. The experiment performed by Campuzano was done using RF confining drift tube. Compounds such as N-ethylaniline, acetaminophen, tetramethylammonium, tetraethylammonium, choline, and acetylcholine were infused in 1 M in 49% (v/v) acetonitrile/water and 0.1% (v/v) formic acid, whereas compounds such as naphthalene, phenanthrene, anthracene, pyrene, triphenylene, C60 and C70 were infused in 100% toluene. [71].

The main principle of this method involves plotting a surface for the values of the

function with respect to the values of σ and ϵ. This generated surface will lead us to the inference of optimized values. We developed the algorithm explained in a later section, to produce multiple Collision Cross Section using multiple σ and ϵ values in order. In the plot ϵ, σ, and F are represented on X-axis, Y-axis, and Z-axis respectively. The bound for which the values of parameters should be taken is the important question to be considered.σ and ϵ can be considered within the specific range. For our initial start, we considered a wide range of values so as to get a complete plot of the surface. This selected range covered most of the values used till now by people in various algorithms. The plot shown in Results section, give a clear idea about the minimums obtained throughout the values. For every value of ϵ, the program runs for a set of given σ and the loop repeats itself for a given number of iterations. In this way, we can cover a wide range of these parameters and calculate Collision Cross Sections to plot the function. Once we have the surface ready, a user can select their own range of this parameter to get rid of other unnecessary time-consuming calculations. IMoS provides an excellent tool to calculate the Collision Cross Section of the structures/molecules in less time as compared to another algorithm for mobility and Collision Cross Section calculation. As discussed in the earlier section, IMoS provide an accurate prediction for Collision Cross Section numerically, for the rigorous Trajectory Method using Lennar-Jones potential parameters. As we need to do thousands of calculations for multiple values of σ and ϵ, IMoS provides a speedy calculation under such conditions.

Optimization can be performed for different bath gases such as N_2, He, air etc, because the value for Collision Cross Section depends hugely on colliding gases. Here we tried to optimize the parameters for N_2 as this bath gas is generally used in the field of Ion Mobility Spectrometry in studying collision effects on the different type of chemical structure. Potential parameters for basic elements (C, H, N, O, F) were initially considered for optimization. We choose C, H, N, O, and F as the order of optimization. This is done as C is found in abundance in almost all chemical structures. Hydrogen is the next element found in abundance. Later O, N, and F were

considered for optimization, using the optimized value of previous elements. Once the first round of optimization for all the five elements are completed, we go for the second round using all the new optimized values for all the elements considered. The process of iteration will be continued till the function converges to minimum for the Lennard-Jones potential parameter for all the elements giving a minimum of a function. Once we arrive at the optimized values for these five elements, we can use these values to optimize further elements provided we have experimental Collision Cross Section for a given structure containing the element we need to optimize. In a similar manner, parameters can be optimized for other bath gases. There are multiple ways this optimization process can be done. One can start the code to optimize Carbon for Fullerene structure, which is composed of only carbons in its rigid state. After getting the minimum value for Carbon, Hydrogen optimization can be performed by taking Hydrocarbon structure into consideration. Later, optimization for O, N, and F can be performed step by step considering the structures containing O, N, and F respectively. One might consider this as a suitable option, but it has two drawbacks:

- Experimental values may not be available for all the selected structures in a given bath gas.
- If carbon is optimized only considering Fullerenes structures (which contains all carbon atoms) there is a chance we will miss out the effect of the presence of other elements in the structure during calculation which might mislead the Collision Cross Sectional value.

In order to perform fair optimization, we considered structures that do not contain a single element in the structure and whose experimental values are easily available.

4.2 Optimization Code: Explanation

Our aim is to minimize the function ($F(\sigma,\epsilon)$) for n number of structures, for which the experimental values are present. Experimental values for 16 molecules were arranged in the above equation and optimization was carried out. A windows executable

was made for the algorithm on MATLAB compiler runtime 901 with another input text file for the program. This text file contains user defined first and end values of σ and ϵ. This value depends upon the range of parameters for which the user wants the program to work and will be read accordingly by the program. Step size is included in the algorithm to advance the values of the parameters. This step size is again a user defined value and can be increased depending on the accuracy required. Generally, the accuracy of not more than four digits after decimal is recommended as it would not affect the Collision Cross Sectional area. In this work, the initial step size was considered to be 0.1 so as to get the largest possible surface. Once the line of the minimum is obtained, the accuracy can be increased further by considering a smaller step size. Once the values for the parameter and step size is declared, the program is ready to execute.

After setting the parameters, the program goes into the for loop. Two for loops are used for ϵ and σ respectively. For every ϵ value, the program considers a set of predefined values for σ in a given step size and generates the Collision Cross Section. Every pair of ϵ and σ is then supplied to the LJ table Excel file which is input for IMoS executable. The System command is then used in order to call the IMoS executable and values for Collision Cross Section are generated. In the next step, it considers another set of parameters and the process of calculation is continued throughout the loop. For every iteration, the IMoS.cla file is updated and the output file name is changed accordingly. This helps us to keep track of the number of outputs IMoS generates. These output files are saved and can be used to cross-check the Collision Cross Section values for the structures. The values of Collision Cross Section are then read through each generated output text file and stored in an Excel file. Later, value of the function (F(σ,ϵ)) is calculated for an individual pair of ϵ and σ by using calculated and experimental Collision Cross Sectional area.

using calculated and experimental Collision Cross Sectional area. Similarly, the code is also compiled on Linux-based MATLAB 2017. The Linux version uses MATLAB compiler runtime 901. As discussed earlier, there are increased amount of calcula-

tions to be performed. As we can see, Collision Cross Sectional area for 16 structures/molecules are calculated for each pair of ϵ and σ. If we consider a large range for ϵ and σ there might be millions of calculation to be performed. As the number of calculations to be performed were higher, the use of supercomputer was a must. We used Indiana University based Big Red II supercomputer system. Working on Big Red II provides the advantage of fast calculation as multiple cores can be used for IMoS for speeding up the calculation. The process is embedded in the flow-chart in figure 4.1.

4.3 Optimization Plots and Discussions

Figure 4.2 shows an interpretation of the optimization technique used to obtain the minimum of the function and its respective values for ϵ and σ. The bowl shaped figure represents the contour for the function at different heights. The bottom of the bowl giving a single point represents the minimum for the function. In order to achieve the minimum of the function, we need to travel on both the planes each representing potential parameters in such a way that the path will lead us to the minimum. As seen from the figure below, we travel along the σ keeping ϵ constant till we get minimum. During this travel, the value for the function will start dipping and at some point it will again start rising notifying us that we have already passed the minimum. Taking a step back and establishing theσ for which we get the minimum of the function, now we start changing ϵ, keeping the established σ constant. In a similar fashion, we go through ϵ till we get the minimum of the function for that height. The process is repeated until we reach the minimum point for the function and then record the values for ϵ and σ.

4.3.1 Round One Optimization

The process of optimization starts with optimizing potential parameters for carbon at first. Before starting the optimization, we need to set the potential parameter

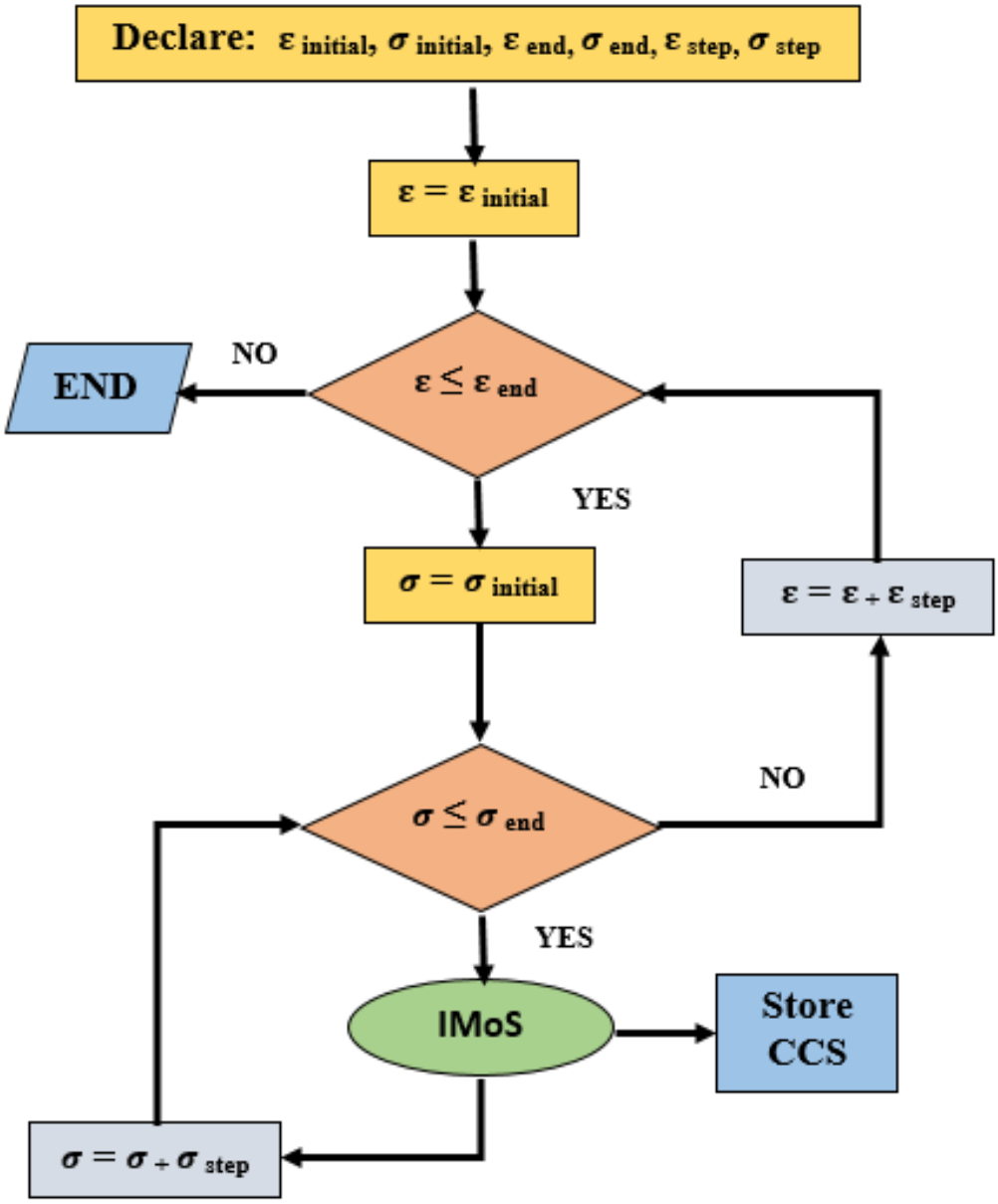

Fig. 4.1. Algorithm layout for optimization process

values of other elements to some values and hence we choose to start with values given by Campuzano et. al. listed in the table 4.1. Carbon is present in abundance in almost all the chemical structures. The experimental structures used here are the ions dominating in Carbon. Hence optimizing Carbon at first was the obvious choice, which will give an idea about the physical appearance of the structure when poten-

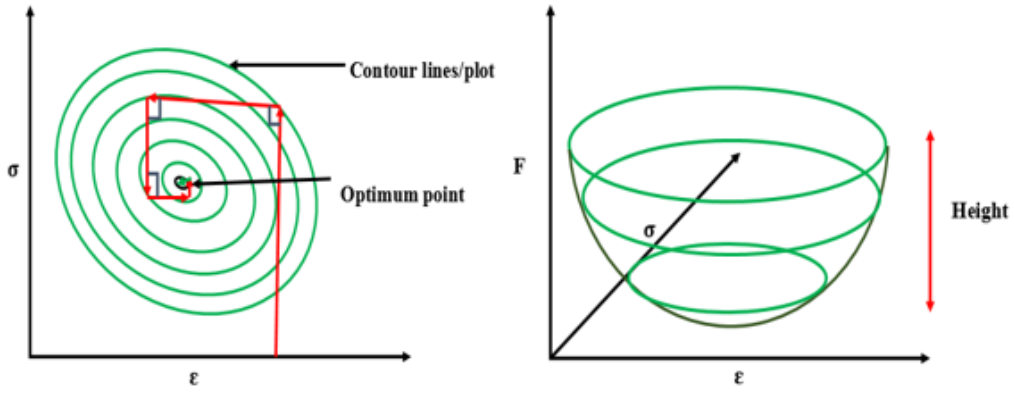

Fig. 4.2. Interpretation of the optimization method

Table 4.1.
Values for LJ potential parameter in N_2

Element	ϵ	σ
H	0.8204	1.2409
C	4.2319	3.5814
O	3.5902	4.3920
N	2.4195	3.2550
F	2.0162	3.1285

tial parameters are modified iteratively. Keeping the above mention fact in mind, it was also necessary to involve other elements such as Hydrogen, Nitrogen, or Oxygen into picture so as to consider the overall effect of these atoms on the gas molecule trajectories. Taking other elements into account will make sure that the process of optimization is carried on a global scale. Placement of the charge on the structure is another important area to be considered as charge will also have an effect on gas molecule trajectories. The structures were singly charged for calculation of Collision Cross Sections in IMoS. Once the structures are wisely chosen, we can start the iterative optimization method.

