



**BERGISCHE
UNIVERSITÄT
WUPPERTAL**

Theoretical Investigations on Cluster Formation in Mass Spectrometry

Dissertation

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von
Alexander Haack

Bergische Universität Wuppertal
Fakultät für Mathematik und Naturwissenschaften
Physikalische und Theoretische Chemie
unter der Anleitung von Prof. Dr. Thorsten Benter

Wuppertal, März 2020

The PhD thesis can be quoted as follows:

urn:nbn:de:hbz:468-20200312-111139-2

[<http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20200312-111139-2>]

DOI: 10.25926/x68p-cx61

[<https://doi.org/10.25926/x68p-cx61>]

Erklärung

„Ich versichere, die Arbeit selbstständig verfasst zu haben, nur die in der Dissertation angegebenen Hilfsmittel benutzt und alle wörtlich oder inhaltlich übernommenen Stellen als solche gekennzeichnet zu haben, und dass die Dissertation in der gegenwärtigen oder einer anderen Fassung noch keiner anderen Fakultät, keiner Gesamthochschule und keiner anderen wissenschaftlichen Hochschule vorgelegen hat.“

Wuppertal, den
11. März 2020



Alexander Haack

One has the impression that modern physics is based on assumptions which somehow are similar to the smile of a cat that is not even there.

Albert Einstein

"Well! I've often seen a cat without a grin," thought Alice; "but a grin without a cat! It's the most curious thing I ever saw in all my life!"

Lewis Carroll - Alice in Wonderland

If quantum mechanics hasn't profoundly shocked you, you haven't understood it yet.

Niels Bohr

Danksagung

Es ist nicht selbstverständlich, eine begonnene Arbeit auch zum Abschluss zu bringen. Neben den eigenen Leistungen denke ich, dass das Umfeld, in dem man sich befindet, einen mindestens genauso großen Beitrag dazu leistet. Gleich einem Nährboden, in dem ein Keim wächst, kann sich eine Arbeit nur entfalten, wenn die Umgebung alle nötigen Voraussetzungen dazu liefert. In diesem Sinne möchte ich gerne ein paar Menschen danken.

Das wissenschaftliche Umfeld, die Physikalische und Theoretische Chemie der Universität Wuppertal (PTC), hat mich stets mit ihrer Arbeitsatmosphäre, ihrem sozialen und wissenschaftlichen Charakter, gestärkt, motiviert und unterstützt. Nicht nur unter exzellenten Wissenschaftler*innen, sondern auch bei ausgezeichneten Freunden habe ich Arbeiten, Lachen, Schaffen und mich zu Hause fühlen dürfen. Das schließt natürlich alle Mitglieder der PTC ein, ich möchte dennoch namentlich Thorsten Benter, Hendrik Kersten, Christine Polaczek, Marco Thinius, Kai Kroll und Stefan Hellhake danken. Diese Namen werde ich am Stärksten mit dieser Zeit meines Lebens verbinden.

Dasselbe gilt für die Arbeitsgruppe der Physical Chemistry an der University of Waterloo, ON, Canada. Auch wenn es nur drei Monate waren, die ich dort verbringen durfte, empfand ich dort das gleiche Gefühl von „zu Hause sein“. Auch hier möchte ich ein paar Namen hervorheben: Scott Hopkins, Jeff Crouse und Christian Ieritano waren mir ebenso gute Begleiter in Wissenschaft und Freundschaft.

Auch das private Umfeld trägt auf seine ganz eigene Weise dazu bei, dass die begonnene Arbeit vollendet wird. Ein unerschütterlicher Glaube, ein „das ist das Richtige für dich“ ist vielleicht sogar überzeugender, wenn die aussprechende Person gerade *nicht* dem wissenschaftlichen Inhalt der Arbeit folgen kann, sondern aus einer rein sozialen Perspektive erkennt, welche Wünsche und Träume in einem stecken. Genau das habe ich immer bei meiner Familie und meinen Freunden erfahren dürfen und danke all diesen Menschen von Herzen für ihre Motivation, ihre Unterstützung und natürlich auch für ihr Vermögen, mich von Zeit zu Zeit abzulenken.

Abstract

The thermochemistry and reaction dynamics of polar gas phase molecules clustering around an ion in atmospheric pressure ionization mass spectrometry (API-MS) dominate many observed effects. While the equilibrium conditions under elevated pressure conditions determine the distribution of the available charges inside different clusters, the reaction dynamics under non-equilibrium conditions downstream of the instrument will determine the ultimate fate of these clusters. Together, this will dictate the final degree of charging (e.g., protonation) of the analyte.

In this work, two examples of the effect of clustering in MS are studied by using Computational Chemistry. In the first example, charge depletion and charge retention in nanoESI-MS for a small peptide, Substance P, is investigated. By modeling the thermochemical data and potential energy paths for proton transfer (PT), it is shown that protic solvent vapors like methanol (MeOH) can abstract a proton from the multiply charged peptide. This is due to their ability to form hydrogen bond networks. The resulting charge dilution upon increasing cluster sizes lowers barrier heights along the PT paths. Aprotic solvent molecules as for example acetonitrile (ACN), however, lack the ability to form cluster networks at the charge site. Thus, there is no trend observed regarding the barrier heights for PT. The charge is thus retained and not depleted. This is in accordance with experimental findings, where Substance P is observed in its 3+ charge state when ACN is added to the gas phase, but is found mainly in its 2+ charge state with MeOH vapor added.

The second example deals with the effect of dynamic clustering in differential mobility spectrometry. This is done by modeling the geometry, temperature dependent collision cross section (CCS), and electrical field dependent ion temperature and mobility of many cluster structures in the framework of the Mason-Schamp equation applying the two-temperature theory. Thermochemical weighting yield the cluster size distribution and ensemble averaged mobility in dependence of the field strength. It is shown that increasing the field strength leads to higher ion temperatures and subsequent evaporation of the solvent molecules, which are firmly bound at low field strengths. This leads to an overall decrease of CCS and thus to an increase of mobility. Calculating dispersion plots from these data and comparison with experimental results, semi-quantitative agreement is found. Especially the good accordance with trends observed regarding different solvents, the background temperature and the solvent vapor concentration highlights the importance of dynamic clustering for differential mobility. The importance of correctly modeling the CCS as a function of temperature and using two-temperature theory to model the "hard-sphere" effect is also emphasized.

From these two case studies, it becomes apparent that the thermochemical stability of cluster systems needs to be considered. However, a strong obstruction in the modeling of these cases is the breakdown of the harmonic approximation (HA) usually used in the calculations. Due to their loosely bound nature, anharmonic effects gain significance. Two new approaches are presented in this work to account for these effects for a more accurate description of the thermochemistry. The first uses the well known Vibrational Perturbation Theory to 2nd order (VPT2) in a hybrid approach. To reduce computational time, only the harmonic contributions are calculated with high accuracy, while the minor anharmonic effects are calculated with faster, less accurate methods. The second approach uses a modified version of the quasi-harmonic approximation (QHA). Instead of using Cartesian or internal coordinates, the configurational distribution, obtained from molecular dynamics (MD) simulations, is represented in normal coordinates. Overall translation, overall rotation and internal rotations are projected out of the MD trajectories by enforcing explicitly the Eckart-Sayvetz conditions for these motions, leaving only true vibrational movement. Both developed methods were tested by calculating the dissociation enthalpies of different proton bound cluster systems. Comparison with standard methods regarding experimental results showed that both methods have similar or better accuracy than the HA. Similar accuracy is only achieved for challenging cases with many internal rotations or very flexible structures. Here, the developed methods as well as the HA exhibit significant deviations from experimental results. Thus, the two methods capture the anharmonic effects to some degree and perform better than the HA for systems of small to medium anharmonicity. For strongly anharmonic systems their performance has room for improvement.

Contents

1	Motivation	1
2	Introduction to Computational Chemistry	7
2.1	Electronic Structure Calculations	8
2.1.1	Hartree-Fock Theory	8
2.1.2	Density Functional Theory	12
2.2	Nuclear Position and Movement	15
2.2.1	Geometry Optimization	15
2.2.2	Vibrations and Rotations	18
2.2.3	Molecular Dynamics Simulations	22
2.3	Thermochemistry	23
2.4	Reaction Dynamics	26
2.5	Collision Cross Section Modeling	30
3	Charge Retention/Charge Depletion in ESI-MS	35
3.1	Abstract	35
3.2	Introduction	36
3.3	Methods	38
3.3.1	Computational Methods	38
3.3.2	Experimental Methods	38
3.4	Results and Discussion	39
3.4.1	Structure and Thermochemistry of the Different Clusters	39
3.4.2	Proton Transfer Paths	46
3.4.3	Multiply Charged Analytes	49
3.5	Conclusion	53
3.6	Supporting Information	55
3.7	Follow-up work	58
3.7.1	Description of the Method	58
3.7.2	Results and Discussion	60
3.7.3	Conclusion	62
4	Modeling of Differential Ion Mobility	63
4.1	Abstract	63
4.2	Introduction	64
4.2.1	Mobility at low field strength	64
4.2.2	Mobility at high field strength	65
4.2.3	Differential Mobility Spectrometry	65
4.3	Computational Methods	67
4.3.1	Ab initio calculations	67
4.3.2	CCS calculations	68
4.3.3	Mobility calculations	69
4.3.4	Boltzmann weighting	69

4.3.5	Dispersion Plot calculation	73
4.4	Results and Discussion	74
4.4.1	Cluster structures	74
4.4.2	CCS fits: MD vs. rigid samples	75
4.4.3	Mobility calculations	79
4.4.4	Dispersion plots	82
4.5	Conclusion	86
4.6	Supporting Information	87
4.7	Follow-up work	88
4.7.1	Methods	89
4.7.2	Results and Discussion	89
4.7.3	Conclusion	93
5	Thermochemistry from MD Simulations	95
5.1	Abstract	95
5.2	Introduction	95
5.3	Theory and Methods	98
5.3.1	The quasiharmonic approximation (QHA)	98
5.3.2	Implementation and special considerations	99
5.3.3	Electronic structure and MD simulations	102
5.4	Results and Discussion	102
5.4.1	$(\text{ACN})_2\text{H}^+$	102
5.4.2	$[\text{MeNH}_3 + \text{MeOH}]^+$	111
5.4.3	QHA discussion	114
5.5	Summary and Conclusions	117
5.6	Appendix A - The Eckart Frame	118
5.7	Appendix B - Separating out Internal Rotation	119
5.8	Supporting Information	120
6	Summary and Outlook	123
6.1	Answering Questions from Experiment	123
6.2	Developing new Calculation Methods	126
6.3	General Conclusion	128
	References	129

List of Figures

2.1	Basis Set Example in HF Calculations	11
2.2	PES and reaction coordinate example	17
2.3	Internal Rotation Example	20
2.4	Reaction Dynamics Considerations	27
2.5	General Collision Schematics	32
3.1	Geometries of the Homogeneous Solvent Clusters	41
3.2	Geometries of the Amine-Solvent Clusters	43
3.3	Thermodynamic Stability of Amine-Solvent Clusters	45
3.4	Proton Transfer Paths for the $[\text{MeNH}_2 + \text{H} + \text{S}_n]^+$ Clusters	47
3.5	Proton Transfer Paths for the $[\text{H} - \text{EDA} + \text{H} + \text{S}_n]^+$ Clusters	50
3.6	Measured Ion Population of Diamines	52
3.7	Mean Charge State of Diamines	53
3.8	Special Modes of the $(\text{ACN})_2\text{H}^+$ Cluster	56
3.9	Additional Structure of the $[\text{MeNH}_2 + \text{H} + 2\text{ACN}]^+$ Cluster	56
3.10	Mass Spectra of Diamine Cluster	57
4.1	New Quantum Harmonic Superposition Approximation	71
4.2	$[\text{Me}_4\text{N} + (\text{ACN})_n]^+$ cluster structures	75
4.3	Temperature dependence of the CCS	77
4.4	Calculated Mobility and Boltzmann weighing parameters	81
4.5	Calculated α function for the $[\text{Me}_4\text{N} + (\text{ACN})_n]^+$ system	82
4.6	Calculated and Experimental Dispersion Plots	83
4.7	Variation of Parameters in Dispersion Plot Calculations	85
4.8	Cluster Geometries for Me_4N with ACE and MeOH	87
4.9	α Function for $[\text{Me}_4\text{N} + (\text{ACE})_n]^+$ and $[\text{Me}_4\text{N} + (\text{MeOH})_n]^+$	88
4.10	Tricarbastannatrane Ion Geometry	89
4.11	Dispersion Plots for Tricarbastannatrane Ion	91
4.12	Population Analysis for Tricarbastannatrane Ion	93
5.1	Geometries of the $(\text{ACN})_2\text{H}^+$ and $[\text{MeNH}_3 + \text{MeOH}]^+$ Clusters	103
5.2	PES for $(\text{ACN})_2\text{H}^+$ Dissociation Pathways	104
5.3	Special Modes of the $(\text{ACN})_2\text{H}^+$ Cluster - revised	105
5.4	NuSol Results for Proton Movement in $(\text{ACN})_2\text{H}^+$	107
5.5	Internal Rotation Correction in $(\text{ACN})_2\text{H}^+$	108
5.6	Diagnostic plots for the QHA	109
5.7	Normal Frequency Assignment in the $[\text{MeNH}_3 + \text{MeOH}]^+$ cluster	113
5.8	Boltzmann plots of \mathcal{E}_{kin} distributions for $(\text{ACN})\text{H}^+$	115

List of Tables

2.1	Partition Functions for Molecules	24
3.1	Validation of Model Chemistry	39
3.2	Stability of the Homogeneous Solvent Clusters	41
3.3	Stability of the Amine-Solvent Clusters	43
3.4	Validation of the Hybrid VPT2 Method for Thermochemistry	60
3.5	Proton Transfer Thermochemistry of the Amine-Solvent Clusters	61
4.1	Parameters for CCS Fitting Functions of Me_4N^+ -solvent clusters	79
4.2	Binding energies of tricarbastannatrane-solvent clusters	90
5.1	Geometry of the $(\text{ACN})_2\text{H}^+$ Cluster	103
5.2	QHA Results for the $(\text{ACN})_2\text{H}^+$ Cluster	111
5.3	QHA Results for the $[\text{MeNH}_3 + \text{MeOH}]^+$ Cluster	114
5.4	Normal Frequencies of ACN and $(\text{ACN})\text{H}^+$	120
5.5	Normal Frequencies of the $(\text{ACN})_2\text{H}^+$ Cluster	121
5.6	Normal Frequencies of the $[\text{MeNH}_3 + \text{MeOH}]^+$ Cluster	122

Abbreviations and Nomenclature

List of Abbreviations

ADMP	Atom-Centered-Density-Matrix-Propagation
AIMD	Ab Initio Molecular Dynamics
AP	Atmospheric Pressure
APCI	Atmospheric Pressure Chemical Ionization
API	Atmospheric Pressure Ionization
B2PLYP	B2PLYP Exchange-Correlation Functional including perturbative corrections
B3LYP	B3LYP Exchange-Correlation Functional
BAT	Bond-Angle-Torsion, Internal Coordinates
BH	Basin Hopping
CASSCF	Complete Active Space Self Consistent Field
CC, CCSD, CCSD(T)	Coupled Cluster Calculations using Singly (S), Doubly (D) or Triply (T) Excited Determinants
CCS	Collision Cross Section
cf.	<i>confer</i> (Latin), compare
CI, CISD, CISD(T) ..	Configuration Interaction using Singly (S), Doubly (D) or Triply (T) Excited Determinants
CID	Collision Induced Dissociation
CoM	Center of Mass
CV	Compensation Voltage
DFT	Density Functional Theory
DMS	Differential Mobility Spectrometry
EI	Electron Ionization
Eq.	Equation
ESI	Electrospray Ionization
etc.	<i>et cetera</i> (Latin), and other similar things
FAIMS	Field-Asymmetric Waveform Ion Mobility Spectrometry
Fig.	Figure
GB	Gas-Phase Basicity
GD3, GD3(BJ)	Grimme's Dispersion Correction #3 w/ or w/o Becke-Johnson damping

GTO	Gaussian-Type Orbital
HA	Harmonic Approximation
HF	Hartree-Fock
i.e.	<i>id est</i> (Latin), that is
IID	Ion-Induced Dipole
IMS	Ion Mobility Spectrometry
IQ	Ion-Quadrupole
IR	Internal Rotation
IRC	Intrinsic Reaction Coordinate
IRMPD	Infra-Red Multiple Photon Dissociation
IVR	Intra-Molecular Vibrational Redistribution
LC	Liquid Chromatography
LCAO	Linear Combination of Atomic Orbitals
MC	Monte Carlo
MD	Molecular Dynamics
MO	Molecular Orbital
MP2, MP3, MP4	Møller-Plesset Perturbation Theory to 2 nd , 3 rd or 4 th Order
MS	Mass Spectrometry, Mass Spectrometer
nESI	nano-ESI
NMA	Normal Mode Analysis
PA	Proton Affinity
PES	Potential Energy Surface
PT	Proton Transfer
QHA	Quasi-Harmonic Approximation
QHSA	Quantum-Harmonic Superposition Approximation
QM	Quantum Mechanical
RIP	Reactant Ion Peak
RRKM Theory	Rice-Ramsperger-Kassel-Marcus Theory
SCA	Supercharging Agents
SCF	Self-Consistent Field
SP	Substance P
SV	Separation Voltage
TIC	Total Ion Current
TS	Transition State
vdW	van der Waals
VPT2	Vibrational Perturbation Theory to 2 nd Order
VSCF	Vibrational Self-Consistent-Field
ZPE	Zero-Point Energy

Important Nomenclature

$\alpha(E)$	Alpha Function in Differential Mobility
σ	Covariance Matrix of a Set of Coordinates
F	Force Constant Matrix
\mathbf{g}	Gradient
\mathbf{G}	Kinetic Energy Matrix
\mathbf{M}	Mass Matrix
\mathbf{q}	Internal Coordinates
\mathbf{Q}	Normal Coordinates
\mathbf{r}	Cartesian Coordinate Vector of Electrons
\mathbf{R}	Cartesian Coordinate Vector of Nuclei
χ	Vibrational Anharmonic Constant
$\Delta_R G$	Gibbs Reaction Enthalpy
η, φ	HF or DFT Molecular Orbital, respectively
\hat{H}, H	Hamilton Operator, Hamiltonian
\hat{T}	Kinetic Energy Operator
\hat{V}, V	Potential Energy Operator, Potential
$\langle \cdots \rangle$	Ensemble or Temporal Average
\mathcal{E}	Energy
\mathcal{H}	Hessian, 2 nd Derivative Matrix
\mathcal{Z}	Superposition Partition Function
μ	Reduced Mass
Ω	Collision Integral, Collision Cross Section
ω, ν	Normal and Fundamental Frequency
ϕ	Basis Function
ψ	General Wavefunction
ρ_e	Electron Density
ρ_S	Degeneracy of a Level or Density of States
θ	Scattering Angle in Collision Event
ζ	Decay Coefficient of a Basis Function
A, B, C	Rotational Constants along Principle Axes
b	Impact Parameter in Collision Events
CV	Compensation Voltage
E	Electrical Field Strength
F	Force Constant
G_{corr}	Thermal Corrections to Gibbs Enthalpy
I	Moment of Inertia
k	Reaction Rate Constant

K	Coefficient of Ion Mobility
m	Mass
M_e	Number of Electrons in a Molecule
M_N	Number of Nuclei in a Molecule
N	Particle Number or Number Density
P	Population, Population Density, Portion
SV_{pp}	Separation Voltage, Peak-to-Peak
T	Absolute Temperature
v_D	Ion Drift Velocity
v_r	Relative Velocity in Collision Events
z	Charge Number of a Nucleus or Molecule
Z	Partition Function

Non-SI Units and Natural Constants

ϵ_0	Vacuum Electric Permittivity, $\epsilon_0 = 8.854\,187\,812\,8 \times 10^{-12} \text{ F m}^{-1}$
\hbar	Reduced Planck Constant, $\hbar = 1.054\,571\,817 \times 10^{-34} \text{ J s}$
c	Speed of Light, $c = 299\,792\,458 \text{ m s}^{-1}$
e	Elementary Charge, $e = 1.602\,176\,634 \times 10^{-19} \text{ C}$
h	Planck Constant, $h = 6.626\,070\,15 \times 10^{-34} \text{ J s}$
k_B	Boltzmann Constant, $k_B = 1.380\,649 \times 10^{-23} \text{ J K}^{-1}$
\AA	Angstrom, $\text{\AA} = 1 \times 10^{-10} \text{ m}$
a_0	Bohr radius, $a_0 = 5.291\,772\,109\,03 \times 10^{-11} \text{ m}$
E_h	Hartree Energy, $E_h = 4.359\,744\,722\,207\,1 \times 10^{-18} \text{ J}$
amu	Atomic Mass Unit, $\text{amu} = 1.660\,539\,066\,60 \times 10^{-27} \text{ kg}$
eV	Electron Volts, $\text{eV} = 1.602\,176\,634 \times 10^{-19} \text{ J}$
Td	Townsend, $\text{Td} = 10^{-21} \text{ V m}^2$

1 Motivation

Mass spectrometry (MS) is a widely used tool in many fields of analytical chemistry: In biochemistry, protein sequencing [1], lipids [2], metabolites [3], carbohydrates [4, 5], protein-drug interactions [6, 7] and many more are studied. Fuel composition [8], environmental analysis [9], food-safety [10], and structure elucidation of small molecules [11] are also important fields in which MS is used, as well as fundamental areas as for example ion-spectroscopy [12, 13], ion-molecule interactions [14], photo- and ion-dissociation [15–17]. This list is by no means complete and consequently the development of improved MS instruments and methods is not only subject to research [18] but also commercially interesting, valued almost 5 billion USD on the global market in 2016 and growing [19].

In general, an analyte molecule is first ionized in an environment (solid, liquid, gaseous, vacuum) with a certain ionization method. Through electrical fields, the ion is guided to an analyzer section of the instrument, in which different ions are separated by their mass-to-charge ratio, m/z . Finally, the ion is detected. Scanning the analyzer to allow for many different m/z and recording the ion intensity at each step yields a mass spectrum. Each stage, i.e., the ionization environment (and the transfer into it), the particular ionization method, the ion guiding, the m/z analysis, and the ion detection is a whole research field on its own and many different methods for all stages have been developed. It is not the scope of this work to give a detailed description of these stages, the reader is referred to one of the many textbooks about mass spectrometry (for example [20, 21]). However, the first stages, i.e., the ion generation and transfer are of great importance for this work and should be reviewed here. Since it is more common, only positive ionization mode will be discussed. The conditions in the ion transfer stages are also of interest when the pressure has to be reduced between ion source and mass analyzer, and will be outlined briefly.

Low Pressure Ionization In the most simple of cases, the analyte A is already in the gas phase and is ionized by means of energetic electrons (70 eV) at low pressures (around 10^{-6} mbar) according to



a process known as *electron ionization* (EI). Because a good amount of energy is transferred to the internal degrees of freedom of the analyte and there are hardly any collisions at that pressure to dissipate the excess energy, the $A^{+\bullet}$ often fragments inside the ion source region. The fragments produced through a cascade of unimolecular

1 Motivation

decay reactions are highly reproducible and can be used for identification and structural elucidation when compared with data base spectra. In this type of instrument, pressure reduction between ionization source and mass analyzer is not that crucial because the ion source pressure is already very low.

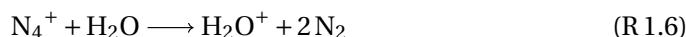
Another ionization technique is called *chemical ionization*. Here, the analyte is diluted in a reactant gas, R, (e.g., CH₄, NH₃, ...) which is primarily ionized through EI. Due to elevated pressures in the ion source (a few mbar), the ionized reactant molecules can transfer their charge to the analyte in a collision event via multiple ways, e.g., proton transfer or charge transfer:

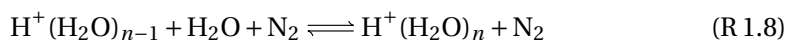


These processes deposit far less excess energy in the ionic products and thus less fragmentation is observed. The information of the molecular mass, often not obtained with EI, is very valuable for analyte identification. It also enhances sensitivity in routine applications because the signal intensity is not distributed over many fragments.

However, since multiple channels for ionization are possible and there may be enough collisions for further reactions, more than one signal can be observed. The rate constants for different processes (kinetics) and - given sufficient time - maybe even the thermodynamic stability of different products (thermochemistry) will influence the observed ion distribution. Already at this stage, ion-molecule reactions are pivotal for either the explanation of an observed mass spectrum or the design of an experiment, e.g., the choice of the reactant gas.

Atmospheric Pressure Ionization There are many reasons why ionization at atmospheric pressure (API) is today one of the most prominent ionization methods. Many analytes in chemistry, especially in biochemistry, are present in aqueous solution. Transferring them into the gas phase in sufficient amounts is rather challenging when the pressure has to remain low. However, if the ionization chamber is tolerant to high pressures, the analyte solution (or the liquid exiting a liquid chromatograph) can be sprayed through a heated capillary into a stream of heated nitrogen, efficiently transferring the analyte to the gas phase. This allows the coupling of liquid chromatography (LC) with MS, which was one of the initial goals of introducing API-MS [22, 23]. Because a heated filament for the production of electrons is not stable under such conditions, new ionization techniques had to be developed. The first instruments used a ⁶³Ni foil to produce highly energetic electrons via β decay. Already in 1973 [22], the authors noticed the following reaction cascade, established only three years earlier [24]:

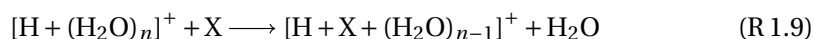




The proton bound water clusters, $\text{H}^+(\text{H}_2\text{O})_n$, were always observed in the spectra, even without any solvent introduced, and it was argued that traces of water are always present on the instrument walls and thus in the ion source gas phase. The ubiquitous presence of water can be stretched even further: Since the ion sources at atmospheric pressure are not required to be completely air-tight, air from the laboratory can enter the chamber. Even an air-tight source will adsorb rather sticky water molecules on all surfaces upon exposure to ambient air. Additionally, the modern ambient ionization methods are operated in the laboratory air directly, being exposed to water in the low percent range [25].

At atmospheric pressure, termolecular reactions (Reactions (R 1.5) and (R 1.8)) are feasible and due to residence times of up to 100 ms in the ionization chamber [23] it can be safely assumed that the system is actually in thermal equilibrium [23, 25] - which is in stark contrast to the vacuum ionization techniques. The thermochemistry of the water cluster system is well established [26] and the cluster size distribution P_n depends on the temperature and the partial pressure of water. In general these clusters are comparably stable: The free enthalpy of reaction for the $\text{H}^+(\text{H}_2\text{O}) + \text{H}_2\text{O} \longrightarrow \text{H}^+(\text{H}_2\text{O})_2$ reaction is $101.7 \text{ kJ mol}^{-1}$ at 300 K [26]. Thus, even for low mixing ratios of 1 ppmV water, the $n = 4$ cluster is the most abundant cluster size [25].

Further ionization techniques, e.g., plasma driven API sources (corona discharges [27] or helium plasmas [28–30]) follow essentially the same reaction cascade (R 1.4)-(R 1.8), yielding protonated water clusters, as reviewed in Ref. [25]. It was shown that API sources using VUV lamps or lasers to ionize dopant molecules prior to analyte ionization, essentially yield water clusters as well [31, 32]. Although it is always possible that certain species along this cascade, e.g., N_4^+ or He^{M} directly ionize the analyte molecule, the water clusters will play an essential role due to Kebarle's water exchange reaction [33]:



Here, X can be a solvent molecule S from the LC injection or the analyte A. As stated before, the system can be assumed to be in thermodynamic equilibrium. Thus, the available charges will be distributed over clusters of the general form $[\text{H} + (\text{H}_2\text{O})_n + (\text{S})_m + (\text{A})_l]^+$ where n , m and l are integers including zero, depending on the gas phase basicity and relative concentration of H_2O , S and A.

In other ionization techniques as for example Electrospray Ionization (ESI), the cluster mechanism will be important as well. In ESI, the analyte solution is sprayed inside a strong electrical field gradient, leading to the formation of a Taylor cone [34, 35]. This in turn leads to highly charged droplets, which eventually produce charged analyte molecules. The mechanism still remains elusive and several models have

1 Motivation

been proposed (ion evaporation model [36, 37], charge residue model [38–40], chain evaporation model [41–43]). However, in all cases, the analyte interacts with solvent molecules in the liquid phase, at the liquid/gaseous interface or in the gas phase (in the high pressure regions). Thus, the charges located at the analyte will be stabilized by interactions with solvent molecules and it is reasonable to assume that this stabilized system will survive the transition to the gas phase, as observed in many different computer simulations by e.g., the Konermann group [42–46] and supported by studies on the evaporation process in ESI [47]. Thus, this process eventually leads to clusters which - at sufficient pressure - will equilibrate with the surrounding environment.

Some mass spectrometers also contain a stage where ions are separated by their mobility (Ion Mobility Spectrometry, IMS) or differential mobility (Differential Mobility Spectrometry, DMS or Field-Asymmetric Waveform Ion Mobility Spectrometry, FAIMS) prior to mass analysis. Since the pressure conditions are often atmospheric pressure or slightly below, (dynamic) clustering with solvent molecules has been shown to be an important factor in these instruments [48–51].

Ion Transfer Since most mass analyzers operate at low pressures ($< 10^{-6}$ mbar) to provide a collision free environment necessary for ion separation, the pressure has to be reduced between the ion source and the analyzer region. This is typically done by multiple pressure reduction stages and differential pumping. In the first pressure reduction down to a few mbar, the ion/matrix mixture is expanded from the exit of a small diameter capillary or orifice into a much lower pressure region, creating a *hydrodynamic flow* or *jet expansion* [52]. There, the velocity distribution becomes much narrower, i.e., the random motion is transformed into the forward motion away from the capillary or orifice exit. As a consequence, also the internal degrees of freedom lose thermal energy and the ion/matrix mixture is effectively cooled. Since still sufficient collisions occur during this process, a new thermal equilibrium, corresponding to the low temperatures and concentrations of the matrix, is established. Sampling from this region can easily produce species as for example $[\text{H}^+(\text{H}_2\text{O})_{12}]^+$ [32]. Downstream of the Mach disc, created by the jet expansion, a new equilibrium is established because the forward kinetic energy is distributed back into all accessible degrees of freedom through collisions. Sampling behind the Mach disc shows considerably smaller cluster size distributions ($n = 4 \pm 1$) under the same conditions [32].

To efficiently separate the ions from the background gas, electrical fields are used to guide the ions while pumping away the neutrals. Under such conditions, the reduced electrical field strength, i.e., the ratio of the field strength and the particle density, can reach quite high values. Since the acceleration by the electrical field is countered only by very few collisions, the mean collision energy is elevated, potentially resulting in fragmentation of the clusters, possibly also of the analyte molecule [25, 32]. It is this stage, at which the dynamics - not the thermochemistry - of the possible proton

transfer reactions inside a cluster determines the final charge carrier observed in the mass analyzer. To reiterate, the thermochemistry at the various *equilibrium* stages determines the cluster size distribution $[H + (H_2O)_n + (S)_m + (A)_l]^+$ entering the ion transfer region and the proton transfer dynamics at the *non-equilibrium* stages dictate the final protonation site. The fact that indeed $[A + H]^+$ is observed with many different API methods, even if deuterated dopants are used [53], supports this reasoning [25]. It is of largest benefit to any analytical chemist that this complex system eventually leads to the singly or multiply protonated analyte, $[A + nH]^{n+}$.

Theoretical Investigations To exploit this benefit as much as possible, the thermochemical and dynamical processes mentioned can be studied in more detail. This might help to design the experiment (e.g., choice of solvent, additives) in favorable ways, increasing sensitivity and even selectivity. While thermochemical data of protonation reactions can be obtained experimentally by van-'t-Hoff plots and have been reported for many different cases [54], the system one is interested in, may have not. More so, reaction dynamics are governed by the shape of the molecular interaction potential surface, i.e., barrier heights along pathways or flatness of certain areas. The complex interplay of the energy needed for a reaction pathway and the state density of critical configurations, quantities that eventually determine the probability of a reaction (see section 2.4), are difficult to access experimentally.

For such cases, Computational Chemistry may aid in overcoming such difficulties as it gives insights into both, thermochemistry and reaction dynamics. Its capabilities and limitations will be discussed in detail in Chapter 2. Briefly, using quantum mechanics or parameters derived from that, the electronic and geometric structure of molecules can be studied. In particular, the potential energy surface, the underlying object that determines both thermochemistry and reaction dynamics, can be computed for - in principle - any given system. Although it is an ongoing field of research and a number of difficulties are still present, many different research fields including mass spectrometry, use Computational Chemistry to understand experimental results, support drawn conclusions and direct future work. Just to name a few examples from the MS community, Computational Chemistry is widely used in the modeling of peptide or protein structure [55–58], has been used to study various ionization mechanisms and characteristics [30, 59–61] and understand unknown ion transformation reactions [62–65], is combined with Infra-Red Multiple Photon Dissociation (IRMPD) experiments to identify isomers and protonation sites [15], was applied to characterize ion transfer conditions [66, 67], was used to study clustering in DMS devices [68, 69], among many others. Thus, it is the aim of this work to use computational methods to study a number of specific questions arising from MS experiments but also to combine, refine and develop new methods of calculation for more accurate data.

1 Motivation

Structure of this Work To address the aim of this work more precisely, Chapter 3 deals with ion transformation processes, i.e., charge depletion or charge retention due to clustering with added solvent molecules. These effects were observed experimentally in nanoESI-MS for a small peptide, Substance P, and the thermochemistry and reaction dynamics of these systems are studied using small model analytes. Chapter 4 focuses on the modeling of an ion's behavior in a DMS cell due to dynamic clustering with different solvent molecules. The ion's differential mobility is computed from first principles and compared with experimental data to support the suggested mechanism of dynamic clustering being a major factor in separating ions in a DMS cell. Chapter 5 deals with a more theoretical topic, i.e., the development of a new method to calculate thermochemical data of highly anharmonic species, such as proton bound clusters. The use of standard methods for the calculation of vibrational partition functions is a large source of error when dealing with such systems and consequently there is a need for more accurate methods, especially in the context of this work.

The main parts of the individual chapters are manuscripts which have been submitted for publication (first author). If the manuscripts are already published, the respective citation is given. Additional sections in each chapter describe supporting or follow-up work, which mostly has been submitted for publication as well (co-author work, also mentioned in the respective chapter).

2 Introduction to Computational Chemistry

This chapter introduces the reader to the general area of Computational or Quantum Chemistry. The topics covered are the foundation of what follows in the remaining chapters and thus, are not necessary to understand the results presented but rather are meant to introduce the general concepts, work flows and objects a quantum chemist deals with. It will also help to put the used and developed methods in to perspective and allows the reader to judge their capabilities and limitations. Basic concepts of quantum mechanics, i.e., wavefunctions, eigenvalues, operators, etc., are not discussed here. For literature suggestions about general quantum mechanics as well as quantum chemistry, the reader is referred to [70–74].

Computational Chemistry uses quantum mechanics to study atoms or molecules. Because the Schrödinger equation for such complex systems is not solvable, simplifications have to be made. A very obvious one is that only the time-*independent* Schrödinger equation is solved since the most systems are considered temporally stable. However, even then the total wavefunction of a molecule with M_N nuclei, $\psi_{\text{tot}}(\mathbf{r}, \mathbf{R})$, depends on the coordinates of the electrons, \mathbf{r} , as well as on the coordinates of the nuclei, \mathbf{R} , rendering the solution of the Schrödinger equation impossible even for simple systems such as H_2^+ . The Born-Oppenheimer approximation separates the total wavefunction into an electronic and a nuclear part and states that the kinetic energy of the nuclei does *not* affect the electronic wavefunction. This is reasoned by the fact that the electrons are moving much faster than the nuclei and thus "instantly" adapt to every new nuclear geometry. The total Schrödinger equation reads:

$$\begin{aligned} \hat{H}_{\text{tot}}\psi_{\text{tot}}(\mathbf{r}, \mathbf{R}) &= \mathcal{E}_{\text{tot}}(\mathbf{r}, \mathbf{R})\psi_{\text{tot}}(\mathbf{r}, \mathbf{R}) \\ (\hat{T}_e + \hat{T}_N + \hat{V}_{ee} + \hat{V}_{eN} + \hat{V}_{NN})\psi_{\text{tot}}(\mathbf{r}, \mathbf{R}) &= \mathcal{E}_{\text{tot}}(\mathbf{r}, \mathbf{R})\psi_{\text{tot}}(\mathbf{r}, \mathbf{R}) \end{aligned} \quad (2.1)$$

where \hat{H} is the Hamilton operator containing kinetic and potential energy operators, \hat{T} and \hat{V} , respectively, which consider the electrons (subscript e) and nuclei (subscript N) and their pairwise interactions, and \mathcal{E}_{tot} is the systems total energy. The Born-Oppenheimer approximation now splits the equation into two parts:

$$\underbrace{(\hat{T}_e + \hat{V}_{ee} + \hat{V}_{eN})}_{\hat{H}_e} \psi_e(\mathbf{r}; \mathbf{R}) = \mathcal{E}_e(\mathbf{r}; \mathbf{R})\psi_e(\mathbf{r}; \mathbf{R}) \quad (2.2a)$$

$$\underbrace{(\hat{T}_N + \hat{V}_{NN} + \mathcal{E}_e(\mathbf{R}))}_{\hat{H}_N} \psi_N(\mathbf{R}) = \mathcal{E}_{\text{tot}}(\mathbf{R})\psi_N(\mathbf{R}) \quad (2.2b)$$

2 Introduction to Computational Chemistry

where the electronic wavefunction ψ_e depends on the nuclear coordinates only parametrically and the electronic energy \mathcal{E}_e enters the nuclear Hamiltonian as simple function, independent from the electron movement. The two potential energy terms in the nuclear Hamiltonian \hat{H}_N , i.e., \mathcal{E}_e and \hat{V}_{NN} , where the latter is readily calculated from the Coulomb repulsion of the nuclei, are combined to the total potential energy of the molecule $\mathcal{E}_{\text{pot}}(\mathbf{R})$ at the nuclear geometry \mathbf{R} . Thus, solving the electronic part (Eq. (2.2a)) for many different but fixed nuclear geometries yields the *Potential Energy Surface* (PES), which in return is used as input for solving the nuclear part (Eq. (2.2b)) of the Schrödinger equation. The next two sections address each of these problems, respectively, the last three sections in this chapter describe how thermochemical parameters, quantities from reaction dynamics and collision cross sections are calculated from those results.