The values of potential parameters for other elements except Carbon were set to constant values as tabulated in table 4.1. The initial and final ϵ values were taken as 2

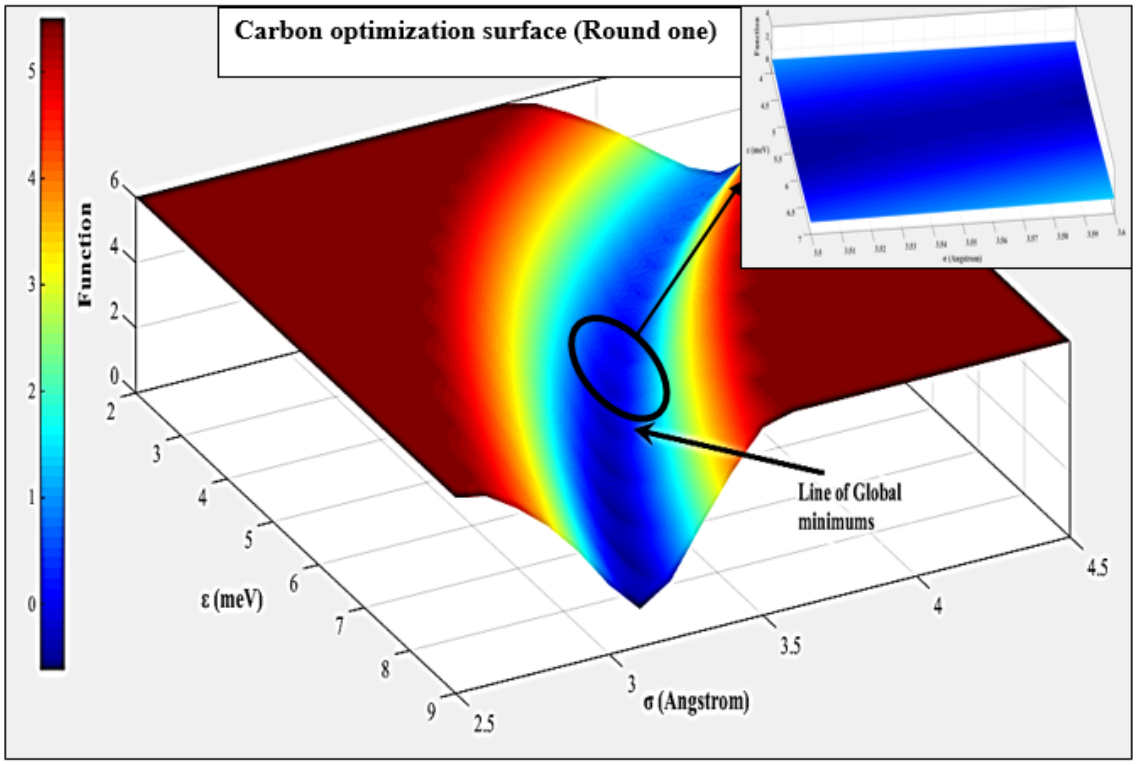

Fig. 4.3. Contour surface for the first round of Carbon optimization.

and 9 respectively. The initial and final σ values were taken as 2.5 and 4.5 respectively and the step size was chosen to be 0.1. We selected a wide range of ϵ and σ to include all possible values for potential parameters used till date to form a contour surface shown in figure 4.3. In this and for all further simulations (unless and until mentioned separately) $9 * e^5$ number of gas molecules were used for the Trajectory Method at temperature of 301 K. The algorithm developed in this work for optimization runs for each ϵ and σ pair to compute the value for the function and is then plotted altogether to give a surface with a line of minimum. The line of minimum thus obtained is the parabolic curve with all the minimums lying along the curve. Although the values on Z-axis are mentioned positive (also, for all the optimization figure in this section), the values for the function are logged to clearly observe the valley for the minimums

and it helps in eliminating the irrelevant data from the graph. This log Z-axis figure in any way does not alter our purpose to demonstrate that the minimum is obtained when the function approaches zero and can be observed clearly in figure 4.3.

As discussed earlier, ϵ is the value which represents the depth of potential well: meaning that this parameter has an ability of attraction. This parameter for any atom in the structure contributes in the attraction for the approaching gas molecule and tends to bend the trajectory of that gas molecule. Similarly, σ contributes in the size of the atom to be modelled computationally. The given value of σ will contribute in the boundary of the effective diameter of a structure, where the gas molecule will tend to collide and will get deflected. Hence greater the value of σ, bigger will be the size of that atom computationally and vice-versa. After understanding the physical meaning of these parameters, it is now easy to understand the nature of the curve. As seen from the curve, higher values of σ yield lower ϵ values for a minimum of the function and the trend continues down the valley of minimums. As ϵ increases up to certain value, the value of σ decreases initially and then it shows very less change even on high increase in ϵ. The curve seems to asymptotically converge for higher values of ϵ giving nearly constant σ (or with a minimum variation). Hence we can say that there is always a counter balance between these parameters to reach the global minimum of the function giving us a well-defined curved valley for minimums.

Once the curve for minimum is obtained, it becomes a necessary step to select the single best pair for ϵ and σ so as to continue the iterative optimization for other elements. The question arises as to what values of potential parameters should be selected from the bunch of pairs, which gives the minimum for the function. We thereby came up with an idea to fix one of the parameter by calculating it externally (not from optimization scheme) and then select the other parameter corresponding to the fixed parameter. With that said, we choose to fix σ by calculating it through summation of Van der Waals radii for Carbon and the colliding Nitrogen atom. The value for σ for carbon is calculated to be 3.52 angstrom units. We performed post processing using MATLAB to collect all the pairs of potential parameters together

which gave the minimum for the function and plotted the graph shown in figure 4.4. The values in the graph are fitted using an exponential curve and the equation is

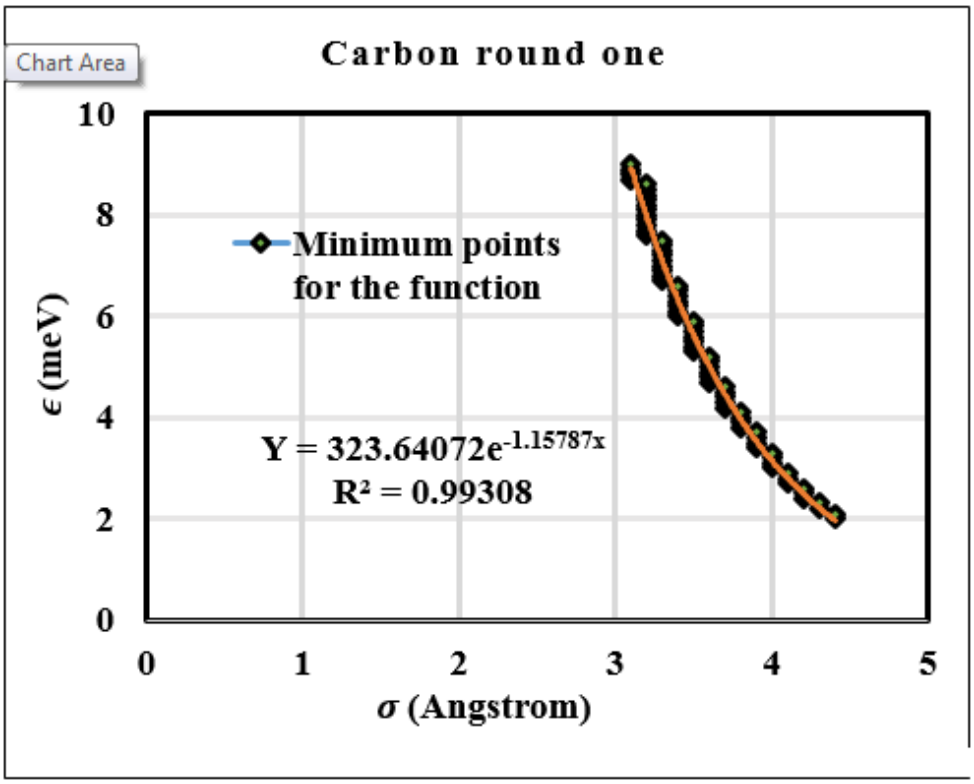

Fig. 4.4. Exponential curve for Carbon first round optimization

displayed on the graph. The regression line is well fitted denoted by coefficient of determination of R-squared equals 0.99308. Using the σ obtained, we determine the valueϵ using the equation to be 5.4929, which is high for a Carbon atom and needs a second round of optimization. The effect of these parameters are important in defining the path of the gas molecule. It now becomes very important to study these parameters to find accurate Collision Cross Sectional values of the structure to avoid any structural misalignment. The figure 4.3 is also cut down to small values as seen from an inset figure so as to compare it with round 2 and 3 optimization and will be recalled later in this section. Slightly dark patches can be observed in the figure 4.3 along the curve. This is due to the fact that the data points are not high enough

to form a smooth surface. A better and a smooth curve can be acquired by taking smaller step sizes and obtaining large number of data points. As this was not the aim of the work, we utilized a step size of 0.1 fulfilling the purpose of our research.

As we have obtained the values forϵ and σ for Carbon atom after first round of optimization, we will move forward to optimize the potential parameters for Hydrogen atom keeping the new values ofϵ and σ for Carbon and other element values from table 4.1 constant in the calculation process. The figure 4.5 demonstrates the valley for the minimum for first round Hydrogen optimization. Hydrogen is the second most

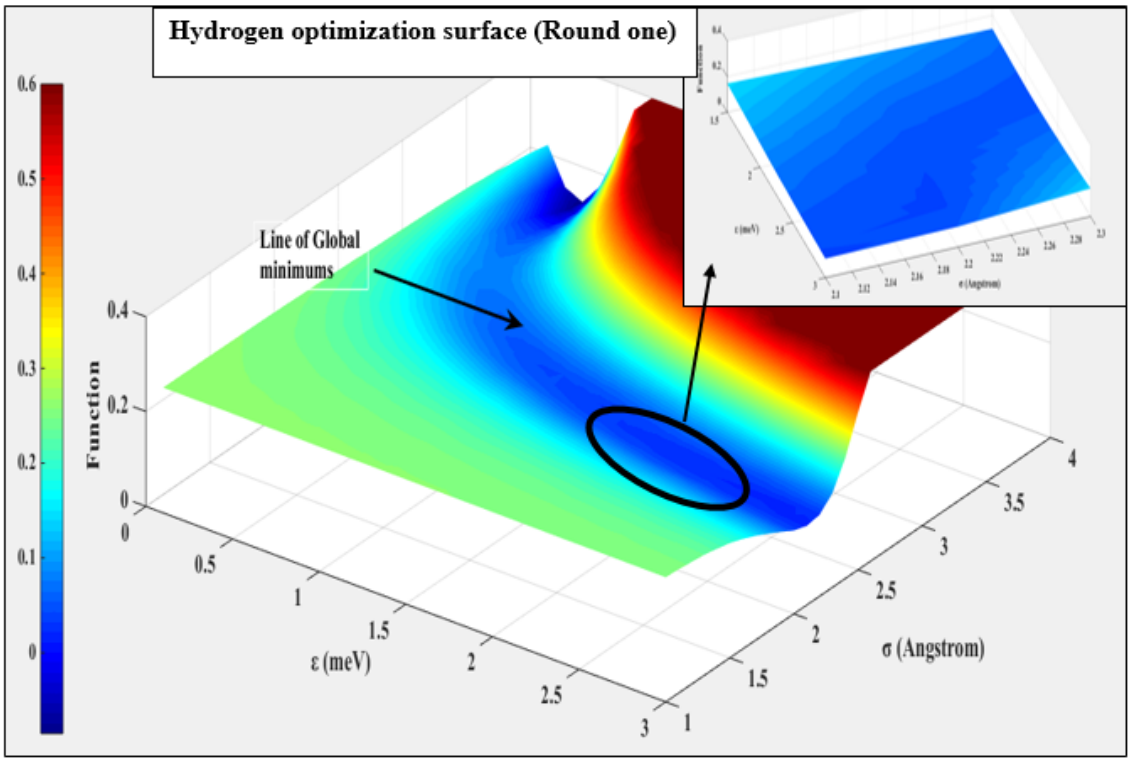

Fig. 4.5. Contour surface for the first round of Hydrogen optimization

fundamental element found in nature. The structure used in this optimization process also dominates in Hydrogen. While trying to optimize the potential parameters for Hydrogen atom, the calculations have already considered the effect of modified potential parameters for Carbon. The range for ϵ was chosen to be from 1 to 3 and for σ was 1 to 4 at step size of 0.1 for both ϵ and σ. Like Carbon, we obtained a

well-defined parabolic curve for Hydrogen as well. As seen from the figure 4.5, the curve is more straight towards higher ϵ meaning that the value for σ does not change very rapidly with increase in the ϵ and asymptotically converges for lower σ. This concludes that the effect of ϵ or the attraction force looses its importance (though for a small interval of ϵ from 1 to 2.5) for Hydrogen in calculation for lower σ values. As noted from the optimization interpretation in figure 4.2, ideally we should get the single optimum point for minimum of the function for both Carbon and Hydrogen. Contrary, we obtained multiple minimum points for Hydrogen and it follows the same explanation as that given for Carbon curve. We can also note from the above figure that there is a sudden shift in values for higher σ for a minimal change in ϵ. This is an artifact occurred during the calculation as values for ϵ less than 0.5 are not acceptable and is not possible computationally, hence we were left with a choice to select the potential parameter pair from the latter half portion for the minimums. We analyzed the values for the minimum in the figure 4.5 The figure 4.6 shows the exponential

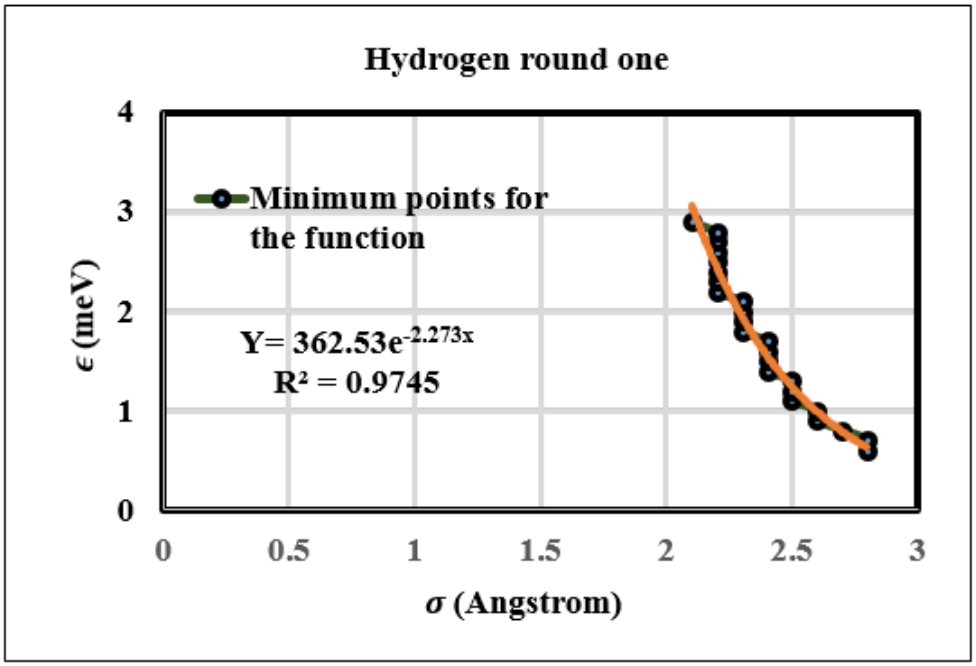

Fig. 4.6. Exponential curve for Hydrogen first round optimization

curve with R-squared value equals 0.9745 which denotes the good fit through the points. Now, the question is to what value for potential parameter should be chosen for next optimization step. We utilize the same idea of using Van der Wall radii and obtained σ to be 2.92 and ϵ was calculated to be 0.4751 from above equation. The calculated ϵ was very less and lie in the region of the artifact created in the figure 24, hence we discarded this value. Instead, we utilize the exponential equation to find the value for ϵ for 2 to 2.92 σ. This range of σ was selected considering the valley of minimum lies within this range. We calculated all the ϵ for the given range of σ and selected the ϵ=2.33 and σ=2.22 , which gave the least minimum for the function.
It is very important to note that the values of Collision Cross Sections will be nearly the same for all the pair of potential parameters that are lying along the curve of minimums. So no matter what values for potential parameters you select from the minimum, you should arrive at the same value of Collision Cross Section. Again, an inset figure is provided so as to compare it with round 2 and 3 optimization and will be recalled later in this section. After establishing the values for potential parameters for Carbon and Hydrogen, we now move forward to optimize potential parameters for Oxygen using the newly modified values for Hydrogen and Carbon and keeping other element values constant. We choose the range for ϵ to be from 1 to 4 and for σ from 2 to 4 at step size of 0.1 for both ϵ and σ. Figure 4.7 shows the minimum for optimization for Oxygen. The curve follows the trend of Hydrogen and Carbon giving a well-defined curve converging toward lowerσ. The graph in figure 4.8 demonstrates the exponential curve for all the minimums of the function for potential parameters with R-squared equals 0.9897. Van der Waals radii calculation gave the value of σ to be 3.34 and the value calculated for ϵ using the exponential equation is 4.584. An important discussion to be made is about the placement of the atoms in the structure, which will make the difference in the calculation and is an important concept to be considered for better understanding the physics of potentials and its application. Imagine you have a structure having a particular atom, say Nitrogen, in the center of the structure for which we have to perform the potential parameter optimization.

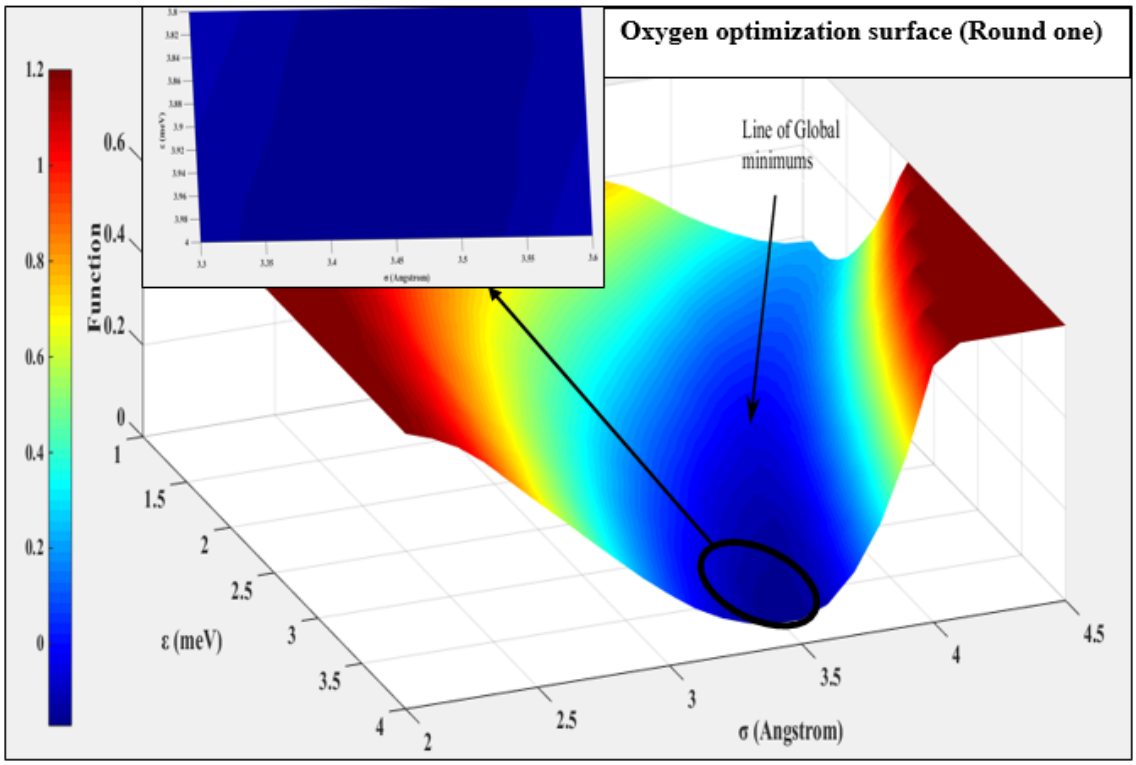

Fig. 4.7. Contour surface for the first round of Oxygen optimization

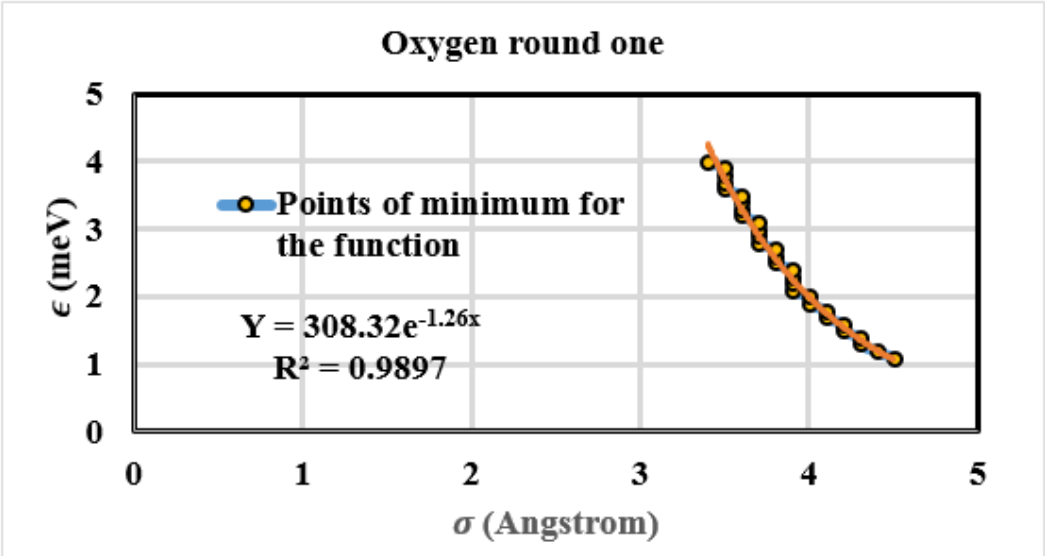

Fig. 4.8. Exponential curve for Oxygen first round optimization

Now consider this atom is surrounded by multiple other atoms forming a structure and is put forth for CCS calculations. As the Nitrogen atom is placed in the center, it will have bare minimum effect on the incoming gas molecule during collision irrespective of the variation in potential parameters. During the close approach, the colliding gas molecules most of the time will be influenced by other atoms (or charges) placed on the surface and not by the core atom as it is far away from the point of collision. Now, if we place this Nitrogen atom on the surface, it will interact on a major scale with the incoming colliding gas molecules resulting in the change of Collision Cross Section. Along with the consideration of the placement of the atoms, we also need to consider the placement of the charge(s). The sizes of the atom is another factor to be considered in the calculations. The larger the size for the atom, greater will be the effect of the potential parameter on the gas molecules. Structures considered here are already studied and the experimental CCS are well established giving us correct information about the co-ordinates of the atoms for each structure. The calculation performed in this optimization method takes into account the overall effect of all the atoms in the structure giving us a good approximation of the minimum function values. Only $\sim 35\%$ of the total 16 structures contained Nitrogen as an atom, which makes the optimization difficult. Due to this lack of Nitrogen molecule in the structures, the curve for the minimum is not observed in figure 4.9. We choose the range for ϵ to be from 2.5 to 6 and for σ from 3.5 to 5 at step size of 0.1 for both ϵ and σ. According to Van der Waals radii calculations the value for σ was about 3.37, which was no the minimum according to figure 4.9. Therefore, to find the pair of potential parameters, we choose to select the least minimum value among the pool of minimum values of the function. The values thus determined were: ϵ =5 and $\sigma = 4$. We then start optimizing with the Fluorine atom and the surface is displaced in figure 4.10. The plot clearly demonstrates no sign for the curve of minimums. The reason for this is already explained. Only two structures: Betamethasone and Dexamethasone, contains Fluorine atom in their structure giving us an inclining planar surface with no curve. From the values of the Z-axis (which is not logged in this case), it can be

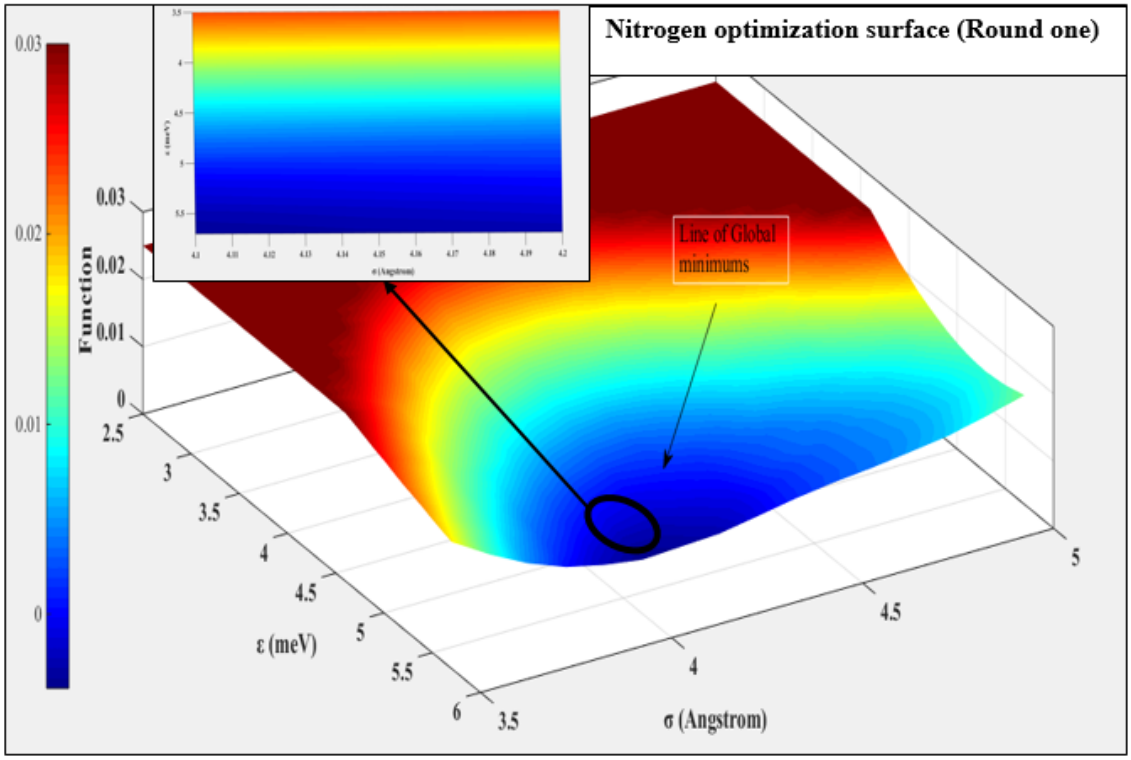

Fig. 4.9. Contour surface for the first round of Nitrogen optimization

clearly seen that there is a very small change in the function and are very close to zero. This explains that there is a minimum effect of Fluorine atom on the calculations performed. An interesting thing to be noted is that unlike other curves inclination, the surface of Fluorine is inclined in opposite direction. This can be considered as an artifact, because less number of structures with Fluorine were considered for the calculation. We can still spot the minimum for the surface above and values of ϵ =1.5 and σ = 3.5 were extracted by the same method used for Nitrogen.

4.3.2 Round Two Optimization

Round two and round one optimization for Carbon and Hydrogen are demonstrated in figure 4.11. We can clearly see the slight shift in the curve of minimums for both Hydrogen and Carbon optimization surface. Round two of Hydrogen utilizes

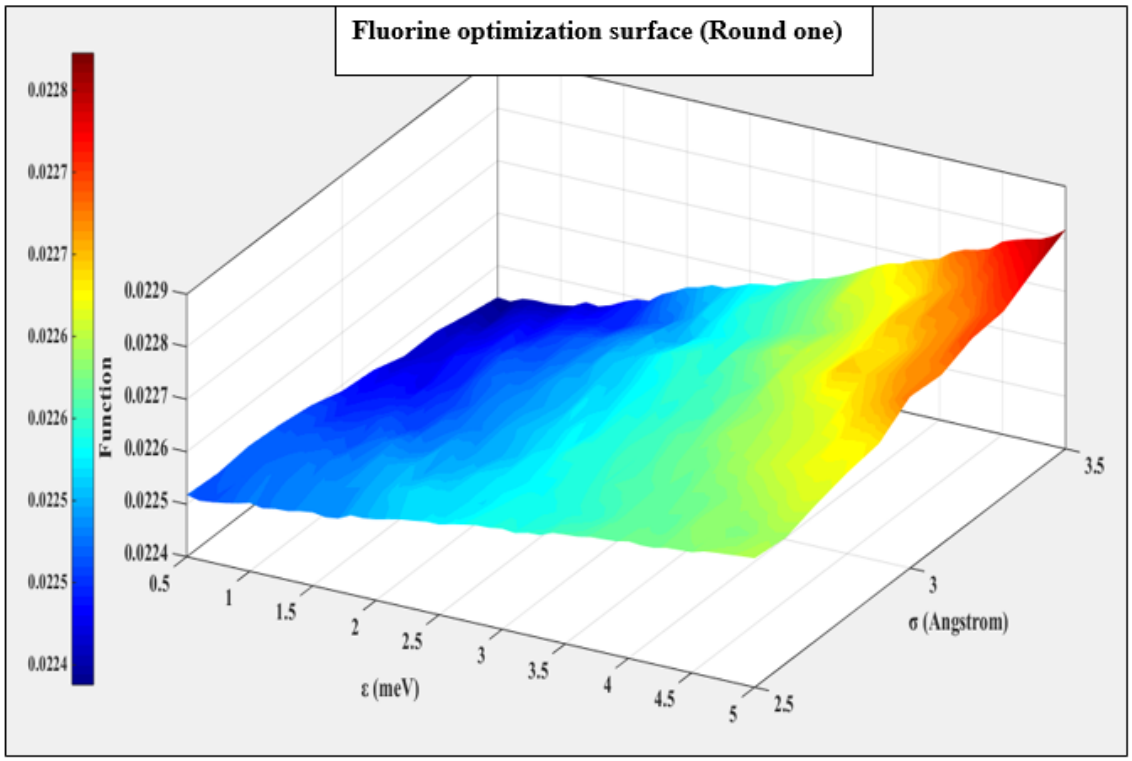

Fig. 4.10. Contour surface for the first round of Fluorine optimization

new values obtained from round two optimization of Carbon. After round two of optimization, the value of ϵ decreased for Carbon to 5 and increased to 3 for Hydrogen. This nature of shifting shows that both the potential parameters are interdependent to each other for combination of different atoms in a structure. The potential parameters for the atoms in the structure counter balance the effect of each other to reach the same CCS. Therefore, the optimization of potential parameters performed is completely based on the structure chosen for optimization and hence can be termed as structure dependent optimization scheme. The round two optimization for Oxygen and Nitrogen are compared in figure 4.12.