2.1 Electronic Structure Calculations

This section illustrates how Computational Chemistry solves the electronic part of the Schrödinger equation (Eq. (2.2a)). Due to the Born-Oppenheimer approximation, the nuclei are always spatially fixed at coordinates \mathbf{R} .

2.1.1 Hartree-Fock Theory

The Equations Since the problem is still not solvable analytically, the electronic wavefunction of an M_e -electron system is approximated through a Slater determinant containing one-electron wavefunctions, called spin orbitals:

$$\psi_e(r_1, \xi_1, \dots, r_{M_e}, \xi_{M_e}) = \frac{1}{\sqrt{M_e!}} \begin{vmatrix} \eta_1(r_1)\alpha(\xi_1) & \eta_1(r_1)\beta(\xi_1) & \eta_2(r_1)\alpha(\xi_1) & \cdots & \eta_m(r_1)\beta(\xi_1) \\ \eta_1(r_2)\alpha(\xi_2) & \eta_1(r_2)\beta(\xi_2) & \eta_2(r_2)\alpha(\xi_2) & \cdots & \eta_m(r_2)\beta(\xi_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \eta_1(r_{M_e})\alpha(\xi_{M_e}) & \eta_1(r_{M_e})\beta(\xi_{M_e}) & \eta_2(r_{M_e})\alpha(\xi_{M_e}) & \cdots & \eta_m(r_{M_e})\beta(\xi_{M_e}) \end{vmatrix} \quad (2.3)$$

The spin orbitals $\eta_j\alpha$ or $\eta_j\beta$ are composed of a spatial orbital $\eta_j(\mathbf{r}_i)$ depending on the spatial coordinates of electron i and a spin part $\alpha(\xi_i)$ or $\beta(\xi_i)$ depending on some spin coordinates ξ_i of electron i . This apparently cumbersome description is necessary to fulfill the Pauli principle, i.e., the wavefunction of a system containing identical fermions changes sign under exchange of two particles. This corresponds to an exchange of two rows of the matrix, which changes the sign of the determinant as required. Since each spatial orbital η_j is combined with two different spin functions, it is occupied with two electrons of opposite spin. This is a well-known empirical fact from the structure of the periodic table of the elements. The shown determinant corresponds to the case that M_e is an even number (closed-shell case) and $m = M_e/2$.

2.1 Electronic Structure Calculations

For a molecule, the spatial orbitals η_j are in general unknown and thus approximated through a linear combination of atomic orbitals (LCAO), ϕ_k :

$$\eta_j = \sum_{k=1}^{M_{\text{basis}}} c_{kj} \phi_k \quad (2.4)$$

These basis functions are located at the nuclei positions and thus mimic the concept of atomic orbitals combining to molecular orbitals (MOs). The type and number of basis functions applied is a crucial part in electronic structure calculations as will be discussed below.

The Slater determinant now consists of known functions but unknown coefficients c_{kj} . For their determination, the variational principle is applied, stating that the energy of any approximate wavefunction ψ_{approx} is greater or equal to the actual (ground state) energy \mathcal{E}_0 :

$$\mathcal{E}_0 \leq \mathcal{E}_{\psi_{\text{approx}}} = \langle \psi_{\text{approx}} | \hat{H}_e | \psi_{\text{approx}} \rangle \quad (2.5)$$

Thus, one has to find the coefficients c_{kj} such that the energy of the corresponding wavefunction becomes minimal. Setting the first derivative to zero:

$$\frac{\partial \langle \psi_{\text{approx}} | \hat{H}_e | \psi_{\text{approx}} \rangle}{\partial c_{kj}} = 0 \quad (2.6)$$

leads to the Roothaan-Hall equations, which are a set of non-standard eigenvalue problems. The coefficients c_{kj} , which one wants to solve for, are also needed as input and thus, the system is solved iteratively in a procedure called *Self-Consistent Field* (SCF). The name results from the fact that the electron density (the field in which the electrons move), determined by the summation over the occupied orbitals of the previous iteration step, influences the solution in the next iteration step and thus, eventually converges to a fixed density.

The output of these calculations are the coefficients c_{kj} , which yield the $M_e/2$ occupied and $(M_{\text{basis}} - M_e)/2$ virtual molecular orbitals η_j (Eq. (2.4)), as well as their corresponding energies ε_j . These type of calculations are called Hartree-Fock (HF) calculations and are the foundation of all ab initio approaches.

Basis Sets While the described formalism is fixed (except maybe the convergence threshold of the SCF equations), the basis set passed on to the HF equations is an important choice the user has to make. The accuracy of the calculations will greatly depend on this choice and thus a deeper understanding of the form of these basis set is beneficial. Since the Schrödinger equation for the H atom is analytically solvable, the general form of the atom orbitals is known: Each orbital is described by three quantum numbers (n, l, m_l) , with the principal quantum number $n = 1, 2, 3, \dots$, the azimuthal quantum number $l = 0, 1, \dots, n - 1$ and its magnetic orientation $m_l = -l, \dots, l$. This structure is reproduced by Gaussian-Type Orbitals (GTO) of the form:

$$\phi^{\text{GTO}} = x^a y^b z^c \exp(-\zeta(x^2 + y^2 + z^2)) \quad (2.7)$$

2 Introduction to Computational Chemistry

where x , y and z are the Cartesian coordinates with respect to the nucleus, the basis function is centered at. The sum $a + b + c$ mimics the azimuthal quantum number where $a + b + c = 0$ corresponds to an s -orbital, $a + b + c = 1$ corresponds to a p -orbital and so forth. ζ is a coefficient describing the radial decay of the wavefunction, small numbers corresponding to a slow decay. A minimal basis set would contain only one such function for each electron. However, such calculations are very inaccurate because the basis set is very inflexible. It is beneficial to apply much more basis functions than electrons to allow for a more flexible description of the resulting molecular orbitals. For example, Figure 2.1a shows a basis set for the H-atom containing two s - and one p -function (i.e., a $2s1p$ basis set). Having two s -orbitals (called *double-zeta* basis set) allows the HF formalism to mix them in the linear combination of Eq. (2.4) according to the variational principle, which will result in a more accurate description of the real wavefunction. Although the electron density around an H-atom is spherically symmetric, this might not be the case for an H-atom bound to, e.g., an oxygen atom. Thus, higher angular momentum functions (called *polarized functions*) are also added to allow for this inhomogeneity, in this case one p -function in each spatial direction. The first s -function is made out of three GTOs with already *fixed* coefficients, called a *contracted* Gaussian. The reason is that the form of the Gaussian, i.e., the $\exp(-\zeta r^2)$ decay, chosen because of easier numerical integration, actually is the wrong decay when compared to the analytical solutions of the H-atom, where a $\exp(-\zeta r)$ dependency is found. Thus, the sum of three GTOs is used to reproduce the correct decay.

Figure 2.1b shows a $3s2p1d$ basis set for oxygen. This basis is also of double-zeta quality because the valence shell ($n = 2$) has two functions for each orbital. However, the $1s$ orbital (core electrons) is only described by one 5-GTO contracted Gaussian. Since the core orbitals are less important for the chemical behavior, the number of basis functions are often reduced to save computational time. Having multiple-zeta quality for the valence shell while having a smaller number of functions for the core shell is called *split-valence* type basis set. Again, a higher angular momentum function (d -function) is given as polarizing function.

To see the results of the HF formalism, Figure 2.1c shows the three lowest occupied molecular orbitals obtained from the solution of the HF equations for the OH radical at 1 Å nuclear separation using the basis set shown in Figures 2.1a and 2.1b. As expected, the lowest MO is basically just the $1s$ -function of the oxygen atom. The second lowest MO is then a combination of the s -orbitals of both the oxygen and the hydrogen atoms, weakly contributing to the chemical bond. The third MO is mainly a combination of the oxygens $2p$ -functions with the $1s$ -function of the hydrogen atom, representing the main binding orbital (since electron density at both nuclei is significant). For clarity, higher orbitals including virtual ones are not shown.

The discussed basis set, known as def2-SVP [75], is a rather small one. For a good accuracy, at least triple-zeta quality with more than one polarization function is recommended [75]. Also *diffuse* functions, having very small ζ , are often added since they are able to describe electron density also far from the nucleus, important for

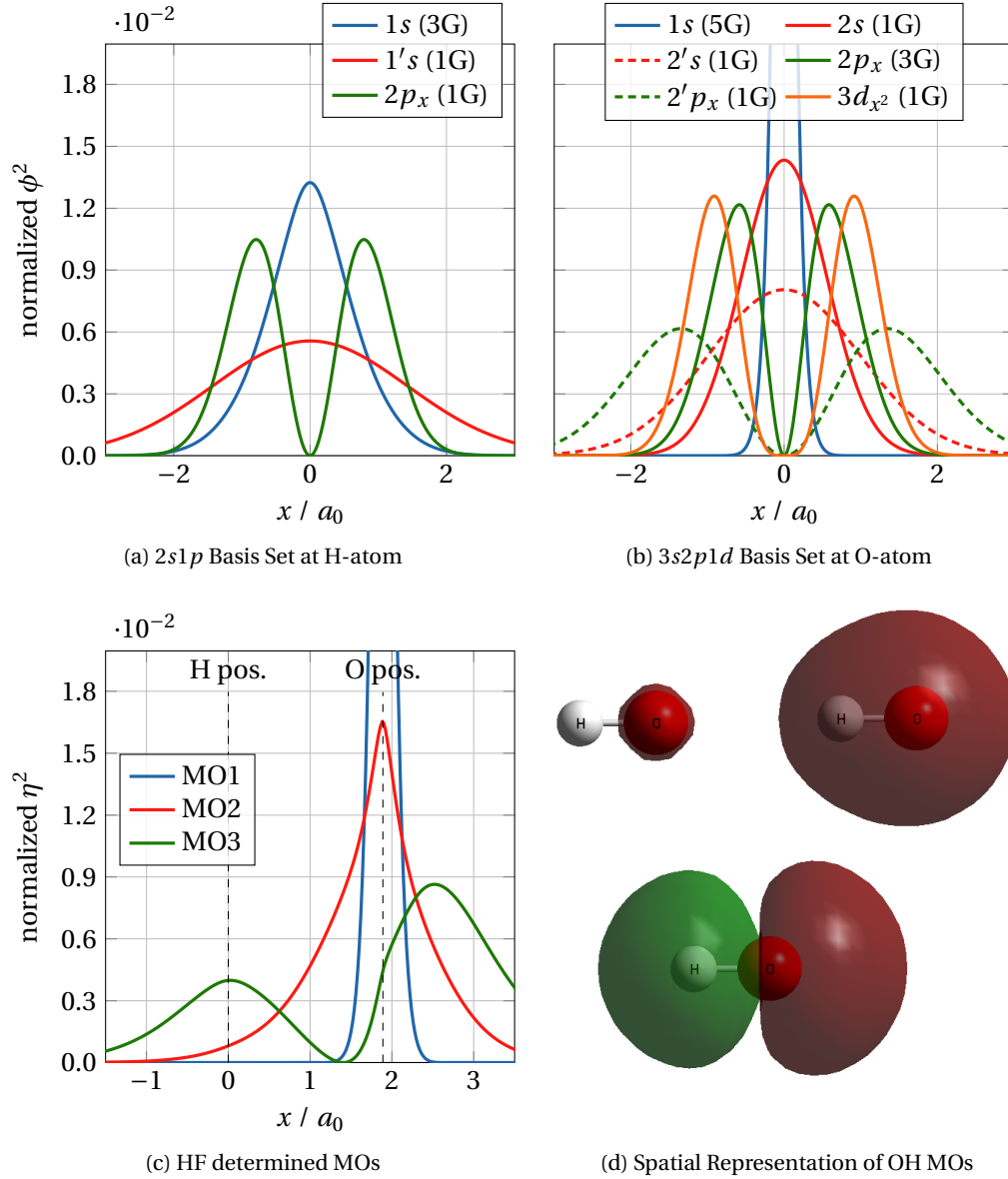


Figure 2.1: Basis set example for the OH radical in HF calculations at 1 \AA nuclear separation. For p - and d -orbitals, only the function in x -direction is shown.

The three lowest MOs are given by (only major contributions shown):

$$\begin{aligned} \eta_1 &= 0.99\phi_{1s}^{\text{O}} - 0.03\phi_{2s}^{\text{O}} \\ \eta_2 &= 0.30\phi_{1s}^{\text{O}} + 0.61\phi_{2s}^{\text{O}} + 0.37\phi_{3s}^{\text{O}} + 0.17\phi_{1s}^{\text{H}} \\ \eta_3 &= 0.18\phi_{2s}^{\text{O}} + 0.27\phi_{2's}^{\text{O}} + 0.54\phi_{2p_x}^{\text{O}} + 0.30\phi_{2'p_x}^{\text{O}} - 0.36\phi_{1s}^{\text{H}}. \end{aligned}$$

2 Introduction to Computational Chemistry

anions or delocalized systems. Popular basis sets are the ones of Dunning [76–79], Ahlrichs [75, 80, 81], and Pople [82, 83]. While a larger basis set will increase accuracy, it also increases computational effort significantly. Thus, a balance between accuracy and cost has to be found - as always in computational approaches.

Post-HF Methods The HF formalism suffers from some significant approximations. Firstly, the electron correlation is not treated directly but through each electron interacting with the total electron density. Secondly, the electronic wavefunction is approximated by one Slater determinant, however, especially for higher spin states, it can be shown that multiple determinants are required to accurately describe the correct wavefunction. *Configuration Interaction* (CI) calculations approximate the ground state wavefunction as linear combination of Slater determinants. Besides the HF determinant, additional ones are introduced, where occupied orbitals are replaced by virtual ones, representing electronic excitation. This is often restricted to singly, doubly or triply excited determinants and thus, methods termed CID, CISD or CISD(T) are used. The same approach but with an exponential expansion of the excited determinants, the *Coupled Cluster* (CC) calculations are commonly used, again limiting the number of excitations (CCSD, CCSD(T)). While in these cases the form of the orbitals η_j are only calculated once through the HF formalism and then the coefficients for the determinant expansion are determined in a decoupled step, *Complete Active Space Self Consistent Field* (CASSCF) calculations vary all coefficients (for MOs and determinants) simultaneously. A different, perturbative approach, termed *Møller-Plesset Perturbation Theory* tries to estimate the electron correlation directly through an expansion of the energy up to a certain term (MP2, MP3, MP4).

These methods always add considerable computational effort to the HF formalism but also increase the accuracy greatly. Since they are not used in this work, the reader is referred to the literature for a more detailed description.

2.1.2 Density Functional Theory

The Kohn-Sham Approach A different formalism, developed parallel to the HF method, is called *Density Functional Theory* (DFT) and relies on the concept that the ground state energy $\mathcal{E}^{(0)}$ is computable from the ground state electron density $\rho_e^{(0)}(\mathbf{r})$ directly. Thus, there exists a functional \mathcal{F} such that

$$\mathcal{E}^{(0)} = \mathcal{F}[\rho_e^{(0)}(\mathbf{r})] + \int \rho_e^{(0)}(\mathbf{r}) V_{eN}(\mathbf{r}) d\mathbf{r} \quad (2.8)$$

holds, where the integral describes the energy resulting from the Coulomb potential between the electron density and the nuclei, V_{eN} . The functional \mathcal{F} has to account for the kinetic energy of the electrons T and their pairwise interaction energy \mathcal{E}_{ee} :

$$\mathcal{F}[\rho_e(\mathbf{r})] = T[\rho_e(\mathbf{r})] + \mathcal{E}_{ee}[\rho_e(\mathbf{r})] \quad (2.9)$$

Assuming all functionals in Eq. (2.9) are known and there is a way of describing the electron density, the only task is to find the particular density that minimizes the energy in Eq. (2.8) by means of the variational principle [84]. In this ideal situation, DFT would actually be exact and outperform all HF approaches [73]. However, neither is the case: The exact form of the functional \mathcal{F} is not known and there is no trivial way of expressing the electron density.

To actually apply DFT, two main concepts are introduced [85]. Addressing the description of the electron density, wavefunction theory is used similar to the HF approach: A Slater determinant similar to the one in Eq. (2.3) is built, having spin orbitals $\varphi_j\alpha$ and $\varphi_j\beta$. The electron density can then be expressed as:

$$\rho_e(\mathbf{r}) = \sum_{j=1}^{M_e} |\varphi_j(\mathbf{r})\sigma_j|^2 \quad (2.10)$$

where σ_j is the appropriate spin function, α or β , of electron j .

Addressing the form of the functional \mathcal{F} , Kohn and Sham [85] separated the functional into parts, which can be calculated exactly and "the rest". The known parts are the kinetic energy of the electrons neglecting their interactions, T_S , and the energy through the Coulomb potential, J . The residual part of the *kinetic* energy, T_C , correcting for the electron correlation and being rather small compared to T_S , as well as the non-classical exchange and correlation *potential* energies, \mathcal{E}_{ncl} , are combined into an unknown term, the so-called exchange-correlation functional \mathcal{E}_{XC} :

$$\mathcal{F}[\rho_e(\mathbf{r})] = T_S[\rho_e(\mathbf{r})] + J[\rho_e(\mathbf{r})] + \underbrace{T_C[\rho_e(\mathbf{r})] + \mathcal{E}_{\text{ncl}}[\rho_e(\mathbf{r})]}_{\mathcal{E}_{\text{XC}}[\rho_e(\mathbf{r})]} \quad (2.11)$$

Combining Eqs. (2.8), (2.11) and (2.10) leads to an expression of the energy in terms of the orbitals φ_j . Applying the variational principle, i.e., which orbitals have to be used to minimize the energy, it can be shown that the resulting density has to form an effective potential

$$V_{\text{eff}}(\mathbf{r}) = \int \frac{\rho_e(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta \mathcal{E}_{\text{XC}}}{\delta \rho_e(\mathbf{r})} - \sum_{k=1}^{M_N} z_k e \int \frac{\rho_e(\mathbf{r})}{|\mathbf{R}_k - \mathbf{r}|} d\mathbf{r} \quad (2.12)$$

such that the orbitals are eigenfunctions to a Hamiltonian applying this particular potential:

$$[\hat{T}_e + V_{\text{eff}}] \varphi_j = \epsilon_j \varphi_j \quad (2.13)$$

In other words, the MOs are determined such that the electron density in Eq. (2.10) approximates the true ground state density $\rho_e^{(0)}(\mathbf{r})$ best. Thus, the main difference between HF and DFT is that in the former case the orbitals are optimized to approximate the exact wavefunction while in the latter case they are optimized to approximate the exact electron density.

As in the HF formalism, these equations have to be solved iteratively since the electron density is determined by the orbitals through Eq. (2.10) but the orbitals rely on

2 Introduction to Computational Chemistry

the electron density through Eq. (2.13) and (2.12). Again, a basis set description of the orbitals is used (cf. Eq. (2.4)), and thus, the coefficients c_{kj} have to be determined. Having an infinite basis set and an exact expression for \mathcal{E}_{XC} this would lead to the exact ground state energy. However, the form of the exchange-correlation functional is unknown and thus, approximations have to be used.

DFT in Practice Over the years, many different functionals, \mathcal{E}_{XC} , have been developed. Some rely on drastic physical simplifications to arrive with an expression, some fit parameters to empirical data. Most functionals are designed as *either* exchange *or* the correlation part, only some are designed for both. In principle, any exchange can be combined with any correlation functional, although the accuracy varies. Thus, some combinations have been found to perform well. For example Becke's three parameter exchange functional [86] combined with the correlation functional developed by Lee, Yang and Parr [87], termed B3-LYP (or B3LYP), is probably the most frequently used functional for organic molecules. A variation of this functional, including a perturbative approach to 2nd-order for electron correlation is termed B2PLYP and considerably increases accuracy at the cost of computational effort though [88].

Dispersion effects, also called van der Waals or London forces, are weak forces acting over large distances, resulting from fluctuations in the electron density. In wavefunction theory these fluctuations are treated in multi-determinant calculations through singly excited determinants. If a rather local orbital is excited, the change in electron density will influence nearby electrons, resulting in weak (attractive) correlations [89, 90]. In DFT, these interactions are often described by parametrized pairwise energy terms of the form $-C_6/R_{ij}^6$ (or higher orders), added to the DFT-energy. The dispersion correction is parametrized for many different functionals and adds no significant computational effort to the calculations, though being a somewhat coarse approximation. Popular dispersion corrections were developed by Grimme, e.g., the GD3 [91] or GD3(BJ) [92] dispersion corrections. These types of interactions are especially important for weakly bound systems (molecular clusters, hydrogen bonds, van der Waals complexes, etc.) because in these cases the long-range forces dominate the total interaction energy.

When choosing a particular method to determine the electronic energy, wavefunction or density functional theory, only the electronic problem of a molecule is "solved". In the Born-Oppenheimer approximation, the position of the nuclei, \mathbf{R} , is a pure matter of choice. The next section will discuss methods to find particularly interesting molecular geometries, such as minima or transition states, pathways connecting these, as well as different methods to describe the movement of the nuclei within the PES.

2.2 Nuclear Position and Movement

Following the Born-Oppenheimer approximation, the potential energy surface (PES), termed $\mathcal{E}_{\text{pot}}(\mathbf{R})$, is defined as the sum of the electronic energy $\mathcal{E}_e(\mathbf{R})$, determined with any method described in the previous section, and the energy resulting from the nuclear Coulomb repulsion potential $V_{\text{NN}}(\mathbf{R})$ at a particular nuclear geometry \mathbf{R} , i.e.,

$$\mathcal{E}_{\text{pot}}(\mathbf{R}) = \mathcal{E}_e(\mathbf{R}) + \frac{1}{4\pi\epsilon_0} \sum_{k=1}^{M_{\text{N}}-1} \sum_{l=k+1}^{M_{\text{N}}} \frac{z_k z_l e^2}{|\mathbf{R}_k - \mathbf{R}_l|} \quad (2.14)$$

This is not the total energy of the molecule, \mathcal{E}_{tot} , since the kinetic energy of the nuclei is not accounted for in Eq. (2.14). Within the Born-Oppenheimer approximation, the PES is the potential, the nuclei move in. Since \mathbf{R} is a $3M_{\text{N}}$ dimensional vector (although only $3M_{\text{N}} - 6$ dimensions are of interest, see below), the PES is a hypersurface and multidimensional mathematical analysis is applied for its description. The first and second derivatives of the potential energy with respect to the nuclear coordinates are thus a vector, the *gradient* \mathbf{g} , and a matrix, the *Hessian* \mathcal{H} , respectively, with their entries

$$g_i = \frac{\partial \mathcal{E}_{\text{pot}}}{\partial R_i} \quad (2.15a)$$

$$\mathcal{H}_{ij} = \frac{\partial^2 \mathcal{E}_{\text{pot}}}{\partial R_i \partial R_j} \quad (2.15b)$$

These two quantities are very helpful for the analysis of the PES and can be computed analytically for most electronic structure methods, in particular, HF and DFT. For some advanced wavefunction theory methods, only numerical approaches exist. Although they are computed with considerable computational cost (especially the Hessian is cumbersome) they are crucial for all questions about the nuclear position and movement as is discussed in the next three sections.

2.2.1 Geometry Optimization

Equilibrium Geometry A common task in Computational Chemistry is the determination of the equilibrium geometry of a molecule. This is represented by a (local or global) minimum of the PES and the concept is that this is how a molecule would look like in the real world. As known from standard mathematical analysis, a minimum is a configuration \mathbf{R}_{min} such that the gradient is zero and the Hessian is positive definite, i.e., has only positive eigenvalues. Since it is impossible to invert this problem, i.e., find the position \mathbf{R}_{min} from only the condition that $\mathbf{g}(\mathbf{R}_{\text{min}}) = 0$, a numerical approach is required. Guessing a reasonable approximation of the equilibrium structure with the users chemical intuition, the actual minimum is found by a gradient descent approach: At the guess geometry $\mathbf{R}_{\text{guess}}$, the energy and its gradient are calculated and then a new geometry is defined by following the gradient downhill from

2 Introduction to Computational Chemistry

the old one:

$$\mathbf{R}_{\text{guess}}^{(k+1)} = \mathbf{R}_{\text{guess}}^{(k)} - \lambda^{(k)} \mathbf{g}(\mathbf{R}_{\text{guess}}^{(k)}) \quad (2.16)$$

where λ is a scaling factor. The energy and its gradient at the new geometry are calculated again and the procedure is repeated until a certain threshold regarding the change in either energy or gradient is met, i.e., until convergence is reached. The equations shown are simple representations. Modern algorithms often take more sophisticated approaches. Strictly speaking, this only yields the next local minimum and there is no guarantee that the global minimum is obtained. This is why chemical intuition is so important and why Computational Chemistry is still chemistry and not informatics.

Additionally, it is even not guaranteed that a minimum is reached at all, since the gradient is zero at many points, i.e., maxima or saddle points. The latter are a minimum in some but a maximum in the other coordinates. Thus, to ensure that a minimum is found, the Hessian is computed and diagonalized. If \mathbf{R} is given in $3M_N$ Cartesian coordinates, six of the eigenvalues will be zero. They correspond to the three overall translations and three rotations of the molecule, which do not change the potential energy in Eq. (2.14). Thus, to decide whether the found geometry is a minimum, only the $3M_N - 6$ (or $3M_N - 5$ for linear molecules) non-zero eigenvalues are studied. In practice, this is more complicated and will be covered in the next section.

PES Scans While the equilibrium geometry of a molecule is of great importance, chemistry deals with molecular *change* and thus, reaction pathways and transition states are also of interest. It is straightforward to perform the gradient descent method with one coordinate fixed: The respective entry in the gradient is set to zero and the coordinate will not change. Since the molecular geometry \mathbf{R} can also be represented by internal coordinates such as bond distances, angles and torsion angles, an optimized geometry under the constrain that, e.g., a certain bond is held fixed at 1.5 Å, may be obtained. When the fixed parameter is then incrementally changed, the minimum energy path along this coordinate is scanned. This can be pictured as changing one parameter, while the remaining geometry is allowed to relax. These PES scans are useful for transition state searches, analysis of binding situations and general analysis of the PES in areas of interest.

When a maximum along a PES scan is found, this geometry can be used as a first guess structure for an actual transition state (TS) search. In principle, this is similar to the gradient descent approach, however, the gradient is modified such that it follows the negative eigenvectors uphill. Compared to a minimum search this is much more difficult and transition state searches, although crucial for chemistry, are still an elaborate task. Again, at the obtained geometry, the Hessian should be calculated and in addition to the six zero and $3M_N - 7$ positive eigenvalues, one should be negative. The corresponding eigenvector, i.e., the change in geometry along which the PES is a maximum, should correspond to the reaction path.

When a TS is obtained, the reaction pathway can be calculated as well. While a PES

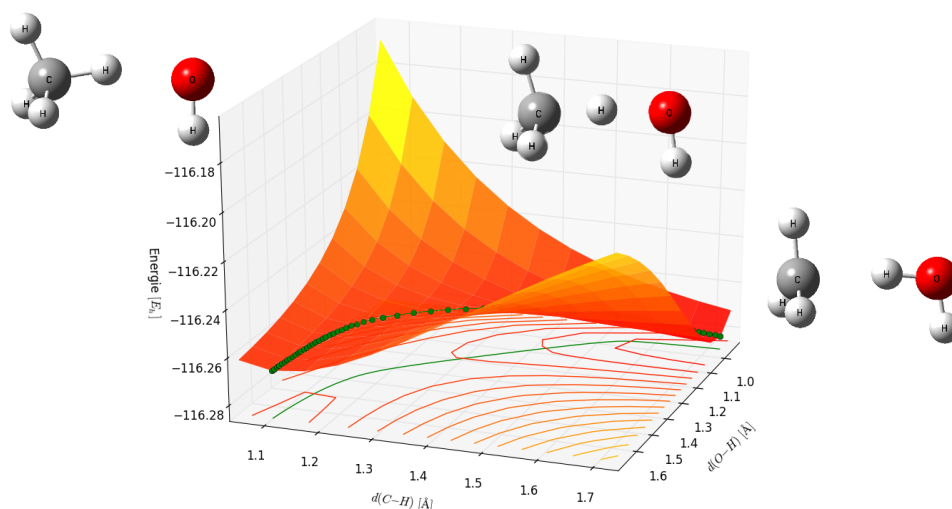


Figure 2.2: H-atom abstraction from methane by an OH radical. The surface is obtained by varying the C–H and C–O distances, while all other parameters are allowed to relax. The green line corresponds to the minimum energy path through this surface, i.e., the reaction coordinate.

scan is already a good approximation, the actual reaction path is generally lower in energy and should be computed whenever possible. In these algorithms, a minimum on a hypersphere around the TS is searched. This corresponds to the geometry with the lowest energy within close vicinity to the TS. The obtained geometry is then again used as center of a hypersphere and the iterative procedure is continued until the center of the sphere is lower in energy than every other point considered, i.e., it is a (local) minimum. This algorithm can be pictured as moving downhill from the TS as steeply as possible, finding the next (local) minimum on the PES and thus the lowest energy path from this minimum to the TS. Calculating this path for both sides of the TS, the lowest reaction path between two minima, connected by this TS is obtained. The path obtained is the *reaction coordinate*, or intrinsic reaction coordinate, IRC [93]. Figure 2.2 shows the reaction path for the H-atom abstraction from methane by an OH radical, superimposed on a 2D PES scan, upon varying the C–H and H–O distances. The contour lines below the surface show that the reaction indeed progresses through a TS.

Today there are fully-automated algorithms that try to find the lowest energy path from a minimum structure, which prevents the elaborate TS search [94]. However, the computational demand of such methods is rather high and chemical intuition and preliminary information about the progression of a reaction can still compete with those approaches. With the ever growing computational power though, automated methods will certainly become more popular in the future.

2.2.2 Vibrations and Rotations

Up to now, the kinetic energy of the nuclei was not considered. Scans and individual points on the PES still represent resting nuclei and the movement of nuclei along a reaction path is an imagination rather than actually dealing with kinetic energy, momentum, and velocities. While a direct treatment of these quantities are covered in the next section, this part deals with solving the nuclear Schrödinger Eq. (2.2b) and thus taking the kinetic energy of the nuclei into account in a time-independent, quantum state approach.

Rigid-Rotor / Harmonic-Oscillator The $3M_N$ degrees of freedom can be separated into three translational, three rotational (if the molecule is not linear) and $3M_N - 6$ internal, or vibrational degrees. While the translational motion is completely independent of the PES, the rotations depend on the *position* of the minimum and the vibrations on the *curvature* of the surface. The moment of inertia tensor \mathbf{I} can be computed and diagonalized using the calculated equilibrium geometry. Its eigenvalues are the principle moments of inertia I_{aa} , I_{bb} and I_{cc} , which correspond to the independent rotation around the principle axes a , b and c . The rotational constants at the equilibrium position A_e , B_e and C_e , appearing in the corresponding Hamiltonian, can then be calculated according to

$$\Lambda = \frac{h}{8\pi^2 c I_{\lambda\lambda}} \quad (\Lambda, \lambda) \in \{(A_e, a), (B_e, b), (C_e, c)\} \quad (2.17)$$

It is assumed that the molecule rotates rigidly (no centrifugal distortion and without any coupling from the vibrational motion, i.e., change of geometry through vibrational broadening). This may not hold true for very loosely bound molecules; a better description is discussed below.

The vibrational motion is more complicated than the rotation. Without going into detail (cf. [95, 96] for a more elaborate discussion), the (mass-weighted) Hessian computed in Cartesian coordinates at the equilibrium geometry is transformed into a Hessian in internal coordinates \mathbf{q} , where the overall translation and rotation have been separated out, i.e., these coordinates are orthogonal to the internal ones. The theoretical foundation of this separation was developed by Eckart [97] and later by Sayvetz [98] and Wilson [99, 100]. This new Hessian is then diagonalized yielding $3M_N - 6$ new coordinates \mathbf{Q} , called normal coordinates. In the molecule, they correspond to the independent normal vibrations. Along these coordinates, the PES is approximated through a Taylor series up to second order, i.e., as parabola. Then, all vibrations are harmonic and independent from each other and the Schrödinger equation can be solved analytically for each vibration. The eigenvalues in the new Hessian correspond to the force constants F_l of the harmonic oscillation and, taking the reduced mass of each vibration μ_l into account, yield the normal frequencies ω_l , observable in infra-red or Raman spectroscopy [95, 96]. The vibrational energy levels

and the zero-point energy (ZPE) are then

$$\mathcal{E}_{v_1, v_2, \dots, v_{3M_N-6}} = \sum_{l=1}^{3M_N-6} \hbar\omega_l \left(v_l + \frac{1}{2} \right) \quad (2.18a)$$

$$\mathcal{E}_{\text{ZPE}} = \mathcal{E}_{0,0,\dots,0} = \sum_{l=1}^{3M_N-6} \frac{1}{2} \hbar\omega_l \quad (2.18b)$$

Reducing the vibrational problem to the harmonic oscillator approximation simplifies the treatment significantly. Basically, only the Hessian at the equilibrium geometry has to be computed and everything else is obtained from the eigenvectors and eigenvalues without any significant additional computational effort. The vibrational Schrödinger equation is not explicitly solved during this procedure since the energy levels are analytically known. The force constants and reduced masses of the vibrations, the only input parameters needed, directly follow from the diagonalization process. However, this approximation can introduce significant errors. For one, anharmonicity is not considered, which can be especially important for flat potentials along certain normal coordinates (e.g., weakly bound systems). Additionally, the normal mode analysis (NMA) yields coordinates, which are only locally independent from each other, i.e., the off-diagonal elements in the (diagonalized) Hessian are zero only at the equilibrium geometry or in close proximity. If a molecule vibrates significantly, either through high temperature, vibrational excitation or if, again, the potential is very flat, the normal coordinates are not independent from each other anymore and couple. More elaborate methods, trying to incorporate these effects, are discussed below.

Another problem in the harmonic oscillator approximation are internal rotations (IRs). While the described geometry optimization yields a minimum, maybe even the global one, there are several other minima easily accessible, if the molecule has internally rotating groups with small barriers (see Fig. 2.3b). Neglecting these additional minima and treating the global one as harmonic potential is only valid when the barriers are high and the molecule is trapped inside the corresponding minimum. Even when the rotating group is completely symmetric and the potential has only one minimum (Fig. 2.3a), the periodicity of the potential changes the energy levels compared to the harmonic oscillator. In the extreme case, where the barrier height is negligible, i.e., a free rotation, the energy levels are given by [101, p. 438]:

$$\mathcal{E}_K = \frac{\hbar^2 K^2}{2I_{\text{IR}}} \quad K = \dots, -2, -1, 0, 1, 2, \dots \quad (2.19)$$

where I_{IR} is the reduced moment of inertia of the rotation and K is a quantum number. In particular, the lowest energy level ($K = 0$) is zero. Thus, internal rotations can introduce significant errors to the harmonic approximation if not treated separately. Automated treatments within the NMA framework have been published [102] but close inspection of the results is still necessary.

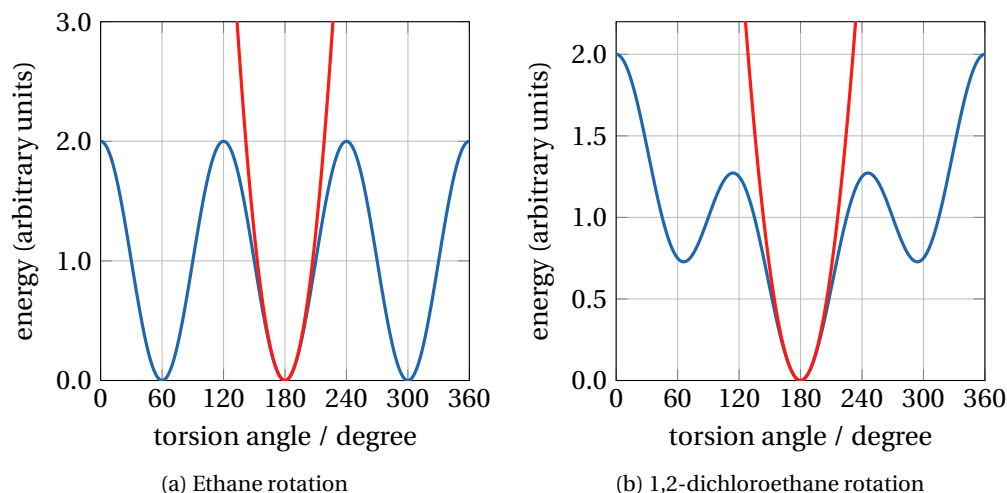


Figure 2.3: Harmonic approximation (red line) for periodic potentials (blue line) occurring in e.g. internal rotations. If the rotating group is completely symmetric (e.g. a methyl group), all periodic minima are the same. However, if the rotating group is asymmetric as a $-\text{CH}_2\text{Cl}$ group, several minima exist along rotation, given that the framework rotated against is also asymmetric.