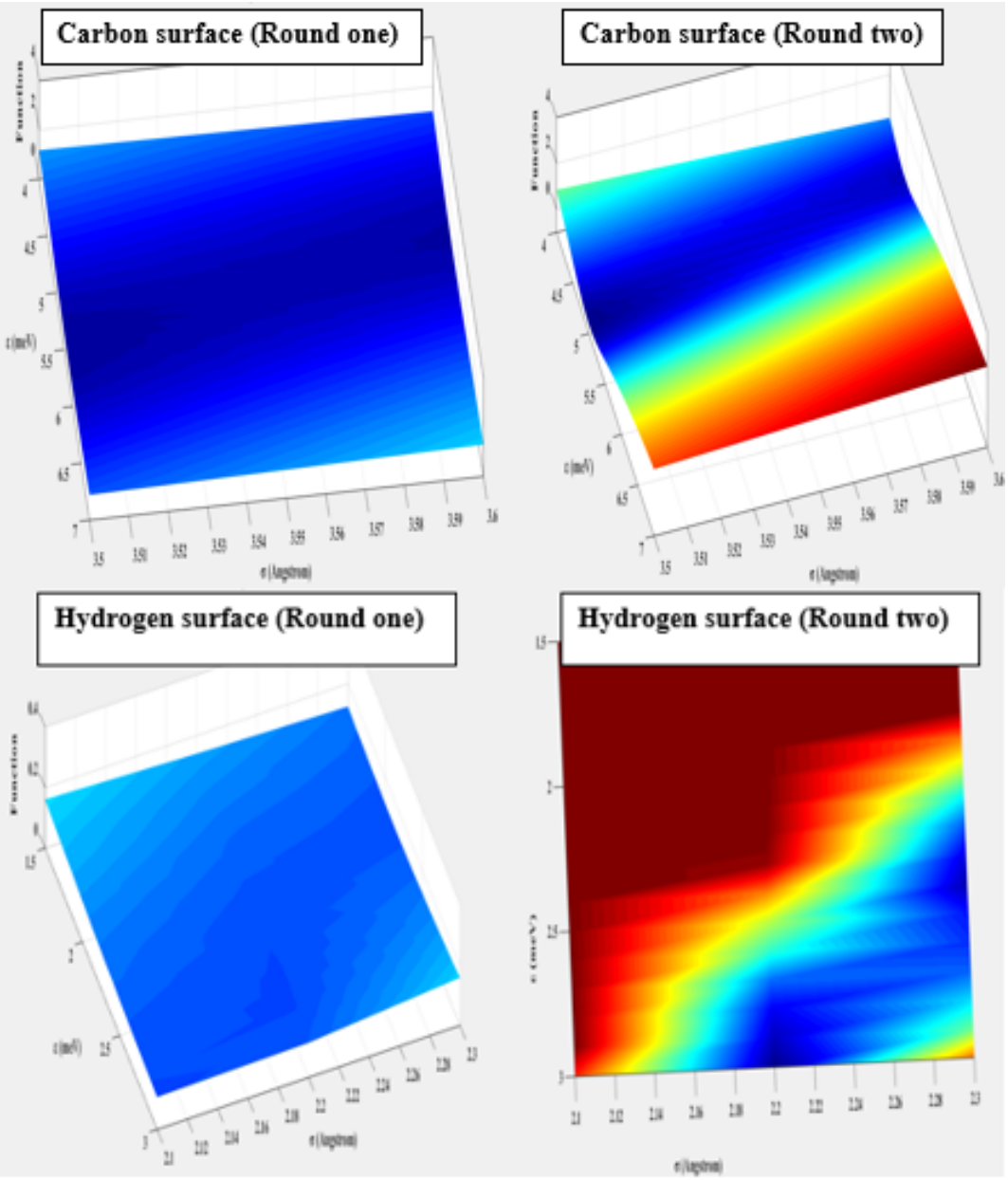

Fig. 4.11. Comparison between round one and two for Carbon and Hydrogen

4.3.3 Round Two Optimization

Again from the comparison of first and second round of optimization for Oxygen and Nitrogen, we can see a slight shift in the line of minimum when the step size was decreased to 0.01. This increases the accuracy of calculations and was able to

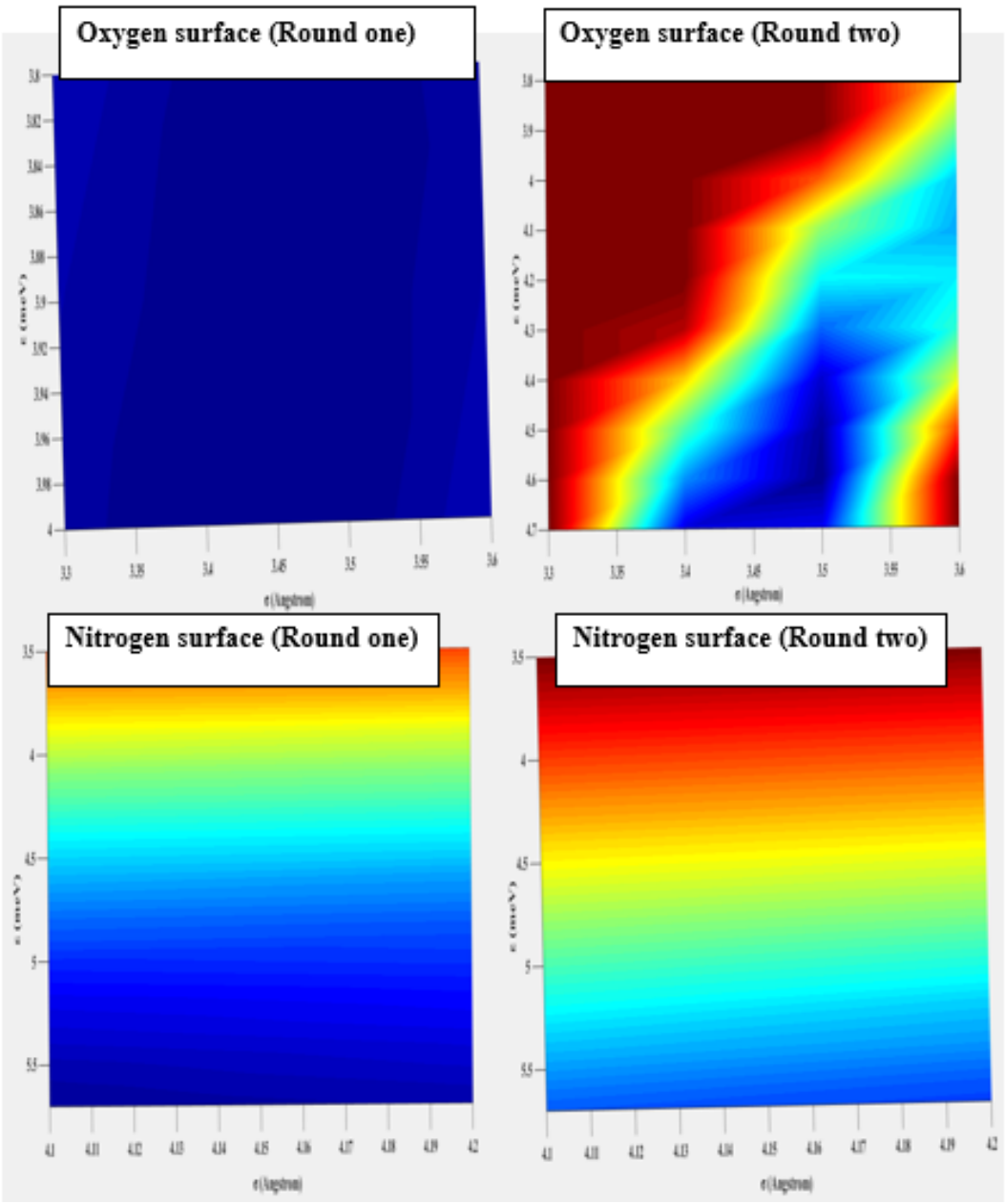

Fig. 4.12. Comparison between round one and two for Oxygen and Nitrogen

attend two places after decimal. As discussed earlier, the valley of minimum do not differ a lot as the structure with Oxygen and Nitrogen are less in number and do not affect the calculation compared to Carbon and Hydrogen. So we selected the previous values for potential parameters for Oxygen and Nitrogen and continue for third round

of optimization. Figure 4.13 is the logged figure for Fluorine for the second round of optimization:

4.3.4 Round Two Optimization

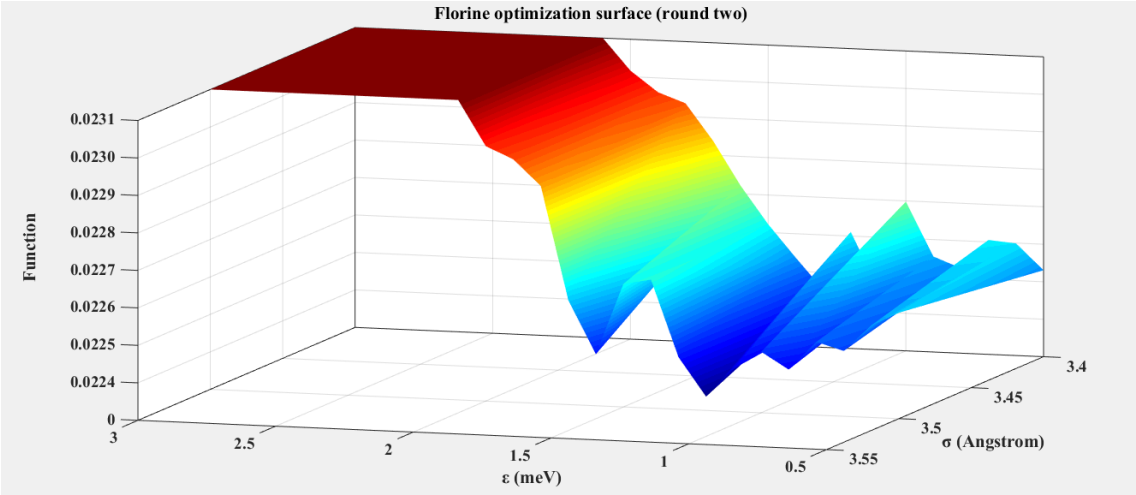

Fig. 4.13. Second round optimization for Fluorine

The figure 4.13 demonstrates the surface for Fluorine atom considering the modified parameters for other elements in the calculation. We can see a slight dip at lower ϵ and a rise again for a higher ϵ. Thus, we got the region of minimum when the step size was decreased to 0.01. The figure above has many ups and down contour because less number of data points were calculated and plotted.

4.3.5 Round Three Optimization

The third round of optimization was carried out for all the atoms considering all the new modified values for potential parameters after round two optimization. We chose to show the third round of optimization only for Carbon and Hydrogen as the surface for other elements did not show a significant change in their valley of minimums. As seen from the figure 4.14 the curve of minimums get narrowed down

more as compared to the second round and we obtained a shift in the least minimum of the function. The new values for all the five atoms are tabulated in table 4.2.