Treating Anharmonicity A popular extension to the harmonic oscillator / rigid rotor approximation uses a perturbative treatment and is called Vibrational Perturbation Theory to 2nd Order (VPT2). In contrast to the harmonic approximation, the Taylor expansion of the PES is not truncated after the quadratic but after the quartic term [103]:

$$\mathcal{E}_{\text{pot}} = \frac{1}{2} \sum_l F_{ll} Q_l^2 + \frac{1}{6} \sum_{lst} F_{lst} Q_l Q_s Q_t + \frac{1}{24} \sum_{lstu} F_{lstu} Q_l Q_s Q_t Q_u \quad (2.20)$$

where F_{ll} are the quadratic, F_{lst} the cubic and F_{lstu} the quartic force constants. Since the potential is expanded with respect to the normal coordinates Q_l and their off-diagonal force constants F_{ls} are zero by construction of the normal coordinates, only the diagonal quadratic force constants F_{ll} appear in the equation. The energy levels of this potential are then also a Taylor series but for VPT2 truncated after the second term:

$$\mathcal{E}_{\text{vib}} = h\chi_0 + \sum_{l=1} \hbar\omega_l \left(v_l + \frac{1}{2} \right) + \sum_{l=1} \sum_{s \leq l} h\chi_{ls} \left(v_l + \frac{1}{2} \right) \left(v_s + \frac{1}{2} \right) \quad (2.21)$$

The vibrational anharmonic constants χ are applied as frequency to be consistent with the vibrational frequencies, ω_l . However, most reports in the literature and computational chemistry programs report them as wavenumbers. Analytical equations for the determination of the χ_{ls} are available using the force constants appearing

in Eq. (2.20). They include also effects such as rotational-vibrational interactions through Coriolis coupling and centrifugal distortion. Consequently, the rotational constants A_e , B_e and C_e are also changing using the vibrationally averaged geometry including centrifugal distortion rather than the equilibrium one. This resembles a far more realistic description of the PES since the vibrational levels are not equidistant but converge towards higher level numbers, rotational-vibrational coupling is incorporated to some degree, and also the rotational levels are adjusted to more realistic values. The energy difference between the lowest and first excited vibrational levels are thus called *fundamental frequencies* and can be calculated as

$$\nu_l = \omega_l + 2\chi_{ll} + \frac{1}{2} \sum_{l \neq s} \chi_{ls} \quad (2.22)$$

Central to this treatment are the force constants F_{ll} , F_{lst} and F_{lstu} of Eq. (2.20). Modern implementations of the VPT2 method calculate them by numerical differentiation along the normal coordinates [104, 105]: Displacing the geometry by a step size δ along both sides of a normal coordinate l and calculating the second derivative matrix at these geometries leads to three Hessians, $\mathcal{H}(-\delta_l)$, $\mathcal{H}(0)$, and $\mathcal{H}(\delta_l)$, which are used to estimate the third and fourth derivative by numerical differences. Recent implementations of VPT2 also include correction necessary through degeneracy, internal rotations, and resonances [106]. The computational cost is significantly higher as compared to the harmonic approximations since $2(3M_N - 6) + 1$ Hessian calculations have to be performed in contrast to only one. Additionally, the accurate determination of the higher force constants through the numerical approach is rather difficult and close inspection of results from the automatized procedures implemented is still necessary. As in all perturbation approaches, the formalism only works well when the perturbative corrections are small compared to the unperturbed (here the harmonic) part. This does not hold true for all systems of interest.

While the above approaches rely on an analytical approximation to the PES, the vibrational Schrödinger equation can also be solved for a potential given by grid points along a coordinate. This is performed with the Fourier-Grid-Hamiltonian Method [107]. From PES scans, the potential is obtained at defined grid points and the algorithm directly yields the amplitudes of the vibrational wavefunctions at these grid points, additionally to the energy levels. While this is an accurate method since it does not rely on any approximation of the PES (and increasing the number of grid points through interpolation will increase accuracy), it is only extended up to three dimensions [108] and thus the coordinates used are intrinsically assumed to be independent from each other. Thus, considering the coupling between all modes as through the anharmonic vibrational constants χ_{ls} is not possible. The best approximation would be scans along the normal coordinates, since they fulfill the requirement of decoupling best and their reduced masses are easily calculated from the NMA. Nevertheless, methods have been developed, which use internal [109, 110] or Cartesian

[111] coordinates as well. Solving the vibrational Schrödinger equation for a given potential is rather fast; the calculation of the PES at the grid points is, however, the most time consuming part of these methods. A sufficient density of grid points is required, even when additional points are created by interpolation, and thus, about 10-100 energy calculations have to be performed for a single coordinate.

2.2.3 Molecular Dynamics Simulations

Solving the rotational-vibrational Schrödinger equation (even with all the mentioned approximations) treats the movement of the nuclei in a quantum fashion. In Molecular Dynamics (MD) simulations, however, the nuclei are treated as classical particles. This is reasoned by their much larger mass, as compared to electrons, and thus slower movement. For classical non-relativistic movement the Newton equations of motion are valid and hence, the movement of the nuclei on the PES can be studied by numerical integration over time [112]. Given the nuclei positions in Cartesian coordinates, \mathbf{R} , and their masses as a $(3M_N \times 3M_N)$ matrix (where each mass appears three times), \mathbf{M} , the equations of motion read

$$\mathbf{M} \cdot \mathbf{a} = \mathbf{M} \cdot \frac{d^2 \mathbf{R}}{dt^2} = -\mathbf{g}(\mathbf{R}) = -\frac{d\mathcal{E}_{\text{pot}}(\mathbf{R})}{d\mathbf{R}} \quad (2.23a)$$

$$d\mathbf{R} = \mathbf{v} dt \quad (2.23b)$$

$$d\mathbf{v} = \mathbf{a} dt \quad (2.23c)$$

where \mathbf{v} and \mathbf{a} are the velocity and acceleration vectors in Cartesian coordinates, respectively. Given a nuclear geometry $\mathbf{R}(t)$ and a velocity vector $\mathbf{v}(t)$ at a time t , the potential energy and its derivative are evaluated through Eq. (2.23a). The former velocity value is used to update the geometry to $\mathbf{R}(t + dt)$ through Eq. (2.23b) and then the velocity is updated to $\mathbf{v}(t + dt)$ through Eq. (2.23c) with the acceleration obtained from the potential gradient (Eq. (2.23a)). In both cases, a finite time difference has to be used, thus $dt \rightarrow \Delta t$. Today, more elaborate numerical integration schemes are applied but they rely on the same fundamentals.

Since the integration scheme is fully determined, the only question is, how the potential and its gradient are evaluated at each time step. Early approaches, termed Born-Oppenheimer-Molecular-Dynamics, just applied the full SCF calculations. This is very time consuming and long molecular trajectories are not feasible in this manner. Thus, most MD methods step away from quantum mechanics completely and use parametrized force fields to calculate the potential. A lot of effort is necessary to develop force fields and their parameter sets, and many different types are available: sometimes specific for defined molecular classes, sometimes rather universal. While they are computationally fast, their accuracy strongly depends on (1) the functional form of the force field, i.e., what type of interactions are covered and how, and (2) the parameter set, i.e., are the force constants, equilibrium bond lengths, and other parameters applicable in the system at hand. For example, the AMBER force field [113]

describes bond stretches and valence angles with a harmonic potential around the equilibrium value, torsion angles as simple periodic functions with equal minima and non-bonded interactions through Lennard-Jones interactions accompanied by Coulomb interactions of partial atomic charges. Due to the harmonic bond stretches, this force field is unable to form and break bonds, and the description of the PES will only be accurate close to the equilibrium geometry. Even in close vicinity to the equilibrium geometry, the description will only be accurate when the parameter set applied to the force field is accurate and applies to the system studied. Inaccurately determined partial charges, for example, will lead to a poor description of the non-bonded interactions. Thus, the major drawback of force field based computational chemistry is the lack of *ab initio* generality.

Already in 1985, an *ab initio* MD (AIMD) code was reported [114], which circumvents the full calculation of the SCF density at each time step by propagating the molecular orbitals within the classical equations. A later implementation, named Atom-Centered-Density-Matrix-Propagation (ADMP) [115–117], propagates the full density matrix rather than the individual orbitals by assigning a fictitious mass to it and thus also a velocity and kinetic energy. This kinetic energy is completely independent from the one appearing in the DFT functional (Eq. (2.11)) and has actually no physical meaning. Relying on the idea that the electron density will vary slowly as the nuclei move, it is adjusted under certain boundary conditions within only one SCF step per time step. This makes this approach feasible for small to medium sized molecules, offering a route to study the movement of the nuclei on the completely anharmonic surface.

2.3 Thermochemistry

Many applications in chemistry are dealing with a thermalized ensemble of molecules, i.e., a number of particles N having a temperature T . If the volume is fixed, such an ensemble is called the canonical ensemble, or NVT ensemble. The internal energy, corresponding to the temperature, is distributed among all degrees of freedom. Thus, zero-point corrected electronic energies are only part of the information regarding the energetics, since the mean internal energy in one molecule has to be taken into account as well. Therefore, statistical thermodynamics is used to predict thermodynamic properties such as internal energies, enthalpies, entropies and free enthalpies. The central quantity in this formalism is the partition function Z . In a general case of N particles, being able to occupy some energy levels with energy \mathcal{E}_i and degeneracy ρ_S , Z is defined as

$$Z = \sum_i \rho_S^{(i)} \exp\left(-\frac{\mathcal{E}_i}{k_B T}\right) \quad (2.24)$$

where the sum runs over all levels and k_B is the Boltzmann constant. The partition function defines the ensemble average of molecules having energy \mathcal{E}^* , $\langle N(\mathcal{E}^*) \rangle$,

2 Introduction to Computational Chemistry

Table 2.1: Translational, rotational, vibrational and electronic partition functions for molecules, cf. [118, p. 104].

Contribution	Formula	Magnitude	Approximation
Translation	$Z_{\text{trans}} = \left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} V$	10^{24}	particle in 3D box
Rotation	$Z_{\text{rot}} = \frac{1}{\sigma_{\text{rot}}} \left(\frac{k_B T}{hc}\right)^{3/2} \left(\frac{\pi}{A_e B_e C_e}\right)^{1/2}$	10^3	rigid rotor
Vibration	$Z_{\text{vib}} = \prod_l \left[1 - \exp\left(-\frac{\hbar\omega_l}{k_B T}\right)\right]^{-1}$	1-10	harm. oscillator
Int. Rot.	$Z_{\text{IR}} = \frac{1}{\sigma_{\text{IR}}} \left(\frac{8\pi^3 I_{\text{IR}} k_B T}{h^2}\right)^{1/2}$	1-10	free rotor
Electronic	$Z_{\text{elec}} = \sum_i \rho_S^{(i)} \exp\left(-\frac{\mathcal{E}_i}{k_B T}\right)$	1	HF/DFT calc.

m : mass of molecule, k_B : Boltzmann constant, T : absolute temperature, h : Planck constant, V : volume, σ : symmetry number, c : speed of light, A_e, B_e, C_e : rotational constants (in cm^{-1}), ω_l : normal frequencies, I_{IR} : reduced moment of inertia, \mathcal{E}_i : electronic energy levels (incl. ZPE) with degeneracy $\rho_S^{(i)}$

through the Boltzmann distribution:

$$\frac{\langle N(\mathcal{E}^*) \rangle}{N} = \frac{1}{Z(T)} \rho_S^* \exp\left(-\frac{\mathcal{E}^*}{k_B T}\right) \quad (2.25)$$

In case of molecules, the internal energy can be distributed among translational, rotational, vibrational and electronic degrees. Assuming these contributions to be independent from each other (which already neglects rotational-vibrational coupling), the total partition function becomes a product of the individual partition functions:

$$Z = Z_{\text{trans}} Z_{\text{rot}} Z_{\text{vib}} Z_{\text{elec}} \quad (2.26)$$

For each of these contributions, closed equations can be found when appropriate approximations are made, as summarized in Table 2.1. For example, if a harmonic oscillator is assumed, each vibration is independent from the others and the vibrational partition function becomes a product of the contribution of each vibration. The partitioning of the energy over the equidistant energy levels within each oscillator can be calculated and yields the presented formula. This equation is used, even when fundamental frequencies ν_l are applied [104], since mostly low lying vibrational levels are populated at normal temperatures and the different spacing towards higher level numbers can be neglected. This approach is called simple perturbation theory [119]. Analytical equations exist for an anharmonic oscillator including the different spac-

ing [120], however, they are much more complicated and still rely on some degree of approximation.

Sometimes rather crude approximations are used to derive these formulas. Nevertheless, they provide a connection between molecular parameters derived from quantum mechanical calculations, such as rotational constants or normal frequencies, and thermodynamic quantities such as the Gibbs free enthalpy, G ,

$$G(T) - G(0) = -Nk_B T \ln \left(\frac{Z}{N} \right) \quad (2.27a)$$

$$G_{\text{corr}} = G(T) - G(0) + \mathcal{E}_{\text{ZPE}} \quad (2.27b)$$

The two quantities, $G(T) - G(0)$ and G_{corr} describe the *change* in Gibbs free enthalpy from 0 K to a temperature T since absolute enthalpies are not accessible in thermodynamics. Having these quantities defined, Gibbs free reaction enthalpies for chemical reactions can be computed, essential to the prediction and understanding of reaction pathways

$$\Delta_R G = \sum_{\text{products}} [\mathcal{E}_{\text{pot}}(\mathbf{R}_{\text{min}}) + G_{\text{corr}}] - \sum_{\text{reactants}} [\mathcal{E}_{\text{pot}}(\mathbf{R}_{\text{min}}) + G_{\text{corr}}] \quad (2.28)$$

In summary, while the potential and kinetic energies of the electrons as well as the Coulomb repulsion of the nuclei are summarized in the total potential energy \mathcal{E}_{pot} , the kinetic energy of the nuclei is subject to a thermal distribution among all degrees of freedom, given through the partition function Z , above the quantum mechanically introduced ZPE, incorporated into G_{corr} .

The partition function also aids to calculate the ensemble distribution among different conformers (e.g., protonation sites, rotamers, etc.). Using the superposition approximation [121–123], each of these local minima on the PES with energy \mathcal{E}_i , is assumed to be independent from the others and its partition function Z_i is calculated in the usual way. The population, P_i , of configuration i is then calculated as

$$P_i = \frac{Z_i \exp\left(-\frac{\mathcal{E}_i - \mathcal{E}_{\text{ref}}}{k_B T}\right)}{\sum_j Z_j \exp\left(-\frac{\mathcal{E}_j - \mathcal{E}_{\text{ref}}}{k_B T}\right)} = \frac{\mathcal{Z}_i}{\sum_j \mathcal{Z}_j} \quad (2.29)$$

where \mathcal{E}_{ref} is the energy of the lowest energy configuration and each \mathcal{E}_j includes the respective ZPE. Comparing this expression to the Boltzmann distribution, (2.25), it can be seen that the partition function, Z_j , of each minimum acts as degeneracy factor, ρ_S . Thus, the normal exponential decay of the population with respect to the energy is applied where each conformer has a state density equal to its individual partition function. The sum in the denominator is thus also a partition function.

2.4 Reaction Dynamics

Reaction dynamics deals with non-equilibrium conditions. Thus, time becomes an important factor. For an elaborate discussion, the reader is referred to textbooks such as [118]. In many reaction mechanisms, unimolecular decay steps play a key role. E.g., the dissociation of a cluster AB into fragments A and B, induced by a collision partner M



can be rewritten as a combination of activation, internal equilibrium, and dissociation [124]:

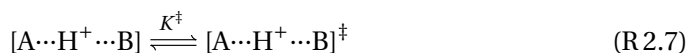


Here, AB^* describes an energized molecule and AB^\ddagger the molecule in its critical configuration, or transition state. K^\ddagger is thus the equilibrium constant of this internal energy redistribution and k^\ddagger is the rate constant from this configuration into the dissociation channel. This separation was introduced by Eyring and Polani [125, 126].

Gas-phase proton transfer reactions



can be expressed in a similar fashion:



The critical configuration AB^\ddagger is often depicted as *loose* TS [62], i.e., there is no actual barrier but the energy along the reaction coordinate rises and asymptotically reaches the product energy. In these cases, *phase-space theory* is used and the TS is located at the centrifugal barrier [127–129]. In contrast, the definition of a TS as discussed above (section 2.2.1) corresponds to a well defined, *tight* transition state. Figure 2.4 pictures both situations.

It is easier to work in a microcanonical framework, i.e., treating AB^* as a molecule with a particular energy \mathcal{E}^* , as defined in Fig. 2.4. The following treatment of the reaction rate of unimolecular dissociation, i.e., $k_{\text{uni}}(\mathcal{E}^*)$, is called statistical rate theory or RRKM theory, named after its authors Rice, Ramsperger, Kassel, and Marcus. The central quantity is the state density ρ_S . It already appeared in section 2.3 as degeneracy of a quantum level in the Boltzmann distribution. Although quantum levels are discrete, the number of states per unit energy is treated as continuous function,

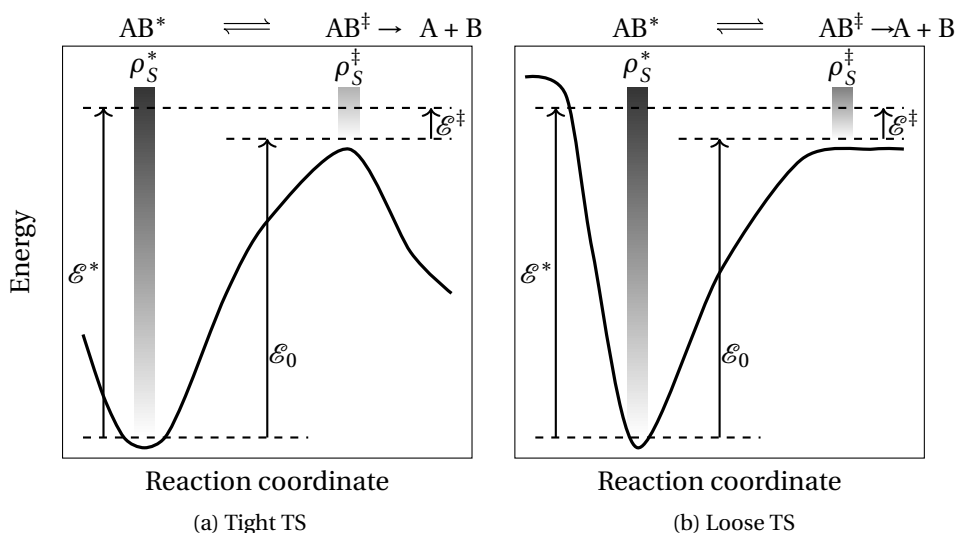


Figure 2.4: Schematics of PES for reaction dynamics [62, 118]. Certain energies (measured between ZPE levels) as well as state densities (ρ_S) are defined. Note, that the transition state both situations have the same height but not the same state density.

the state density. This approximation loses its validity for small molecules (less than four nuclei) or small energies \mathcal{E}^* . For a molecules with many vibrational degrees of freedom (even methane already has nine), the different combination of quantum numbers in Eq. (2.18a) results in numerous discrete energy levels per unit energy, validating the assumption of a continuous state density. For a multidimensional harmonic oscillator, a semi-classical equation for the state density can be given

$$\rho_S(\mathcal{E}) = \frac{1}{(s-1)!} \frac{\mathcal{E}^{s-1}}{\prod_{l=1}^s \hbar \omega_l} \quad (2.30)$$

although better approximations exist [130, 131]. Here, ω_l are the normal frequencies and s is the number of vibrations, which is $3M_N - 6$ for a stable structure but $3M_N - 7$ for a (tight) TS, since the reaction coordinate is not a stable vibration.

As Eq. (2.30) shows, the state density is proportional to \mathcal{E}^{s-1} , thus growing steeply with \mathcal{E} . If the rotational levels, which are even more closely spaced, are taken into account as well, it becomes apparent that a continuous approximation is valid. This also supports another important assumption in RRKM theory: The energy is distributed fast (compared to k_{uni}) over all vibrational degrees of freedom, a process called intra-molecular vibrational redistribution (IVR). Even when a specific vibrational mode is excited by, e.g., the collision in Reaction (R2.2), this energy is distributed over all modes through coupling of the densely spaced vibrational levels.

2 Introduction to Computational Chemistry

The integral of ρ_S from zero up to a certain energy \mathcal{E} is called the state or level number N_S and, following Eq. (2.30) is approximated by

$$N_S(\mathcal{E}) = \int_0^{\mathcal{E}} \rho_S(\tilde{\mathcal{E}}) d\tilde{\mathcal{E}} = \frac{1}{s!} \frac{\mathcal{E}^s}{\prod_{l=1}^s \hbar\omega_l} \quad (2.31)$$

This number grows even faster with energy than the state density.

The state density ρ_S determines the equilibrium between the excited molecule AB^* and the TS AB^\ddagger through

$$K^\ddagger = \frac{\rho_S^\ddagger(\mathcal{E}^\ddagger)}{\rho_S^*(\mathcal{E}^*)} \quad (2.32)$$

Note that the state densities are generally different functions for the two configurations AB^* and AB^\ddagger since the number and also the frequencies of the vibrations change. The energy at which they are determined is also different (cf. Fig. 2.4), in particular $\mathcal{E}^\ddagger < \mathcal{E}^*$. Thus, the state density at the TS is generally much smaller, i.e., $\rho_S^\ddagger(\mathcal{E}^\ddagger) \ll \rho_S^*(\mathcal{E}^*)$ and thus $K^\ddagger \ll 1$. This explains why a molecule is "stable", or better meta-stable, even when the internal energy \mathcal{E}^* is higher than the barrier height \mathcal{E}_0 : Assuming fast IVR, it is rather unlikely that all energy \mathcal{E}^* is concentrated in the reaction coordinate, especially, if $\rho_S^*(\mathcal{E}^*)$ becomes high and there is a high state density for energy storage.

Once the critical configuration is reached, it is assumed that half the molecules dissociate while the other half reacts back to the excited molecule AB^* . The movement along the reaction coordinate is treated as a translational rather than a vibrational movement. Thus, the excess energy \mathcal{E}^\ddagger can be distributed along vibrational, rotational and this particular translational degree of freedom. Hence, it is possible that a high vibrational level is populated in the TS but the kinetic energy along the reaction coordinate is only small. Or a low vibrational level is populated in the TS and most of the energy \mathcal{E}^\ddagger is concentrated in the translational movement. This is why the state number N_S^\ddagger and not the state density of the TS appears in equation (2.33), which represents the final result of RRKM theory:

$$k_{\text{uni}}(\mathcal{E}^*) = d_{\text{rxn}} \frac{N_S^\ddagger(\mathcal{E}^\ddagger)}{h\rho_S^*(\mathcal{E}^*)} \quad (2.33)$$

All levels up to \mathcal{E}^\ddagger can contribute to the reaction, leaving the left-over energy to the translation along the reaction coordinate. d_{rxn} is the degeneracy of the reaction channel.

Following RRKM theory, the microcanonical rate constant of the unimolecular dissociation solely depends on the ratio of the state number in the TS and the state density of the stable molecule. As stated before, the state density in the TS is much smaller than in the stable molecule since the energy and the number of vibrations

is lower (as pictured in Fig. 2.4 through the different shading of the state densities). However, since the state number increases faster with energy than the state density ($N_S \propto \mathcal{E}^s$ vs. $\rho_S \propto \mathcal{E}^{s-1}$), the reaction rate effectively *increases* with energy.

When comparing the state density of a loose and a tight transition state, it is seen that a loose TS has a higher state density than a tight TS [62]. The reason is that a tight TS has $3M_N - 7$ vibrational degrees of freedom, a loose TS only has $3M_N - 12$ where the remaining five are rotations of the fragments against each other in the phase-space limit [128]. Because rotational levels are much more closely spaced, the state density is higher in a loose TS, as pictured in Fig. 2.4 by the different shading of the state densities in the loose and tight TS. If a molecule can decompose over a loose or a tight TS, both having the same barrier height \mathcal{E}_0 , the loose TS is favored. This is an important consequence of RRKM theory and is a key issue for the discussion in Chapter 3.

RRKM theory is part of Computational Chemistry because all necessary input parameters, i.e., geometries of stable molecules and TS, their (ZPE corrected) energy difference, and their state densities through their vibrational frequencies and rotational constants are available from calculations.

The microcanonical rate constant is useful, when the molecule is excited to a defined energy \mathcal{E}^* through a laser or some well defined collision event in molecular beam studies. However, a canonical rate, i.e., a rate at a fixed temperature rather than at fixed energy is required when the molecule is thermally excited, i.e., through multiple collisions. The microcanonical rate is transformed into a canonical rate by weighting the microcanonical rates at each \mathcal{E}^* by their respective thermal population, given by the Boltzmann distribution in Eq. (2.25). Hence,

$$\begin{aligned}
 k_{\text{uni}}(T) &= \int_0^{\infty} k_{\text{uni}}(\mathcal{E}^*) \frac{1}{Z^*(T)} \rho_S^*(\mathcal{E}^*) \exp\left(-\frac{\mathcal{E}^*}{k_B T}\right) d\mathcal{E}^* \\
 &= d_{\text{rxn}} \frac{1}{hZ^*(T)} \int_0^{\infty} \frac{N_S^\ddagger(\mathcal{E}^\ddagger)}{\rho_S^*(\mathcal{E}^*)} \rho_S^*(\mathcal{E}^*) \exp\left(-\frac{\mathcal{E}^*}{k_B T}\right) d\mathcal{E}^* \\
 &= d_{\text{rxn}} \frac{1}{hZ^*(T)} \int_0^{\infty} N_S^\ddagger(\mathcal{E}^\ddagger) \exp\left(-\frac{\mathcal{E}^*}{k_B T}\right) d\mathcal{E}^* \\
 &= d_{\text{rxn}} \frac{k_B T}{hZ^*(T)} \int_0^{\infty} \rho_S^\ddagger(\mathcal{E}^\ddagger) \exp\left(-\frac{\mathcal{E}^*}{k_B T}\right) d\mathcal{E}^* \\
 &= d_{\text{rxn}} \frac{k_B T}{hZ^*(T)} \exp\left(-\frac{\mathcal{E}_0}{k_B T}\right) \int_0^{\infty} \rho_S^\ddagger(\mathcal{E}^\ddagger) \exp\left(-\frac{\mathcal{E}^\ddagger}{k_B T}\right) d\mathcal{E}^\ddagger \\
 k_{\text{uni}}(T) &= d_{\text{rxn}} \frac{k_B T}{h} \frac{Z^\ddagger(T)}{Z^*(T)} \exp\left(-\frac{\mathcal{E}_0}{k_B T}\right)
 \end{aligned} \tag{2.34}$$

Here, it is used that $\mathcal{E}^* = \mathcal{E}_0 + \mathcal{E}^\ddagger$ and that $\int_0^{\infty} \rho_S^\ddagger(\mathcal{E}^\ddagger) \exp(-\mathcal{E}^\ddagger/k_B T) d\mathcal{E}^\ddagger$ is just the

partition function of the TS, $Z^\ddagger(T)$. Note the factor of $k_B T$ appearing from line three to four. For more details on this derivation, see [118, p. 238f].

The last equation is the Eyring equation [125, 126], which was derived independently to RRKM theory for calculations of a thermal rate constants. The consistency of both theories is very reassuring and thus both are important tools in Computational Chemistry.

2.5 Collision Cross Section Modeling

Collision Dynamics The modeling of collision cross sections (CCS) is not traditionally part of Computational Chemistry. Nevertheless, it is closely related and an important part of this work. Especially in modern mass spectrometry, where ions are often either generated at atmospheric pressure or separated at elevated pressure conditions through their (differential) ion mobility, their motion through gases, determined by their CCS, is an important factor to consider. While the fundamentals of ion mobility [132] will not be discussed in this section, since they are not directly related to Computational Chemistry, the modeling of CCSs in general is covered.

In a gas at temperature T and particle density N , the gas particles of mass m_{bg} have many different velocities v , described by the Maxwell-Boltzmann distribution:

$$P(v)dv = 4\pi \left(\frac{m_{bg}}{2\pi k_B T} \right)^{3/2} v^2 \exp\left(-\frac{m_{bg}v^2}{2k_B T}\right) dv \quad (2.35a)$$

$$\langle v \rangle = \sqrt{\frac{8k_B T}{\pi m_{bg}}} \quad (2.35b)$$

The kinetic energy of a single particle is $1/2mv^2$. However, the energy of a collision between a particle of interest, having mass m_p , and the background gas, is not the sum of their kinetic energies. Upon an elastic collision, the velocity of the center of mass (CoM) of the two particles is not changing (neither in magnitude nor in direction) and thus, this part of the kinetic energy is not available for the collision. If the total kinetic energy is rewritten as

$$\begin{aligned} \mathcal{E}_{\text{kin}} &= \frac{1}{2}m_{bg}v_{bg}^2 + \frac{1}{2}m_p v_p^2 \\ &= \frac{1}{2}\mu v_r^2 + \frac{1}{2}(m_{bg} + m_p) v_{\text{CoM}}^2 \end{aligned} \quad (2.36)$$

it becomes clear that the collision energy (first term in the latter expression) is defined by the reduced mass, $\mu = m_p m_{bg} / (m_p + m_{bg})$, and the *relative* velocity $\mathbf{v}_r = \mathbf{v}_p - \mathbf{v}_{bg}$. If the particle of interest is diluted in the background gas and thus, can be treated as ensemble with the same temperature, the distribution of the relative velocities is also a Maxwell-Boltzmann distribution applying the reduced mass. Consequently, an

ensemble average can be defined as well:

$$P(v_r)dv_r = 4\pi \left(\frac{\mu}{2\pi k_B T} \right)^{3/2} v_r^2 \exp\left(-\frac{\mu v_r^2}{2k_B T}\right) dv_r \quad (2.37a)$$

$$\langle v_r \rangle = \sqrt{\frac{8k_B T}{\pi\mu}} \quad (2.37b)$$

A single collision event is fully defined by the reduced mass of the collision partners, their relative velocity and its *impact parameter* b , i.e., the minimum distance of the initial trajectories [133, p. 49]. While the total collision energy is given by the initial kinetic energy, $1/2\mu v_r^2$, the interaction potential V between the collision partners will influence the energetics as well, as shown in Fig. 2.5, which depicts a collision event in the CoM frame. As a result of an (elastic) collision, the trajectories change by an angle θ , the *scattering angle*. In principle, it can be computed from the known parameters of the collision, assuming specular scattering [134], through [133, p. 51]

$$\theta(\mathbf{v}_r, b) = \pi - 2b \int_{R_m}^{\infty} \left[R^2 \sqrt{1 - \frac{b^2}{R^2} - \frac{V(R)}{1/2\mu v_r^2}} \right]^{-1} dR \quad (2.38)$$

where the integration is over the distance R of the collision partners from the distance of closest encounter, R_m (see Fig. 2.5) to infinity.

The scattering angle is the central quantity when dealing with collisions because it is directly related to the amount of momentum transferred between the collision partners along \mathbf{v}_r , i.e., $\mu \mathbf{v}_r (1 - \cos\theta)$ [132, p. 145]. Integrating over a circle around \mathbf{v}_r yields the *momentum transfer* (or diffusion) cross section Q_D [132–135]

$$Q_D(\mathbf{v}_r, b) = 2\pi \int_0^{\infty} (1 - \cos\theta(\mathbf{v}_r, b)) b db \quad (2.39)$$

For a simple, spherical symmetric hard-sphere potential, Eq. (2.39) yields a cross section of πb_{\min}^2 [136], where b_{\min} is the hard-sphere distance, which is the known collision cross section from standard kinetic gas theory. However, for more realistic potentials, the scattering angle depends not only on the distance (defined through b), but also on the collision energy (defined through the relative velocity \mathbf{v}_r). Thus, for an ensemble average, taking into account many collisions with different relative velocities, instead of just using $\langle v_r \rangle$ from Eq. (2.37b), integration of Q_D over \mathbf{v}_r , weighted by the respective Maxwell-Boltzmann distribution (cf. Eq. (2.37a)), has to be performed:

$$\begin{aligned} \Omega^{(1,1)} &= \left(\frac{\mu}{2k_B T} \right)^3 \int_0^{\infty} v_r^5 \exp\left(-\frac{\mu v_r^2}{2k_B T}\right) Q_D(\mathbf{v}_r, b) dv_r \\ &= \frac{\pi}{8} \left(\frac{\mu}{k_B T} \right)^3 \int_0^{\infty} v_r^5 \exp\left(-\frac{\mu v_r^2}{2k_B T}\right) \int_0^{\infty} 2b(1 - \cos\theta(\mathbf{v}_r, b)) db dv_r \end{aligned} \quad (2.40)$$

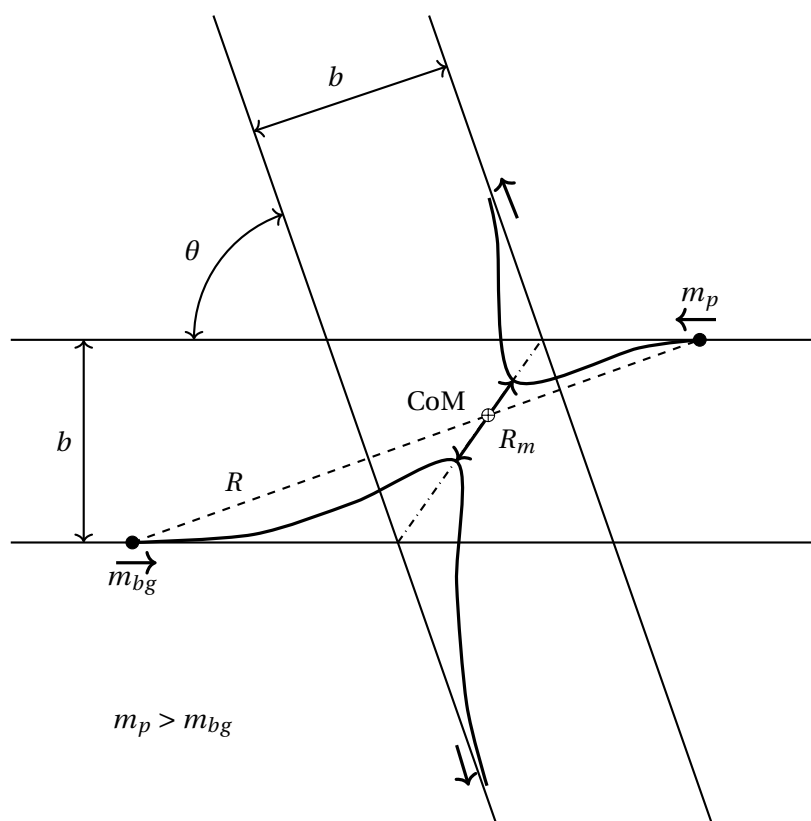


Figure 2.5: General collision schematics with impact parameter b , closest encounter R_m and scattering angle θ . The plane shown is oriented perpendicularly to the velocity of the center of mass (CoM). Note that b does not change upon an elastic collision [133, p. 448].

This quantity is called the *collision integral*. It is noted that a term of v_r^5 appears in the integral, which is an additional factor of v_r^3 compared to the Maxwell-Boltzmann distribution. This takes into account that the collision frequency (and thus the probability of a collision) is proportional to v_r and that the plane of the encounter can move with the center of mass velocity ($\propto v_r^2$) [135]. The normalization factors are chosen such that $\Omega^{(1,1)} = \pi b_{\min}^2$ for a hard-sphere collision [132, 136, 137]. The superscript (1,1) indicates that this integral is a special case of a more general definition. While this is sufficient for this work, i.e., dealing with mobility, it is noted that for other transport coefficients, such as viscosity and the mean free path, other forms or even combinations of different integrals are necessary [133, p. 484, 524].

When addressing collisions between atoms and molecules, the orientation of the collision has to be taken into account as well, since the potential is not spherically symmetric anymore. An atom hitting the molecule with the same impact parameter and relative velocity will be scattered differently, when colliding with a different site

of the molecule. Thus, $\Omega^{(1,1)}$ has to be averaged over all possible impact orientations, given by the three Euler angles α , β , and γ [136, 137]:

$$\Omega_{\text{avg}}^{(1,1)} = \frac{1}{8\pi^2} \int_0^{2\pi} d\alpha \int_0^\pi \sin\beta d\beta \int_0^{2\pi} d\gamma \quad (2.41)$$

$$\frac{\pi}{8} \left(\frac{\mu}{k_B T} \right)^3 \int_0^\infty v_r^5 \exp\left(-\frac{\mu v_r^2}{2k_B T}\right) dv_r \int_0^\infty 2b(1 - \cos\theta(\alpha, \beta, \gamma, \mathbf{v}_r, b)) db$$

α and β are the spherical coordinates, which give the orientation of one axis of the molecule with respect to the fixed coordinate system: Hence, the integration is the same as over the surface of a sphere. γ describes the rotation around this axis: Hence, the integration is over a circle [138]. The latter was already mentioned, when the integration around \mathbf{v}_r was performed to yield Q_D . The additional normalization factor of $1/8\pi^2$ accounts for the integration over all three angles. The integrals are separated since the integration variables are independent from each other. However, the scattering angle now depends on all five parameters, i.e., $\theta(\alpha, \beta, \gamma, \mathbf{v}_r, b)$.

Numerical Approaches to the CCS Solving the integral in Eq. (2.41) is rather difficult. For atom-atom collisions, full ab initio potentials can be calculated. Due to the spherical symmetry, these can be used to analytically solve Eqs. (2.38) and (2.41). For molecule-atom interactions, however, such ab initio PESs are far too complex to be calculated. Even if, analytically solving Eq. (2.38) for all combinations of parameters α , β , γ , \mathbf{v}_r , and b , on which the scattering angle depends, would be impossible and thus is to solve Eq. (2.41). One possible numerical solution to this problem is called the *trajectory method* [136, 137]. Here, a number of particles are initialized with different starting conditions, i.e., orientations, impact parameters, and relative velocities. Then, their trajectories upon collision are explicitly calculated (via Eqs. (2.23a)-(2.23c)) for a given potential V (see below) and the scattering angle is measured after the collision event. Given enough trajectories, the averages in Eq. (2.41) are approximated to a sufficient degree.

The potential necessary for the trajectory calculation may be obtained through a full HF or DFT treatment at each time step. However, this is computationally very demanding since many different trajectories have to be evaluated. Thus, it is expressed by simple functional forms defined for pair-wise interactions between the colliding atom and each nucleus in the molecule. Different forms of these potentials exist [137, 139–141]. They are often simple Lennard-Jones 12-6 or Exp-6-type potentials for the van der Waals (vdW) interaction, combined with possible ion-induced dipole

2 Introduction to Computational Chemistry

(IID) and ion-quadrupole (IQ) potentials:

$$V_{\text{tot}} = V_{\text{vdW}} + V_{\text{IID}} + V_{\text{IQ}} \quad (2.42a)$$

$$V_{\text{vdW}} = \sum_{i=1}^{M_N} \epsilon_i \left[1.84 \times 10^5 \exp\left(\frac{12R_i}{R_i^*}\right) - 2.25 \left(\frac{R_i^*}{R_i}\right)^6 \right] \quad (2.42b)$$

$$V_{\text{IID}} = -\frac{\alpha e^2}{2} \left[\left(\sum_{i=1}^{M_N} \frac{z_i X_i}{R_i^3} \right)^2 + \left(\sum_{i=1}^{M_N} \frac{z_i Y_i}{R_i^3} \right)^2 + \left(\sum_{i=1}^{M_N} \frac{z_i Z_i}{R_i^3} \right)^2 \right] \quad (2.42c)$$

$$V_{\text{IQ}} = \sum_{i=1}^{M_N} \sum_{j=1}^3 \frac{z_i z_j e^2}{R_{ij}} \quad (2.42d)$$

At every time step during each individual trajectory calculation, the potentials are evaluated for the distance R_i from the collision atom to each nucleus i in the molecule and then summed over all M_N nuclei. The parameters of the vdW interaction, i.e., the equilibrium distance R_i^* and the potential well depth ϵ_i have to be taken from force fields or parametrized to reproduce experimental CCS. In Eq. (2.42c), α is the polarizability of the collision gas, z_i are the partial charges of the nuclei in the molecule and X_i , Y_i and Z_i are their Cartesian coordinates. For IQ interactions to occur, the collision gas has to have a quadrupole moment. The latter becomes important, when molecular collision gases, as for example N_2 , are considered. For simplicity reasons, N_2 is treated as atom for the vdW and IID potentials, but for the IQ potential, a charge separation of $z_j = -0.4825e$ on the two nuclei, countered by a point charge of $z_j = +0.965e$ in between, is used to model the quadrupole moment. The second sum in Eq. (2.42d) runs over these three charges [141].

The input needed for such calculations is thus the geometry of the molecule, for which the CCS is to be calculated, as well as the partial charges of each nucleus. These data are readily obtained by standard ab initio calculations (HF, post-HF or DFT), which explains the close relation of CCS modeling to Computational Chemistry. This relation is also highlighted through the use of force fields describing the vdW interactions.

It is noted that this definition of the (averaged) collision integral, $\Omega_{\text{avg}}^{(1,1)}$ seems to be rigorous. However, it neglects orientational preferences, which might appear for ions in electric fields, as well as inelasticity of collisions, where momentum conservation does not hold anymore. The nuclei of the molecule are fixed in space at positions \mathbf{X} , \mathbf{Y} and \mathbf{Z} and do not move upon collision. An even more rigorous approach to the calculation of CCS, i.e., by actually performing MD simulations of a molecule in a gas, explicitly modeling inelastic collisions as been introduced recently [142]. Additionally, it is noted that the described formalism is purely classical. Thus, quantum effects of the scattering process [133, p. 69, 668] are completely ignored.

3 Charge Retention/Charge Depletion in ESI-MS

Preceding work to this chapter has been published in

M. Thinius, C. Polaczek, M. Langner, S. Bräkling, A. Haack, H. Kersten, and T. Benter, “Charge Retention/Charge Depletion in ESI-MS: Experimental Evidence”, [Journal of The American Society for Mass Spectrometry](#), DOI: [10.1021/jasms.9b00044](#) (2020)

and is summarized in the [Introduction](#) section below. The main part of this chapter is reprinted with permission from

A. Haack, C. Polaczek, M. Tsolakis, M. Thinius, H. Kersten, and T. Benter, “Charge Retention/Charge Depletion in ESI-MS: Theoretical Rationale”, [Journal of The American Society for Mass Spectrometry](#), DOI: [10.1021/jasms.9b00045](#) (2020). Copyright 2020 American Chemical Society.

3.1 Abstract

Gas phase modification in ESI-MS can significantly alter the charge state distribution of small peptides and proteins. The preceding paper presented a systematic experimental study on this topic using Substance P and proposed a charge retention/charge depletion mechanism, explaining different gas- and liquid-phase modifications. In this work we aim to support this rationale by theoretical investigations on the proton transfer processes from (multiply-)charged analytes towards solvent clusters. As model systems we use small (di-)amines as analyte and methanol (MeOH) and acetonitrile (ACN) as gas phase modifiers. The calculations are supported by a set of experiments using (di-)amines, to bridge the gap between the present model system and Substance P used in the preceding study. Upon calculation of the thermochemical stability as well as the proton transfer pathways, we find that both ACN and MeOH form stable adduct clusters at the protonation site. MeOH can form large clusters through a chain of H-bridges, eventually lowering the barriers for proton transfer to an extent that charge transfer from the analyte to the MeOH cluster becomes feasible. ACN, however, cannot form H-bridged structures due to its aprotic nature. Hence, the charge is retained at the original protonation site, i.e., the analyte. The investigation confirms the proposed charge retention/charge depletion model. Thus, adding aprotic solvent vapors to the gas phase of an ESI source more likely yields higher charge states than using protic compounds.

3.2 Introduction

High charge states of peptides or proteins generated with electrospray ionization (ESI) are favorable for structural investigations in mass spectrometry. Increasing the charge state lowers the m/z ratio so that larger proteins can be analyzed by standard mass analyzers [145, 146]. Furthermore, dissociation efficiencies are enhanced, which is relevant to sequence analysis [147]. Adding supercharging agents (SCAs) such as *m*-nitrobenzyl alcohol to the solution phase increases the mean as well as the maximum charge state of proteins [148, 149]. The mechanism of this supercharging is rationalized in terms of a local enrichment of the SCAs around the protein caused by rapid water/solvent loss. Unfavorable charge stabilization in the bulk liquid phase forces the protons (or metal cations) to remain at the protein site, which in turn leads to a higher charge state [150].

However, other methods add polar solvent vapors to the gas phase of the ion source to change the charge state distribution [151–153]. In the work described in the preceding article of this journal issue, we investigated the charge state distribution of Substance P (SP) generated with a nano-ESI (nESI) source with different gas phase modifiers [143]. The fact that adding acetonitrile (ACN) mainly yielded $[\text{SP} + 3\text{H}]^{3+}$ ion signals whereas methanol (MeOH) exclusively produced the $[\text{SP} + 2\text{H}]^{2+}$ charge state ion signals indeed strongly suggests that gas phase modifiers can significantly change the observed charge state distribution. Due to Coulombic repulsion (as discussed in [143]; see also [Supporting Information](#)), it is very unlikely that an additional charge is transferred to the $[\text{SP} + 2\text{H}]^{2+}$ to form the $[\text{SP} + 3\text{H}]^{3+}$ in the gas phase, e.g., via proton transfer from a protonated solvent or modifier molecule. Hence, the only pathway that leads to the $[\text{SP} + 3\text{H}]^{3+}$ is the initial production of the 3+ charge state followed by charge retention throughout the ion transfer from the ion source to the mass analyzer. Depending on their properties, gas phase modifiers seem to either retain the initial analyte charge(s) or remove/deplete them via proton transfer reactions.

Work from the Russell group [47, 154, 155] shows that the initial, kinetically trapped structure of $[\text{SP} + 3\text{H}]^{3+}$ in the gas phase is probably as follows: Two of the protons are H-bonded intramolecularly and are thus protected from removal. The third charge however resides at the guanidine group of arginine and is located at the outside of the polypeptide. This would explain why the charge state can only be altered between 3+ and 2+. Gas phase proton transfer from multiply-charged proteins to bases such as dimethylamine or even water has been observed before and even used to deliberately reduce the charge state [156–159]. It is also known that MeOH can remove the charge from protonated molecules with lower proton affinity (PA) through gas phase proton transfer even in medium pressure regions [160]. However, the PA of methylguanidine (acting as simple model for the active amino acid in SP) of about 1000 kJ mol^{-1} [161] is higher than the PAs of ACN ($779.2 \text{ kJ mol}^{-1}$) and MeOH ($754.3 \text{ kJ mol}^{-1}$) [162]. Even when the proton of interest resides at the amine group of lysine, the PA would still be higher (methylamine, again acting as model substance, has a PA of $899.0 \text{ kJ mol}^{-1}$

[162]). From these simple considerations, both ACN and MeOH should not be able to remove the charge via proton transfer.

In the preceding publication in this journal [143] as well as in other reports [163], it is argued that one of the two following mechanisms potentially leads to charge depletion: i) the abundant gas phase modifiers cause an enrichment of these modifiers *in* the ESI generated droplets at late evaporation stages. This may even slow down the evaporation rate, which is supported by our CID experiments conducted: Adding gas phase modifiers did not lead to any fragmentation of SP, whereas fragmentation was observed in the absence of the modifiers [143]. When the evaporation process is slowed down significantly, a small droplet may enter the collision region and collisional activation does not lead to fragmentation of SP but only to further evaporation of the droplet. Similar observations with non-covalent protein complexes were rationalized in the same way [164]. In these droplets, proton transfer can occur due to a stabilizing effect for solvated protons. ii) Another plausible mechanism is the growth of small modifier clusters at the protonation site in the gas expansion in the first pressure reduction stage of the mass spectrometer: Since the local temperature drops significantly due to the adiabatic expansion, the abundant modifier molecules can "condense" at the charge site. When the cluster size is sufficiently large, the PA of the cluster exceeds the PA of the protonated amino acid. From a molecular point of view, these two mechanisms are similar: In both cases the modifier molecules concentrate around the charge site - in the liquid or gas phase - and may remove the proton in a proton transfer reaction forming protonated modifier clusters. From a macroscopic point of view, there is a continuous transition between small, modifier-rich droplets, and clustering of modifier around the charge site.

In this contribution, we model the proton transfer reaction by bringing a number of modifier molecules in close proximity to the charge site, which, for simplicity reasons, will be called "clusters" from here on.

Proton bound MeOH clusters are known since long [165] and their PAs have been determined for different n [166]. For $n = 3$, the PA of the MeOH cluster is $937.2 \text{ kJ mol}^{-1}$ and thus exceeding that of methylamine. Because of its structure, MeOH can form large stable clusters through hydrogen bond networks as shown by theoretical investigations [167–169]. For ACN, however, this is more difficult. ACN can stabilize a proton only with one H-bond without offering another H-bonding site. Any additional ACN molecule can only form weaker interactions as for example $\text{CH}\cdots\text{N}$, dipole-dipole or electron-donating interactions (see below). Therefore, $(\text{ACN})_2\text{H}^+$ with a PA of about 895 kJ mol^{-1} [170–172] is the largest closed solvent shell cluster (in the literature the term "closed shell cluster" is also used) [173]. Larger clusters $(\text{ACN})_n\text{H}^+$ with $n \geq 3$ are rather unstable and are only rarely observed [174–176]. Mechanistically both modifiers S must initially stabilize/solvate the charge through formation of a hydrogen bond of the type $\text{A}^+ - \text{H}\cdots\text{S}$. It is argued that ACN will remain a ligand, MeOH however can remove the charge via proton transfer, when a sufficiently large cluster network is formed. These theoretical considerations are supported by experimental results [143]: Using ACN as gas phase modifier, $[\text{SP} + 3\text{H} +$

3 Charge Retention/Charge Depletion in ESI-MS

ACN]³⁺ was observed as a significant signal, whereas the [SP + *n*H + MeOH]^{*n*+} (*n* = 2, 3) clusters were not observed. Instead, MeOH formed protonated clusters of the type (MeOH)_{*n*}(H₂O)_{*m*}H⁺, where *n* = 10 – 30 and *m* = 0 – 3. In addition to protonated ACN monomers (ACN)H⁺ and dimers (ACN)₂H⁺, no other ACN clusters were observed. For more details cf. [143].

In this work the mechanism of "supercharging" (or much more appropriate: charge retention) through gas phase modifiers is investigated, i.e., if and how protons residing at the basic sites of amino acids are transferred to modifier molecules or clusters via proton transfer reactions. Because the mechanisms and fundamental reasons are difficult to access with experimental methods, this work mainly contains theoretical investigations on primary amines acting as model systems for lysine. Only selected experimental data are presented to validate the theoretical findings regarding these model systems.

3.3 Methods

3.3.1 Computational Methods

All calculations are carried out with the Gaussian09 [177] program package. Density functional theory (DFT) is used with the B3LYP [86, 87] functional and def2-TZVPP [75] basis set for every geometry optimization. Empirical dispersion corrections (-D) [178] are always applied for a better description of loosely bound clusters. Scans of bond distances or - in case of observed transition states (TS) - intrinsic reaction coordinate calculations [93] are conducted at the same level of theory.

For thermochemical data the B3LYP-D/def2-TZVPP optimized geometries are used to conduct frequency calculations for obtaining zero-point energy and thermodynamic functions at the same level of theory [179]. Additionally, single point energy calculations are performed with the more accurate B2PLYP-D [88] functional using the larger def2-QZVPP basis set to correct for electron correlation and to reduce basis set superposition error. This method, using the B2PLYP-D/def2-QZVPP//B3LYP-D/def2-TZVPP electronic energy combined with the B3LYP-D/def2-TZVPP thermodynamic functions, is validated by calculating the proton affinity (PA) and gas phase basicity (GB) of ACN, MeOH and methylamine (see Table 3.1). It is found that using the refined electronic energy increases the accuracy, whereas the computational time remains reasonable. Especially for ACN, the experimentally measured PA or GB are in much better agreement. It is noted that the B2PLYP-D/def2-QZVPP//B3LYP-D/def2-TZVPP electronic energy deviates only in the 10 – 100 μE_h range from the value obtained with a full geometry optimization on B2PLYP-D/def2-QZVPP.

3.3.2 Experimental Methods

Experiments were performed with a Bruker micrOTOF instrument (Bruker Daltonics, Bremen, Germany). Ion transfer stage voltages, optimized for the low mass range

Table 3.1: Proton affinities (PA) and gas phase basicities (GB) in kJ mol^{-1} at 298.15 K calculated at B3LYP-D/def2-TZVPP and by combining B2PLYP-D/def2-QZVPP//B3LYP-D/def2-TZVPP electronic energies with B3LYP-D/def2-TZVPP thermal energy.

Method	MeOH		ACN		MeNH ₂	
	PA	GB	PA	GB	PA	GB
B3LYP	759.8	729.4	792.9	761.8	904.0	869.6
B2PLYP//B3LYP	754.8	724.4	784.6	753.5	899.2	864.8
Lit. (exp.) [162]	754.3	724.5	779.2	748	899.0	864.5

and minimum fragmentation, are as followed: capillary exit: 50 V, skimmer 1: 35 V, skimmer 2: 23 V, hexapole 1: 22 V, hexapole 2: 20.6 V, hexapole RF: 80 V, lens transfer time: 25 μs , lens pre-pulse storage time: 5 μs , lens 1 storage: 30 V, lens 1 extraction 21.5 V, lens 2: 9 V, lens 3: -20 V, lens 4: 0 V, lens 5: -27.5 V. The experimental setup has been described previously [143]. Briefly, the ions are generated from a custom nano-electrospray ionization source, which is mounted directly onto the inlet capillary of the mass spectrometer. The source chamber is flushed with nitrogen. The gas phase modifiers are introduced to the ion source via the nitrogen gas line. The resulting mixing ratio of the modifiers is varied between 0.1 and 15 %V. Solvents used for sample preparation and gas phase modification all have HPLC grade purity. 1,9-diaminononane, 1,8-diaminooctane, 1,5-diaminopentane, 1,3-diaminopropane, ethylenediamine, 1,2-phenylenediamine, *n*-hexylamine, *n*-butylamine were used as analytes. The electrospray solutions are composed of acetonitrile/water (1/1) and methanol/water (1/1), respectively, with 0.1 % formic acid ($\geq 98\%$) and 0.1 mM/1 mM analyte concentrations. All chemicals were purchased from Sigma-Aldrich Chemie (Munich, Germany) and are used without further purification.

3.4 Results and Discussion

3.4.1 Structure and Thermochemistry of the Different Clusters

The observed differences of the behavior of MeOH and ACN regarding charge retention or charge depletion must have something to do with their structures and their ability to form larger clusters, because their PAs are comparable. Although the structures of the homogeneous MeOH clusters were recently calculated [169], structural analysis of the ACN clusters used restrictions during geometry optimization and has become obsolete [171]. Here, we present full geometry optimizations at the B3LYP-D/def2-TZVPP level of theory for homogeneous MeOH clusters up to $n = 4$ and ACN clusters up to $n = 3$. Additionally, the analyte adduct clusters, i.e., the protonated amine (or diamine) with n ligands $[\text{A} + \text{H} + \text{S}_n]^+$, are geometrically optimized as well.

Homogeneous Cluster Figure 3.1 shows the geometries of the homogeneous protonated solvent clusters, S_nH^+ . Full Cartesian coordinates are given in the *Supporting Information* of the original publication. The structures shown always correspond to the global minimum structures with respect to the Gibbs enthalpy.

As already mentioned in the literature for MeOH [169], when n is odd, there is an oxygen atom with two closely bound protons and when n is even, the additional proton is located between two oxygen atoms. The entire cluster structure is bound through a linear chain of hydrogen bonds, and thus, the charge formally resides in the center of the H-bond chain but is effectively distributed over the entire cluster. A small asymmetry is discernible for $n = 4$: the two hydrogen bonds binding the proton are not equally long (1.135 Å and 1.271 Å). This was already mentioned in the literature [169]. A cyclic structure exists for $n = 4$, but is less stable than the linear structure [169]. The thermodynamic stabilities of the MeOH clusters calculated in the present work are summarized in Table 3.2, with good agreement with experimental data.

The ACN clusters behave differently: Because ACN can only act as a donor for one hydrogen bond without offering a new acceptor site as MeOH does, direct interaction with the charge is possible only via direct binding to the proton. For $n = 2$, the most stable configuration is a linear arrangement of the two ACN molecules, which was already proposed in the literature [171]. However, in contrast to this report, the present results suggest that the symmetric dimer, where both N–H distances are equal, is more stable than the asymmetric dimer. This is probably due to the lack of polarization and diffuse functions applied in the former study. Using the B3LYP-D/def2-TZVPP model chemistry, the asymmetric configuration has no minimum on the potential energy surface. For $n = 3$ however, neither the symmetric cluster with a C_3 axis, nor a T-shaped system appears to have a minimum with the present method applied. Instead, the third ACN binds to a mostly unchanged dimer through an interaction of the CN lone pair with the π -system of another ACN molecule. The proton does not seem to be able to bind three ACN molecules. The binding energy of the third ACN is consequently low ($\Delta(\mathcal{E}_{\text{elec}} + \mathcal{E}_{\text{ZPE}})_{2 \rightarrow 3} = 0.442$ eV compared to 0.931 eV for MeOH $n = 2 \rightarrow 3$), supporting the experimental findings that $(\text{ACN})_2H^+$ is the largest closed solvent shell cluster [173]. A configuration where the third ACN binds in a head-to-tail fashion is possible but less stable by 18 kJ mol⁻¹. The calculated thermodynamic stability of the ACN clusters is always larger than the experimentally derived values (cf. Table 3.2). Especially for the trimer the difference is comparably large with the calculated energy being 1.5 times the experimental value. This may be due to the failure of the harmonic approximation for the calculation of the vibrational partition function of the clusters, caused by significant anharmonic effects. For example, a vibration corresponding to the proton motion between the two ACN nitrogen atoms in the dimer is very poorly described by the harmonic potential around the minimum (leading to a normal frequency of 300 cm⁻¹). A scan of the energy along this normal coordinate reveals a potential being rather flat around the equilibrium but rising steeply for smaller N–H distances (see Fig. 3.8b). The fundamental frequency is thus expected to be much higher than the harmonic, leading to a smaller partition func-

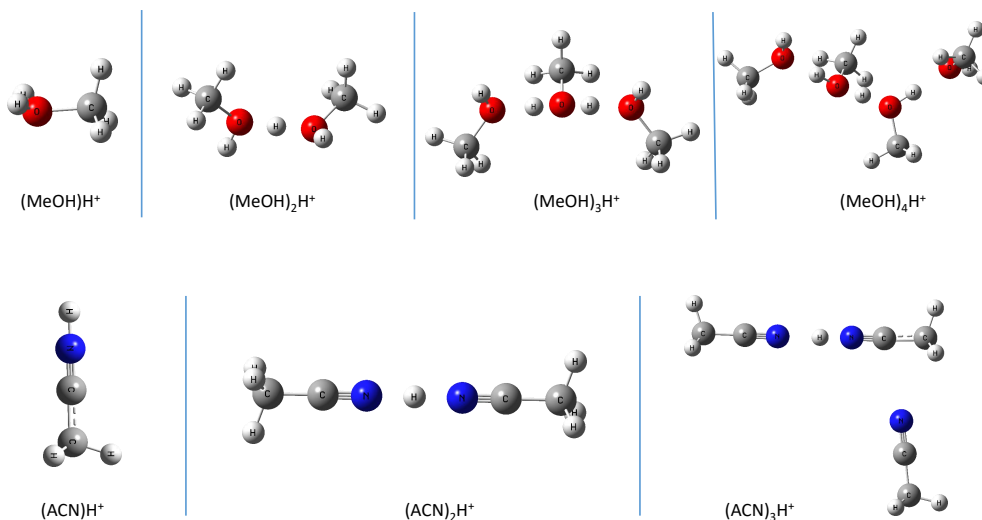


Figure 3.1: Optimized geometries of the $(\text{MeOH})_n\text{H}^+$ ($n = 1, \dots, 4$) and $(\text{ACN})_n\text{H}^+$ ($n = 1, 2, 3$) clusters obtained at the B3LYP-D/def2-TZVPP level of theory. Although there are more conformers possible, only the global Gibbs enthalpy minimum structures are shown.

Table 3.2: Stability of the homogeneous clusters by combining B2PLYP-D/def2-QZVPP//B3LYP-D/def2-TZVPP electronic energies with B3LYP-D/def2-TZVPP thermal energy. Values are $-\Delta_{\text{hom}}G$ in kJ mol^{-1} for the reactions $\text{S}_{(n-1)}\text{H}^+ + \text{S} \longrightarrow \text{S}_n\text{H}^+$ at 298.15 K.

	S=MeOH				S=ACN		
	$n = 1$	2	3	4	$n = 1$	2	3
This work	724.4	105.8	54.9	30.2	753.5	108.2 ^c	21.3
Lit. (exp.) ^a	724.5	100.1	55.0	30.2	748	92.6	14.1
Lit. (calc.) ^b		116.7	48.5	38.1			

^a $n = 1$ from [162], all other values calculated with measured data from Ref. [54]

^b calculated on M062X/6-31++G(d,p) [168]

^c including free rotor correction to the partition function

3 Charge Retention/Charge Depletion in ESI-MS

tion. Strong anharmonicity in proton-sharing systems is long known from theory and experiments [180–182]. Also, the internal rotations of the methyl groups around the N–C–C axis are essentially barrier free. The barrier height in the dimer is estimated to be 3.43 μeV (see Fig. 3.8a). Thus the partition function was corrected assuming a free internal rotor [101, 102]:

$$Z_{\text{IR}} = \frac{1}{\sigma_{\text{IR}}} \left(\frac{8\pi^3 I_{\text{IR}} k_B T}{h^2} \right)^{1/2} = 2.62 \quad (3.1)$$

where $\sigma_{\text{IR}} = 3$ is the symmetry number of the rotation, k_B the Boltzmann constant, h the Planck constant, T the absolute temperature and $I_{\text{IR}} = 2.66 \times 10^{-47} \text{ kg m}^2$ the reduced moment of inertia. This reduces the difference to the experimental value to 2.4 kJ mol^{-1} . Hence, anharmonic effects seem to have a greater contribution to the error. A free rotor correction was thus not performed for the trimer.

Amine-Solvent Adduct Clusters Fig. 3.2 shows stable geometries of the adduct clusters $[A + H + S_n]^+$. Again, full Cartesian coordinates are given in the *SI* of the original publication. For the MeOH clusters, it is clearly visible that each cluster is formed through a chain of hydrogen bonds, where each additional MeOH binds to the free hydrogen of the OH group of the former MeOH. Although there are other structures, e.g. a cyclic one for $n = 4$, or cage structures, where all MeOH molecules bind to different hydrogens of the amine, these linear configurations are most likely promoting proton transfer because the solvent cluster moiety at the charged analyte site is reminiscent of the respective isolated protonated solvent cluster (cf. Fig. 3.1). This is the reason why we presented these conformations, despite the fact that they may not represent the global energy minimum. Although the PA of the trimer is already larger than that of methylamine, the proton still resides at the amine and is not incorporated into the MeOH cluster. This is in agreement with an investigation of protonated trimethylamine/MeOH clusters [183]. In this work, the authors report that the first linear $[\text{MeNH}_2 + H + (\text{MeOH})_n]^+$ cluster with both an amine and a MeOH ion core is found for $n = 6$.

The situation is different for ACN. The $[\text{MeNH}_2 + H + (\text{ACN})_2]^+$ cluster has its global minimum for a configuration, where the two ACN molecules bind to different protons of the amine. These geometries are also possible for MeOH, but in case of ACN, *only* this cluster is expected to exist. We tried to optimize a structure with the second ACN binding to the first one in a similar fashion as the third ACN binds to the $(\text{ACN})_2\text{H}^+$ cluster, namely via a weak interaction of the CN lone pair to the π -system of the ACN (cf. Figure 3.1). However, there is no or only a very shallow minimum, which is most probably not stable under thermal conditions; furthermore, both ACN molecules binding to the same proton does not result in a stable geometry. There is probably a stable although much weaker configuration, where the second ACN binds to the tail of the first. This configuration however does not lead to the global minimum of the $(\text{ACN})_2\text{H}^+$ cluster via proton transfer and the proton transfer energetics is

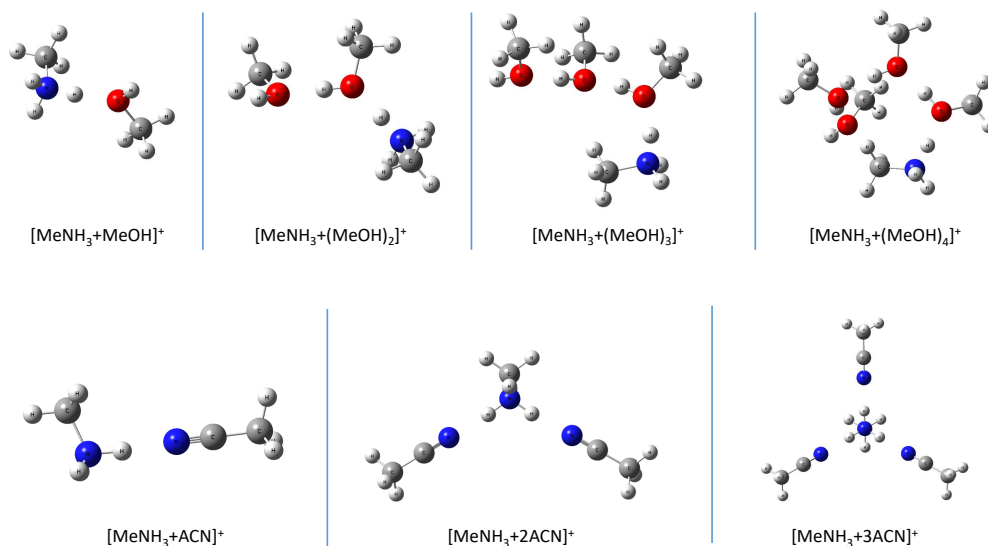


Figure 3.2: Optimized geometries of the $[\text{MeNH}_2 + \text{H} + (\text{MeOH})_n]^+$ ($n = 1, \dots, 4$) and $[\text{MeNH}_2 + \text{H} + (\text{ACN})_n]^+$ ($n = 1, 2, 3$) clusters obtained at the B3LYP-D/def2-TZVPP level of theory.

Table 3.3: Stability of the adduct clusters by combining B2PLYP-D/def2-QZVPP//B3LYP-D/def2-TZVPP electronic energies with B3LYP-D/def2-TZVPP thermal energy. Values are $-\Delta_{\text{mix}}G$ in kJ mol^{-1} for the reactions $[\text{MeNH}_2 + \text{H} + \text{S}_{(n-1)}]^+ + \text{S} \longrightarrow [\text{MeNH}_2 + \text{H} + \text{S}_n]^+$ at 298.15 K.

	S=MeOH				S=ACN		
	$n = 1$	2	3	4	$n = 1$	2	3
This work	54.7	24.5	5.6	-4.9	79.4	54.4	42.0
Lit. (exp.)	49.3 ^a				77.8 ^b	48.4 ^b	29.9 ^b

^a Ref. [184]

^b Ref. [185]

3 Charge Retention/Charge Depletion in ESI-MS

most likely similar to the situation of only one ACN present. Hence, the bound ACN molecules present at the charge site of the analyte do structurally not resemble the (most stable) homogeneous proton bound ACN cluster - which is in contrast to the MeOH case. The resulting effects are discussed in the following.

Table 3.3 shows the thermodynamic stability of the adduct clusters. For MeOH, each additional ligand binds further and further away from the charge site - because the proton remains located at the amine. The Gibbs enthalpy gained by growing thus diminishes rapidly. For $n = 4$, further increase of the cluster size even seems to be thermodynamically unfavorable, i.e., the gain of binding energy is not large enough to compensate the entropy loss. In case of ACN, however, the energy decreases far less pronounced, because every new ACN molecule binds directly to the charge site. Experimental data for ACN indicate that the stability of the larger clusters is overestimated in the calculations, most probably due to anharmonic effects of the NH–N bridges. However, the principal conclusion still holds.

With the GB of methylamine, the cluster and the adduct stabilities, the Gibbs enthalpy of the adduct formation $\Delta_{\text{Ass}}G$ ($\text{MeNH}_3^+ + nS \longrightarrow [\text{MeNH}_2 + \text{H} + S_n]^+$), their dissociation towards the products $\Delta_{\text{Diss}}G$ ($[\text{MeNH}_2 + \text{H} + S_n]^+ \longrightarrow \text{MeNH}_2 + S_n\text{H}^+$), and the overall reaction $\Delta_{\text{Rxn}}G$ ($\text{MeNH}_3^+ + nS \longrightarrow \text{MeNH}_2 + S_n\text{H}^+$) is calculated using a thermodynamic cycle:

$$\Delta_{\text{Ass}}G(n) = \sum_{i=1}^n \Delta_{\text{mix}}G(i-1 \rightarrow i) \quad (3.2)$$

$$\Delta_{\text{Diss}}G(n) = - \sum_{i=1}^n \Delta_{\text{mix}}G(i-1 \rightarrow i) + \text{GPB}(\text{MeNH}_2) + \sum_{i=1}^n \Delta_{\text{hom}}G(i-1 \rightarrow i) \quad (3.3)$$

$$\Delta_{\text{Rxn}}G(n) = \text{GPB}(\text{MeNH}_2) + \sum_{i=1}^n \Delta_{\text{hom}}G(i-1 \rightarrow i) \quad (3.4)$$

where $\Delta_{\text{mix}}G(i-1 \rightarrow i)$ are the Gibbs enthalpies of growth of the adduct clusters as listed in Table 3.3, and $\Delta_{\text{hom}}G(i-1 \rightarrow i)$ are the Gibbs enthalpies of growth of the homogeneous clusters as listed in Table 3.2. $\Delta_{\text{Rxn}}G$ incorporates the energy gain through the proton transfer and also the contribution from the solvent-solvent bindings. This cycle was chosen because we believe that it reflects the actual steps, i.e., the solvent-solvent binding occurs sequentially at the charge site (and not a neutral solvent cluster S_n binding to the charge site). The three reaction steps are shown in Fig. 3.3 for MeOH and ACN. For both modifiers the overall reaction is thermodynamically favorable for $n \geq 3$. This supports the notion that with increasing number of solvent molecules, the proton transfer from an analyte is possible, even when its PA or GB is higher than that of the solvent. However, in both cases the adduct clusters $[\text{MeNH}_2 + \text{H} + S_n]^+$ are constantly lower in energy as compared to their dissociation products, i.e., $\Delta_{\text{Diss}}G(n)$ is positive for all n . It follows that in thermodynamic equilibrium a distinct distribution of the adduct clusters should be present.

There are two main differences discernible for MeOH and ACN: First, the mixed MeOH clusters are not as stable and not much energy is gained by adding more and

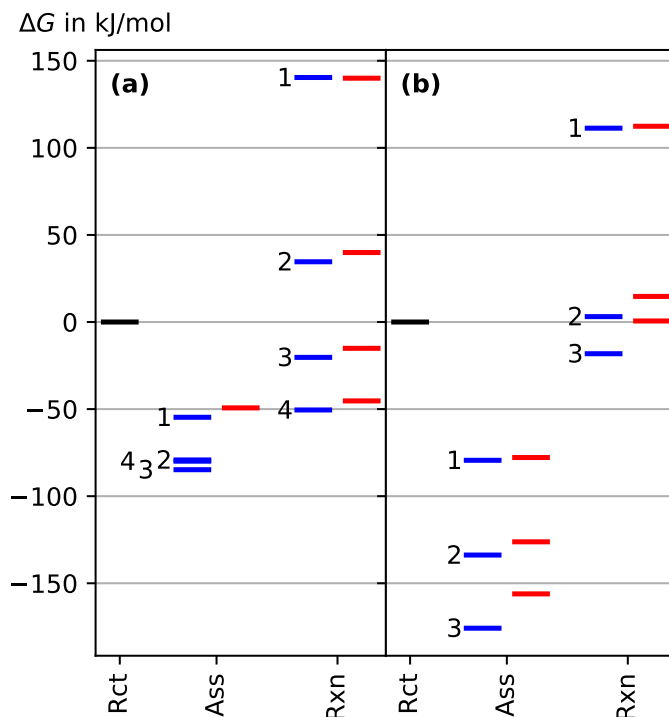


Figure 3.3: Gibbs enthalpy (Eqs. (3.2)-(3.4)) through association, Ass ($\text{MeNH}_3^+ + n\text{S} \longrightarrow [\text{MeNH}_2 + \text{H} + \text{S}_n]^+$) and for the overall reaction, Rxn ($\text{MeNH}_3^+ + n\text{S} \longrightarrow [\text{MeNH}_2 + \text{H} + \text{S}_n]^+ \longrightarrow \text{MeNH}_2 + \text{S}_n\text{H}^+$) relative to the reactants, Rct ($\text{MeNH}_3^+ + n\text{S}$, black) for (a) MeOH and (b) ACN for different n . Blue: this work, red: experimental data from literature (see Tables 3.2 and 3.3).

more MeOH. In case of ACN, the energy is significantly lower and every additional solvent adds a significant amount of binding energy. Second, the dissociation energy for MeOH decreases with n because the energy gain in the homogeneous clusters is still large. This results in a comparably small value for $\Delta_{\text{Diss}}G(4; \text{MeOH})$ of 29.4 kJ mol^{-1} . In case of ACN, however, $\Delta_{\text{hom}}G(i-1 \rightarrow i)$ becomes rather small due to the unfavorable binding situation in the trimer, which decreases the smallest dissociation energy to $\Delta_{\text{Diss}}G(2; \text{ACN}) = 136.9 \text{ kJ mol}^{-1}$. In summary, because the energy of the mixed clusters for MeOH is not lowered significantly, whereas the homogeneous cluster are still gaining energy, the dissociation energy is becoming smaller with n . For ACN, the mixed clusters are lower in energy but the homogeneous clusters are not becoming much more stable, which results in a considerably higher dissociation energy for every n .

Due to the high collision energies and steep concentration gradients in the different pressure reduction stages of the mass spectrometer, the chemical system is not in thermodynamic equilibrium during the ion transfer to the mass analyzer. Clus-

ter dissociation processes occur and lead to either the protonated amine and neutral solvent or the protonated solvent cluster and the neutral amine. The kinetics of the latter reaction path, the proton transfer to the homogeneous solvent clusters, is investigated in the next section.

3.4.2 Proton Transfer Paths

The potential energy surfaces (PES) of the different adducts were investigated with respect to the corresponding proton transfer reactions. Starting from the optimized geometries of the $[A + H + S_n]^+$ clusters, the H-bonded proton was moved away from the amine nitrogen with a step size of 0.05 \AA . Whereas this distance was held constant for each step of the scans, all other internal parameters were allowed to relax. According to [186], the parameter δ may be defined as a measure of the proton transfer progression

$$\delta = d(N_{\text{amine}}, H) - d(S, H) \quad (3.5)$$

where $d(X, Y)$ is the distance between atoms X and Y and S is either the methanol oxygen or the acetonitrile nitrogen. When a transition state (TS) is found on the PES, the intrinsic reaction coordinate (IRC) is calculated.