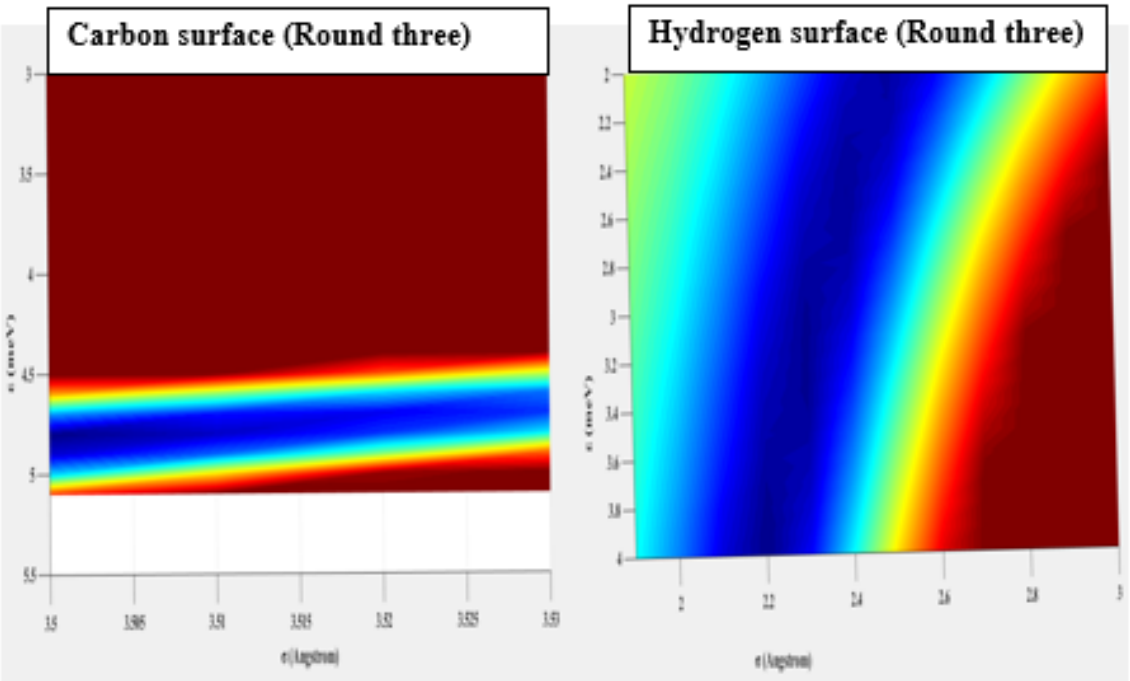

Fig. 4.14. The third round optimization for Carbon and Hydrogen

Table 4.2.
Newly modified LJ potential parameters in N2 N_2

Element	ϵ	σ
H	3	2.2
C	4.8	3.5
O	5	4.1
N	4.58	3.34
F	1.5	3.5

4.4 Optimization Considering the Case for Optimizing Hydrogen at First

After the optimization for the first case (starting optimization with Carbon) is done, we obtained a reliable set of parameters, which can be used in the calculation for the Trajectory Method. An obvious question arises at this point: What if the optimization started with some other element instead of Carbon? Will the potential parameters be same as that of optimization done with Carbon at first? These questions become important to answer in order to identify the global relationship between the LJ potential parameters. The question arises due to the fact that every atom in the molecule interacts with colliding gas molecule up till certain level. Now if we optimize the Carbon at first, there might be a possibility that it will reduce the effect of other elements to interact with the colliding gas molecules because of the fact that the Carbon was modified at first. This optimization process could be useful and will be correct if it gives similar values if we started the optimization with another atom. In order to perform the optimization, we choose to start the optimization with Hydrogen. As discussed earlier the optimization has a very little effect on the elements like N, O, and F, as they are found less in number as compared to H and C elements, we kept same LJ potential parameter for these elements as given in Table 4.2.

4.4.1 Round One Optimization

The optimization was run for ϵ from 1.5 to 4 and σ from 1.5 to 3 for Hydrogen and ϵ from 3.5 to 5 and σ from 3 to 4.5. The results are demonstrated in the figure 4.15 and 4.16. The trend is maintained and gives us the valley for minimums. The value chosen from the list of minimums were: For Hydrogen ϵ =2.8 σ=2.4 and for Carbon $\epsilon = 4.6$ and $\epsilon = 3.5$. These value are close bu not close enough to come to a conclusion. Hence we performed second and third round of optimization in order to get to correct values.

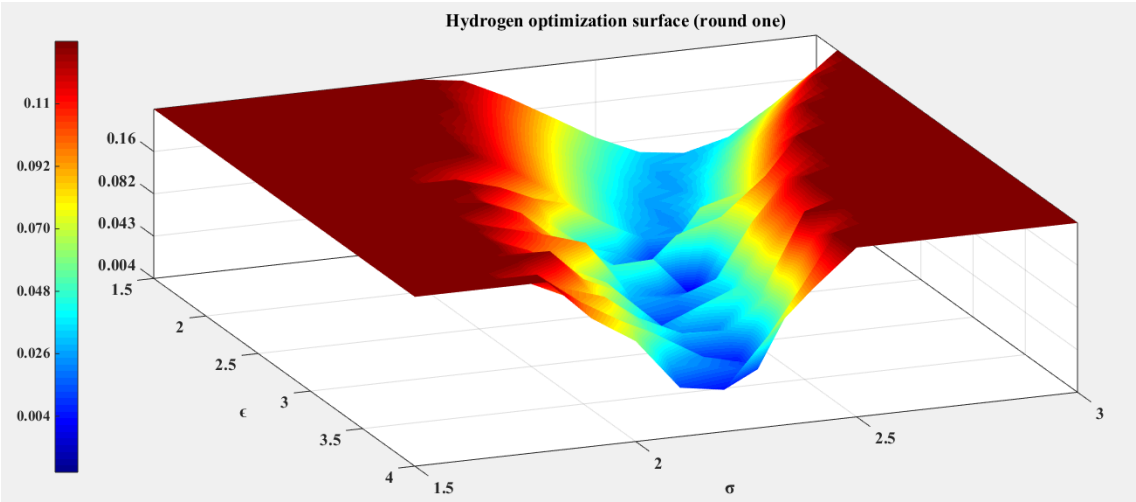

Fig. 4.15. The first round optimization for Hydrogen

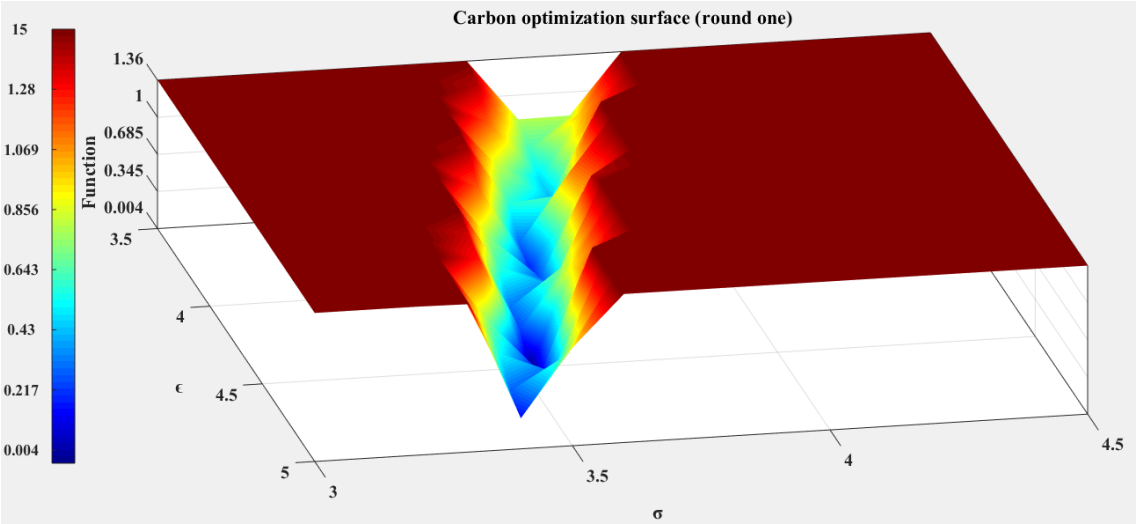

Fig. 4.16. The first round optimization for Carbon

4.4.2 Round Two and Three Optimization

From figure 4.17 and 4.18 above we can clearly see that there is very small change in the valley and we can say that the optimization has achieved stability and we have reached the optimal solution. The values are tabulated in table 4.3 and are similar to that for case 1 (optimization with Carbon at first).

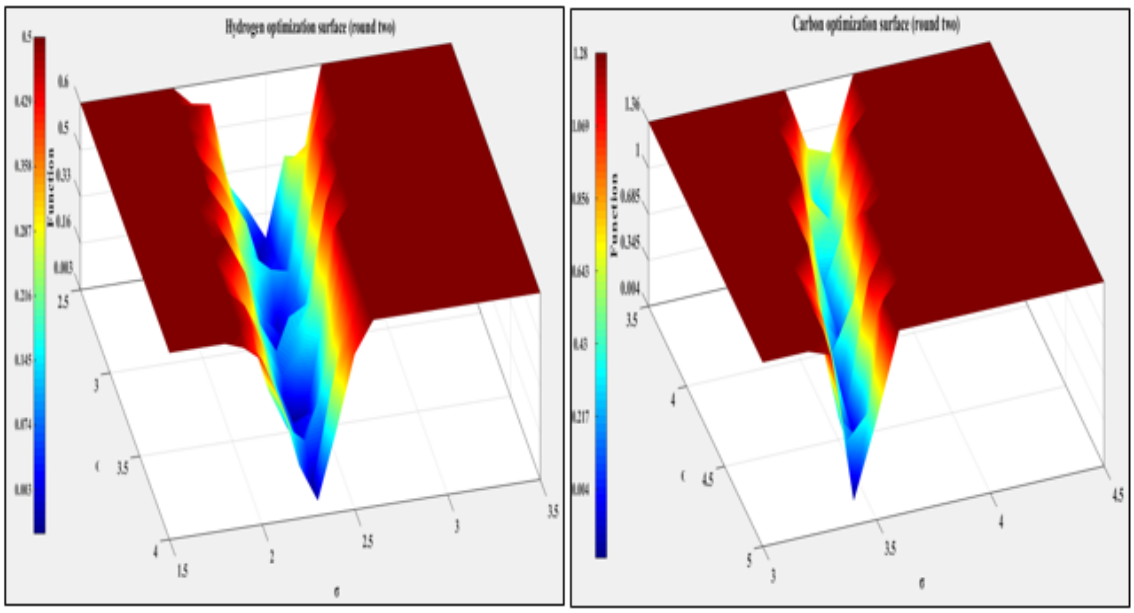

Fig. 4.17. The second round optimization for Carbon and Hydrogen

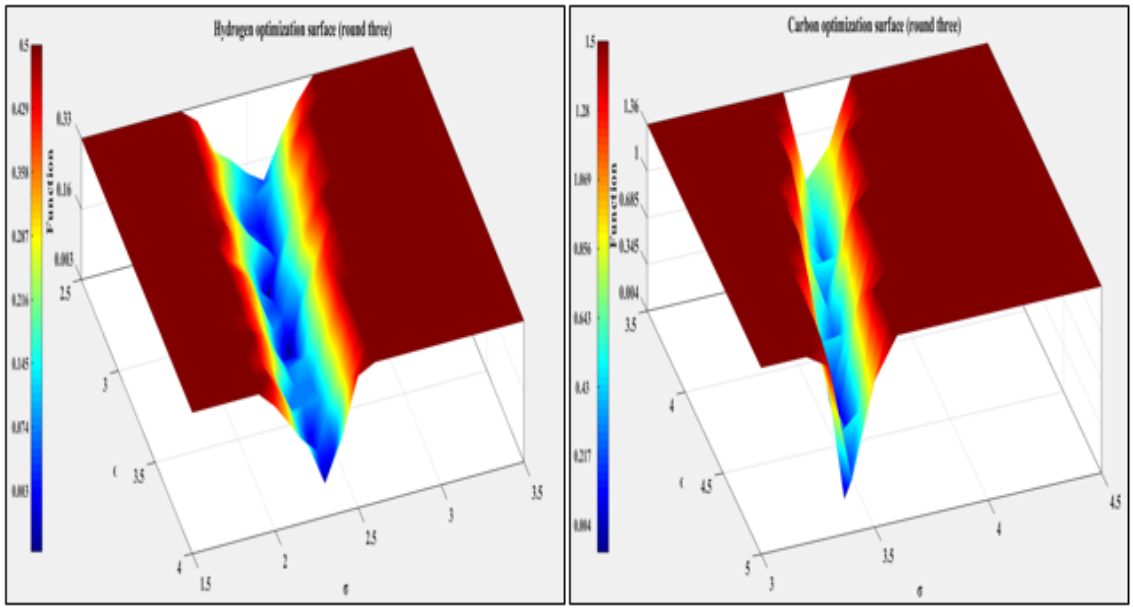

Fig. 4.18. The third round optimization for Carbon and Hydrogen

4.5 Validation of Results

Here we validated the newly obtained Lennard-Jones potential parameter by calculating the Collision Cross Sections for all 16 molecules seen from table 4.4. We

Table 4.3.
Newly modified LJ potential parameters in N2 N_2

Element	ϵ	σ
H	3	2.4
C	4.5	3.5
O	5	4.1
N	4.58	3.34
F	1.5	3.5

also calculated the % error for both CCS calculated with old LJ parameters and CCS calculated with new optimized LJ parameters. The error was seen to be within 4% for all the structures calculated considering new optimized Lennard-Jones potential parameters. Whereas, the error reaches to 11% in case of CCS calculated using old Lennard-Jones potential parameter and is tabulated in table 4.5. This validation proves that the newly calculated Lennard-Jones potential parameters gives Collision Cross Section closer to experimental results.

Table 4.4.
CCS calculation for old and new LJ values.

Structure	**Old LJ: CCS values**	**New Optimized LJ: CCS values**	**Experimental CCS**
Triphenylene	140.44	145.17	143.3
Nethylaniline	112.95	119.54	124.5
Dexamethasone	176.77	188.89	190.7
Acetaminophen	120.45	128.50	131.1
Betamethasone	176.25	188.58	189.6
Anthracene	127.38	132.00	129.6
Choline	104.17	113.68	115.4
Phenanthrene	126.74	131.35	129.1
Acetylcholine	117.80	127.64	127.8
c60	214.04	211.24	212.9
c70	231.70	228.73	231.4
Naphtalene	111.28	115.74	115.8
Paracetamol2	121.25	128.36	131.1
Pyrene	131.78	136.47	135
TtEA	113.43	122.40	122.2
TMA	95.31	103.27	107.4

Table 4.5.
% error for new optimized values and old values w.r.t. experimental CCS

% error for optimized LJ values	**% error for old LJ values**
1.30	1.99
3.99	9.28
0.95	7.30
1.98	8.12
0.54	7.04
1.85	1.72
1.49	9.74
1.74	1.83
0.12	7.83
0.78	0.54
1.15	0.13
0.06	3.90
2.09	7.52
1.09	2.38
0.17	7.18
3.85	11.25

5. SUMMARY AND CONCLUSIONS

5.1 Summary

Computational methods employed to calculate Ion Mobility and Collision Cross Section are widely used today to study the structural geometry and its behavior in gas phase. A detailed study for benchmark comparison for two widely used mobility calculator: IMoS and MOBCAL was carried out in this work. The test was conducted for the same input parameters, so that the comparison made is fair enough. Although IMoS and MOBCAL are inherently two differently written algorithms, Chapman-Enskog simplification for IMoS makes the CCS result comparable for both the algorithms. The algorithms were compared for its performance and efficiency, and inferences from major simplification employed to the CCS calculation method are discussed in section 5.2 in detail. Other part of the work talks about the Lennard-Jones potential parameters optimization method used in calculating CCS in the Trajectory Method. Millions of calculations were performed using IMoS in order to find the globally optimized parameters, which can be used in mobility calculator. Optimization was performed for the potential parameters using N2 as the buffer gas. The developed iterative optimization method plots the surface for the minimum of the function and gives the optimized parameter. A detailed list of inferences of Lennard-Jones optimized values are made in section 5.3.