Amine-methanol Clusters In Figure 3.4 the calculated PES scans for the linear $[\text{MeNH}_2 + H + (\text{MeOH})_n]^+$ ($n = 1, \dots, 4$) proton transfer reactions are shown. As discussed in the structure section, there is only one minimum structure for each cluster and the corresponding geometry shows a protonated amine and the MeOH molecules acting as ligands (Fig. 3.2). Hence, moving the proton towards the MeOH cluster, the energy increases and finally approaches the value of the reaction products. The terminal points of each scan are located in the product channel of the PES, i.e., the O–H distance does not change significantly anymore and only the amine nitrogen is moved away from the proton.

The binding energy of additional H-bonded MeOH molecules is similar for each cluster, as seen in the energy differences of the four equilibrium geometries. However, throughout the scans, the energy contribution caused by the more favorable charge stabilization in the larger clusters becomes apparent. This is shown by the different gradients of the PES': Moving the proton from the amine towards the methanol molecule ($\delta < 0$), the PES' have positive curvatures. When the proton is closer to the MeOH oxygen than to the amine nitrogen ($\delta > 0$), the curvature changes its sign for higher n : The initial incline is lowered and a plateau forms around 0.25 \AA , caused by the initial OH bond of the first MeOH being stretched towards its acceptor oxygen and thus delocalizing the charge from the first oxygen into the cluster. For $n = 3, 4$ the second proton transfer, with the second H atom moving along the H-bond chain towards its acceptor oxygen atom, is discernible in the scans. Hence, the H-bond network between the MeOH molecules is a crucial characteristic for the lower steepness of the PES'.

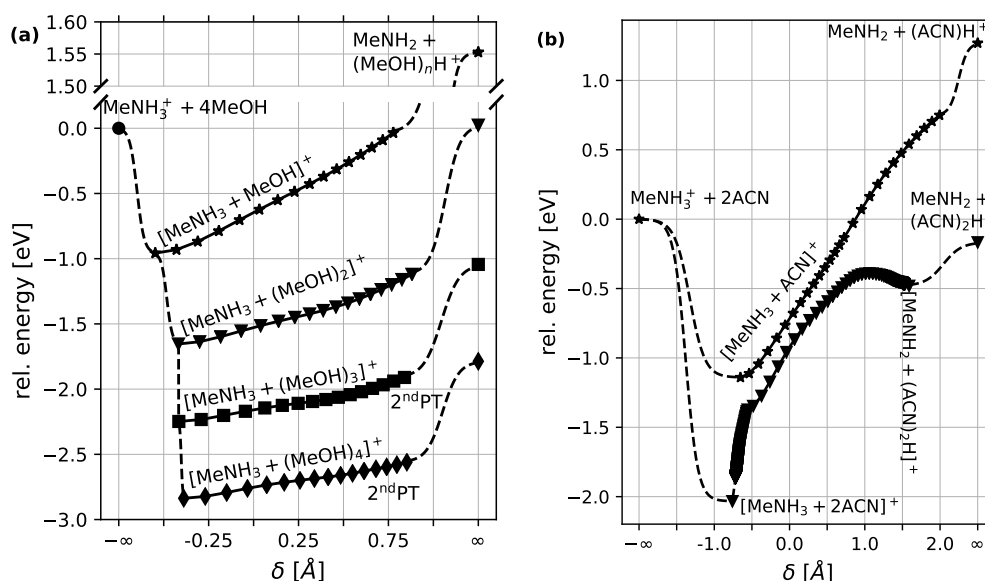


Figure 3.4: PES scans for the $[\text{MeNH}_2 + \text{H} + \text{S}_n]^+$ clusters with S being (a) MeOH and (b) ACN. Energies are obtained at the B3LYP-D/def2-TZVPP level of theory. Dashed lines connect bound and unbound configurations. The first point for each scan corresponds to the minimum structure, the H-bound proton is then moved towards the MeOH oxygen or ACN nitrogen, respectively. All energies are given relative to that of the reactants.

All of the reactions shown have single-well potentials, and thus there are no transition states to surpass. Assuming a $[\text{MeNH}_2 + \text{H} + (\text{MeOH})_n]^+$ cluster is gaining energy in the reaction coordinate through collisions, it can decompose into $\text{MeNH}_2 + (\text{MeOH})_n\text{H}^+$ or release one MeOH molecule. The latter reaction path is always lower in energy for the four PES' calculated, although the state density of both channels should be similar, because they are dissociative channels. Hence it does not represent the favored reaction channel and the mixed clusters should still be detectable [187].

Amine-acetonitrile Clusters Figure 3.4 also shows the PES scans for the ACN clusters. Whereas for $n = 1$ the potential curve is similar to the MeOH case, for $n = 2$ there are significant differences. The most obvious difference is a second minimum with an ACN ion core, which resembles the homogeneous ACN dimer having a MeNH_2 ligand weakly binding to the proton (see Fig. 3.9). Both minima are connected via a TS located 1.64 eV above the lower minimum, i.e., the cluster with amine ion core. A TS for the reaction of the proton bound ACN dimer with methylamine was already considered in an experimental study [188]. This TS exists, because for the

3 Charge Retention/Charge Depletion in ESI-MS

proton transfer to occur, the second ACN molecule has to move away from its proton to bind to the transferred one. While the proton does not move initially (cf. the steep increase in energy caused by small change in δ), the second ACN moves away from its proton resulting in a configuration, where both ACN bind to the same proton. This configuration, where three electron rich nitrogen atoms bind to the same proton, is energetically not favorable. The proton transfer occurs with the proton moving from the amine-ACN bridge to the ACN-ACN bridge and, with subsequent loss of the amine, the energy decreases until the second minimum is reached. Consequently, the TS exists because the two ACN molecules, unlike MeOH, cannot form a moiety at the charge site, which is structurally similar to the isolated homogeneous cluster.

For the proton transfer to occur, the system has to pass through a TS, which is energetically favored but kinetically hindered. When the $[\text{MeNH}_2 + \text{H} + 2\text{ACN}]^+$ cluster is formed and subsequently collisionally activated, the cleavage channel towards $[\text{MeNH}_2 + \text{H} + \text{ACN}]^+ + \text{ACN}$ is much lower in energy than the TS, which yields $(\text{ACN})_2\text{H}^+$. Additionally, the state density of the TS is much lower than that of the cleavage reactions [62], which renders the passing of the TS statistically unfavorable. Thus, proton transfer is very unlikely to occur with ACN as modifier because it cannot form H-bond networks.

Comparison with Experimental Results Due to the low-mass cut off of the mass spectrometers used, methylamine was replaced with larger primary amines, i.e., *n*-butylamine and *n*-hexylamine. It is assumed that the length of the hydrocarbon chain has a less pronounced effect on the amine-solvent interaction than the solvent itself. No significant differences between *n*-butylamine and *n*-hexylamine results were noticed. In the following thus the proxy $\text{R}-\text{NH}_2$ is used for both species.

The effect of MeOH and ACN vapor added to the ion source gas on the observed ion population is investigated. When the amine is electrosprayed from ACN/ H_2O solution, the cluster $[\text{R}-\text{NH}_2 + \text{H} + \text{ACN}]^+$ is abundantly present in the mass spectra. Adding ACN to the ion source gas results in detection of the second cluster $[\text{R}-\text{NH}_2 + \text{H} + 2\text{ACN}]^+$. Increasing the ACN addition promotes the intensity of the second ACN cluster but even when the ion source gas is saturated with ACN vapor, no higher amine-ACN clusters are observed. When using MeOH as gas phase modifier, the $[\text{R}-\text{NH}_2 + \text{H} + \text{MeOH}]^+$ is the only amine-MeOH cluster ion present in the mass spectrum. Thus with both solvents, amine-solvent clusters are observed, which is consistent with the theoretical findings. The maximum cluster number is $n = 2$ for ACN and $n = 1$ for MeOH. The difference in the maximum cluster number is rationalized by their thermodynamic stability. Regarding declustering as well as proton transfer processes, Figure 3.3 clearly shows that the amine-ACN clusters are more stable and have larger dissociation energies, so that the second ACN cluster is surviving the elevated ion temperatures during transfer in contrast to the second MeOH cluster.

The ion signal intensity is strongly affected by the gas phase modification. In-

creasing the MeOH gas phase mixing ratio continuously reduces the total ion current (TIC). Additionally, the sum of all analyte related intensities (bare and clustered protonated amine) normalized to the TIC, in the following referred to as normalized analyte intensity, decreases as well. In contrast, the addition of ACN enhances the TIC and the absolute analyte intensity equally; the normalized analyte intensity remains unchanged. Based on the theoretical investigations it is concluded that proton transfer in the presence of MeOH is likely to occur but proton transfer in the presence of ACN is very unlikely; the experimental results fully support this conclusion. Proton transfer to MeOH clusters can explain the observed loss of protonated analyte with MeOH addition. Presumably, due to the low-mass cut off of the mass spectrometer the protonated MeOH-clusters are not observed resulting in a reduced TIC. On the contrary, the increasing analyte signal intensities with ACN addition indicates that proton transfer to ACN clusters does not occur but rather charge depletion is reduced. However, it is noted that the changes in signal intensities are not solely due to charge retention and charge depletion processes, but in addition due to altered spraying and ionization efficiencies due to the gas phase modification. Because proton transfer from a singly protonated amine to a solvent molecule/cluster is not experimentally observed, also doubly protonated amines are studied. The results are discussed in the following section.

3.4.3 Multiply Charged Analytes

In the experiments [143], Substance P was recorded as a triply and doubly charged ion. Although the distance and the electrostatic shielding through the peptide between the different charge sites [154] is large, the effect of a spatially close charge on the proton transfer reactions was thus investigated with a corresponding proxy: Doubly protonated ethylene diamine ($[\text{H}-\text{EDA}-\text{H}]^{2+}$), considered as a rather stressed model system, was studied. The distance between the formally charged nitrogen atoms is only 3.84 Å. In addition, the two CH₂ spacer groups offer little electrostatic shielding. This model system is thus regarded as an upper limit scenario regarding effects caused by a second charge site within an ion.

Theoretical Findings Figure 3.5 shows the PES scans for the EDA–MeOH ($n = 1, 2, 3$) and –ACN ($n = 1, 2$) clusters. Potential energy due to Coulomb repulsion is acquired because proton transfer and subsequent increase of the fragment distance leads to separation of the two charges. Consequently, all reaction products are lower in energy as compared to structures where the two charges are in close vicinity. This observation explains why analytes exhibit charge retention or depletion and never charge increase in the gas phase. There is a stable adduct structure of the composition $[\text{H}-\text{EDA} + \text{H} + \text{S}_n]^{2+}$ for all clusters studied. Thus, there always has to be a TS, which connects this local minimum with the decreasing Coulomb-repulsion dominated energy path towards the products.

3 Charge Retention/Charge Depletion in ESI-MS

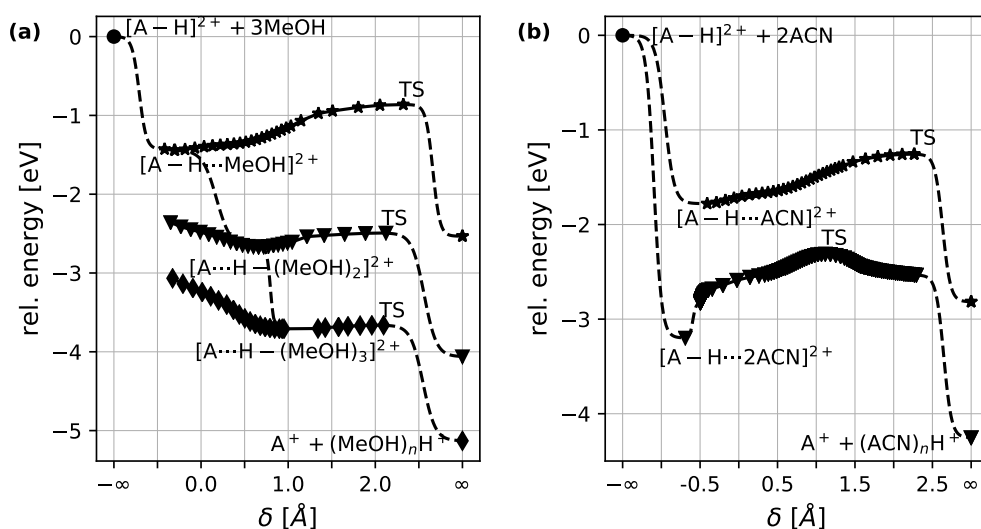


Figure 3.5: PES scans for the $[A + H + S_n]^{2+}$ ($A = \text{H-EDA}$) clusters with S being (a) MeOH or (b) ACN. Energies are obtained at the B3LYP-D/def2-TZVPP level of theory. Dashed lines connect bound and unbound configurations. The minimum structure of each cluster is marked with a label; dotted lines indicate the hydrogen bond and solid lines the shorter, covalent bond.

Focusing on the MeOH clusters (cf. Fig. 3.5a), for $n = 1$ there is an amine ion core and the energy initially rises along δ . This is analogous to the methylamine systems and is due to the higher basicity of nitrogen as compared to oxygen. Only very late in the product channel, i.e., for an O–H distance near equilibrium and at a long N–H distance, there is a transition state (0.58 eV above the minimum), indicating that the charge stabilization through the ligand (in this case nitrogen) becomes smaller than the Coulomb repulsion of the two charges. Adding another MeOH to the system reveals an interesting effect: For EDA the MeOH ion core ($\delta > 0$), i.e., a proton residing within the MeOH cluster and H–EDA⁺ acting as ligand (highlighted as dotted and solid lines in the labels of Fig. 3.5) is the only minimum with a bound configuration, whereas for methylamine the amine ion core ($\delta < 0$) is still the favorable configuration (even for $n = 4$). The separation of the charges by moving one proton along the hydrogen bond chain within the MeOH cluster is energetically favored. Note that there is no second bound-state minimum with an amine ion core. Progressing along δ , i.e., moving the H–EDA⁺ away from the MeOH cluster (the proton already is near its equilibrium position at the MeOH oxygen) the energy rises to a TS as already observed for $n = 1$. In addition to the favored configuration of the MeOH ion core the energy difference between the minimum and the TS is much lower (0.18 eV). This results in a better charge distribution because the H-bond chain promotes the proton

transfer. This effect is even more pronounced for $n = 3$, where the energy barrier for the proton transfer is < 0.05 eV.

For $n = 1$ the ACN system behaves similarly to the MeOH system with an energy barrier of 0.52 eV. However, for $n = 2$ the most stable configuration is still the amine ion core. Particularly the second minimum observed for methylamine (ACN ion core) vanishes in this case and the energy decreases monotonically from the observed TS towards the products. The small contribution of an amine binding to the $(\text{ACN})_2\text{H}^+$ cluster, as already observed through the shallow minimum in the methylamine case (Fig. 3.4), is overcome by the Coulomb repulsion of the two charges. Moreover, the energy barrier for $n = 2$ is 0.89 eV and thus even higher than for $n = 1$. This is due to the same issues regarding forming the homogeneous ACN cluster in the methylamine case. ACN lacks the ability to form solvent cluster moieties at the charge site with similar configurations as the homogeneous isolated ACN clusters, which requires large geometrical changes and a high energy increase for the proton transfer.

In summary, the second charge strongly influences the proton transfer to attached ligands. In this model, proton bound solvent structures similar to the free $(\text{MeOH})_n\text{H}^+$ clusters are readily formed at the charge site, because they offer much better charge stabilization than the doubly charged amine. For $n \geq 2$ only a MeOH ion core is observed. The small barriers leading to dissociation of the protonated ligand structure are then easily overcome. ACN struggles with the formation of the homogeneous clusters at the charge site and a high energy TS hinders the proton transfer. Although the TS is lower in energy than the cleavage channel of losing one ACN, the state density is much lower. Additionally, when the two charges do not interact as strongly due to increased electrostatic shielding, this energy difference may be reversed. Hence, proton transfer to larger ACN clusters is unlikely even in multiply-charged molecules.

Comparison with Experimental Results Larger and higher-mass analytes than EDA were experimentally studied as both singly protonated and doubly protonated species are of interest. The terminal diamines 1,2-diaminoethane, 1,3-diaminopropane, 1,5-diaminopentane, 1,8-diaminooctane and 1,9-diaminononane are investigated. Qualitatively no differences were observed in the results. Therefore, in the following, diamine ($\text{H}_2\text{N}-\text{R}-\text{NH}_2$), is used as a proxy for all analytes studied.

The effect of MeOH and ACN on the observed ion population is illustrated in Figure 3.6. Representative mass spectra are given in Figure 3.10. When the diamine is electrosprayed from an ACN/ H_2O solution and then exposed to MeOH vapor, the bare singly protonated diamine, $[\text{H}_2\text{N}-\text{R}-\text{NH}_2 + \text{H}]^+$ dominates the spectrum. Diamine-ACN clusters $[\text{H}_2\text{N}-\text{R}-\text{NH}_2 + 2\text{H} + (\text{ACN})_n]^{2+}$ are hardly observed. The bare doubly protonated diamine is observed with low abundance (cf. Fig. 3.6a). Clustering with MeOH is very rarely observed; only diaminononane and -octane yield clusters of the type $[\text{H}_2\text{N}-\text{R}-\text{NH}_2 + 2\text{H} + (\text{MeOH})_n]^{2+}$ with $n = 1 - 3$, in addition to the mixed solvent cluster $[\text{H}_2\text{N}-\text{R}-\text{NH}_2 + 2\text{H} + (\text{ACN})(\text{MeOH})]^{2+}$. These diamine-MeOH clusters are much lower in intensity than the bare $[\text{H}_2\text{N}-\text{R}-\text{NH}_2 + \text{H}]^+$ signal. The averaged

3 Charge Retention/Charge Depletion in ESI-MS

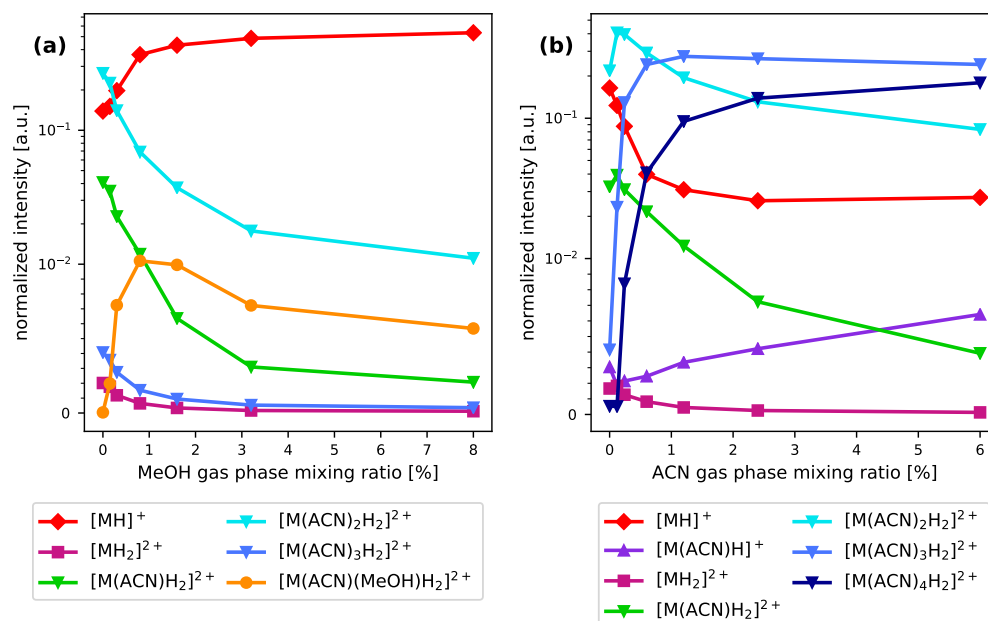


Figure 3.6: Influence of (a) MeOH and (b) ACN on the observed ion population of 1,8-diaminooctane (M), which serves as a model system for all diamine measurements. The intensities (normalized to TIC) are plotted on a linear ordinate-scale between 0 and 10⁻² and on a logarithmic ordinate-scale above.

charge state approaches 1 with MeOH addition, as shown in Fig. 3.7. Since electro-spraying from a MeOH/H₂O solution already leads to an average analyte charge state close to one, the average charge state is slightly reduced even further by MeOH vapor addition. It is concluded that the addition of MeOH changes the initial ion population towards the singly protonated bare diamine (charge depletion), which is readily rationalized by the proposed proton transfer reaction.

In the presence of ACN - regardless of being supplied via spraying an analyte/ACN solution or via addition to the gas phase, or both - the doubly protonated diamine-ACN clusters, $[H_2N-R-NH_2 + 2H + (ACN)_n]^{2+}$ with $n = 1-6$, dominate the spectrum, where the averaged cluster number is dependent on the ACN gas phase mixing ratio (cf. Fig. 3.6b) and the analyte used. The doubly protonated bare diamine is observed with lower abundance. Because the doubly protonated cluster species are generated to such an extent and wide range, the averaged charge state is shifted towards 2 with increasing ACN vapor addition (see Fig. 3.7). In contrast, the singly protonated species is solely clustered with one ACN molecule. Both singly protonated species, $[H_2N-R-NH_2 + H]^+$ and $[H_2N-R-NH_2 + H + ACN]^+$, have low abundances, especially when the ACN mixing ratio is high (> 2%). It is concluded that ACN promotes

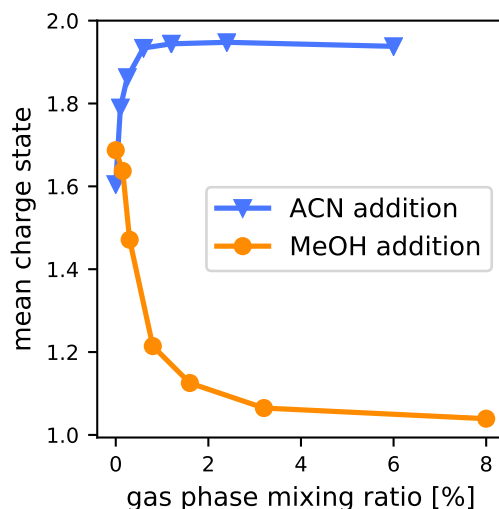


Figure 3.7: Mean charge state of 1,8-diaminooctane in dependence of the gas phase mixing ratio of ACN and MeOH.

the doubly protonated species via clustering (charge retention). The experiments provide no indication for proton transfer to ACN ligands. Therefore, the experimental observations regarding the interactions between ACN and the diamines strongly support the theoretical findings. Furthermore, the experimental results strongly suggest that the MeOH clusters $[\text{H}_2\text{N}-\text{R}-\text{NH}_2 + 2\text{H} + (\text{MeOH})_n]^{2+}$ are significantly less stable than the ACN clusters $[\text{H}_2\text{N}-\text{R}-\text{NH}_2 + 2\text{H} + (\text{ACN})_n]^{2+}$, which is in good agreement with the calculated PES scans.

3.5 Conclusion

The thermochemical stabilities of proton transfer intermediates and products of protonated methylamine reacting with either methanol or acetonitrile as modifiers, i.e., $[\text{MeNH}_2 + \text{H} + \text{S}_n]^+$ and S_nH^+ , where $\text{S} = \text{MeOH}, \text{ACN}$ were studied. The methanol clusters exhibit a trend toward better stabilization with increasing ligand number n through the hydrogen bond network they can establish, although proton transfer is thermochemically unfavorable for all cluster sizes investigated. Acetonitrile cannot form larger homogeneous clusters and thus there is no trend observed.

This rationale is reinforced by the potential energy surfaces calculated for the corresponding proton transfer reactions. Not only the energy difference between the mixed clusters but also the steepness of the potential along the reaction coordinate is lowered when larger MeOH clusters are considered. This is due to charge dilution along the hydrogen bond network in the cluster. ACN molecules clustered at the charge site need to undergo extensive geometrical changes and thus high activation

3 Charge Retention/Charge Depletion in ESI-MS

barriers to transfer a proton away from the amine. This conclusion can be extended regarding other gas phase modifiers, as studied in the preceding publication: When the modifier is aprotic, as for example diethyl ether or acetone and thus cannot form a hydrogen bond network, retention of higher charge states is observed because proton transfer is hindered. Protic modifiers, for example ethanol, iso-propanol or ammonia, respectively, can form hydrogen bond networks and thus can remove a charge from the analyte molecule.

The effect of a second charge present on the analyte was also investigated. Doubly protonated ethylene diamine was studied with methanol and acetonitrile as modifiers. For methanol it is observed that for $n \geq 2$ the proton is not stabilized at the amine but moves toward the methanol ligands. Then the small barrier towards the proton transfer reaction products (further decreasing with increasing cluster size) is easily overcome. This is consistent with the experimentally observed shift in the average charge state of diamines down to +1 upon MeOH vapor addition. ACN has a large barrier for $n = 2$ and proton transfer is not becoming favorable with increasing cluster size, as was supported by experimental observations of stable diamine-ACN clusters. The observed trends fully support the conclusion that protic solvents can remove a proton from the analyte, whereas aprotic solvents will only bind to a proton without removing it.

These findings have implications regarding the charge state distributions of small peptides or proteins in ESI-MS. Adding aprotic compounds to the gas phase of the ionization source can lead to higher charge state distributions than using protic (or no) substances - provide no other processes as for example protein unfolding or dissociation of protein complexes occur. We explicitly move away from the term "supercharging" *in this context*, normally referring to liquid phase modification. Our studies clearly demonstrate that in the gas phase no additional charge is transferred to an analyte but only existing charge states are retainable. Besides retention, only depletion of higher charge states through polar gas phase modifiers is possible, which will transfer protons into small solvent clusters. Currently, it remains unclear, whether the proton transfer occurs in a small droplet or within the gas phase cluster forming at the charge site. This issue is subject to ongoing experiments and calculations. The present results support both pathways.

Acknowledgment

The authors would like to thank Bruker Daltonics, Bremen, Germany, in particular Dr. Andreas Brekenfeld, for the continuous and considerable support regarding both technical and scientific questions. We also thank the reviewers for their extensive and very helpful comments and suggestions, which significantly improved the quality of this paper.

3.6 Supporting Information

Coulomb Repulsion Consideration Charging an already charged molecule in the gas phase via e.g., proton transfer (PT), is unlikely. For large distances, i.e., before the electron densities overlap significantly, the interaction potential between two singly charged molecules can be approximated through the Coulomb potential

$$\mathcal{E}_{\text{pot}} = \frac{1}{4\pi\epsilon_0} \frac{e^2}{R} \quad (3.6)$$

If a distance of, e.g., $R = 4 \text{ \AA}$ is assumed before the potential flattens due to interactions of the electron densities, the potential energy would be 3.6 eV. To be that close in a collision, the center of mass collision energy

$$\mathcal{E}_{\text{kin}} = \frac{1}{2} \mu v_r^2 \quad (3.7)$$

would need to be greater or equal to that value. For a heavy analyte as for example SP colliding with a light acid as H_3O^+ , the reduced mass, μ , is approximately equal to the mass of the lighter collision partner, thus, 19 amu. Thus, the relative velocity would need to be

$$v_r = \sqrt{\frac{2 \cdot 7.2 \text{ eV}}{19 \text{ amu}}} = 6050 \text{ m s}^{-1} \quad (3.8)$$

or larger. Assuming a Maxwell-Boltzmann distribution of the relative velocities as in Eq. (2.37a) at a temperature of 300 K, the portion of velocities $\geq 6050 \text{ m s}^{-1}$ can be calculated as

$$\begin{aligned} P(v_r \geq 6050 \text{ m s}^{-1}) &= \int_{6050 \text{ m s}^{-1}}^{\infty} 4\pi \left(\frac{\mu}{2\pi k_B T} \right)^{3/2} v_r^2 \exp\left(-\frac{\mu v_r^2}{2k_B T}\right) dv_r \\ &= 3.8 \times 10^{-60} \end{aligned} \quad (3.9)$$

Thus, protonating an already charged molecule in the gas phase essentially impossible.

Additional Figures More details as for example Cartesian coordinates and energies can be found in the *SI* of the original publication [144].

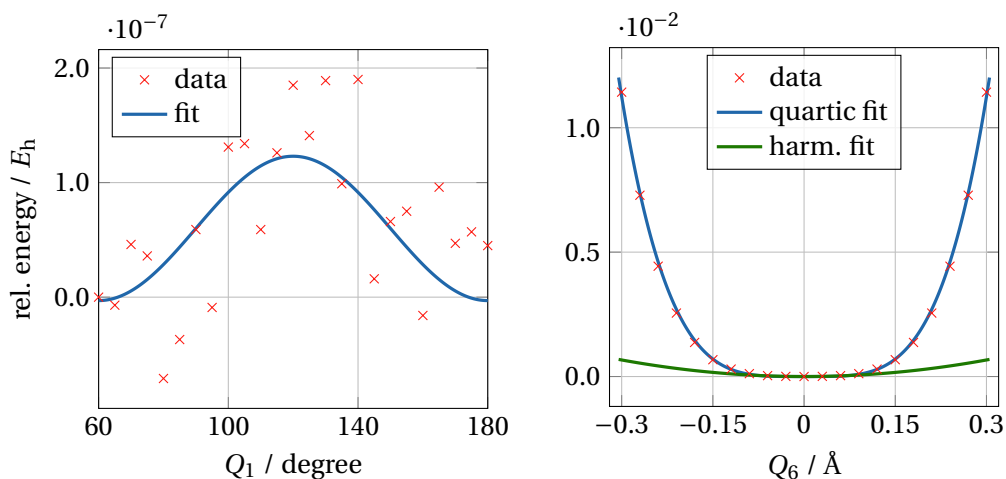


Figure 3.8: PES scans and subsequent fits for the internal rotation (Q_1) and the H^+ -movement along the $N-H^+-N$ bond (Q_6) of the $(ACN)H^+$ cluster. The strong fluctuation along Q_1 is because the accuracy of the SCF equations is of same magnitude as the barrier of the IR. The harmonic fit along Q_6 considers only the curvature at the minimum as usually done by the NMA.

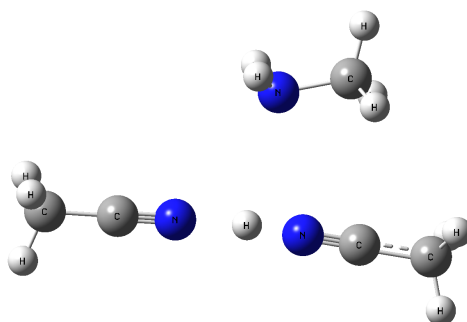


Figure 3.9: Additional structure of the $[MeNH_2 + H + 2ACN]^+$ cluster, corresponding to the second minimum in Fig. 3.4b.

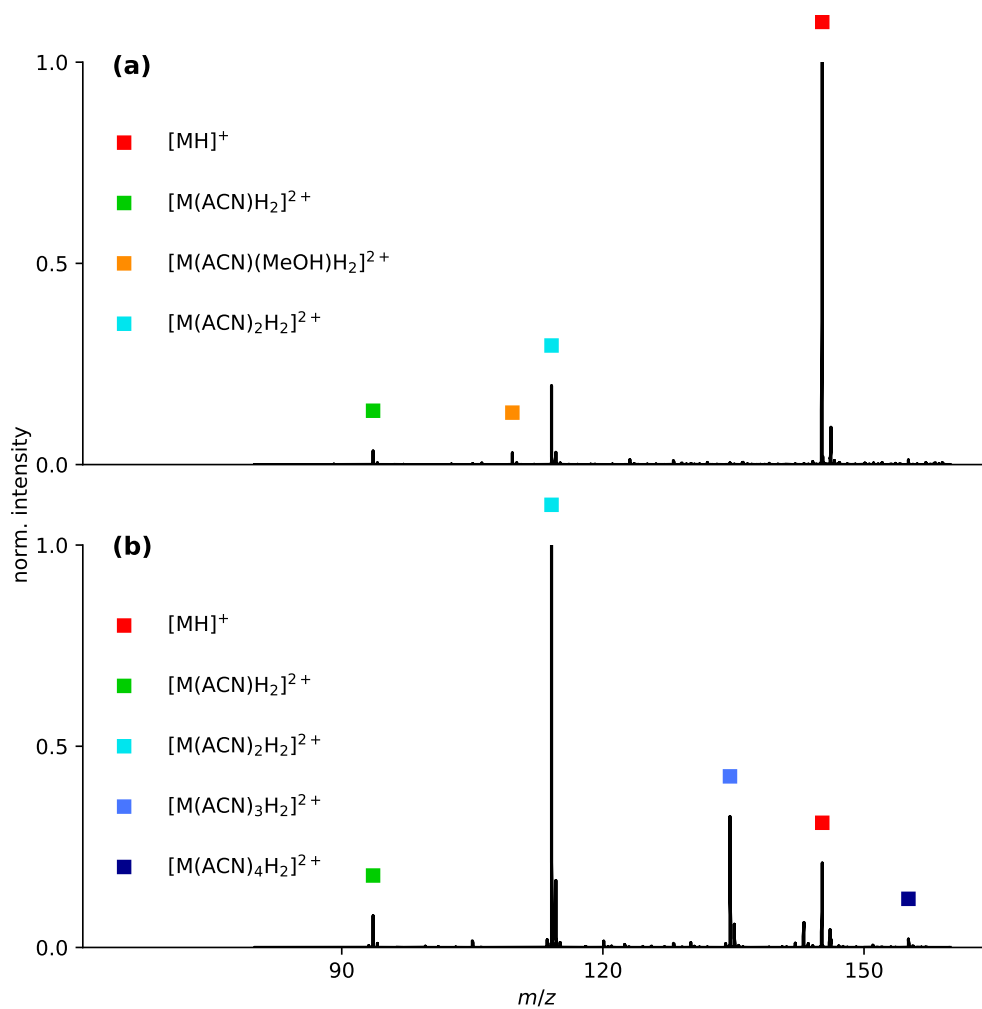


Figure 3.10: Mass spectrum of 1,8-diaminooctane (M) electrospayed from ACN/H₂O solution with (a) 0.8%V MeOH and (b) 0.25%V ACN present in the ion source. The signal intensities are normalized to the most intense signal. Addition of MeOH vapor promotes the intensity of the singly protonated amine while the addition of ACN increases the amount and intensity of doubly protonated signals.

3.7 Follow-up work

As stated in Section 3.4.1, the inherent anharmonicity of loosely bound clusters introduces errors to the thermochemical data. To tackle this issue, the capabilities of the VPT2 method (see Section 2.2.2) in this regard are investigated. Because a full VPT2 treatment on the chosen level of theory is computationally demanding, a new hybrid method is developed. This section gives a thorough description, the obtained results for the studied clusters, i.e., the $(\text{MeOH})_n\text{H}^+$ ($n = 1, 2, 3, 4$), $(\text{ACN})_n\text{H}^+$ ($n = 1, 2, 3$), $[\text{MeNH}_3 + (\text{MeOH})_n]^+$ ($n = 1, 2, 3, 4$) and $[\text{MeNH}_3 + (\text{ACN})_n]^+$ ($n = 1, 2, 3$) clusters, and an evaluation of the used method. More details regarding this follow-up work can be found in Ref. [189].

3.7.1 Description of the Method

Balancing the computational demand and the accuracy of a method is one of the most challenging tasks in Computational Chemistry. While the harmonic approximation to the PES is rather simple to compute (only one Hessian calculation at the equilibrium geometry), large errors can be introduced by neglecting anharmonic effects. In contrast, VPT2 is designed to manage anharmonicity in a perturbative manner but it needs $2(3M_N - 6) + 1$ Hessian calculations and is thus very time consuming. A general path of solving such issues is to define hybrid methods: Calculate the time consuming parts on a low level of theory but increase the accuracy when possible. This is already used in this chapter when geometry optimization and Hessian calculations for normal frequencies on B3LYP-D/def2-TZVPP are combined with a more accurate description of the electronic energy by using B2PLYP-D/def2-QZVPP.

The basic idea of this follow-up work is to calculate the anharmonic corrections to the harmonic approximation on a lower level of theory and thus save computing time. Recalling from Section 2.2.2, the vibrational frequencies and zero-point energy can be corrected by the anharmonic constants χ_0 and χ_{kl} (see eqs. (2.21) and (2.22)).

$$\nu_k = \tilde{\omega}_k + 2\chi_{kk} + \frac{1}{2} \sum_{k \neq l} \chi_{kl} \quad (3.10)$$

$$\mathcal{E}_{\text{ZPE}} = \chi_0 + \frac{1}{2} \sum_k \left(\tilde{\omega}_k + \frac{1}{2} \chi_{kk} + \frac{1}{2} \sum_{l > k} \chi_{kl} \right) \quad (3.11)$$

Similar to that, the rotational constants at the equilibrium position A_e , B_e , and C_e can be corrected using the vibro-rotational matrix $(\alpha_k^a, \alpha_k^b, \alpha_k^c)$ (to obtain the rotational constants at the vibrationally averaged geometry, A_0 , B_0 and C_0) and the centrifugal distortion tensor $\tau_{\alpha\beta\gamma\delta}$ (to account for centrifugal distortion), where k and l are indices for the normal coordinates and α , β , γ , δ are indices for the principal axes a , b

and c [103, 105]:

$$A'_0 = \underbrace{\tilde{A}_e - \frac{1}{2} \sum_k \alpha_k^a}_{A_0} + \frac{1}{4} \underbrace{(3\tau_{bcbc} - 2\tau_{caca} - 2\tau_{abab})}_{\text{centrifugal distortion term}} \quad (3.12)$$

$$B'_0 = \tilde{B}_e - \frac{1}{2} \sum_k \alpha_k^b + \frac{1}{4} (3\tau_{caca} - 2\tau_{abab} - 2\tau_{bcbc}) \quad (3.13)$$

$$C'_0 = \tilde{C}_e - \frac{1}{2} \sum_k \alpha_k^c + \frac{1}{4} (3\tau_{abab} - 2\tau_{bcbc} - 2\tau_{caca}) \quad (3.14)$$

All quantities in the above equations marked with $\tilde{}$, i.e., the harmonic parts of ν_k , A'_0 , B'_0 and C'_0 are calculated with greater accuracy and only the small correction terms are calculated with lower accuracy. Since the anharmonic constants are not very sensitive to the level of theory [190], this combination is valid.

Following simple perturbation theory [104, 119], the more accurate frequencies and rotational constants, ν_k , A'_0 , B'_0 and C'_0 , are then used for the calculation of partition functions according to the standard formula derived for the harmonic case (see Table 2.1)

$$Z_{\text{vib}} = \prod_k \left[1 - \exp\left(-\frac{h\nu_k}{k_B T}\right) \right]^{-1} \quad (3.15)$$

$$Z_{\text{rot}} = \frac{1}{\sigma_{\text{rot}}} \left(\frac{k_B T}{hc} \right)^{3/2} \left(\frac{\pi}{A'_0 B'_0 C'_0} \right)^{1/2} \quad (3.16)$$

and to calculate the thermodynamic correction to the Gibbs enthalpy, G_{corr} , the anharmonic corrected ZPE is used (Eq. (3.11)).

To actually conduct this hybrid anharmonic treatment, first, a geometry optimization and frequency calculation is conducted on the medium B3LYP-D/def2-TZVPP level of theory and the harmonic frequencies ω_k and equilibrium rotational constants A_e , B_e and C_e are saved. Using that geometry, a single-point energy calculation is performed on the high B2PLYP-D/def2-QZVPP level to refine the electronic energy $\mathcal{E}_{\text{elec}}$. After that, the geometry is reoptimized on the lower B3LYP-D/def2-SVP level of theory and a full VPT2 treatment is performed. The reoptimization is necessary, because the minimum geometry on the medium level, $\mathbf{R}_{\text{min}}^{\text{med}}$, might not be exactly the minimum on the lower level, $\mathbf{R}_{\text{min}}^{\text{low}}$. The anharmonic constants, χ_0 , χ_{kl} , $\alpha_k^{a,b,c}$ and $\tau_{\alpha\beta\gamma\delta}$, calculated from the numerical determination of the force constants (see Eq. (2.20)), are saved and combined with the harmonic values calculated on the medium level according to Eqs. (3.10)-(3.14). Thus, the most time consuming part, i.e., the calculation of $2(3M_N - 6) + 1$ Hessians, is conducted on the lowest level of theory.

For a test of this hybrid method, the calculated partition functions can be compared with those obtained from a VPT2 treatment fully performed on the medium level of theory. This comparison is shown for the vibrational partition function of some small test cases in Table 3.4, additional to the harmonic result on the medium

3 Charge Retention/Charge Depletion in ESI-MS

level. First of all, it should be noted that even for the molecules chosen, which are rather strongly bound, there is a significant difference between the harmonic and VPT2 results, highlighting the importance of anharmonic effects. The newly developed hybrid method is able to correct the harmonic partition functions with the anharmonic constants computed on the lower level of theory, so that the difference to the medium level VPT2 becomes much smaller. Additionally, the time needed for VPT2 (hybrid) compared to VPT2 (med) is a factor of 5 less for the chosen levels of theory and thus a significant speed up can be achieved.

Table 3.4: Vibrational partition function, $\ln(Z_{\text{vib}})$, calculated with different methods. *med* is B3LYP-D/def2-TZVPP, *hybrid* the method described in the text.

	MeOH	ACN	(ACN)H ⁺	MeNH ₃ ⁺
harmonic (med)	-53.5	-47.2	-58.4	-83.0
VPT2 (hybrid)	-52.7	-46.6	-57.5	-82.0
VPT2 (med)	-52.7	-46.5	-57.5	-81.5

3.7.2 Results and Discussion

The results shown in Table 3.4 are very promising and the developed hybrid method is applied to the clusters studied above in this chapter, i.e., the (MeOH)_nH⁺ ($n = 1, 2, 3, 4$), (ACN)_nH⁺ ($n = 1, 2, 3$), [MeNH₃ + (MeOH)_n]⁺ ($n = 1, 2, 3, 4$), and [MeNH₃ + (ACN)_n]⁺ ($n = 1, 2, 3$) clusters. In particular, the thermochemistry of the proton transfer reactions (Eqs. (3.2)-(3.4), Fig. 3.3) is studied.

Table 3.5 shows the calculated Gibbs energies of the association, dissociation and overall reaction, i.e., $\Delta_{\text{Ass}}G$, $\Delta_{\text{Diss}}G$ and $\Delta_{\text{Rxn}}G$ for both the original method used earlier in this chapter and the newly developed hybrid method. For $n = 1$, both methods perform comparably well and show only small deviations to experimental data. However, for increasing n the differences to the available experimental data increases for both methods. In particular, the hybrid method does not show a clear superiority over the original method, i.e., the deviations to the experimental values are sometimes larger, sometimes smaller than with the original method.

There are multiple reasons, why this could be the case. One important fact to notice is that the DFT functionals B3LYP as well as B2PLYP are parametrized to fit thermochemical data most accurately *within* the harmonic approximation. Thus, the good performance is often due to the wrong reasons: Calculate the electronic structure such that within the harmonic approximation, good thermochemical data is produced - not such that the real electronic structure is reproduced best. Introducing anharmonic corrections will thus lift this error cancellation and this better description of the PES will reveal the flaws in the electronic structure calculation. While the magnitude of this effect is difficult to estimate, it is important to notice.

Table 3.5: Proton Transfer thermochemistry of the Amine-Solvent clusters at 298.15 K. Here, *harm* describes the B3LYP-D/def2-TZVPP thermo- plus B2PLYP-D/def2-QZVPP electronic energy, whereas *hybrid* is the method described in this section, i.e., the same as *harm* but with anharmonic corrections computed on B3LYP-D/def2-SVP.

		MeOH				ACN		
	Method	$n = 1$	2	3	4	$n = 1$	2	3
$\Delta_{\text{Ass}}G$	<i>harm</i>	-54.7	-83.6	-82.3	-101.6	-79.4	-133.8	-175.8
	<i>hybrid</i>	-53.0	-77.7	-88.6	-120.2	-84.2	-139.3	-192.7
	exp ^a	-49.4				-77.8	-126.2	-156.1
$\Delta_{\text{Diss}}G$	<i>harm</i>	195.1	118.2	62.0	51.1	190.7	136.9 ^b	157.5
	<i>hybrid</i>	191.4	112.1	69.2	89.5	194.4	144.4	158.5
	exp ^c	189				190	141	157
$\Delta_{\text{Rxn}}G$	<i>harm</i>	140.4	34.6	-20.3	-50.5	111.2	3.0 ^b	-18.3
	<i>hybrid</i>	138.4	34.4	-19.4	-30.6	110.1	5.1	-34.1
	exp ^d	140.0	39.9	-15.1	-45.3	112.4	14.7	0.6

^a Ref. [184] for MeOH and Ref. [185] for ACN

^b including free rotor correction for the internal rotation

^c calculated as difference between $\Delta_{\text{Rxn}}G$ and $\Delta_{\text{Ass}}G$ (see Eqs. (3.2)-(3.4))

^d Ref. [184] for MeOH and Ref. [54, 191–193] for ACN

Nevertheless, there are inherent problems with the new hybrid method. The most important one is due to the different harmonic frequency calculations, i.e., the different Hessians, $\mathcal{H}^{\text{med}}(\mathbf{R}_{\text{min}}^{\text{med}})$ vs. $\mathcal{H}^{\text{med}}(\mathbf{R}_{\text{min}}^{\text{low}})$. Not only are the equilibrium geometries slightly different, also the Hessians may differ due to the different levels of theory. Thus, also the normal frequencies and even the normal coordinates can deviate. This leads to serious problems: The assignment of the anharmonic corrections of mode k^{low} to the "correct" mode k^{med} is not a straight forward procedure since the order of the modes (by frequency) may differ between the two levels of theory. In this work, the assignment is made according to the normal coordinates, i.e., where Q_k^{low} and Q_k^{med} have the largest overlap. However, especially with low frequency modes in cluster structures, this can be rather difficult.

There are also general problems with VPT2. While the PES is not approximated as a simple parabola anymore, it is still assumed to be a single well. However, especially if internal rotations (free or hindered) exist, the thermochemistry will be inaccurate. Moreover, the low normal frequencies can become as small as the anharmonic corrections, sometimes leading to negative fundamental frequencies. This can happen, if the numerical determination of the higher order force constants, by calculating

3 Charge Retention/Charge Depletion in ESI-MS

Hessians at displacements along the normal coordinates, fails due to the multi-well character of an internal rotation. This is problematic especially for weakly bound cluster structures.

As already mentioned earlier in this chapter (Section 3.4.1), the potential along a normal coordinate can sometimes differ largely from a parabola rendering a perturbative treatment rather difficult. For example, harmonic frequency of the proton movement inside the $(\text{ACN})_2\text{H}^+$ cluster is calculated to be 300 cm^{-1} (medium level). As can be seen in Fig. 3.8b, the potential can better be described by a quartic fit and a VTP2 treatment (also on the medium level) leads to a fundamental frequency of 5772 cm^{-1} , which is surely not realistic. Perturbative approaches perform well only if the corrections are small compared to the unperturbed value, which is not the case for this vibration.

3.7.3 Conclusion

To evaluate the newly developed hybrid method, it can be stated that a significant speed up of the VPT2 treatment with comparably results can be achieved (Table 3.4). However, this will only work well when covalently bound molecules are studied, where VPT2 can be applied well. Loosely bound clusters with strong anharmonic modes, many internal rotations and even coupling between modes are difficult to treat with VPT2 in general. Hence, also the hybrid method will fail.

Even for molecules not as difficult, errors can arise through the difference in the calculated Hessians at the slightly different equilibrium geometries leading to a difficult assignment of the normal coordinates. Two solutions are proposed: (1) The original normal coordinates could be retained and only the additional Hessian calculations at the displaced geometries (for the determination of the force constants) could be conducted at the lower level of theory. (2) An automatized procedure could be implemented for the comparison of the normal coordinates calculated at the two levels, i.e., maximizing the dot product and thus ensure maximum overlap. How well those approaches work is difficult to judge but worth trying in the future.

4 Modeling of Differential Ion Mobility

The main part of this chapter has been reprinted with permission from

A. Haack, J. Crouse, F.-J. Schlüter, T. Benter, and W. S. Hopkins, "A First Principle Model of Differential Ion Mobility: the Effect of Ion-Solvent Clustering", [Journal of The American Society for Mass Spectrometry](#) **30**, 2711–2725 (2019). Copyright 2019 American Chemical Society.

The [Follow-up work](#) section is adapted with permission from

J. Crouse, A. Haack, T. Benter, and W. S. Hopkins, "Understanding non-traditional differential mobility behavior: a case study of the tricarbostannatrane cation, $N(CH_2CH_2CH_2)_3Sn^+$ ", [Journal of the American Society for Mass Spectrometry](#), DOI: [10.1021/jasms.9b00042](#) (2020). Copyright 2020 American Chemical Society.

Preceding work can be found in [196].

4.1 Abstract

The use of Differential Mobility Spectrometry (DMS) as a separation tool prior to mass analysis has increased in popularity over the years. However, the fundamental principles behind the difference between high- and low-field mobility is still a matter of debate - especially regarding the strong impact of solvent molecules added to the gas phase in chemically-modified DMS environments. In this contribution we aim to present a thorough model for the determination of the ion mobility over a wide range of field strengths and subsequent calculation of DMS dispersion plots. Our model relies on first principle calculations only, incorporating the modeling of the "hard-sphere" mobility, the change in CCS with field strength and the degree of clustering of solvent molecules to the ion. We show that all three factors have to be taken into account to qualitatively predict dispersion plots. In particular, Type A behavior (i.e. strong clustering) in DMS can only be explained by a significant change of the mean cluster size with field strengths. The fact that our model correctly predicts trends between differently strong binding solvents, as well as the solvent concentration and the background gas temperature highlights the importance of clustering for differential mobility.

4.2 Introduction

4.2.1 Mobility at low field strength

The movement of ions through a collision gas at elevated pressures induced by an electrical field is of great interest in mass spectrometry and related areas: Not only is ion mobility important in modern atmospheric pressure ionization sources [197–199], in Ion Mobility Spectrometry (IMS) it is used to separate ions according to their size prior to mass analysis [200–202]. The mobility of ions has also been subject to theoretical investigations [132, 203–205] since it offers insights into fundamental physics of collisions and ion-neutral interactions, e.g., kinetic theory in the context of the Boltzmann transport equation [206].

The acceleration of the ion ensemble through the electrical field is countered by collisions with the background gas, leading to a constant drift velocity v_D proportional to the applied field strength, E :

$$v_D = K(E) \cdot E \quad (4.1)$$

The proportionality constant K is called the ion mobility and, for low field strengths¹ (as applied in IMS), can be regarded as constant. In this zero-field limit, the mobility can accurately be described by the Mason-Schamp equation:[132]

$$K(0) = \frac{3}{16} \left(\frac{2\pi}{\mu k_B T} \right)^{1/2} \frac{ze}{N\Omega(T)} \quad (4.2)$$

Here, μ is the reduced mass of the ion-neutral pair, k_B is the Boltzmann constant, T the gas temperature, ze the ions charge, N is the neutral particle density and Ω is the collision cross section (CCS) of the ion in the particular collision gas. This equation is used, for example, to determine the CCS from mobility measurements, opening the opportunity to identify compounds when compared to theoretical determination of the CCS (see for example [207–210]). The latter can be performed by numerically solving the following equation:[136, 137]

$$\Omega^{(1,1)} = \frac{1}{8\pi^2} \int_0^{2\pi} d\alpha \int_0^\pi \sin\beta d\beta \int_0^{2\pi} d\gamma \quad (4.3)$$

$$\frac{\pi}{8} \left(\frac{\mu}{k_B T} \right)^3 \int_0^\infty v_r^5 \exp\left(-\frac{\mu v_r^2}{2k_B T}\right) dv_r \int_0^\infty 2b [1 - \cos\theta(\alpha, \beta, \gamma, v_r, b)] db$$

¹Since the collision frequency of the ions with the background gas depends on the applied field strength E and the particle density N , changing both parameters by the same factor does not change the amount of energy and momentum gained by the ions. For that reason, the reduced field strength $E_r = E/N$, given in Townsend, Td ($1 \text{ Td} = 10^{-21} \text{ Vm}^2$) is typically used to describe the experimental conditions. This way, the categorization of "low" and "high" (reduced) field strengths is independent of pressure (and temperature) and actually reflects the mean collision energy experienced by the ions [135].

Here, the angles α , β and γ describe the orientation of the impact between ion and collision gas, b is the impact parameter, v_r the relative velocity and θ the resulting scattering angle, which depends on all these parameters and the interaction potential between the collision partners. We note that this equation provides only a first order approximation to the collision cross section (hence the (1,1) superscript), but it is commonly used and yields sufficiently accurate results [137].

4.2.2 Mobility at high field strength

If the field strength is further increased, the gas phase ion mobility starts to change [132]. From a theoretical standpoint, the change of the mobility becomes significant when the velocity distribution of the ion ensemble starts to differ from the velocity distribution of the collision gas, i.e., when the acceleration through the electrical field significantly increases the collision frequency over the thermal collision rate [135]. One popular approach to describe the two different velocity distributions is the so-called two-temperature theory [211–213] in which the velocity distribution of the ions (in the direction of the electric field) is described by a different temperature than the distribution of the collision gas. While still assumed to be a 1-D Maxwellian distribution, the higher temperature, T_{ion} , broadens the distribution, allowing for higher ion velocities. The collisions between ions and neutrals are then determined by the relative velocity distribution, which depends on the reduced mass of the ion-neutral pair and an effective temperature, T_{eff} , which can be expressed as:[211–213]

$$T_{\text{eff}} = T_{\text{bath}} + \frac{m_{\text{bath}} v_D^2}{3k_B} \quad (4.4)$$

Here, m_{bath} is the mass of the collision gas. Although this formula is used for any type of potential with good accuracy [132, p. 154], it is only exact when a repulsive R^{-4} potential between ion and neutral is assumed [214].

Applying this effective temperature instead of the background gas temperature to the Mason-Schamp equation (Eq. (4.2)) greatly improves the accuracy of the equation at higher field strengths [214]. It can also be applied to the theoretical description of the CCS by calculating $\Omega^{(1,1)}$ at T_{eff} instead of T . With that, the Boltzmann weighting of the relative velocities v_r (exponential in fourth integral of Eq. (4.3)) then accounts for the correct shape of the distribution by applying not only the reduced mass but also the effective temperature. If the ions are polyatomic, the high impact collisions will also deposit energy into the internal degrees of freedom of the molecule. While there is some theoretical work on the influence of inelastic collision on the mobility [132, chapter 6-6], their contributions are often ignored for simplicity reasons.

4.2.3 Differential Mobility Spectrometry

The electric field dependency of the mobility is exploited in Differential Mobility Spectrometry (DMS) [215–219], also known as Field-Asymmetric Waveform Ion Mobil-

4 Modeling of Differential Ion Mobility

ity Spectrometry (FAIMS). Briefly, the ions are introduced to a small gap between two parallel plate electrodes at atmospheric pressure. On the plates, an asymmetric waveform is applied to produce an oscillating electric field orthogonal to the gas flow stream. The so-called "separation field" (peak-to-peak voltage in the order of 1-4 kV), creates a high-field phase in one and a low-field phase in the opposite direction such that the integral over one waveform cycle is zero [220]. Ideally, this would be a rectangular waveform, where the low field portion has $1/n$ -th the amplitude and lasts n times as long. However, in practice other, more practical waveforms are applied, e.g., a two-harmonics or double sine waveform (see below). Because of the field dependency of the ion mobility, the overall displacement towards the electrodes is not canceled and ions are dragged towards one of the plates, eventually being neutralized and precluding their detection. This drift is compensated by a constant voltage (called compensation voltage; 0-100 V in magnitude) applied on the electrodes. By scanning this compensation voltage, different ions are able to pass through the DMS cell. Thus, while in traditional IMS ions are mainly separated by their CCS, in DMS the difference in high- and low-field mobility is the characteristic which separates the ions.

Adding solvent vapor in low concentrations to the gas phase often increases the differential mobility effect and thus the separation capability [217]. Three main trends are observed in experiments [217, 221]: the ions mobility appear larger in high- than in low-field (Type A), larger in the low-field than in the high-field (Type C) and an intermediate case, where the high-field mobility is larger at first, but eventually the low-field mobility exceeds that of the high-field condition (Type B). These types can be easily identified by varying the separation field amplitude and recording the compensation voltage required for optimal transmission, yielding a so-called dispersion plot. It should be noted that these types are not intrinsically different. For example, any Type A case would become a Type B case for high enough field strengths since the low-field mobility eventually overcomes the high-field one (see below). Also, Type C can be seen as limiting case of Type B, where the initial excess of the high-field mobility over the low-field mobility vanishes.

Besides the theoretical improvements with regard to the mobility at higher field strengths mentioned above, no thorough model yet exists to explain differential mobility under various experimental conditions. The effect of solvent vapor in particular is still a matter of debate; while its effect through the change of the interaction potential and subsequent change in CCS has been discussed [50], other literature reports suggest a dynamic clustering/declustering mechanism [49–51]. In the high-field, where the ions internal energy increases due to collisional heating, firmly bound solvent molecules will boil off, decreasing the overall average CCS and thus increasing the ensemble mobility (as observed for Type A ions). Even in IMS it has been shown that (dynamic) clustering has to be considered when modeling the reactant ion peak (RIP), e.g., the proton bound water cluster system $[H + (H_2O)_n]^+$ [48].

In this contribution we present a thorough model for the calculation of an ion's mobility under different conditions and over a wide range of field strengths, incorporat-

ing two-temperature theory, the temperature dependence of the CCS, and the effect of ion-solvent clustering. To test our model, we simulate dispersion plots of tetramethyl ammonium (Me_4N^+) in a pure nitrogen atmosphere and seeded with methanol (MeOH), acetonitrile (ACN) and acetone (ACE). Experimental DMS data for those conditions are available in the literature [68] and show a range of different behavior, i.e., from Type C, very weak Type B (no modifier added), strong Type B (MeOH added) and Type A (ACN, ACE) behavior. We reproduce these data by applying the same formalism for each system studied and having the results solely determined by the nature of the system (i.e. not through fitting to experimental data or introduction of system specific parameters). Our model strongly suggests the importance of clustering/declustering reactions for differential mobility and yields the first thorough method for the estimation of trends in of dispersion plots from first principles.

4.3 Computational Methods

4.3.1 Ab initio calculations

To explore the large configurational space accessible to larger cluster sizes (different binding sites/configurations of the solvents to the ion), we performed basin hopping (BH) simulations to find candidate structures for further optimization. Details of the basin hopping algorithm can be found elsewhere [222, 223]. Briefly, to generate random input structures, different moieties (ion and n solvents) were each rotated by a random angle of $-10^\circ \leq \phi \leq 10^\circ$ around their body-fixed x , y or z axis and randomly translated along the overall x , y or z axis by $-0.5 \text{ \AA} \leq l \leq 0.5 \text{ \AA}$. If dihedral angles were suspected to influence the potential energy (e.g., methyl group rotation in MeOH compared to a fixed ion-HO framework), they were also randomly altered with an angle $-5^\circ \leq \varphi \leq 5^\circ$. These random structures were optimized using the AMBER force field [113] accompanied by partial charges calculated with the Merz-Singh-Kollman scheme [224, 225] for a "guess structure" optimized at the B3LYP-GD3/6-31++G(d,p) [83, 86, 87, 91] level of theory. This procedure yielded up to 20000 structures per cluster. These structures were subsequently combined into energy bins of a certain size and the lowest energy structure in each bin were used for geometry comparison. Describing the geometry of a molecule by a vector of mass-weighted distances to the center of mass, two geometries can be compared by cosine similarity, i.e., measuring the angle between these two vectors. If the angle was smaller than a predefined threshold, the structures were categorized as being identical. The BH algorithm would typically find 10 to 50 unique candidate structures.

Each of the unique candidate structures was then re-optimized at the B3LYP-GD3 / 6-31++G(d,p) level of theory, as suggested for the MobCal-MPI code [141] (see below). Many candidate structures identified by molecular mechanics yielded the same structure following DFT treatment. DFT optimizations typically produced 1-4 structures for which ESP charges [224, 225] and frequency calculations were also performed to obtain input for partition function calculations (see below). Additional

4 Modeling of Differential Ion Mobility

single-point energy calculations were performed at the B2PLYP-GD3BJ/def2-TZVPP [75, 88, 92] level of theory for a better description of the electronic energy. Each of those structures is defined by the number of solvent molecules $n = 0, 1, 2, \dots$, and an index to name different conformers $\alpha = a, b, c, \dots$, e.g., $[\text{Me}_4\text{N} + (\text{ACN})_2]_b^+$.

For some test cases, molecular dynamic (MD) simulations were performed to test the influence of vibrational broadening (due to elevated ion temperatures) on the CCS. Because simple force fields neglect the anharmonic nature of loosely bound clusters, we conducted ab initio MD simulations with the Atom Centered Density Matrix Propagation (ADMP) [115–117] model using, again, the B3LYP-GD3/6-31++G(d,p) level of theory. This version of a Car-Parinello [114] density matrix propagation is much faster than to actually solve the SCF equations at each time step, allowing us to perform 5000×0.2 fs trajectories on a full quantum mechanical (QM) potential energy surface (PES). To model the temperature effect, a kinetic energy of $(3M_{\text{N}} - 6)k_{\text{B}}T_{\text{ion}}$, with M_{N} being the number of atoms, k_{B} the Boltzmann constant and T_{ion} the ions temperature, was applied, randomly distributed over the nuclei. This yields a mean kinetic energy of $1/2k_{\text{B}}T_{\text{ion}}$ per vibrational degree of freedom, as should be the case for a thermalized multidimensional harmonic oscillator.

All ab initio and molecular mechanics calculations were conducted with the GAUSSIAN 16 program package [226].

4.3.2 CCS calculations

Collision cross section (CCS) calculations (Eq. (4.3)) were performed with the recently developed MobCal-MPI code [141] using the trajectory method [136, 137], which allows for fast calculations of the CCS through parallelization of the different trajectories. We slightly modified the code to be able to adjust the temperature which is used for the Boltzmann weighting of the relative velocities (see Eq. (4.3)) and set it to the effective temperature studied (Eq. (4.4)). All calculations were performed in molecular nitrogen and using 10 cycles, 48 velocity parameter integrations and 512 impact parameter integrations as suggested by the authors [141]. We used either the equilibrium geometry or 100 random samples from the MD simulations (conducted at the corresponding temperature), all accompanied by the determined ESP charges and MMFF94 van der Waals parameters [227, 228], to calculate the CCS over a range of effective temperatures. These datapoints were subsequently fitted to a simple function of the form $a T_{\text{eff}}^b + c$. We have no complete physical meaning to this fitting function. It is stated in the literature [50, 132], that the CCS is proportional to $T_{\text{eff}}^{-1/2}$ for low temperatures / field strengths and an ion/neutral potential proportional to R^{-4} . However, the corresponding fitting function $a T_{\text{eff}}^{-0.5} + c$ showed systematic deviation from the data, probably due to elevated (effective) temperatures. Thus, we chose to also adjust the exponent for a more accurate interpolation. In terms of physical meaning, the offset c can be seen as CCS in the infinite temperature limit, when all long-range interactions become negligible and only the hard-sphere core is contributing to the CCS [229].

4.3.3 Mobility calculations

The individual drift velocity v_D of each structure (n, α) considered is calculated with Eq. (4.1) using the Mason-Schamp expression for the mobility, K (Eq. (4.2)), the applied field strength, E and two-temperature theory, i.e., setting $T \rightarrow T_{\text{eff}}$ (Eq. (4.4)). Since the effective temperature is not known a priori, we start by using the background gas temperature, T_{bath} , use the obtained drift velocity to calculate the effective temperature with Eq. (4.4) and again, calculate the drift velocity with Eq. (4.1) and (4.2). This is done iteratively until the effective temperature obtained converges to a predefined threshold of 10^{-4} K. In this self-consistent calculation, $\Omega(T)$ is also updated at every iteration step to the effective temperature applied using the analytical fitting function determined beforehand. Having an analytical function for $\Omega(T)$ dramatically speeds up the calculations since otherwise the CCS would have to be computed at every step of this self-consistent calculation explicitly using MobCal-MPI. As a consequence of the different CCS values (viz. their fitting functions), each cluster has a different effective temperature and a different drift velocity.

4.3.4 Boltzmann weighting

For simplicity, we assume complete equilibration of the system at any given field strength. A detailed discussion of this assumption is given below. To obtain the population distribution of a thermalized system with different local minima, e.g. different conformers, the quantum-harmonic superposition approximation (QHSA) is typically used [121–123]. In this formalism, the density of states in each minimum is assumed to be independent from each other and thus, the relative population P_i of minimum i can be calculated from the vibrational partition function $Z_{\text{vib},i}$ and the zero-point corrected electronic energy $\mathcal{E}_{0,i}$, both readily available from ab initio calculations:

$$P_i = \frac{\mathcal{Z}_i}{\sum_j \mathcal{Z}_j} \quad \text{with} \quad \mathcal{Z}_i = Z_{\text{vib},i} \exp\left(-\frac{\mathcal{E}_{0,i} - \mathcal{E}_{0,\text{ref}}}{k_B T}\right) \quad (4.5)$$

where the sum in the denominator runs over all minima considered.

We have to modify this description in two ways. First, we are considering dissociation equilibria (viz. solvent molecules evaporating from the ion) and thus the total partition function has to be used (including rotational and translational contributions). Also, the concentration of the solvent $[S]$ has to be considered, because it determines the collision frequency between ions and solvent molecules which, in turn, determines the equilibrium constant. In particular, if the solvent concentration is zero, the population of any cluster has to be zero as well, independent from any energy differences. This leads to

$$\mathcal{Z}_i = \left(\frac{[S]}{N}\right)^{n_i} Z_{\text{tot},i} Z_{\text{tot},S}^{\bar{n}-n_i} \exp\left(-\frac{\mathcal{E}_{0,i} - \mathcal{E}_{0,\text{ref}}}{k_B T}\right) \quad (4.6)$$

where N is the total particle density and n_i the number of solvent molecules firmly attached to the ion in structure i . The partition function for the system is thus a

4 Modeling of Differential Ion Mobility

product of the (total) partition function of the ionic cluster i and the (total) partition function of each unbound solvent, with \tilde{n} being the largest cluster size considered. For example, if $\tilde{n} = 1$, the partition function of the bound cluster is just $Z_{\text{tot},i}$ whereas the partition function of the dissociated system is $Z_{\text{tot},i-1} Z_{\text{tot},S}$. This ensures that the partition functions are being calculated in the same configurational space, treating the bound and dissociated states as two "local minima" on the same PES. The contributions of the additional translational degrees of freedom in Z_S under dissociation is describing an increase in entropy when forming two particles out of one. A similar, although not as thorough, formalism was already used in the literature for the determination of field dependent mobilities under consideration of clustering [50].

The second modification is necessary due to the different ion temperatures of each structure considered, assuming that the internal degrees of freedom are also heated and thus also have a temperature of T_{ion} . Initially, we want to deduce the appropriate formula for a simple system containing two local minima, A and B, as shown in Figure 4.1, and then move on to the generalization of multiple minima. For simplicity, we drop the index on the energy ($\mathcal{E}_0 \rightarrow \mathcal{E}$). The total partition function of the system can then be written as an integral over the total density of states $\rho(\mathcal{E})$ and the Boltzmann weighting function $f(\mathcal{E})$:

$$\mathcal{Z} = \int_0^{\infty} \rho(\mathcal{E}) f(\mathcal{E}) d\mathcal{E} \quad (4.7)$$

The normal superposition approximation now states:

$$\mathcal{Z} = \int_{\mathcal{E}_A=0}^{\infty} \rho_A(\mathcal{E}) f(\mathcal{E}) d\mathcal{E} + \int_{\mathcal{E}_B}^{\infty} \rho_B(\mathcal{E}) f(\mathcal{E}) d\mathcal{E} = \mathcal{Z}_A + \mathcal{Z}_B \quad (4.8)$$

While the first term, \mathcal{Z}_A , is just the partition function Z_A of minimum A, the second term can be written in a different way:

$$\begin{aligned} \int_{\mathcal{E}_B}^{\infty} \rho_B(\mathcal{E}) \exp\left(-\frac{\mathcal{E}}{k_B T}\right) d\mathcal{E} &= \int_{\mathcal{E}_B}^{\infty} \rho_B(\mathcal{E}) \exp\left(-\frac{\mathcal{E} - \mathcal{E}_B}{k_B T}\right) \exp\left(-\frac{\mathcal{E}_B}{k_B T}\right) d\mathcal{E} \\ &= \exp\left(-\frac{\mathcal{E}_B}{k_B T}\right) \int_{\mathcal{E}_B}^{\infty} \rho_B(\mathcal{E}) \exp\left(-\frac{\mathcal{E} - \mathcal{E}_B}{k_B T}\right) d\mathcal{E} \end{aligned} \quad (4.9)$$

If we now set $\mathcal{E} - \mathcal{E}_B = \tilde{\mathcal{E}}$ and recognize that $\rho_B(\mathcal{E}) = 0$ for $\mathcal{E} < \mathcal{E}_B$, the integral becomes the partition function of B:

$$\exp\left(-\frac{\mathcal{E}_B}{k_B T}\right) \int_0^{\infty} \rho_B(\tilde{\mathcal{E}}) \exp\left(-\frac{\tilde{\mathcal{E}}}{k_B T}\right) d\tilde{\mathcal{E}} = Z_B \exp\left(-\frac{\mathcal{E}_B}{k_B T}\right) \quad (4.10)$$

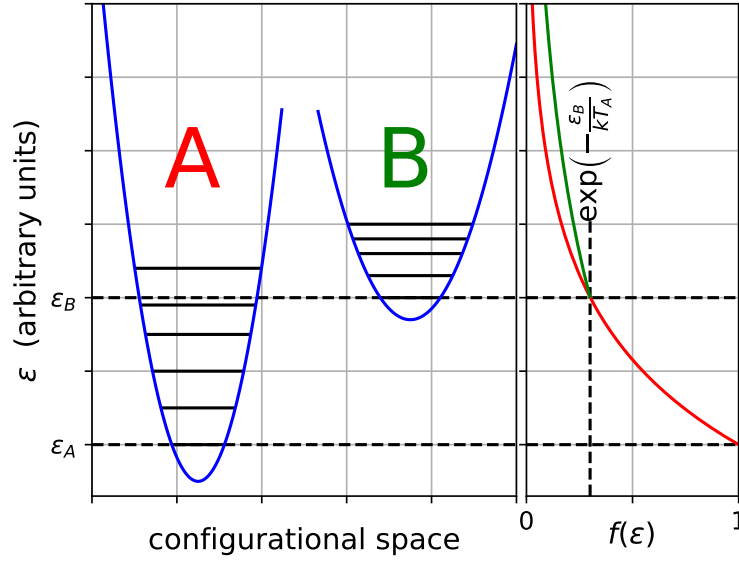


Figure 4.1: Quantum Harmonic Superposition Approximation (QHSA). The population density in minimum A is according to $\exp(-\mathcal{E}/k_B T_A)$ (red) and minimum B is $\exp(-(\mathcal{E} - \mathcal{E}_B)/k_B T_B)$ (green) scaled by the constant value of the exponential of minimum A at the energy of minimum B, i.e., $\exp(-\mathcal{E}_B/k_B T_A)$. See text for details.

In total this yields:

$$\mathcal{Z} = \mathcal{Z}_A + \mathcal{Z}_B = Z_A \underbrace{\exp\left(-\frac{\mathcal{E}_A}{k_B T}\right)}_{=1, \text{ since } \mathcal{E}_A=0} + Z_B \exp\left(-\frac{\mathcal{E}_B}{k_B T}\right) \quad (4.11)$$

which is the normal superposition approximation described in Eq. (4.5).

If we consider Figure 4.1, it can be seen that the constant $\exp(-\mathcal{E}_B/k_B T)$ is a scaling factor for the population density function for minimum B with respect to minimum A because the population density of minimum A declines to this value at that particular energy. For the case of different temperatures in minima A and B, their Boltzmann functions will become $f_A(\mathcal{E}) = \exp(-\mathcal{E}/k_B T_A)$ and $f_B(\mathcal{E}) = \exp(-(\mathcal{E} - \mathcal{E}_B)/k_B T_B)$. However, one must still scale the population distribution in minimum B with the exponential of minimum A since that describes the overall population density at the given energy. Hence, we set

$$\begin{aligned} \mathcal{Z} &= \int_0^\infty \rho_A(\mathcal{E}) f_A(\mathcal{E}) d\mathcal{E} + \exp\left(-\frac{\mathcal{E}_B}{k_B T_A}\right) \int_{\mathcal{E}_B}^\infty \rho_B(\mathcal{E}) f_B(\mathcal{E}) d\mathcal{E} \\ &= Z_A(T_A) + Z_B(T_B) \exp\left(-\frac{\mathcal{E}_B}{k_B T_A}\right) \end{aligned} \quad (4.12)$$

4 Modeling of Differential Ion Mobility

Thus, the partition function of each minimum is calculated at their respective temperature, however, the scaling factor of the partition function in minimum B is determined by the exponential decay of the population density in the next lower energy minimum. For any system with multiple, energy sorted minima (i.e. $\mathcal{E}_{0,i+1} > \mathcal{E}_{0,i}$ for all i), the scaling factors of all lower energy minima have to be taken into account. With the lowest energy set to zero ($\mathcal{E}_{0,1} = 0$), it follows that:

$$\mathcal{Z}_1 = Z_1(T_1) \quad (4.13a)$$

$$\begin{aligned} \mathcal{Z}_i &= Z_i(T_i) \prod_{j=2}^i \exp\left(-\frac{\mathcal{E}_{0,j} - \mathcal{E}_{0,j-1}}{k_B T_{j-1}}\right) \\ &= Z_i(T_i) \exp\left(-\sum_{j=2}^i \frac{\mathcal{E}_{0,j} - \mathcal{E}_{0,j-1}}{k_B T_{j-1}}\right) \quad \text{for } i \geq 2 \end{aligned} \quad (4.13b)$$

It is worth mentioning that this weighting reduces back to the standard superposition approximation if all T_i are equal.

Assuming that the dissociated solvent molecules are not heated by the electric field, their partition function is calculated at the background gas temperature. Hence, we set $Z_i Z_S \rightarrow Z_i(T_i) Z_S(T_{\text{bath}})$. Including the concentration dependency, this leads to

$$\mathcal{Z}_1 = \left(\frac{[S]}{N}\right)^{n_1} Z_1(T_1) Z_S(T)^{\bar{n}-n_1} \quad (4.14a)$$

$$\mathcal{Z}_i = \left(\frac{[S]}{N}\right)^{n_i} Z_i(T_i) Z_S(T)^{\bar{n}-n_i} \exp\left(-\sum_{j=2}^i \frac{\mathcal{E}_{0,j} - \mathcal{E}_{0,j-1}}{k_B T_{j-1}}\right) \quad \text{for } i \geq 2 \quad (4.14b)$$

We calculate all partition functions using standard formula (see Table 2.1) since they have a relatively simple functional dependence of the temperature and all other input data (e.g., mass, rotational constants, normal frequencies; read in from the ab initio output) have to be determined only once, independent from the temperature.