5.2 Benchmark Comparison for Two Mobility Calculator: IMoS and MOCBAL

- Values for Collision Cross Section obtained from IMoS and MOBCAL are similar (within 1% difference) concluding the validity of a new algorithm: IMoS, even

though both the programs have a different approach for calculation. The CCS calculation methods applied here are PA, EHSS, and TM, for He and N_2 for ion-quadrupole potentials involved for N_2.

- The efficiently written IMoS algorithm provide a faster performance compared to MOBCAL. For the given same number of gas molecules used in the calculation of CCS for both the programs, IMoS performed one order of magnitude higher than that of its MOBCAL counterparts in all the cases studied.

- Due the flexibility of IMoS to be easily parallelize, higher number of cores can be used in calculations, The relation between increasing in speed is nearly linear to the increase in number of cores.

- The attempt was made to demonstrate the efficiency of the simplification employed for CCS calculation. It was observed that, both the algorithms are not efficient enough for studying large biomolecules (>10000 atoms). Hence simplifications were required to be incorporated in the algorithm, which can be easily done with the help of IMoS.

- With that said, EHSS gives the better simplification for larger biomolecules when He is used for calculations. EHSS is a simple and time efficient method considering the effect of scattering providing a real collision environment

- The TDHSS model was established as a simplification in N_2 gas for a structure with more than 200 atoms. The effect of with and without partial charges were also considered in the study. A corrected Projection Area method described in equation 3 was studied and was found to be in a good agreement with other methods. Simplifications used in N_2 could be extrapolated to coarse grained models due to its inherent simplifications of collisions and potentials.

5.3 Lennard-Jones Optimized Parameters for Trajectory Method in N_2

The process of iterative optimization was carried out for Carbon, Hydrogen, Oxygen, Nitrogen, and Fluorine atoms using an in-house developed MATLAB algorithm to find the optimized effective Lennard-Jones potential parameters when N2 gas is considered for calculations. Sixteen pharmaceutically small molecule structures with established experimental Collision Cross Sections were chosen to carry out the optimization process. The major conclusion for optimization work are listed below:

- A new representation was put forth showing the relationship between Lennard-Jones potential parameters for an atom giving us a well-defined curve of minimums for the elements presented here.

- A new optimization algorithm was developed, which gave a new set of values for LJ potential parameters that can be used and tested in the future studies in Ion Mobility calculators.

- The well-established relation for these potential parameters seen from the plots can lead us to simplify the methods used for calculating Collision Cross sections. One can opt to modify either of the parameters so as to get rid of the long computational time required for such kind of calculations.

- As this optimization was done considering the Collision Cross Sections (structure dependent), the optimized parameters presented here are more useful for the research work that includes CCS and ion mobility calculations.

- We also checked by starting the optimization from Hydrogen and it yield very similar result to that of case 1 optimization (starting with carbon first). Hence we can say that the values obtained for Lennard-Jones potential parameter are correct and can be used for calculation purposes.

5.4 Future Work

A future work is recommended considering the vastness of this project. Some important points are discussed below to be considered for future studies.

- As noted from the conclusion, the optimization performed was a structure dependent optimization process. The work presented here encourages us to look into deeper aspects of such kind of optimization by considering more number of structures into calculations. The structures must have an experimental CCS value, which becomes a challenging job. More number of experiments are recommended to be performed in order to get the CCS values. The structures considered should have majority of elements required for optimization, so as to get the global effective optimized Lennard-Jones potential parameters.

- As seen from chapter 4, performing optimization on elements such as Nitrogen is a difficult task due to less availability in the structures. Therefore, structures that are dominating in Nitrogen such as Amino Acids should be selected for optimization scheme. This consideration will give the best possible contour curve for Nitrogen given that we have correct experimental CCS values.

- A smaller step size can provide a more detailed and smooth curve because then we will have more number of data points. With this condition comes the difficulty of higher simulation time required for calculations. We need to use supercomputers with more number of cores to perform the optimization process faster.

- The optimization process described in this work is done considering N2 as a buffer gas. There are other buffer gases, which are used in Ion Mobility Spectrometry such as He, Ar, and Air for which the optimization must be performed.

REFERENCES

REFERENCES

[1] Bohrer, B.C., Mererbloom, S.I., Koeniger, S.L., Hilderbrand, A.E., Clemmer, D.E.: Biomolecule Analysis by Ion Mobility Spectrometry. Annu Rev Anal Chem. 1, 293-327 (2008).

[2] Larriba, C., de la Mora, J.F.: The Gas Phase Structure of Coulombically Stretched Polyethylene Glycol Ions. Journal of Physical Chemistry B. 116, 593-598, (2012).

[3] Larriba, C., Hogan, C.J., Attoui, M., Borrajo, R., Garcia, J.F., de la Mora, J.F.: The Mobility-Volume Relationship below 3.0 nm Examined by Tandem Mobility-Mass Measurement. Aerosol Sci Tech. 45, 453-467, (2011).

[4] Larriba-Andaluz, C., Fernandez-Garcia, J., Ewing, M.A., Hogan, C.J., Clemmer, D.E.: Gas molecule scattering & ion mobility measurements for organic macro-ions in He versus N-2 environments. Phys Chem Chem Phys. 17, 15019-15029, (2015).

[5] Li, M.D., Mulholland, G.W., Zachariah, M.R.: Understanding the mobility of nonspherical particles in the free molecular regime. Phys Rev E. 1-2, (2014).

[6] Oberreit, D., Rawat, V.K., Larriba-Andaluz, C., Ouyang, H., McMurry, P.H., Hogan, C.J.: Analysis of heterogeneous water vapor uptake by metal iodide cluster ions via differential mobility analysis-mass spectrometry. J Chem Phys. 143, 1-2, (2015).

[7] Ruotolo, B.T., Benesch, J.L.P., Sandercock, A.M., Hyung, S.J., Robinson, C.V.: Ion mobility-mass spectrometry analysis of large protein complexes. Nat Protoc. 3, 1139-1152, (2008).

[8] Trimpin, S., Clemmer, D.E.: Ion mobility spectrometry/mass spectrometry snapshots for assessing the molecular compositions of complex polymeric systems. Anal Chem. 80, 9073-9083, (2008).

[9] Zhang, C.L., Thajudeen, T., Larriba, C., Schwartzentruber, T.E., Hogan, C.J.: Determination of the Scalar Friction Factor for Nonspherical Particles and Aggregates Across the Entire Knudsen Number Range by Direct Simulation Monte Carlo (DSMC). Aerosol Sci Tech. 46, 1065-1078, (2012).

[10] Clemmer, D.E., Pierson, N., Shi, H.L., Ewing, M.: Developing IMS-MS techniques as a means of following structural transitions of biopolymers in solution. Abstr Pap Am Chem S. 247, 1-2 ,(2014).

[11] Criado-Hidalgo, E., Fernandez-Garcia, J., de la Mora, J.F.: Mass and Charge Distribution Analysis in Negative Electrosprays of Large Polyethylene Glycol Chains by Ion Mobility Mass Spectrometry. Analytical Chemistry. 85, 2710-2716, (2013).

[12] Ewing, M.A., Glover, M.S., Clemmer, D.E.: Hybrid ion mobility and mass spectrometry as a separation tool. J Chromatogr A. 1439, 3-25, (2016).

[13] Fenn, L.S., McLean, J.A.: Simultaneous glycoproteomics on the basis of structure using ion mobility-mass spectrometry. Mol Biosyst. 5, 1298-1302, (2009).

[14] Fernandez-Garcia, J., de la Mora, J.F.: Measuring the Effect of Ion-Induced Drift-Gas Polarization on the Electrical Mobilities of Multiply-Charged Ionic Liquid Nanodrops in Air. J Am Soc Mass Spectr. 24, 1872-1889, (2013).

[15] Gaye, M.M., Nagy, G., Clemmer, D.E., Pohl, N.L.B.: Multidimensional Analysis of 16 Glucose Isomers by Ion Mobility Spectrometry. Analytical Chemistry. 88, 2335-2344, (2016).

[16] Glover, M.S., Dilger, J.M., Acton, M.D., Arnold, R.J., Radivojac, P., Clemmer, D.E.: Examining the Influence of Phosphorylation on Peptide Ion Structure by Ion Mobility Spectrometry-Mass Spectrometry. J Am Soc Mass Spectr. 27, 786-794, (2016).

[17] Larriba, C., de la Mora, J.F., Clemmer, D.E.: Electrospray Ionization Mechanisms for Large Polyethylene Glycol Chains Studied Through Tandem Ion Mobility Spectrometry. J Am Soc Mass Spectr. 25, 1332-1345, (2014).

[18] Matz, L.M., Hill, H.H.: Evaluating the separation of amphetamines by electrospray ionization ion mobility spectrometry/MS and charge competition within the ESI process. Analytical Chemistry. 74, 420-427, (2002).

[19] McLean, J.A.: The Mass-Mobility Correlation Redux: The Conformational Landscape of Anhydrous Biomolecules. J Am Soc Mass Spectr. 20, 1775-1781, (2009).

[20] Trimpin, S., Tan, B., Bohrer, B.C., O'Dell, D.K., Merenbloom, S.I., Pazos, M.X., Clemmer, D.E., Walker, J.M.: Profiling of phospholipids and related lipid structures using multidimensional ion mobility spectrometry-mass spectrometry. International Journal of Mass Spectrometry. 287, 58-69, (2009).

[21] Valentine, S.J., Kurulugama, R.T., Bohrer, B.C., Merenbloom, S.I., Sowell, R.A., Mechref, Y., Clemmer, D.E.: Developing IMS-IMS-MS for rapid characterization of abundant proteins in human plasma. International Journal of Mass Spectrometry. 283, 149-160, (2009).

[22] Martinez-Lozano, P., Labowsky, M.: An experimental and numerical study of a miniature high resolution isopotential DMA. J Aerosol Sci. 40, 451-462, (2009).

[23] Kumar, A., Kang, S., Larriba-Andaluz, C., Ouyang, H., Hogan, C.J., Sankaran, R.M.: Ligand-free Ni nanocluster formation at atmospheric pressure via rapid quenching in a microplasma process Nanotechnology. 25,1-2, (2014).

[24] Graves, B., Olfert, J., Patychuk, B., Dastanpour, R., Rogak, S.: Characterization of Particulate Matter Morphology and Volatility from a Compression-Ignition Natural-Gas Direct-Injection Engine. Aerosol Sci Tech. 49, 589-598, (2015).

[25] Flagan, R.C.: Continuous-Flow Differential Mobility Analysis of Nanoparticles and Biomolecules. Annu Rev Chem Biomol. 5, 255-279, (2014).

[26] Eiceman, G.A., Krylov, E.V., Nazarov, E.G., Miller, R.A.: Separation of ions from explosives in differential mobility spectrometry by vapor-modified drift gas. Analytical Chemistry. 76, 4937-4944, (2004).

[27] Eiceman, G.A., Feng, Y.: Limits of separation of a multi-capillary column with mixtures of volatile organic compounds for a flame ionization detector and a differential mobility detector. J Chromatogr A. 1216, 985-993, (2009).

[28] Carbone, F., Beretta, F., D'Anna, A.: Size Distribution Functions of Ultrafine Ashes From Pulverized Coal Combustion. Combust Sci Technol. 182, 668-682, (2010).

[29] Carbone, F., Attoui, M., Gomez, A.: Challenges of measuring nascent soot in flames as evidenced by high-resolution differential mobility analysis. Aerosol Sci Tech. 50, 740-757, (2016).

[30] Attoui, M., de la Mora, J.F.: Flow driven transmission of charged particles against an axial field in antistatic tubes at the sample outlet of a Differential Mobility Analyzer. J Aerosol Sci. 100, 91-96, (2016) .

[31] Webb, I.K., Garimella, S.V.B., Norheim, R.V., Baker, E.S., Ibrahim, Y.M., Smith, R.D.: A Structures for Lossless Ion Manipulations (SLIM) Module for Collision Induced Dissociation. J Am Soc Mass Spectr. 27, 1285-1288, (2016).

[32] Ridgeway, M.E., Kaplan, D.A., Park, M.A.: Trapped ion mobility spectrometry. Abstr Pap Am Chem S. 244, 1-2, (2012).

[33] Fernandez-Lima, F.A., Kaplan, D.A., Park, M.A.: Note: Integration of trapped ion mobility spectrometry with mass spectrometry. Rev Sci Instrum. 82,1-2, (2011).

[34] Chen, T.C., Ibrahim, Y.M., Webb, I.K., Garimella, S.V.B., Zhang, X., Hamid, A.M., Deng, L.L., Karnesky, W.E., Prost, S.A., Sandoval, J.A., Norheim, R.V., Anderson, G.A., Tolmachev, A.V., Baker, E.S., Smith, R.D.: Mobility-Selected Ion Trapping and Enrichment Using Structures for Lossless Ion Manipulations. Analytical Chemistry. 88, 1728-1733, (2016).

[35] Larriba-Andaluz, C., Hogan, C.J.: Collision cross section calculations for polyatomic ions considering rotating diatomic/linear gas molecules. J Chem Phys. 141,1-2, (2014).

[36] Ouyang, H., Larriba-Andaluz, C., Oberreit, D.R., Hogan, C.J.: The Collision Cross Sections of Iodide Salt Cluster Ions in Air via Differential Mobility Analysis-Mass Spectrometry. J Am Soc Mass Spectr. 24, 1833-1847, (2013).

[37] Larriba-Andaluz, C., Hogan, C.: Novel interfaced approach to mobility calculations with diffuse scattering and Maxwell rotational distributions for diatomic gases in the free molecular regime. Abstr Pap Am Chem S. 246,1-2, (2013).