For the determination of the ensemble mobility, we just weight the individual mobilities by the now accessible relative populations:

$$\langle K \rangle_{\text{ens}} = \sum_i K_i P_i \quad (4.15)$$

We want to point out that this approach assumes fast thermal equilibrium compared to any change in the field strength. This means that we assume an ionic cluster, as soon as it adopts a new structure through dissociation or change of conformation, equilibrates to the new effective temperature determined by its CCS instantly. For fast varying fields (compared to collision frequencies), this may not hold true and kinetic simulations with RRKM theory determined rate constants would be needed to accurately determine the relative population of each cluster structure. This would also circumvent the rather strange assumption of thermal equilibrium between systems having different temperatures, which clearly is in contrast to the zeroth law of

thermodynamics. One would then calculate the rate constants for association $A^+ + S \longrightarrow [A + S]^+$ with the collision frequency applying T_{eff} of A^+ , since this temperature describes the relative velocity distribution between ion and neutral. The rate constant for dissociation $[A + S]_a^+ \longrightarrow A^+ + S$ or rearrangement $[A + S]_a^+ \longrightarrow [A + S]_b^+$ would be calculated applying T_{ion} of $[A + S]_a^+$, since this temperature determines the internal energy and thus state density at the dissociation or transition state energy. This approach will be subject to future work.

4.3.5 Dispersion Plot calculation

With the described formalism the ensemble mobility of a cluster-solvent system (Eq. (4.15)) can be calculated for any given field strength E , which is introduced in the Mason-Schamp equation for the self-consistent determination of the drift velocity and effective temperature. By varying the field strength E , the field dependency of the (ensemble) mobility can be calculated, typically described by the α function, defined as

$$\alpha(E) = \frac{\langle K(E) \rangle_{\text{ens}}}{\langle K(0) \rangle_{\text{ens}}} - 1 \quad (4.16)$$

Just to repeat, at any given field strength, the CCS, the mobility, the partition function and the Boltzmann weight for all structures are re-calculated at the effective temperature determined at this field strength. In case of the zero-field mobility $\langle K(0) \rangle_{\text{ens}}$, this reduces to the background gas temperature.

The α function and its derivative with respect to the field strength, $\alpha' = d\alpha/dE$, can now be used to calculate the compensation voltage (CV) for a known separation field. Given the peak-to-peak separation voltage SV_{pp} and using a double sine waveform with maximum amplitude $D = 2/3 SV_{pp}$ [218]:

$$E(t) = \frac{D}{d} \left(\frac{2}{3} \sin(\omega t) + \frac{1}{3} \sin\left(2\omega t - \frac{\pi}{2}\right) \right) \quad (4.17)$$

in which d is the gap size in the DMS cell, the CV can be calculated according to [50, 215]:

$$CV = - \frac{\langle \alpha(E(t)) \cdot E(t) \rangle_{\text{wf}}}{1 + \langle \alpha(E(t)) \rangle_{\text{wf}} + \langle \alpha'(E(t)) \cdot E(t) \rangle_{\text{wf}}} \cdot d \quad (4.18)$$

In this equation, the averages $\langle \dots \rangle_{\text{wf}}$ are taken over one cycle of the waveform. Doing this for multiple values of SV_{pp} , a full dispersion plot can be calculated from first principles. We want to point out that through our assumption of fast equilibration and thus "equilibrium" α functions, the actual waveform frequency, ω , does not appear in our calculations. Our results thus correspond to the limiting case where ω is small compared to all chemical processes. Although the calculated α curves could be compared to experimentally determined α functions (available from experimental dispersion plots through a least square fit procedure also relying on Eq. (4.18) [215, 217]), we wanted to provide a method to produce dispersion plots since they are more readily available from experiment and thus easier for comparison.

4.4 Results and Discussion

We applied the above formalism to tetramethyl ammonium (Me_4N^+) in gaseous molecular nitrogen seeded with either methanol (MeOH), acetonitrile (ACN) or acetone (ACE) vapor. These systems are a good test of theory, because experimental DMS data is readily available [68], the measured dispersion plots show very different behavior (Type A and Type B, depending on the solvent-modified environment), and the system size is relatively small, allowing for faster computation.

4.4.1 Cluster structures

Figure 4.2 shows all structures n, α for the $[\text{Me}_4\text{N} + (\text{ACN})_n]^+$ ion-solvent cluster systems found by the BH search outlined above. Structures for the ACE and MeOH systems and energies can be found in Fig. 4.8. For $n = 1$ only one unique conformer was found, i.e., where the solvent binds to the triangular face of the tetramethyl ammonium ion. Multiple structures were found for $n = 2$ where the two solvents either bind to two different faces of the tetrahedron ($\alpha = a$ structures) or interact with each other. For example, MeOH shows a structure where the two solvents are forming a hydrogen bond chain ($[\text{Me}_4\text{N} + (\text{MeOH})_2]^+$). Since we found $[\text{Me}_4\text{N} + (\text{ACN})_2]^+$ to be highly populated for low field strengths (see below, especially Fig. 4.5), we also added the lowest-energy conformer $n = 3$ cluster to the formalism. Hence, for all solvents, we ensured that the highest cluster numbers considered are only slightly populated at low field strengths.

Taking a look at the calculated zero-point energy corrected binding energies $\Delta\mathcal{E}_0$, it becomes clear that MeOH is a rather weak binding solvent ($\Delta\mathcal{E}_0 = 0.45$ eV), whereas ACN and ACE interactions with Me_4N^+ are much stronger ($\Delta\mathcal{E}_0 = 0.63$ eV and $\Delta\mathcal{E}_0 = 0.64$ eV, respectively). This drives the cluster stability and the field dependency of the cluster size distribution, ultimately leading to different α functions and dispersion plots (see below).

The cluster geometries reveal an inherent problem with the described formalism; since loosely bound, the cluster structures are not very rigid and describing them as isolated states whose partition functions are calculated in the rigid-rotor / harmonic-oscillator approximation is a source of error. For example, internal rotations may have small barriers and should thus be treated as a free rotation rather than a vibration. While this could be done in principle – although it would take some work – other problems are more difficult to deal with. The inherent anharmonicity of the dissociation coordinates, the number of low frequency modes, and the (presumably) low barrier transition between two cluster conformers of the same size all introduce error within the harmonic-oscillator approximation to the PES and the superposition approximation. We chose our approach for the sake of having closed formula for the partition function in which the (ion) temperature can be applied in a straight forward way and we expect that this should be qualitatively correct for relatively rigid systems. However, our approach is not expected to perform well for fluxional, highly

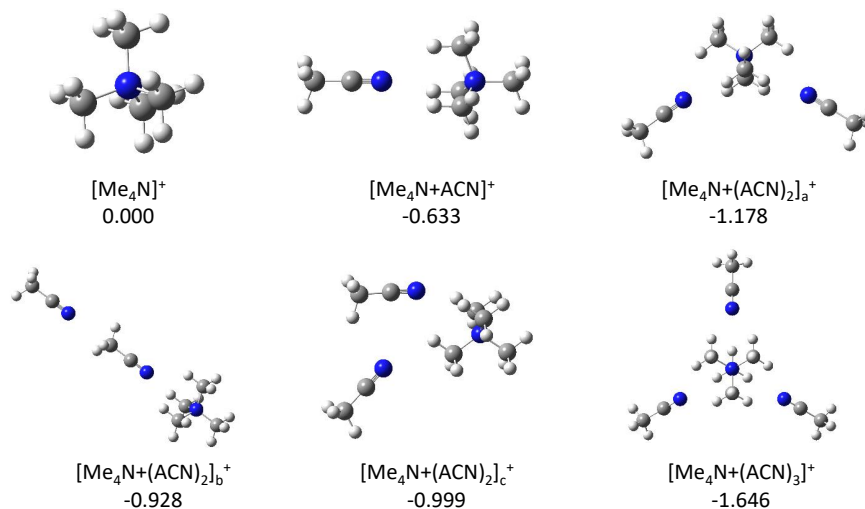


Figure 4.2: Geometries of all $[\text{Me}_4\text{N} + (\text{ACN})_n]^+$ clusters optimized at the B3LYP-GD3/6-31++G(d,p) level of theory. Zero-Point corrected binding energies, $\Delta\mathcal{E}_0$, are given in eV relative to the fully dissociated state. Solvents prefer interacting via the triangular faces of the tetramethyl ammonium ion.

anharmonic species.

We should also note that there is always the possibility, that some local minima were missed with the BH workflow and hence were not included in the formalism. We tried to minimize this source of error by optimizing additional manually generated cluster structures not provided by BH (e.g. solvents binding to the edges of the tetrahedron) but none of those were found to be stable minima.

4.4.2 CCS fits: MD vs. rigid samples

The collision cross section (CCS) of a molecule is temperature dependent in two ways: (1) the effective temperature (following two-temperature theory), which describes the relative velocity distribution between ions and collision gas, enters Eq. (4.3) directly, and (2) the internal energy of a molecular ion, corresponding to the ion temperature, will increase the vibrational motion and thus the occupied configurational space. Point (2) is not taken into account by Eq. (4.3). Thus, performing MD simulations at a certain ion temperature and taking random sample geometries to calculate the CCS for each geometry with Eq. (4.3) at the respective effective temperature should incorporate both effects. We note, though, that inelastic collisions are not modeled this way because the geometries employed for MobCal-MPI calculations are treated as rigid. New approaches to this problem have been published recently [142] and it would be interesting to compare the results of the method presented here when modeling the inelasticity of collisions.

4 Modeling of Differential Ion Mobility

To conduct our treatment, we took 100 random geometries from each MD simulation, hence performing 100 CCS calculations for each temperature. We assume $T_{\text{ion}} \approx T_{\text{eff}}$ as should be true for an atomic collision gas and molecular ions [132, p. 358]. Although our calculations are performed in molecular nitrogen, we assume this relation holds true, rationalized by the rather stiff nature of the N_2 bond.

Figure 4.3 shows the results for Me_4N^+ , $[\text{Me}_4\text{N} + \text{ACE}]^+$, and $[\text{Me}_4\text{N} + (\text{ACE})_2]_a^+$. We only studied the temperature range up to 700 K because of the time-consuming MD simulations. Generally, it can be noted that $\overline{\Omega_{\text{MD}}} \geq \overline{\Omega_{\text{rigid}}}$, i.e., the rigid CCS are always within the MD distributions but towards the lower end, hinting at the anharmonic nature of the PES. For a completely harmonic oscillator, these averages should be equal. Taking a look at the results for Me_4N^+ (Fig. 4.3a), it is safe to say that the contribution due to vibrational extension is negligible compared to the error of the CCS calculation, i.e., the distributions are very narrow and easily within the error of the rigid calculations. To some degree, this still holds true for the $[\text{Me}_4\text{N} + \text{ACE}]^+$ cluster (Fig. 4.3b), although a significant increase in the width of the distributions is visible. In both cases, though, the difference between the fitting functions is negligible. In the case of the $[\text{Me}_4\text{N} + (\text{ACE})_2]_a^+$ cluster (Fig. 4.3c), the width of the CCS distributions from MD sampling becomes relatively large. We conducted an in-depth analysis of the MD simulation performed at 450 K, where the distribution covers a range of more than 10 \AA^2 . During the trajectory, one of the ACE molecules moved from one face of the tetramethyl ammonium to another. When ACE passes over the edge, the ion-solvent cluster increases in size, explaining the broad distribution (skewed towards higher CCS) observed. This, again, hints that the harmonic oscillator treatment for the partition function could be improved by explicitly correcting for anharmonicity. While the mean values still agree within error, the error of the rigid structure CCS does not encompass the entire MD distribution.

We chose to perform rigid CCS calculation only and ignore the vibrational broadening for two reasons. First, a single fitting function, even when considering all MD data, cannot account for the width of the CCS distribution and we are only using the average value of the CCS in the Mason-Schamp equation. Insufficient sampling of the accessible configurational space due to missed ergodic mixing is also a major problem of the MD approach and is likely the reason for the nonphysical set of parameters obtained for $[\text{Me}_4\text{N} + (\text{ACE})_2]_a^+$ (c should be positive, b is very different compared to the other functions, see Fig. 4.3c). Second, there are two counteracting effects regarding the temperature. Although the actual size of the molecule increases with temperature due to vibrational broadening, the overall CCS decreases because the long-range interactions with the background gas become less important for higher relative velocities [229]. The fact that the two fitting functions in the $[\text{Me}_4\text{N} + (\text{ACE})_2]_a^+$ case seem to meet again towards higher temperatures could be interpreted in this manner. This might not be the case for all systems: Especially peptides or proteins can show significant unfolding and thus overall increase of CCS with (effective) temperature [229]. Relying only on a single geometry optimization (for CCS calculations at all temperatures) in contrast to a set of rather time-consuming MD simulations (for

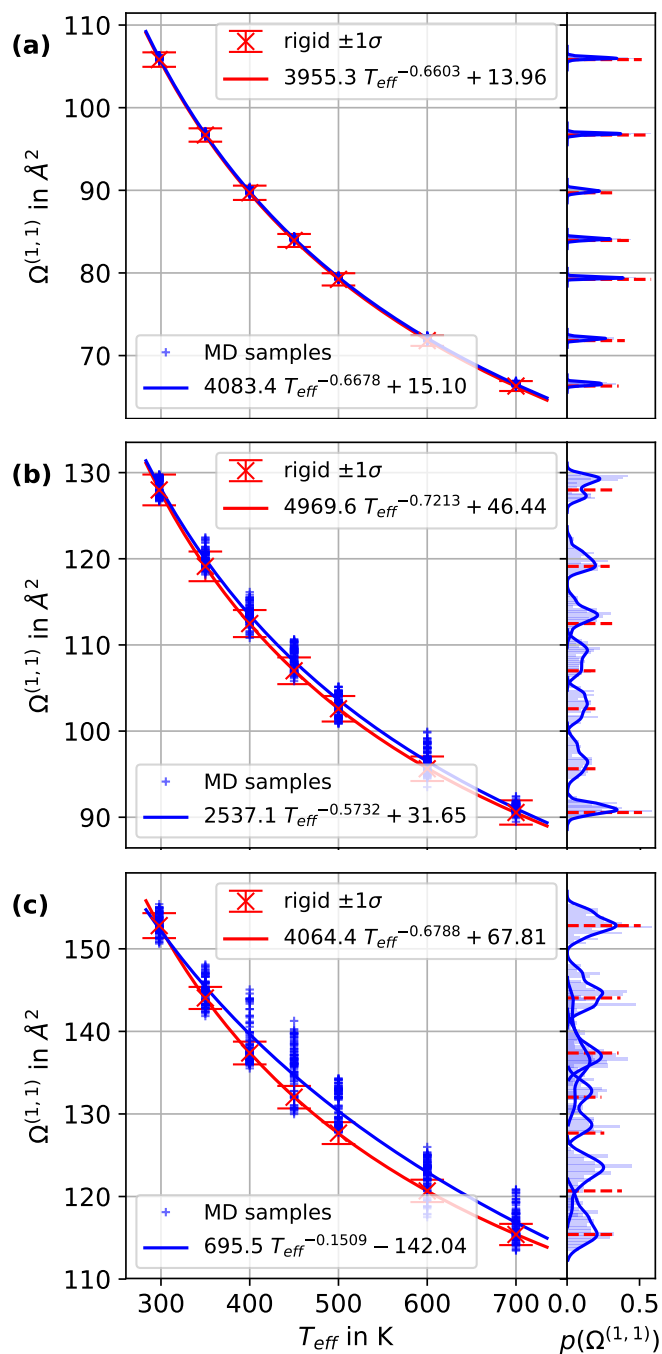


Figure 4.3: Calculated CCSs and respective fitting functions for the MD simulations (blue) and the rigid ion/cluster structures (red) over a range of effective temperatures. (a) Me_4N^+ (b) $[\text{Me}_4\text{N} + \text{ACE}]^+$ (c) $[\text{Me}_4\text{N} + (\text{ACE})_2]_a^+$. The CCS distributions for each MD simulation are also shown (right) as histogram and smooth Gaussian kernel distribution.

4 Modeling of Differential Ion Mobility

each different temperature) dramatically speeds up the whole formalism. The use of MD simulations for the CCS fit and an opportunity to calculate the partition function from the trajectories is discussed below.

We now focus on the fitting functions themselves, which are given for all molecules in Table 4.1. Functions in Table 4.1 differ from those shown in Figure 4.3 owing to the extended temperature range. The ranges were extended because mobility calculation for some species yielded higher temperatures than 700 K; extended fits ensure that the CCS at a specific temperature was always interpolated and never extrapolated. Because the exponent can vary, the fitting parameters are rather insensitive, i.e., two different sets of parameters can yield a very similar fitting function in the temperature range studied. However, those two functions can differ significantly outside of that temperature range, rendering extrapolation of the CCS rather difficult. It would be beneficial to fix the exponent to a value derived from theoretical considerations, since it might well be a constant. While this is beyond the scope of this work, we do note that the average exponent is around -0.72 and no trend regarding cluster size is visible. Also, it is always lower than the value of -0.5 , which was previously described in the literature [50].

Regarding the other two parameters, it can be seen that there is a positive correlation of the offset c with size of the cluster geometry. This is expected since the offset can be interpreted as CCS in the infinite temperature limit, i.e., when long-range interactions are minimized and the CCS is determined only by hard-sphere collisions [229]. The proportionality constant a seems to decrease with size of the ion. Another benefit of knowing the exponent, b , from theory would be that a and c could be determined from just two CCS calculations, covering the needed temperature range.

We should note that the CCS calculations were performed in nitrogen gas only. All interaction with the solvent vapor is modeled only by the stable clusters. However, having about 1.5 mol% solvent in the gas phase, should have an influence on the CCS not only because they are larger than N_2 , but also because they have a permanent dipole moment and thus different interaction potentials ($V \propto R^{-2}$ for ion-permanent dipole compared to $V \propto R^{-4}$ for ion-induced dipole). Blanc's law [230] could be used to account for this additional effect by calculating the CCS in N_2 and in the appropriate solvent vapor, then averaging the two according to their respective mole percent. However, we have no means to calculate the CCS in a pure, e.g., MeOH environment with comparable accuracy as for N_2 . Also, Blanc's law would be a rather rough approximation for the mixture at higher field strengths. Thus, it is rather difficult to estimate the error but we expect it to be small enough for the goal of this paper. Note also that the interaction potential parameters implemented in the MobCal-MPI code were trained to reproduce experimental CCS at 298.15 K. Although these parameters could be different at different temperatures, we assumed them to be constant at all temperatures. Since long-range interactions become less important with increasing effective temperature, the error introduced by using fixed parameters should vanish for higher temperatures [140].

Table 4.1: Parameters for CCS fitting functions ($\Omega^{(1,1)} = a T_{\text{eff}}^{-b} + c$). For each ion a different temperature range is considered, depending on the observed maximum effective temperature. The smaller ions can move faster and thus can have a higher effective temperature.

Cluster	T range in [K]	a in [$\text{\AA}^2 \text{K}^{-b}$]	b	c in [\AA^2]
Me_4N^+	298 – 1300	6250.7	-0.7587	23.15
$[\text{Me}_4\text{N} + \text{ACE}]^+$	298 – 900	6006.8	-0.7619	49.63
$[\text{Me}_4\text{N} + (\text{ACE})_2]_a^+$	298 – 700	4284.0	-0.6903	68.90
$[\text{Me}_4\text{N} + (\text{ACE})_2]_b^+$	298 – 700	4751.3	-0.7152	66.92
$[\text{Me}_4\text{N} + (\text{ACE})_2]_c^+$	298 – 700	4697.6	-0.7073	65.64
$[\text{Me}_4\text{N} + (\text{ACE})_2]_d^+$	298 – 700	4543.0	-0.6911	69.46
$[\text{Me}_4\text{N} + \text{ACN}]^+$	298 – 900	5956.0	-0.7719	44.63
$[\text{Me}_4\text{N} + (\text{ACN})_2]_a^+$	298 – 700	4556.2	-0.7227	61.95
$[\text{Me}_4\text{N} + (\text{ACN})_2]_b^+$	298 – 700	4928.9	-0.7139	61.88
$[\text{Me}_4\text{N} + (\text{ACN})_2]_c^+$	298 – 700	4380.7	-0.6952	56.05
$[\text{Me}_4\text{N} + (\text{ACN})_3]_a^+$	298 – 700	3725.6	-0.6739	77.73
$[\text{Me}_4\text{N} + \text{MeOH}]^+$	298 – 900	5601.3	-0.7481	36.56
$[\text{Me}_4\text{N} + (\text{MeOH})_2]_a^+$	298 – 700	4076.4	-0.6894	44.77
$[\text{Me}_4\text{N} + (\text{MeOH})_2]_b^+$	298 – 700	5367.5	-0.7467	50.25
$[\text{Me}_4\text{N} + (\text{MeOH})_2]_c^+$	298 – 700	5690.0	-0.7479	50.61

4.4.3 Mobility calculations

Having determined the geometries, CCS fits, partition function parameters, and energies, we can now conduct the mobility calculations. Figure 4.4 shows exemplary results for the $[\text{Me}_4\text{N} + (\text{ACE})_n]^+$ system for two different field strengths, i.e., 25 Td and 76 Td. A background gas temperature of 373 K and a solvent concentration of 1.5 mol% was applied.

In Figures 4.4a-c, which show the self-consistent mobility results, it can be seen that a larger CCS correlates with a lower effective temperature and also a lower drift velocity, as one would expect. During the low-field condition the effective temperature is only slightly increased compared to the background gas, whereas substantial heating occurs for the small Me_4N^+ ion under high-field conditions. This behavior arises due to the increase in drift velocity owing to the higher electric field and concomitant decrease in CCS (see Fig. 4.4c and Fig. 4.3). The effective temperatures of the larger clusters also increase with field strength, but not as much as it does for the small bare ion.

The data from Fig. 4.4b-c gives the opportunity to calculate collision frequencies

4 Modeling of Differential Ion Mobility

according to two-temperature theory [132, p. 275]. For the smallest ion, i.e., Me_4N^+ , we estimate the collision frequency with N_2 to be $1.5 \times 10^{10} \text{ s}^{-1}$ at 25 Td. While the larger clusters have a higher CCS, their relative velocities are smaller due to the lower effective temperature and the higher reduced mass. The resulting collision frequency is in the same order of magnitude, e.g. for the $[\text{Me}_4\text{N} + (\text{ACE})_2]_a^+$ cluster at 25 Td about $2.1 \times 10^{10} \text{ s}^{-1}$. Typical DMS devices operate with a waveform frequency in the MHz regime, thus, the collision frequency is four orders of magnitude higher. This validates our assumption of instant thermalization since there is a sufficient number of collisions per cycle of the waveform.

Figure 4.4d shows the Boltzmann weighting factors derived from our implementation of the superposition approximation, again assuming $T_{\text{ion}} \approx T_{\text{eff}}$ [132, p. 359]. Since the $[\text{Me}_4\text{N} + (\text{ACE})_2]_a^+$ cluster is energetically lowest, its weight factor is always 1, independent of temperature (cf. Eqs. (4.13a), (4.14a)). In contrast, the bare ion has the lowest weighting because it is the highest in energy (no contributions from bound solvents). Multiplying these weights with the partition functions plotted in Fig 4.4e result in the population distribution shown in Fig 4.4f. Although the bare ion has the lowest weight, its partition function is orders of magnitude larger due to the additional translational degrees of freedom (or in other words, the entropy gain due to complete dissociation). Changing the field strength results in a change in the Boltzmann weights and the associated partition functions. The careful interaction of those two values ultimately determines the relative population of each cluster structure. Since the variation in weights and partition functions are not the same for all structures, the resulting population distribution changes with field strength. As expected, the mean cluster size decreases with higher field strength, ultimately leading to a higher ensemble mobility (Type A behavior).

It should be noted that, while the self-consistent calculation is an elegant way of calculating drift velocity and effective temperature from just the CCS, the use of the Mason-Schamp equation is still a source of some error. Even when applying two-temperature theory, there are known deviations in the order of 10% at higher field strengths [214]. Although there are higher order corrections possible to two-temperature theory, as well as a different approach called momentum-transfer theory [135, 214], to date there is no closed formula available which correctly predicts the mobility of an ion for high field strengths.

We also want to stress that the error in the partition function calculations increases with temperature. This is due to: (1) neglecting anharmonic contributions, which become more important at higher temperatures (as apparent in the CCS distributions, see Fig. 4.3), and (2) neglecting ro-vibrational coupling, which increases with temperature and is especially significant for low frequency dissociative modes.

Hence, both factors in Eq. (4.15), i.e., the individual mobility and the relative population are source to an increasing error with increasing field strength.

Figure 4.5a shows the α function and its derivative, as determined with our method, for the $[\text{Me}_4\text{N} + (\text{ACN})_n]^+$ system in N_2 with $T_{\text{bath}} = 373 \text{ K}$ and a solvent vapor con-

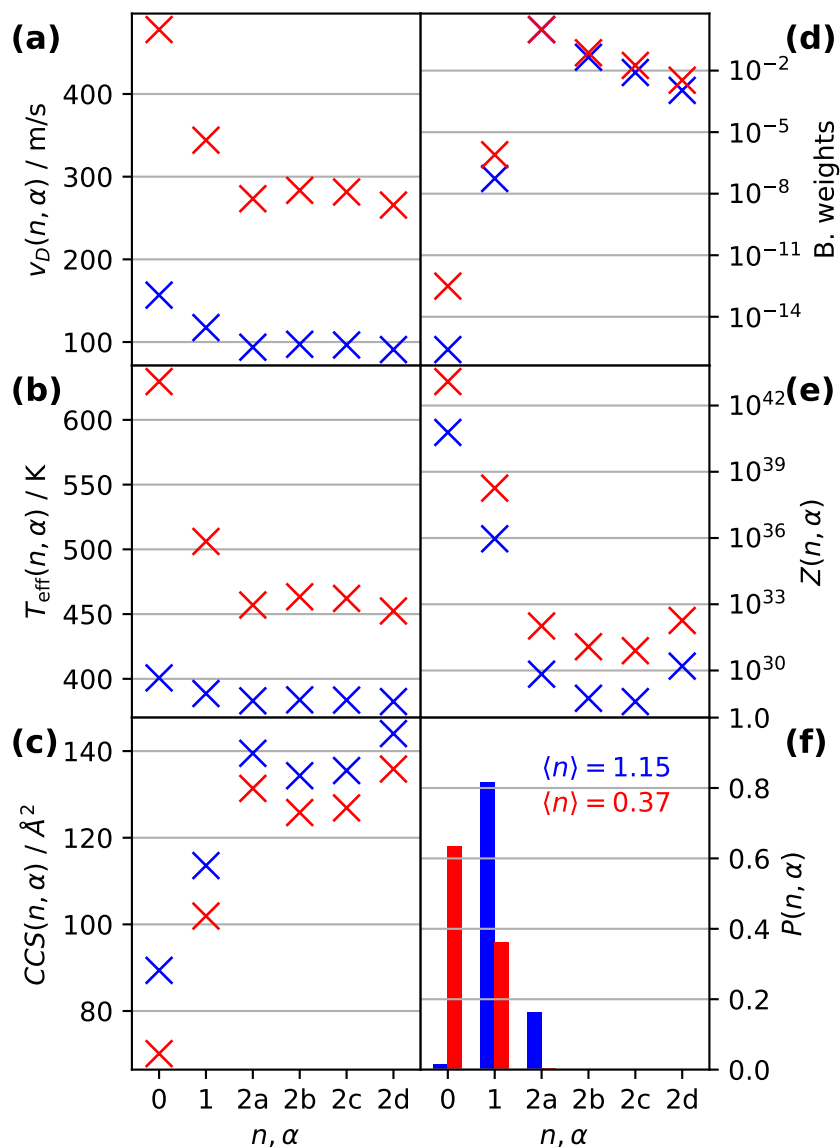


Figure 4.4: (a) drift velocities, (b) effective temperatures, (c) collision cross sections, (d) Boltzmann weights, (e) partition functions and (f) cluster size distributions for 25 Td (500 V/mm at 1 atm and 373 K, blue data) and 76 Td (1500 V/mm at 1 atm and 373 K, red data) for the [Me4N+(ACE)n]+ system. As expected, the increase in field strength, and thus effective temperature, shifts the cluster size distribution to smaller cluster sizes, increasing the overall mobility of the ensemble (Type A). For a detailed explanation of the panels, see text.

4 Modeling of Differential Ion Mobility

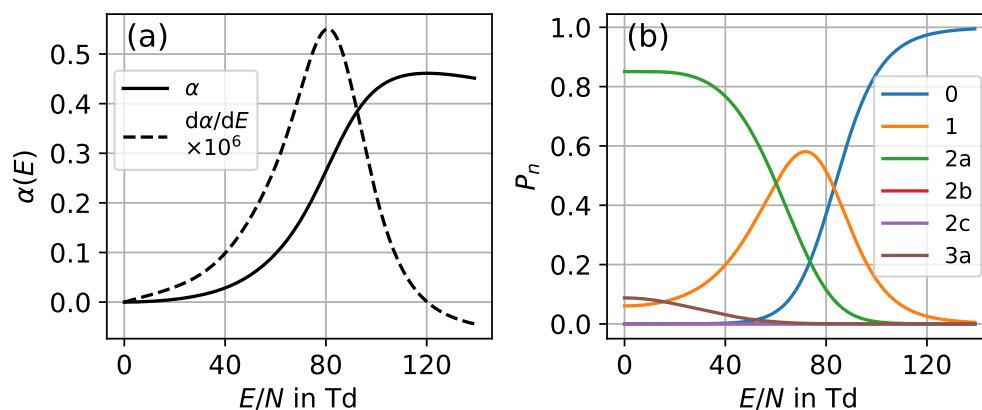


Figure 4.5: (a) α curve and its derivative (b) relative ion-solvent cluster population as a function of reduced field strength (E/N) for the $[\text{Me}_4\text{N} + (\text{ACN})_n]^+$ system. The strong increase in α corresponds to a Type A behavior and is reasoned by the overall decrease of CCS due to a smaller average cluster size. Only for high field strengths, where only the bare ion is populated, the hard-sphere effect becomes significant and the α curve returns to smaller values. Similar figures for MeOH and ACE can be found in Fig. 4.9

centration of 1.5 mol%. Figure 4.5b shows the corresponding relative population of each cluster structure. The small changes of the distribution up to 20 Td lead to an almost constant mobility (or α function), which can be interpreted as the well-known field independency of the mobility in the low-field limit. Further increasing the field strength leads to a gradual decrease of the mean cluster size and thus mean CCS, which in turn increases the mobility of the ensemble (Type A). At very high field strengths, when there is no larger change in cluster size distribution ($P_{n=0} \rightarrow 1$) and the individual CCS has almost converged to its asymptote c , the mobility actually decreases due to the fact that $K \propto 1/\sqrt{T_{\text{eff}}}$ (cf. Eq. (4.2)). Thus, three major factors contribute to the overall mobility of an ensemble: (1) the general $K \propto 1/\sqrt{T_{\text{eff}}}$ effect (sometimes called "hard-sphere effect" [217]), (2) the change of CCS when the background gas is polarizable and (3) the change in the cluster distribution when there are polar solvents able to form stable clusters with the ion.

4.4.4 Dispersion plots

Figure 4.6 shows the calculated and experimental dispersion plots for Me_4N^+ in pure N_2 and seeded with 1.5 mol% of MeOH, ACN and ACE. Although the authors of the experimental paper report a temperature of 423 K inside the DMS cell [68], using a thermocouple for a coarse measurement of the gas temperature inside the DMS cell, we found the temperature to be only around 373 K. Consequently, we applied

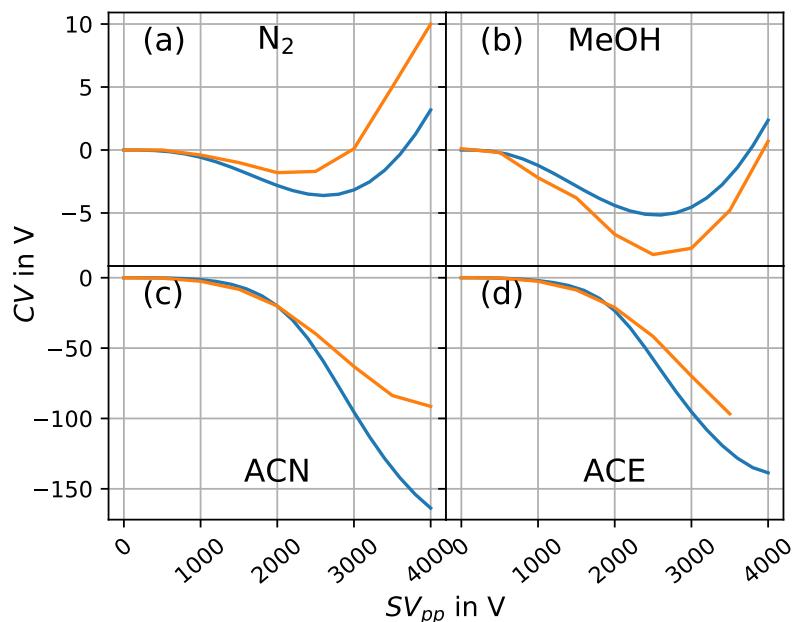


Figure 4.6: Calculated (blue) and experimental (orange) dispersion plots for Me_4N^+ in (a) N_2 at 1 atm and 373 K, and (b) N_2 seeded with MeOH, (c) ACN, and (d) ACE as modifiers (1.5 mol% at 373 K). The experimental data are taken from [68].

$T_{bath} = 373$ K in the calculations. Variations of the temperature used for the calculations showed some noticeable change in the dispersion plots within ± 20 degrees but no qualitative difference within the aim of this work. A gap size of $d = 1$ mm is used in Eq. (4.17) and (4.18) and we discretized the waveform into 1000 time steps for the averages in Eq. (4.18), which was sufficient with regard to convergence.

First, we discuss the dispersion plot of the pure nitrogen environment. Because the CCS of the bare ion decreases with effective temperature, the mobility increases slightly at lower field strengths, thus requiring negative CVs to correct trajectories. Further increasing the separation field strength leads to hard-sphere scattering, which requires increasingly positive CV values for optimal ion transmission. Thus, a very weak Type B curve is generated. The calculated dispersion plot nicely reproduces the experimental data and the two curves match qualitatively.

Upon adding a weakly-interacting solvent vapor to the system, i.e., MeOH, where only $n = 1$ is populated even at low field strengths (see Fig. 4.9b), the Type B curve becomes more pronounced (i.e., the minimum CV becomes more negative and shifts to higher SV values). Our model is able to reproduce this behavior qualitatively, as well as the overall shape of the curve. In particular the SV where the CV has a minimum, called $SV@CV_{min}$ in the literature and reported to correlate with binding energy and

4 Modeling of Differential Ion Mobility

other physicochemical parameters such as pK_a , pK_b and Hammett parameters [231], is nicely reproduced. Introducing a more strongly-interacting solvent vapor (ACN, ACE) yields Type A behavior for the range of SV studied. This indicates that the ion mobility under the high-field conditions is larger than under low-field conditions. Again, this behavior is qualitatively reproduced by our model. Note, that the formalism is the same in all cases; no system specific parameters are introduced and every result arises from the (modeled) nature of the system itself.

At $SV_{pp} > 2000\text{V}$ significant deviations occur between calculated and experimental dispersion curves, especially for ACN. There are several sources that could give rise to these errors. As mentioned previously, the description of the partition functions for high ion temperatures and the use of the Mason-Schamp equation for very high field strengths introduce errors. In addition of those sources of error, our assumption of instant equilibration could break down. While the N_2 collision frequency in the order of 10^{10} s^{-1} should be sufficient to thermalize the ions fast enough, the clustering reaction depends on the collision frequency with the solvents rather than with N_2 . This number is difficult to calculate since we don't have access to $\Omega_{\text{solvent}}^{(1,1)}$. A rough estimate could be given by just scaling the N_2 collision frequency by the mol% of the solvents which would give a frequency in the order of 10^8 s^{-1} . Thus, there are only about 100 collisions per cycle of the waveform with solvent molecules. Fig. 4.5 shows that at low-field conditions the $n = 2$ cluster dominates the population, while at high-field conditions only the bare ion is populated ($SV_{pp} = 4000\text{V}$ corresponds to a maximum reduced field strength of 135 Td under the given conditions). Thus, having a waveform frequency of $\omega = 3\text{ MHz}$, the cluster formation reaction might not be fast enough due to insufficient collision numbers to populate the $n = 2$ state to its equilibrium value. Additionally, the rate of evaporation given by the unimolecular decay of the heated ions in the high-field portion could also be a limiting factor. Especially a change from $n = 2$ to $n = 0$ could be difficult to achieve on these time scales. Indeed, the more negative CV values, compared to the experimental values, suggest a larger dynamic change in clustering in the calculations than in the experiment. Kinetic simulations with theoretically determined rate constants could circumvent this problem and lead to a more accurate population distribution. The deviations in the ACE case are significantly smaller and indeed the change in cluster size between high- and low-field is not as pronounced (see Fig. 4.9a), supporting the suggested error due to non-equilibrium conditions. For MeOH, an underestimation instead of an overestimation is visible and the deviations are smaller overall (see Fig. 4.6b). Since MeOH shows only weak clustering compared to ACN and ACE (see Fig. 4.9b), we don't expect kinetic effects to play a major role. Thus, the deviations are solely due to other errors already mentioned, e.g., neglecting anharmonicity of the PES and the use of two-temperature theory for very high field strengths.

It is important to stress again the critical role of the CCS fitting function. Applying a $\Omega \propto 1/\sqrt{T_{\text{eff}}}$ fitting function as stated in the literature [50, 132] would actually cancel the two-temperature effect in the Mason-Schamp equation, as already mentioned in the literature [50]. This would lead to a constant mobility for the pure N_2 environment

leading to no differential mobility ($CV = 0V$ for all SV_{pp}). This is clearly not observed in the experimental data (see Fig. 4.6a). The determined exponent of $b < -0.5$ is thus critical for reproducing the correct form of the dispersion plot.