[38] Larriba, C., Hogan, C.J.: Ion Mobilities in Diatomic Gases: Measurement versus Prediction with Non-Specular Scattering Models. J Phys Chem A. 117, 3887-3901, (2013).

[39] Larriba, C., Hogan, C.J.: Free molecular collision cross section calculation methods for nanoparticles and complex ions with energy accommodation. J Comput Phys. 251, 344-363, (2013).

[40] Mesleh, M.F., Hunter, J.M., Shvartsburg, A.A., Schatz, G.C., Jarrold, M.F.: Structural information from ion mobility measurements: Effects of the long-range potential (vol 100, pg 16082, 1996). J Phys Chem A. 101, 968-968 ,(1997).

[41] Shvartsburg, A.A., Jarrold, M.F.: An exact hard-spheres scattering model for the mobilities of polyatomic ions. Chem Phys Lett. 261, 86-91, (1996).

[42] Mason, E.A., McDaniel, E.W. John Wiley & Sons., New York, 1-2, (1988).

[43] Chan, P., Dahneke, B.: Freemolecule drag on straight chains of uniform spheres. Journal of Applied Physics. 52, 3106-3110, (1981).

[44] Dahneke, B.E.: Slip correction factors for nonspherical bodiesI Introduction and continuum flow. J Aerosol Sci. 4, 139-145, (1973).

[45] Dahneke, B.E.: Slip correction factors for nonspherical bodiesII free molecule flow. J Aerosol Sci. 4, 147-161, (1973).

[46] Epstein, P.S.: On the Resistance Experienced by Spheres in their Motion through Gases. Physical Review. 23, 710-715, (1924).

[47] Garcia-Ybarra, P.L., Castillo, J.L., Rosner, D.E.: Drag on a large spherical aggregate with self-similar structure: An asymptotic analysis. J Aerosol Sci. 37, 413-428, (2006).

[48] Mackowski, D.W.: Calculation of Total Cross-Sections of Multiple-Sphere Clusters. J Opt Soc Am A. 11, 2851-2861, (1994).

[49] Mackowski, D.W.: Monte Carlo simulation of hydrodynamic drag and thermophoresis of fractal aggregates of spheres in the free-molecule flow regime. J Aerosol Sci. 37, 242-259, (2006).

[50] de la Mora, J.F.: Free-molecule mobility of polyhedra and other convex hard-bodies. J Aerosol Sci. 33, 477-489, (2002).

[51] Bleiholder, C., Wyttenbach, T., Bowers, M.T.: A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (I). Method. International Journal of Mass Spectrometry. 308, 1-10, (2011).

[52] Millikan, R.A.: The General Law of Fall of a Small Spherical Body through a Gas, and its Bearing upon the Nature of Molecular Reflection from Surfaces. Physical Review. 22, 1-23, (1923).

[53] Onsager, L.: Reciprocal relations in irreversible processes. I. Physical review. 37, 1-2, (1931).

[54] Onsager, L.: Reciprocal relations in irreversible processes. II. Physical Review. 38, 1-2, (1931).

[55] Chapman, S., Cowling, T.G. Cambridge university press, 1-2, (1970).

[56] Li, Z.G., Wang, H.: Drag force, diffusion coefficient, and electric mobility of small particles. II. Application. Phys Rev E. 68,1-2, (2003).

[57] Li, Z.G., Wang, H.: Drag force, diffusion coefficient, and electric mobility of small particles. I. Theory applicable to the free-molecule regime. Phys Rev E. 68,1-2, (2003).

[58] Clemmer, D.E., Jarrold, M.F.: Ion mobility measurements and their applications to clusters and biomolecules. Journal of Mass Spectrometry. 32, 577-592, (1997).

[59] Dugourd, P., Hudgins, R.R., Clemmer, D.E., Jarrold, M.F.: High-resolution ion mobility measurements. Rev Sci Instrum. 68, 1122-1129, (1997).

[60] Dugourd, P., Hudgins, R.R., Jarrold, M.F.: High-resolution ion mobility studies of sodium chloride nanocrystals. Chem Phys Lett. 267, 186-192, (1997).

[61] Hudgins, R.R., Dugourd, P., Tenenbaum, J.M., Jarrold, M.F.: Structural transitions in sodium chloride nanocrystals. Phys Rev Lett. 78, 4213-4216, (1997).

[62] Hudgins, R.R., Woenckhaus, J., Jarrold, M.F.: High resolution ion mobility measurements for gas phase proteins: correlation between solution phase and gas phase conformations. International Journal of Mass Spectrometry. 165, 497-507, (1997).

[63] Mesleh, M.F., Hunter, J.M., Shvartsburg, A.A., Schatz, G.C., Jarrold, M.F.: Structural information from ion mobility measurements: Effects of the long-range potential. J Phys Chem-Us. 100, 16082-16086, (1996).

[64] Shelimov, K.B., Clemmer, D.E., Hudgins, R.R., Jarrold, M.F.: Protein Structure in Vacuo: Gas-Phase Conformations of BPTI and Cytochrome c. J Am Chem Soc. 119, 2240-2248, (1997).

[65] Shelimov, K.B., Jarrold, M.F.: "Denaturation" and refolding of cytochrome c in vacuo. J Am Chem Soc. 118, 10313-10314, (1996).

[66] Shelimov, K.B., Jarrold, M.F.: Conformations, unfolding, and refolding of apomyoglobin in vacuum: An activation barrier for gas-phase protein folding. J Am Chem Soc. 119, 2987-2994, (1997).

[67] Shvartsburg, A.A., Hudgins, R.R., Dugourd, P., Jarrold, M.F.: Structural characterization of the fullerene coalescence products using high-resolution ion mobility measurements. Abstr Pap Am Chem S. 214, 1-2, (1997).

[68] Shvartsburg, A.A., Hudgins, R.R., Dugourd, P., Jarrold, M.F.: Structural elucidation of fullerene dimers by high-resolution ion mobility measurements and trajectory calculation simulations. J Phys Chem A. 101, 1684-1688, (1997).

[69] Shvartsburg, A.A., Schatz, G.C., Jarrold, M.F.: Mobilities of carbon cluster ions: Critical importance of the molecular attractive potential. J Chem Phys. 108, 2416-2423, (1998).

[70] Woenckhaus, J., Mao, Y., Jarrold, M.F.: Hydration of gas phase proteins: Folded +5 and unfolded +7 charge states of cytochrome c. Journal of Physical Chemistry B. 101, 847-851, (1997).

[71] Campuzano, I., Bush, M.F., Robinson, C.V., Beaumont, C., Richardson, K., Kim, H., Kim, H.I.: Structural Characterization of Drug-like Compounds by Ion Mobility Mass Spectrometry: Comparison of Theoretical and Experimentally Derived Nitrogen Collision Cross Sections. Analytical Chemistry. 84, 1026-1033, (2012).

[72] Happel, J., Brenner, H. Springer Netherlands,, Dordrecht:Low Reynolds number hydrodynamics, 1-2, (1981).

[73] Landau, L.D., Lifshitz, E.M. Pergamon Press, Oxford, England; New York: Theory of Elasticity, 1-2, (1987).

[74] Kim, H.I., Kim, H., Pang, E.S., Ryu, E.K., Beegle, L.W., Loo, J.A., Goddard, W.A., Kanik, I.: Structural Characterization of Unsaturated Phosphatidylcholines Using Traveling Wave Ion Mobility Spectrometry. Analytical Chemistry. 81, 8289-8297, (2009).

[75] Kim, H., Kim, H.I., Johnson, P.V., Beegle, L.W., Beauchamp, J.L., Goddard, W.A., Kanik, I.: Experimental and theoretical investigation into the correlation between mass and ion mobility for choline and other ammonium cations in N-2. Analytical Chemistry. 80, 1928-1936, (2008).

[76] Kinnear, B.S., Kaleta, D.T., Kohtani, M., Hudgins, R.R., Jarrold, M.F.: Conformations of unsolvated valine-based peptides. J Am Chem Soc. 122, 9243-9256, (2000).

[77] Konijnenberg, A., Yilmaz, D., Ingolfsson, H.I., Dimitrova, A., Marrink, S.J., Li, Z.L., Venien-Bryan, C., Sobott, F., Kocer, A.: Global structural changes of an ion channel during its gating are followed by ion mobility mass spectrometry. P Natl Acad Sci USA. 111, 17170-17175, (2014).

[78] C. Becker, F.A. Fernandez-Lima, D.H. Russell, Ion mobility-mass spectrometry: a tool for characterizing the petroleome, Spectroscopy 24 ,38-42,(2009).

[79] Z. Li, S.J. Valentine, D.E. Clemmer, Complexation of amino compounds by 18c6 improves selectivity by IMSIMSMS: application to petroleum characterization, J. Am. Soc. Mass Spectrom. 22 817-827,(2011).

[80] Vaibhav Shrivastav, Minal Nahin, Christopher J. Hogan & Carlos Larriba Andaluz,Benchmark Comparison for a Multi-Processing Ion Mobility Calculator in the Free Molecular Regime, Journal of The American Association of Mass Spectrometry,1540-1551, (2017).

[81] MFJ Research Group (2016, March 20)
http://www.indiana.edu/∼nano/software/

[82] Mann, M., Meng, C.K. and Fenn, J.B. Interpreting mass spectra of multiply charged ions. Anal. Chem., 61, 1-8,(1989).

[83] Fenn, J.B., Mann, M., Meng, C.K. et al.: Electrospray ionization mass spectrometry: How it all began, 246,1-2 (1989).

[84] Chi-Kit Siu, Yuzhu Guo, Irine S. Saminathan, Alan C. Hopkinson, and K. W. Michael Siu: Optimization of Parameters Used in Algorithms of Ion-Mobility Calculation for Conformational Analyses, J. Phys. Chem., 114, 1204-1212, B (2010).

APPENDICES

APPENDIX A

Table 7.1.
Contains list of molecules used for CCS calculations. MOBCAL and IMoS calculations were performed for PA, EHSS and TM methods listed below for Helium.

Structures	Number of atoms	MOBCAL CCS (A2)			IMoS CCS (A2)		
		PA	EHSS	TM	PA	EHSS	TM
		[I]	[II]	[III]	[IV]	[V]	[VI]
Tryphenylene	30	89.99	92.65	87.03	90.39	92.45	86.28
NEthylaniline	21	66.43	69.51	64.65	66.59	69.82	64.84
Dexamethasone	58	122.65	131.52	124.28	122.84	131.11	123.66
Acetaminophen	21	72.77	74.88	70.24	72.62	76.35	71.21
Betamethasone	58	121.97	130.50	123.25	122.19	130.37	122.86
Anthracene	24	78.65	81.10	75.77	78.98	81.17	75.67
Choline	21	61.00	63.48	59.01	61.03	63.58	58.92
Phenanthrene	24	77.69	80.01	75.08	77.98	80.18	74.96
Acetylcholine	26	73.51	76.96	71.46	73.58	77.56	71.62
C60	60	118.62	120.94	124.66	118.71	121.18	124.23
C70	70	130.29	133.07	137.50	130.35	133.31	137.10
Naphtalene	18	64.10	65.89	61.72	64.36	66.11	61.67
Paracetamol2	21	72.49	76.00	71.27	72.92	74.86	70.06
Pyrene	26	81.37	83.93	79.05	81.70	83.98	78.59
TtEA	29	70.37	73.67	67.76	70.39	73.70	67.49
TMA	17	51.75	53.56	49.68	51.81	53.69	49.67
C20	20	64.84	65.56	67.68	64.86	65.74	66.56
C40	40	93.64	95.32	99.38	93.67	95.51	97.30
C100	100	163.04	166.96	175.10	163.19	167.16	172.77
C180	180	239.42	245.79	259.60	239.54	246.14	257.26

Table 7.2.
Contains list of molecules used for CCS calculations. MOBCAL and IMoS calculations were performed for PA, EHSS and TM methods listed below for Helium.

Structures	**Number of atoms**	**MOBCAL CCS (A2)**			**IMoS CCS (A2)**		
		PA	**EHSS**	**TM**	**PA**	**EHSS**	**TM**
		[I]	**[II]**	**[III]**	**[IV]**	**[V]**	**[VI]**
C180	180	239.42	245.79	259.60	239.54	246.14	257.26
C240	240	300.98	310.20	326.08	301.11	310.92	323.64
C500	500	517.46	533.88	562.70	517.72	534.92	550.27
C540	540	550.06	568.12	598.02	550.30	569.50	581.39
C720	720	696.85	720.21	757.07	697.09	721.73	729.64
16EMINCN2	384	365.23	464.12	498.06	366.55	463.64	497.42
32EMINCN2	768	613.06	786.43	845.79	613.04	788.43	839.01
64EMINCN242	1612	953.65	1252.68	1306.70	954.66	1253.79	1304.11
TPA	39	95.58	102.76	94.61	95.59	103.29	95.07
TBA	53	119.97	131.34	120.69	120.11	131.67	120.55
TDA	125	271.88	303.02	272.08	271.24	301.68	273.47
TDDA	149	325.19	361.04	326.69	325.24	361.68	327.28
TPA4	41	95.97	103.21	94.96	96.03	103.59	95.33
THA2	89	185.08	207.11	189.58	185.27	206.78	188.36
PEG70Cs4	497	753.31	865.01	822.50	753.07	859.96	814.60
PEG90Cs4	630	793.82	929.73	887.87	794.65	926.74	889.83
PEG115Cs43	805	916.74	1079.30	1044.40	919.46	1079.70	1036.70
PEG144Cs4	1008	859.65	1034.30	1013.80	857.48	1033.33	1009.89

Table 7.3.
Ratios of CCS (MOBCAL/IMoS) for each method is presented here for Helium.

Structure	Ratio			Structures	Ratio		
	PA	EHSS	TM		PA	EHSS	TM
	[I]/[IV]	[II]/ [V]	[III]/ [VI]		[I]/[IV]	[II]/ [V]	[III]/ [VI]
Tryphenylene	1.00	1.00	0.99	**C180**	1.00	1.00	0.99
NEthylaniline	1.00	1.00	1.00	**C240**	1.00	1.00	0.99
Dexamethasone	1.00	1.00	1.00	**C500**	1.00	1.00	0.98
Acetaminophen	1.00	1.02	1.01	**C540**	1.00	1.00	0.97
Betamethasone	1.00	1.00	1.00	**C720**	1.00	1.00	0.96
Anthracene	1.00	1.00	1.00	**16EMINCN2**	1.00	1.00	1.00
Choline	1.00	1.00	1.00	**32EMINCN2**	1.00	1.00	0.99
Phenanthrene	1.00	1.00	1.00	**64EMINCN242**	1.00	1.00	1.00
Acetylcholine	1.00	1.01	1.00	**TPA**	1.00	1.01	1.00
C60	1.00	1.00	1.00	**TBA**	1.00	1.00	1.00
C70	1.00	1.00	1.00	**TDA**	1.00	1.00	1.01
Naphtalene	1.00	1.00	1.00	**TDDA**	1.00	1.00	1.00
Paracetamol2	1.01	0.99	0.98	**TPA4**	1.00	1.00	1.00
Pyrene	1.00	1.00	0.99	**THA2**	1.00	1.00	0.99
TtEA	1.00	1.00	1.00	**PEG70Cs4**	1.00	0.99	0.99
TMA	1.00	1.00	1.00	**PEG90Cs4**	1.00	1.00	1.00
C20	1.00	1.00	0.98	**PEG115Cs43**	1.00	1.00	0.99
C40	1.00	1.00	0.98	**PEG144Cs4**	1.00	1.00	1.00
C100	1.00	1.00	0.99	**PEG254Cs4**	1.00	1.00	1.00

Table 7.4.
Lists the values for CCS and time calculated in IMoS and MOBCAL in N2. Simulation considered 9e5 and 4e5 gas molecules for both the programs.

Structures	**Number of atoms**	**CCS (A2)**			
		MOBCAL TM (4E5)	**IMoS TMLJ (4E5)**	**IMoS TMLJ (9E5)**	**IMoS EHSS (4E5)**
		[I]	**[II]**	**[III]**	**[IV]**
TMA	17	107.22	105.57	107.22	64.23
Naphtalene	18	121.17	119.98	121.96	78.88
N Ethylaniline	21	124.58	123.47	124.88	81.81
Acetaminophen	21	131.69	127.20	128.86	88.07
Choline	21	115.92	113.24	115.48	74.42
Paracetamol2	21	131.21	128.28	129.73	87.17
Anthracene	24	136.21	134.65	136.57	95.11
Phenanthrene	24	135.71	134.22	135.73	94.60
Acetylcholine	26	128.01	125.70	127.42	89.44
Pyrene	26	139.95	138.65	140.37	98.22
TtEA	29	122.26	119.33	121.00	85.92
Tryphenylene	30	149.28	146.83	148.30	107.51
Dexamethasone	58	190.4	183.20	185.04	145.99
Betamethasone	58	189.47	183.48	184.78	145.16
C60	60	212.98	209.61	210.41	140.26
C70	70	229.85	226.25	227.90	153.26

Table 7.5.
Ratio TM/EHSS for CCS is provided for Nitrogen.

Structures	CCS Ratio		
	[I]/[II]	[I]/[III]	TM/EHSS
TMA	1.02	1.00	1.64
Naphtalene	1.01	0.99	1.52
N Ethylaniline	1.01	1.00	1.51
Acetaminophen	1.04	1.02	1.44
Choline	1.02	1.00	1.52
Paracetamol2	1.02	1.01	1.47
Anthracene	1.01	1.00	1.42
Phenanthrene	1.01	1.00	1.42
Acetylcholine	1.02	1.00	1.41
Pyrene	1.01	1.00	1.41
TtEA	1.02	1.01	1.39
Tryphenylene	1.02	1.01	1.37
Dexamethasone	1.04	1.03	1.25
Betamethasone	1.03	1.03	1.26
C60	1.02	1.01	1.49
C70	1.02	1.01	1.48

Table 7.6.
Displays CCSs calculated for protein molecules (>1000 atoms). Calculations were performed in all the three methods.

Structures*	Number of atoms	Excluding partial charge(A2)			
		PA	DHSS	TM(L-J)	TDHSS
ADH	20653	6003.84	8453.65	8641.69	8780.01
CytC	1744	1090.98	1527.34	1635.95	1651.90
Lyso	1959	1174.62	1647.67	1710.66	1768.33
Ub	1229	862.80	1196.33	1267.10	1282.48
BSA	9225	3736.38	5194.92	5344.34	5454.59
ConA	14305	4608.45	6474.73	6701.06	6775.79
NIST	20276	7887.36	11016.58	11235.39	11493.23
PK	32300	8233.56	11633.60	11800.56	11966.76
SAP	16400	5609.87	7853.82	7906.23	8012.34

Table 7.7.
Displays CCSs calculated for protein molecules (>1000 atoms). Calculations were performed in all the three methods. Values for $L\xi PA$ are also included, as discussed in the manuscript.

Structures*	**Number of atoms**	**Including partial charge(A2)**				$L\xi PA(A)$
		PA	**DHSS**	**TM(L-J)**	**TDHSS**	
ADH	20653	6010.09	8451.61	8590.50	8813.31	8493.56
CytC	1744	1089.56	1524.36	1629.26	1664.63	1626.84
Lyso	1959	1176.21	1644.49	1710.59	1790.22	1732.48
Ub	1229	861.83	1199.42	1270.63	1304.16	1264.26
BSA	9225	3729.43	5210.70	5370.90	5482.18	5283.62
ConA	14305	4609.98	6480.35	6665.86	6785.22	6565.85
NIST	20276	7894.49	11023.76	11095.01	11343.93	10996.42
PK	32300	8232.32	11611.60	11866.93	12096.87	11576.82
SAP	16400	5609.87	7853.82	7875.04	7916.74	7971.03

APPENDIX B

Table 8.1.
Consists of time calculation for structure in MOBCAL and IMoS. Values for IMoS calculated on different number of threads (1, 5, and 7) are present here.

Structures	Number of atoms	MOBCAL	IMoS 1 Core	IMoS 5 Core	IMoS 7 Core
		[I]	[II]	[III]	[IV]
Tryphenylene	30	1659	140.04	40.93	34.65
NEthylaniline	21	1312	113.84	34.00	26.99
Dexamethasone	58	2450	230.22	61.85	50.66
Acetaminophen	21	1074	112.22	31.90	26.20
Betamethasone	58	2540	220.15	62.61	49.30
Anthracene	24	1250	122.39	34.87	28.44
Choline	21	1257	115.55	32.84	26.23
Phenanthrene	24	1350	123.13	34.16	27.50
Acetylcholine	26	1473	133.82	37.53	29.76
C60	60	3201	240.99	64.46	51.55
C70	70	1876	265.98	73.38	58.23
Naphtalene	18	1142	116.10	31.74	25.08
Paracetamol2	21	1187	122.02	32.41	25.86
Pyrene	26	1389	139.94	37.78	30.02
TtEA	29	1678	172.94	42.22	32.59
TMA	17	1146	116.32	32.58	25.78
C20	20	903	44.96	12.96	11.79
C40	40	1646	63.76	18.58	16.24
C100	100	2494	133.22	38.08	34.04

Table 8.2.
Consists of time calculation for structure in MOBCAL and IMoS. Values for IMoS calculated on different number of threads (1, 5, and 7) are present here.

Structures	**Number of atoms**	**MOBCAL**	**IMoS 1 Core**	**IMoS 5 Core**	**IMoS 7 Core**
		[I]	**[II]**	**[III]**	**[IV]**
C180	180	6341	247.51	70.00	64.07
C240	240	8670	350.47	97.87	90.30
C500	500	15377	813.27	227.04	209.82
C540	540	20655	893.43	260.87	235.69
C720	720	32817	1269.15	363.02	332.64
16EMINCN2	384	11069	1249.17	251.66	198.09
32EMINCN2	768	25655	2055.84	557.44	440.17
64EMINCN242	1612	52836	4003.12	1262.35	997.15
TPA	39	1557	113.58	24.14	18.57
TBA	53	1825	114.53	30.77	23.47
TDA	125	2884	228.05	71.30	53.66
TDDA	149	3114	206.15	62.70	48.88
TPA4	41	1534	78.41	24.49	18.50
THA2	89	2251	169.20	51.40	39.46
PEG70Cs4	497	9777	563.01	168.26	142.55
PEG90Cs4	630	14141	1189.49	245.73	209.41
PEG115Cs43	805	14442	1509.69	313.44	267.77
PEG144Cs4	1008	27671	2256.31	437.97	379.74
PEG254Cs4	1778	69109	4220.83	1225.03	1054.31

Table 8.3.
Ratios for time (MOBCAL to different threads of IMoS) are provided for comparison purposes.

Structures	**Number of atoms**	**CCS Ratio**		
		[I]/[II]	**[I]/[III]**	**TM/EHSS**
Tryphenylene	30	11.85	40.53	47.89
NEthylaniline	21	11.52	38.59	48.61
Dexamethasone	58	10.64	39.61	48.36
Acetaminophen	21	9.57	33.67	41.00
Betamethasone	58	11.54	40.57	51.52
Anthracene	24	10.21	35.85	43.95
Choline	21	10.88	38.27	47.92
Phenanthrene	24	10.96	39.52	49.08
Acetylcholine	26	11.01	39.25	49.50
C60	60	13.28	49.66	62.10
C70	70	7.05	25.56	32.22
Naphtalene	18	9.84	35.98	45.53
Paracetamol2	21	9.73	36.62	45.89
Pyrene	26	9.93	36.77	46.27
TtEA	29	9.70	39.74	51.48
TMA	17	9.85	35.18	44.45
C20	20	20.08	69.68	76.60
C40	40	25.82	88.57	101.33
C100	100	18.72	65.49	73.27

Table 8.4.
Ratios for time (MOBCAL to different threads of IMoS) are provided for comparison purposes.

Structures	**Number of atoms**	**CCS Ratio**		
		[I]/[II]	**[I]/[III]**	**TM/EHSS**
C180	180	25.62	90.59	98.98
C240	240	24.74	88.59	96.01
C500	500	18.91	67.73	73.29
C540	540	23.12	79.18	87.64
C720	720	25.86	90.40	98.66
16EMINCN2	384	8.86	43.98	55.88
32EMINCN2	768	12.48	46.02	58.28
64EMINCN242	1612	13.20	41.86	52.99
TPA	39	13.71	64.51	83.86
TBA	53	15.94	59.32	77.76
TDA	125	12.65	40.45	53.75
TDDA	149	15.11	49.66	63.71
TPA4	41	19.56	62.64	82.93
THA2	89	13.30	43.79	57.04
PEG70Cs4	497	17.37	58.11	68.59
PEG90Cs4	630	11.89	57.55	67.53
PEG115Cs43	805	9.57	46.08	53.93
PEG144Cs4	1008	12.26	63.18	72.87
PEG254Cs4	1778	16.37	56.41	65.55

Table 8.5.
Lists the values for CCS and time calculated in IMoS and MOBCAL in N2.

Structures	**Number of atoms**	**Time (In Seconds)**			
		MOBCAL TM (4E5)	**IMoS TMLJ (4E5)**	**IMoS TMLJ (9E5)**	**IMoS EHSS (4E5)**
		[I]	**[I]**	**[I]**	**[IV]**
TMA	17	9720	49.86	108.99	1.98
Naphtalene	18	10800	53.47	115.59	1.58
N Ethylaniline	21	12480	58.37	128.65	1.82
Acetaminophen	21	12720	57.18	127.58	1.94
Choline	21	12420	58.27	129.31	1.92
Paracetamol2	21	12000	56.45	126.95	1.61
Anthracene	24	14400	61.09	135.95	1.74
Phenanthrene	24	13686	61.55	138.42	1.66
Acetylcholine	26	15600	67.83	150.41	1.68
Pyrene	26	15660	66.80	147.61	1.71
TtEA	29	17040	69.43	156.76	2.12
Tryphenylene	30	18180	75.90	171.25	1.62
Dexamethasone	58	31920	129.46	292.94	2.06
Betamethasone	58	31860	129.33	289.40	2.05
C60	60	38880	113.05	252.77	2.63
C70	70	45120	133.96	301.67	2.69

Table 8.6.
Time ratio for N2

Structures	Number of atoms	Time Ratios	
		[I]/[II]	[I]/[III]
TMA	17	194.96	89.18
Naphtalene	18	202.00	93.43
N Ethylaniline	21	213.81	97.01
Acetaminophen	21	222.47	99.70
Choline	21	213.16	96.05
Paracetamol2	21	212.57	94.52
Anthracene	24	235.73	105.92
Phenanthrene	24	222.34	98.87
Acetylcholine	26	230.00	103.72
Pyrene	26	234.44	106.09
TtEA	29	245.42	108.70
Tryphenylene	30	239.51	106.16
Dexamethasone	58	246.56	108.96
Betamethasone	58	246.35	110.09
C60	60	343.91	153.81
C70	70	336.81	149.57

Table 8.7.
Below displays the time required to calculate CCS for protein molecules.

Structures	Number of atoms	Excluding partial charge(seconds)			
		PA	DHSS	TM(L-J)	TDHSS
ADH	20653	69.23	195.11	2725.16	464.31
CytC	1744	8.25	9.86	122.07	20.87
Lyso	1959	8.14	10.10	143.89	20.23
Ub	1229	5.88	5.36	78.94	16.24
BSA	9225	34.61	51.02	1079.32	225.78
ConA	14305	49.27	147.44	1668.86	270.81
NIST	20276	89.09	192.39	4044.34	779.93
PK	32300	140.43	346.64	6636.44	966.23
SAP	16400	56.03	189.51	1924.85	310.87

Table 8.8.
Below displays the time required to calculate CCS for protein molecules.

Structures	**Number of atoms**	**Including partial charge(seconds)**			
		PA	**DHSS**	**TM(L-J)**	**TDHSS**
ADH	20653	68.88	201.50	13426.27	9851.43
CytC	1744	11.22	12.41	572.77	439.26
Lyso	1959	10.90	12.56	650.30	491.53
Ub	1229	8.02	0.23	364.34	289.23
BSA	9225	48.15	56.24	4525.04	3479.86
ConA	14305	69.39	158.65	7713.84	5752.63
NIST	20276	119.89	191.82	11016.80	7163.20
PK	32300	124.74	342.74	29171.89	18977.07
SAP	16400	72.60	193.67	9178.26	6144.02