To further test our model, we exploited varying the background gas temperature and the solvent concentration (see Fig. 4.7). Experimental data [50, 215, 219, 232] show that with increasing concentration of the solvent vapor, more negative CV values are needed to transmit the ions through the DMS cell. This results in a pronounced Type B behavior that evolves into Type A as concentration increases. This can be explained by a higher degree of clustering and thus larger change in CCS upon evaporation since an increase in gas phase concentration will shift the equilibrium towards more clustered ions. This is modeled by the concentration dependency of the relative population (cf. Eq. (4.6) and (4.14)) and our model nicely reproduces this trend (see Fig. 4.7a).

Increasing the background gas temperature reduces clustering due to an increased internal energy and subsequent shift of the equilibrium towards higher degree of dissociation. Consequently, less pronounced Type B behavior is observed experimentally as T_{bath} increases [51, 219, 233]. Again, our model reproduces this trend correctly (see Fig. 4.7b). When T_{bath} is raised to the point at which ion-solvent clustering is prevented even at low-field conditions, the dispersion plot converges to that of the pure N_2 atmosphere (cf. Fig. 4.7b dashed line) as expected.

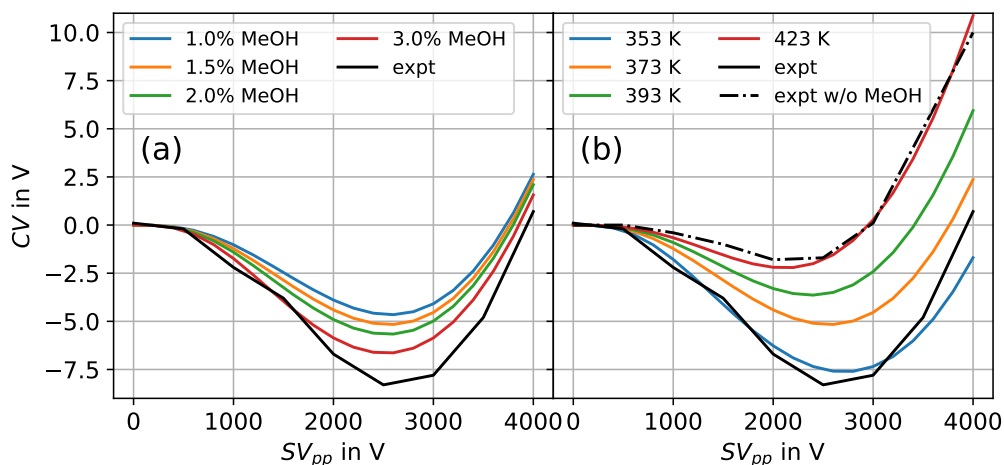


Figure 4.7: (a) Variations of modifier concentration and (b) background gas temperature for Me_4N^+ in N_2 (1 atm) seeded with MeOH (expt conditions are at 1.5 mol% MeOH and 373 K). Increasing the concentration as well as decreasing the temperature leads to pronounced Type B behavior due to higher degree of clustering.

4.5 Conclusion

The ability of our model to reproduce dispersion plots for different separation conditions, i.e., different (or no) solvents added, different concentrations and different background gas temperatures, strongly suggests that we capture the most important effects leading to differential mobility, namely clustering, the hard-sphere effect and the change in individual CCS with effective temperature.

Using two-temperature theory we model the overall dependence of the ion mobility with field strength. This is important to account for the overall decrease of mobility with field strength, known as the hard-sphere effect. With no solvent added and having a very weakly polarizable gas like Helium, this effect is dominant, leading to Type C behavior. While two-temperature theory covers most of this effect, it is only an approximation and known to show deviations for very high field strengths [214]. Higher order terms or a different approach called momentum-transfer theory [135, 214] could improve this description.

Having a polarizable collision gas like N_2 , it is important to model the temperature dependence of the CCS. The reported dependency of $\Omega^{(1,1)} \propto T_{\text{eff}}^{-0.5}$ [50, 132] not only shows systematic deviations from calculated CCS values, in combination with two-temperature theory it also fails to reproduce Type B or Type C systems. We model the functional dependency by fitting a more general function to CCS values of the equilibrium structure at different temperatures. Doing this, we neglect vibrational broadening of the molecule and inelasticity of collisions.

If solvents are added to the gas phase, clustering plays an important role and only when this is taken into account, Type A dispersion plots can be predicted. We model the cluster size distribution by assuming fast thermal equilibrium and using a special version of the superposition approximation, applying harmonic approximations to the partition functions. While this seems to model the data qualitatively, two main errors are introduced: (1) the equilibrium condition might not hold for fast oscillating fields. Kinetic simulations with theoretically determined rate constants would give a better description of the population distribution. (2) the simple description of the partition functions neglects anharmonicity of the PES, as well as ro-vibrational coupling and easy conversion between cluster structures. This could be partially improved by using e.g. the quasi-harmonic approximation [234–236], where the vibrational partition function is determined from MD simulations on the anharmonic surface. Conducting MD simulations for a set of temperatures (ensuring sufficient sampling of the configurational space) could be used to obtain a more accurate description of the temperature dependence of the vibrational partition function as well as more accurate CCS fits (accounting for vibrational broadening).

Despite the simplifications made, our model is able to predict trends in DMS dispersion plots from first principles for different experimental conditions, i.e., different (or no) solvents added, different solvent concentration and different background gas temperature. To our knowledge, this is the first time a thorough model is presented that can account for all these effects. The importance of different contributions to

overall mobility, especially clustering with solvent molecules, is highlighted.

Acknowledgement

We gratefully acknowledge the high-performance computing support from the SHARCNET consortium of Compute Canada. We also want to acknowledge Christian Ieritano for his help with the basin hopping algorithm.

4.6 Supporting Information

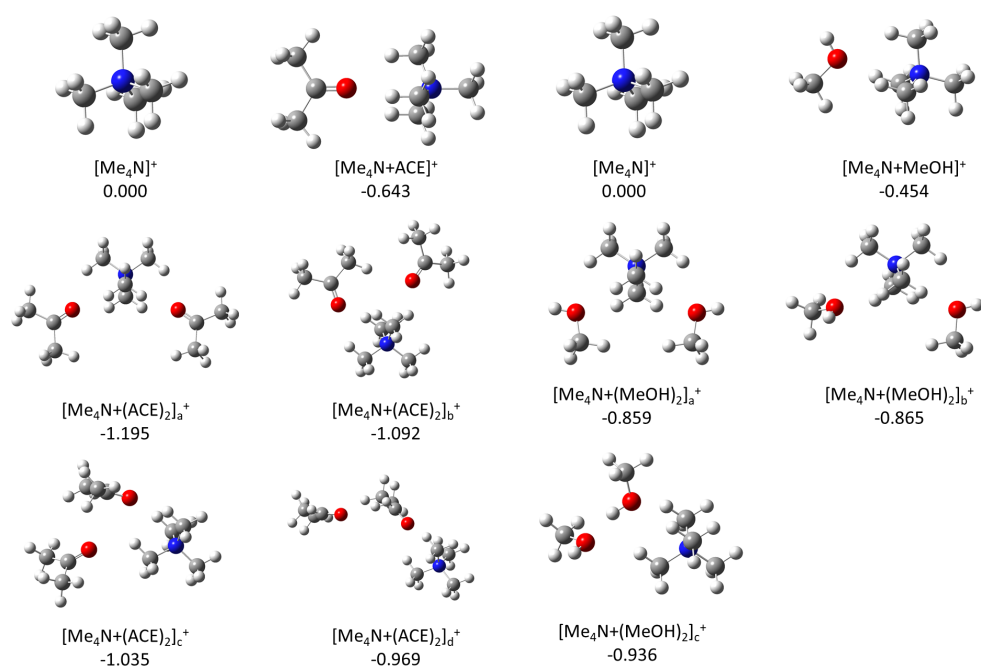


Figure 4.8: Geometries of all $[\text{Me}_4\text{N} + (\text{ACE})_n]^+$ and $[\text{Me}_4\text{N} + (\text{MeOH})_n]^+$ clusters optimized at the B3LYP-GD3/6-31++G(d,p) level of theory. Zero-Point corrected binding energies, $\Delta\mathcal{E}_0$, are given in eV relative to the dissociated state.

4 Modeling of Differential Ion Mobility

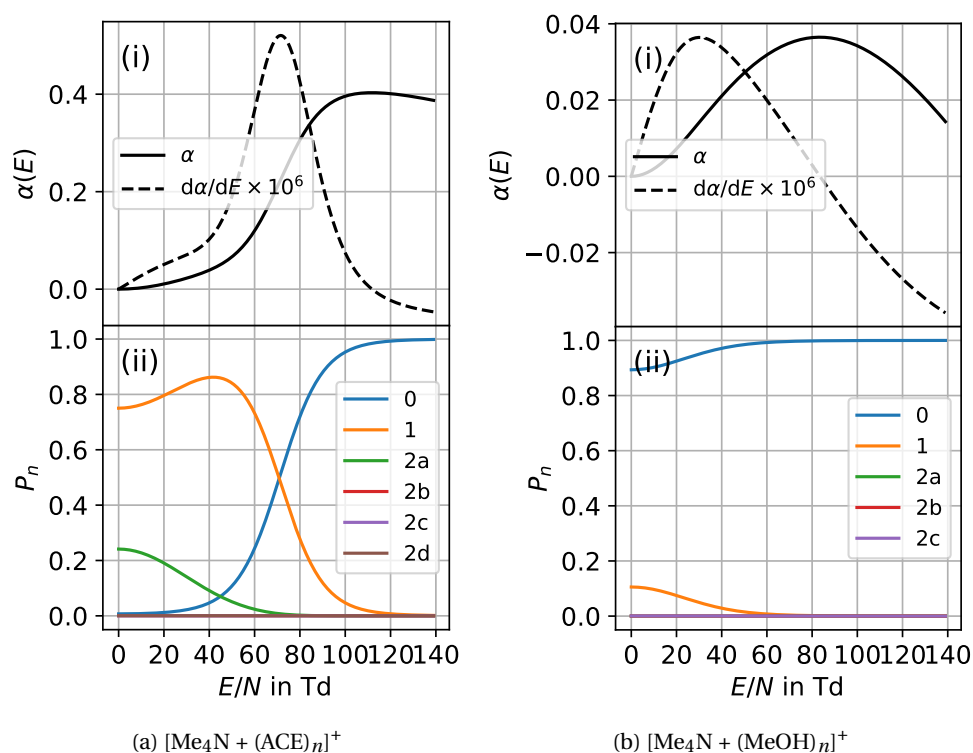


Figure 4.9: (i) α curve and its derivative and (ii) relative ion-solvent cluster population as a function of reduced field strength (E/N) for the $[\text{Me}_4\text{N} + (\text{ACE})_n]^+$ and $[\text{Me}_4\text{N} + (\text{MeOH})_n]^+$ systems.

4.7 Follow-up work

To further test the finding that dynamic clustering is a major contribution to differential mobility, as well as testing the capabilities of the developed model for the prediction of trends in dispersion plots, unusual experimental results can be used. One such case are Type D dispersion plots [237, 238], where the ions show Type C behavior initially but switch to Type A above a certain SV value. Hypothesizing very strong but inflexible clustering, a solvent would bind to the ion and would be stable throughout the whole waveform cycle. In absence of dynamic clustering, the differential mobility would be negative, leading to the (initial) Type C behavior. Only when the field becomes strong enough, dynamic clustering starts, eventually leading to Type A behavior, as explained above in this chapter.

As an example case for ions with Type D behavior, the tricarbostannatrane cation, $\text{N}(\text{CH}_2\text{CH}_2\text{CH}_2)_3\text{Sn}^+$ (see Fig. 4.10), is chosen. IRMPD studies [239] have shown that the tin atom offers a very potent binding site to solvent molecules, rendering the tricarbostannatrane ion a good example to study strong clustering ions.

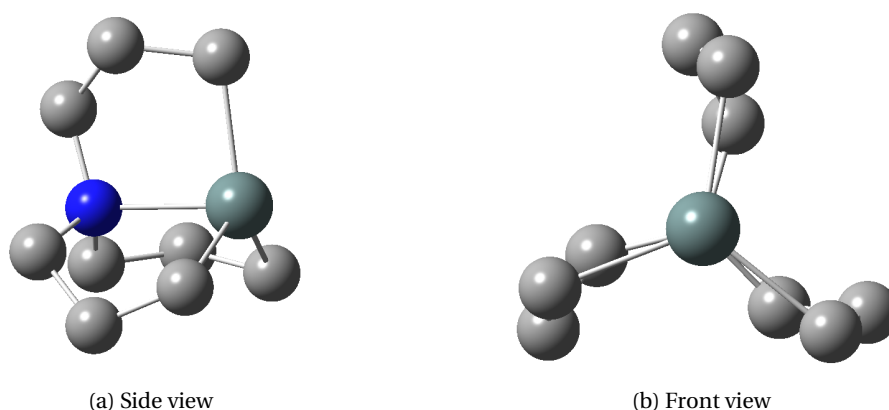


Figure 4.10: Geometry of the tricarbastannatrane ion. The N–Sn bond inside the cage resembles a C_3 symmetry axis. Hydrogen atoms are removed for clarity.

4.7.1 Methods

Experimental For a detailed description of the experimental methods, the reader is referred to the original paper [195]. Briefly, tricarbastannatrane chloride is dissolved in a 99:1 acetonitrile/methanol mixture leading a 100 ngmL^{-1} ESI solution. The differential mobility is studied with a DMS coupled to a hybrid triple quadrupole linear ion trap (SCIEX; Concord, ON, Canada), either in pure N_2 environment or N_2 seeded with 1.5 mol% of ACN or ACE. The bath gas temperature is also varied, setting the DMS plate temperatures to 150°C , 225°C , and 300°C . The peak-to-peak SV is varied from 0 – 4000V.

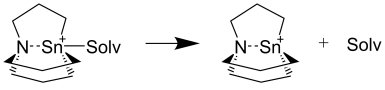
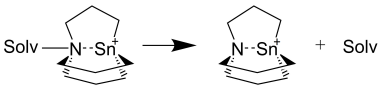
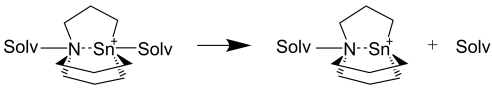
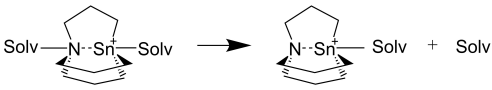
Computational The computational methods are essentially the same as described above in this chapter. However, for a better description of the electronic energy, the single-point energy calculations are performed on the B2PLYP-GD3BJ/def2-TZVPP level of theory including counterpoise correction [240] to account for basis set superposition error. To calculate the partial charges on tin, an atomic radius of 2.17 \AA was chosen [241], the atomic polarizability, necessary for the MMFF94 force field, was set to 7.809 \AA^3 [242].

4.7.2 Results and Discussion

The suggested strong clustering ability of the tricarbastannatrane ion is investigated by analyzing the mass spectrum observed under the different DMS conditions (N_2 pure and seeded with ACN or ACE). While the first quadrupole (Q1) is scanned, the third (Q3) is set to transmit only the bare tricarbastannatrane ion. For the pure N_2 environment in the DMS cell, the mass spectrum only show the A^+ signal, suggesting that no significant clustering due to the ESI solution occurred. However, seeding the

4 Modeling of Differential Ion Mobility

Table 4.2: Gibbs enthalpies of cluster dissociation of the tricarbastannatrane-solvent clusters with different binding motifs and up to two solvent molecules. Negative values suggest spontaneous dissociation.

Reaction	Solv	Binding Energy / kJ mol^{-1}		
		150 °C	225 °C	300 °C
	ACE	39.2	30.7	22.4
	ACN	48.7	41.9	35.2
	ACE	6.2	0.3	-5.3
	ACN	7.7	2.1	-3.3
	ACE	28.2	19.5	11.0
	ACN	32.8	26.0	19.2
	ACE	-4.9	-10.8	-16.6
	ACN	-8.2	-13.8	-19.3

DMS cell with solvents, the $[\text{A} + \text{ACE}]^+$ and $[\text{A} + \text{ACN}]^+$ clusters are the main signals, although the unclustered tricarbastannatrane signal is observed as well. These results suggest (1) that the clusters form to a significant amount only when solvents are added to the DMS cell, (2) that they are stable enough to survive the ion transfer into Q1 but (3) fragment between Q1 and Q3. The missing signals of the multiple clustered species suggest that the binding energies of additional solvents is much lower.

These findings are supported by the theoretical calculations of Gibbs enthalpies of cluster dissociation (cf. Eq. (2.28)), as shown in Table 4.2. The binding motifs where the solvent binds to the tin are much more stable as shown by the large positive dissociation enthalpies. The N-binding is much weaker, sometimes even spontaneous dissociation (negative $\Delta_R G$) is predicted. This renders tricarbastannatrane a good example case of very strong and exclusive clustering, i.e., one solvent binds strongly to the ion and essentially prevents further clustering.

Experimental and theoretical dispersion plots are shown in Fig. 4.11. Regarding the pure N_2 environment, Fig. 4.11a, the experimental results show Type C behavior as expected for a rigid, non-clustering ion (hard-sphere effect). This behavior is, as well as the dependency on the bath gas temperature, correctly predicted by the model described in this chapter, as can be seen in Fig. 4.11d. Note though, that the absolute compensation voltages are overestimated.

In N_2 environment seeded with 1.5 mol% of ACE or ACN, Type D behavior is observed (Fig. 4.11b and 4.11c). I.e., for low SV_{pp} values, positive CV values are needed

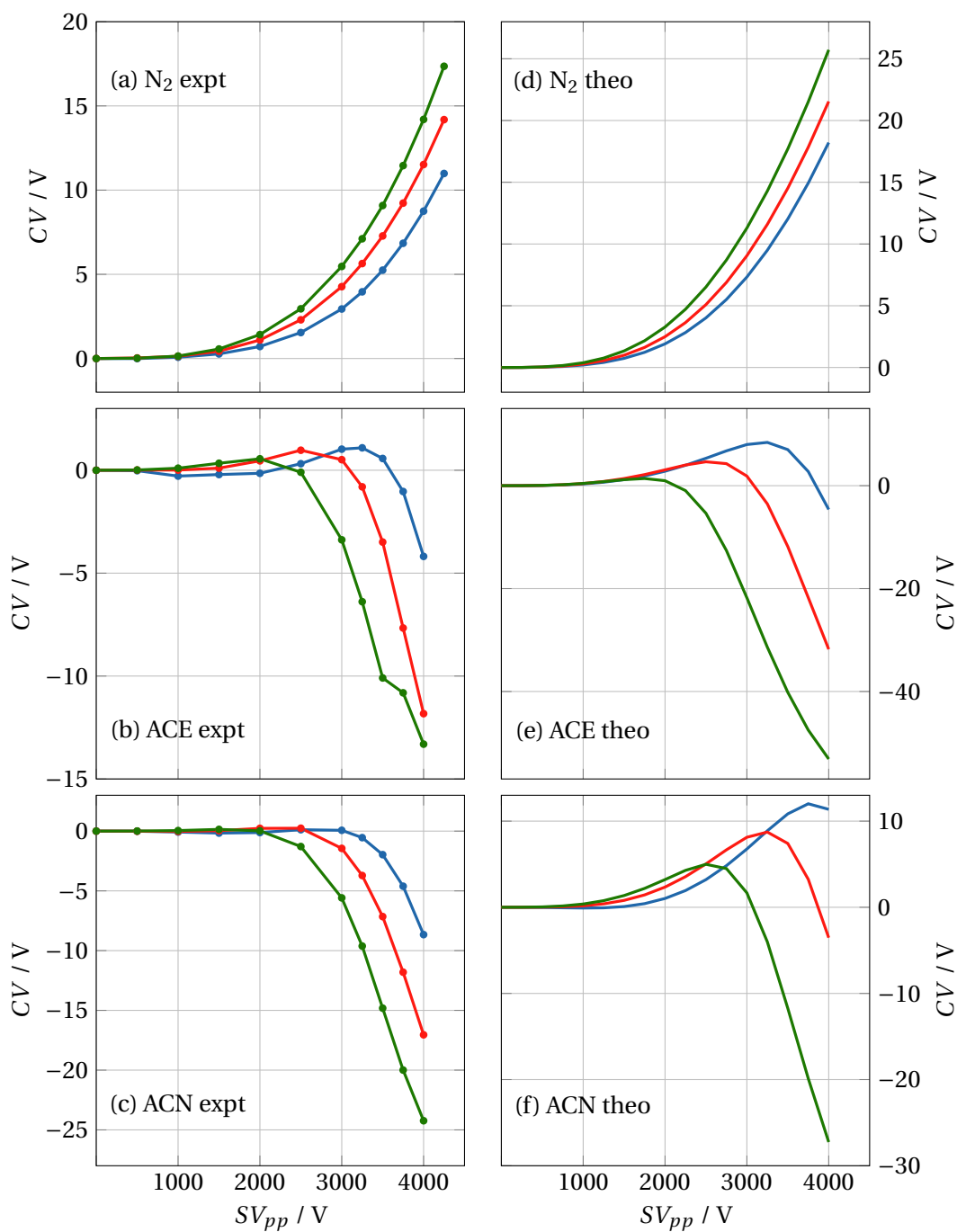


Figure 4.11: Experimental (a-c) and theoretical (d-f) dispersion plots of $\text{N}(\text{CH}_2\text{CH}_2\text{CH}_2)_3\text{Sn}^+$ in a pure N_2 environment and N_2 seeded with 1.5 mol% of ACE and ACN. All panels show three temperatures corresponding to DMS plate temperatures of 150 °C (—), 225 °C (—) and 300 °C (—).

4 Modeling of Differential Ion Mobility

for ion transmission. This is the behavior of a rigid, non-clustering ion. Only when a certain separation amplitude is reached, the dispersion plots rapidly move to negative CV values, normally hinting to dynamic clustering. Furthermore, increasing the bath gas temperature, the transition from positive to negative CV values appears at lower SV_{pp} values. This is very unusual since the increase of the collision gas temperature normally leads to more positive CV values (see for example Fig. 4.7b).

Following the hypothesis stated above, the strong bound $N(\text{CH}_2\text{CH}_2\text{CH}_2)_3\text{Sn}^+\cdots\text{S}$ cluster, where S is either ACE or ACN, survives the whole waveform cycle for low separation field amplitudes. Furthermore, since the binding to one S essentially prevents additional clustering (see Table 4.2), there is no dynamic clustering, leading to the observed Type C behavior. Only when the separation amplitude reaches high enough values that dynamic clustering starts, the CV values shift to the negative. Increasing the background gas temperature, and thus the internal energy, increases T_{eff} (see Eq. (4.4)) and thus dissociation / dynamic clustering starts at lower separation field amplitudes. Thus, the observed Type D behavior is explained by the hypothesized strong clustering.

To further support this model, the theoretical dispersion plots are determined with the method outlined above. As can be seen in Fig. 4.11e and 4.11f, the observed Type D behavior as well as the unusual dependency of the dispersion curves on the background gas temperature is correctly predicted. It should be noted that the agreement is only semi-quantitative in that CV values are overestimated (both in positive and negative direction). Especially for the ACN environment, a much larger Type C maximum is predicted than observed. Reasons for the discrepancy are likely the inaccuracy of the Mason-Schamp equation at high field strengths even with two-temperature theory applied, the negligence of anharmonic effects for the partition functions, neglecting the effect of the added solvents on the collision cross section, and possible non-equilibrium conditions due to the fast changing field - as already discussed above. Nevertheless, this agreement allows to further analyze the calculated ion populations for different experimental conditions (SV_{pp} , T_{bath}).

This is done in Fig. 4.12, where the population, P_n , of the clustered and unclustered ion of the $N(\text{CH}_2\text{CH}_2\text{CH}_2)_3\text{Sn}^+\cdots\text{ACE}$ system, averaged over one waveform cycle, is plotted against the respective separation field amplitude, SV_{pp} . It is clearly visible that the clustered species dominates the ion distribution for low SV_{pp} values. $\langle P_{n=1} \rangle_{\text{wf}} \approx 1$ means that even during the high field phase, no significant declustering occurs. This supports the hypothesis of the clustered ion acting as rigid molecule without dynamic clustering for small SV_{pp} values. Only then the field induced heating becomes large enough (larger separation field amplitudes), dynamic clustering starts: $\langle P_{n=1} \rangle_{\text{wf}}$ falls significantly below unity while $\langle P_{n=0} \rangle_{\text{wf}}$ rises by the same amount, meaning that during the high field portion of the respective separation field cycle, the cluster dissociates. Increasing the effective temperature through heating of the background gas promotes this effect and the point at which dynamic clustering starts shifts to lower SV_{pp} values. It is noted that these points accord well with the region in the (measured) dispersion plots, where the initial Type C behavior switches to Type A.

These critical points in the Type D behavior are thus well captured by the presented model.

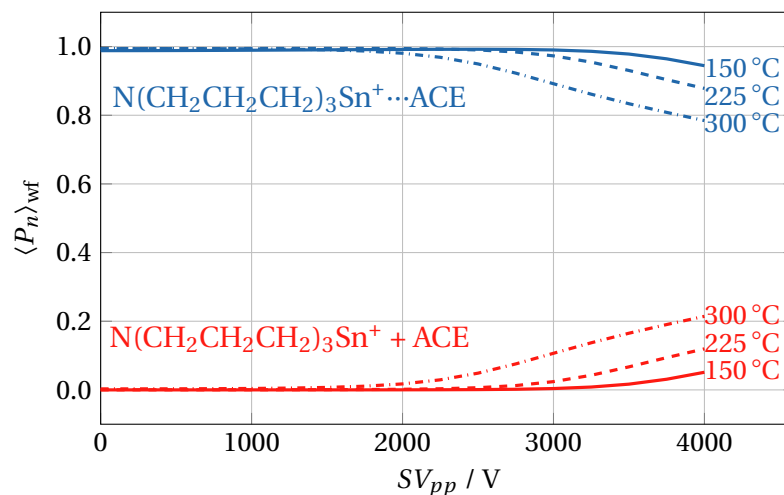


Figure 4.12: Population of the clustered and unclustered tricarbastannatrane ion, averaged over the whole waveform, for multiple separation field amplitudes and background gas temperatures. The clustered species dominates the distribution throughout the whole waveform. Only when SV_{pp} becomes large enough, the average population falls significantly below unity because the cluster dissociates during the high field portion of the waveform.

4.7.3 Conclusion

Tricarbastannatrane cations in a N_2 environment seeded with 1.5 mol% of ACE or ACN show unusual Type D behavior in dispersion plot measurements. This follow-up work confirms the hypothesized explanation for the Type D behavior, i.e., extremely strong and also exclusive clustering of solvents to the ion. The observed clustered species in the mass spectrum (where only $n = 0, 1$ is recorded) as well as the temperature dependency of the Type D curves, i.e., higher bath gas temperatures lead to more negative CV values, are strong evidences for this model since they readily can be explained by the suggested strong clustering. Theoretical modeling of the measured dispersion plots, on the basis of dynamic clustering, show good semi-quantitative agreement with respect to the form of the dispersion curves as well as the bath gas temperature dependency. The modeled binding energies and resulting cluster population distribution, showing strong and exclusive clustering, further support the hypothesized model.

This study resembles yet another strong evidence that dynamic clustering of ions and solvents during the asymmetric waveform inside the DMS cell is a major fac-

4 Modeling of Differential Ion Mobility

tor contributing to the field dependency of the ion mobility. Moreover, the ability to (semi-quantitatively) predict unusual Type D behavior and its background gas temperature dependency with the presented theoretical model further suggests that the major factors of differential mobility are incorporated in the calculations, i.e., hard-sphere effect, the dependency of the CCS on the effective temperature and dynamic clustering with solvent molecules during the waveform cycle.

5 Thermochemistry from MD Simulations

Nomenclature of coordinates in this chapter vary from the rest of this work to be consistent with the nomenclature used in the literature. Some parts of this chapter were published by John Wiley & Sons Ltd under a CC-BY license in

A. Haack, T. Benter, and H. Kersten, “Computational analysis of the proton bound acetonitrile dimer, $(\text{ACN})_2\text{H}^+$ ”, *Rapid Communications in Mass Spectrometry*, DOI: [10.1002/rcm.8767](https://doi.org/10.1002/rcm.8767) (2020)

5.1 Abstract

The theoretical treatment of loosely bound protonated gas phase clusters is rather challenging, since common ab-initio approaches do not correctly reflect the distinct anharmonic conformation of the potential energy surface. This contribution develops a theoretical framework, based on the quasi-harmonic approximation (QHA), to calculate vibrational partition functions and thermochemical properties of small molecular clusters. Multiple molecular dynamics simulations of all studied molecules provide configurational distributions and ensure for sufficient ergodic mixing. For the description of the configurational distribution we use normal coordinates from standard ab initio normal mode analysis, rather than Cartesian or internal coordinates. Possible erroneous effects from large amplitude motion are avoided by transforming the molecule into the Eckart frame for each time step along the trajectory. Furthermore, the Eckart-Sayvetz condition is explicitly enforced for internal rotations, which are well known to falsify QHA results. Partition functions of these motions are then calculated with standard formulas. We apply our treatment to two examples, namely the formation of the proton bound acetonitrile dimer $(\text{ACN})_2\text{H}^+$ and a mixed amine-methanol cluster $[\text{MeNH}_3 + \text{MeOH}]^+$. While in the case of ACN our method performed better than the standard harmonic approximation, it was slightly worse in the latter. Inherent errors and possibilities for improvement are discussed.

5.2 Introduction

Proton bound clusters are ubiquitously present in atmospheric pressure (AP) mass spectrometry as multiple publications, reviews and book articles [25, 29, 31, 244–

246] have shown. Many ionization techniques generate abundant primary ions, e.g., the initially formed N_2^+ in APCI [27, 247, 248], which are prone to subsequent transformation into cluster structures with any dipole moment holding molecule present in the ion source. Main actors are water from ambient air and solvent molecules from a liquid sample injection. Consequently, the presence of charged water [26, 32], methanol [165, 187, 188] and acetonitrile [170, 174–176] cluster distributions have been reported, just to name a few. These somewhat stable ions can further react with neutral analyte (or dopant) molecules to form mixed solvent-analyte clusters, also reported in the literature [32]. Thermodynamics essentially determines whether the analyte is protonated to obtain a $[M + H]^+$ signal, the mixed cluster remains or the analyte is not protonated at all [31, 143, 144]. This includes parameters such as solvent concentration, pressure, ion temperature through electric field strength and ion transfer to the analyzer [32], which always need to be considered and carefully assessed for a comprehensive description. Also the ESI process is largely determined by cluster formation, either due to shrinking droplets (charge residue model [38, 39]) or ejection from the droplets (ion evaporation model [36]). There is also strong evidence in the differential mobility (DMS, FAIMS) community that the periodic formation and decomposition of solvent clusters around the analyte with changing ion temperature in high and low field phases highly contributes to the different observed mobilities [68, 69, 194, 217, 231]. In previous publications [143, 144] we have shown the pivotal role of prolonged cluster chemistry with gas phase additives that strongly enhance or decrease ion signals and alter the observed charge state distribution of peptides or small multiple charged diamines. Despite their importance, in particular regarding ion generation and suppression, often little attention is paid to the cluster chemistry in AP ion sources [25].

A comprehensive theoretical framework to predict dominant cluster species and their thermodynamic properties would be of great benefit for more tailored experimental designs. Indeed, theoretical investigations of reaction mechanisms or thermochemistry are increasingly used owing to increasing computing power, improving method accuracy and not least the possibility to employ black-box program packages, making it available to a broader spectrum of users. For proton bound clusters, however, standard methods often fail. This is mainly attributed to the commonly applied harmonic approximation (HA) for the description of the potential energy surface (PES). If the potential is highly anharmonic, has two or more wells or would better be described by a quartic, rather than a quadratic function - all long known characteristics for H-bonded systems [186, 249] - then a HA based PES will introduce large errors to the derived energy levels, partition functions and thermochemical properties like enthalpy, entropy and Gibbs enthalpy. Additionally, internal rotations around the H-bond often show very low barriers opening the configurational distribution to occupy multiple wells. In a recent publication [144] we stumbled upon an especially difficult case, the proton bound acetonitrile dimer $(ACN)_2H^+$. Firstly, the course of potential energy for the oscillating proton between the two ACN nitrogen atoms essentially followed a quartic form. Secondly, the energy barrier for the internal rota-

tion of the two methyl groups with respect to each other was reported to be as little as $1.26 \times 10^{-7} E_h$, which makes it a free rotor rather than a vibration. Redoing these calculations with an increased accuracy even lowered the barrier to $9.6 \times 10^{-8} E_h$ (see Figure 5.3a). Different methods have been introduced over the years to explicitly tackle these issues. One famous approach to the anharmonicity of the potential is the vibrational perturbation theory to 2nd order (VPT2), where the potential along the normal coordinates is described by the harmonic term and an additional anharmonic one, also incorporating the pair-wise coupling of the modes [103]. These vibrational anharmonic constants are computed from numerical 3rd and 4th derivatives along the normal coordinates [104, 105] and corrections are made to the normal frequencies, the zero point energy (ZPE) and the rotational constants since ro-vibrational coupling, like centrifugal distortion, can be incorporated in this treatment. An improved theory (HDCPT2), paying special attention to symmetry, resonances and internal rotation, has been reported [106] and is implemented in the GAUSSIAN 16 program package [226]. This treatment, as all perturbative approaches, only works well with a relatively small anharmonic contribution compared to the harmonic part. However, this is not the case for a quartic potential and double well potentials cannot be accounted for either. For a small number of problematic modes, the potential can be scanned and the nuclear Schrödinger equation can be explicitly solved for these modes, e.g., by means of the Fourier Grid Hamiltonian method [107]. This approach is fairly accurate for any shape of potential energy curve, however, it neglects the coupling between modes - although the Fourier Grid Hamiltonian method has been extended up to three dimensions [108] - and it is rather cumbersome, since it requires a scan for each single mode. It is worth noting that this procedure can also be done by scanning any internal coordinate, not only the normal coordinates [109, 110]. With the free available NuSol program [111] the wavefunction of an oscillating proton can readily be calculated in three-dimensional space, provided that its motion is completely decoupled from any other degree of freedom. A more general treatment of the nuclear motion is the Vibrational Self Consistent Field (VSCF) approach [250], which has been combined with electronic structure calculations and perturbation-theoretic extensions [251]. Herein the nuclear Schrödinger equation is solved around the equilibrium from a full grid or polynomial representation of the PES under the assumption that the total wavefunction is a product of the wavefunctions of each mode. In this way, one obtains a set of single-mode Schrödinger equations based on an effective potential, which takes into account the impact of all other modes. Hence, the wavefunctions are iteratively determined until this effective potential converges. In principle, this is a fairly accurate approach, especially with the pairwise interactions between the normal modes. The bottleneck however, is the representation of the PES through grid points in one to four dimensions. This is very time consuming even for fairly small systems (< 10 atoms).

In this work we estimate partition functions and derive thermochemical data for proton bound clusters with the quasiharmonic approximation (QHA). This formalism was introduced by Karplus and Kushick [234–236], first in 1981, and was origi-

nally designed for macromolecules. Pros and cons of the QHA approach for the specific case of proton bound clusters will be discussed with respect to computational cost and accuracy.

5.3 Theory and Methods

5.3.1 The quasiharmonic approximation (QHA)

Following closely the description of Levy *et al.* [235], the *classical* partition function of a molecule containing M_N atoms can be expressed as

$$Z = \frac{1}{h^{3M_N}} \int \exp(-\beta H) d\mathbf{p}_x d\mathbf{x} \quad (5.1)$$

with h as the Planck constant, $\beta = 1/k_B T$ the inverse temperature with k_B as the Boltzmann constant and $H = \mathcal{E}_{\text{kin}}(\mathbf{p}_x) + \mathcal{E}_{\text{pot}}(\mathbf{x})$ the classical Hamiltonian depending on $3M_N$ Cartesian coordinates \mathbf{x} and their conjugate momenta \mathbf{p}_x . Separating out translational and overall rotational motion (neglecting ro-vibrational coupling) leaves the classical vibrational partition function

$$Z_{\text{vib}} = \frac{1}{h^{3M_N-6}} \int \exp(-\beta H_{\text{int}}) d\mathbf{p}_q d\mathbf{q} \quad (5.2)$$

with integration over the internal coordinates \mathbf{q} and their conjugate momenta \mathbf{p}_q . The internal Hamiltonian H_{int} reads

$$H_{\text{int}} = \frac{1}{2} \mathbf{p}_q^T \cdot \mathbf{G}^{-1} \cdot \mathbf{p}_q + \frac{1}{2} \mathbf{q}^T \cdot \mathbf{F} \cdot \mathbf{q} \quad (5.3)$$

with the kinetic and potential energy matrices \mathbf{G} and \mathbf{F} , respectively. \mathbf{G} is also known as Wilson's G-matrix [100] and includes the transformation between Cartesian and internal coordinates

$$G_{ij} = \sum_{k=1}^{3M_N} \frac{1}{m_k} \frac{\partial q_i}{\partial x_k} \frac{\partial q_j}{\partial x_k} \quad (5.4)$$

The derivatives in Eq. (5.4) depend on \mathbf{q} , consequently the G-matrix changes with the geometry of the molecule [110]. Nevertheless, for simplicity this matrix is usually treated as a constant, taking the derivatives at the equilibrium geometry \mathbf{q}_0 . Up to here, the formalisms for the HA and the QHA are the same. The essential difference of both approaches lies in the definition of the force constant matrix \mathbf{F} . For the harmonic approximation \mathbf{F} is defined as

$$F_{ij}^{\text{HA}} = \left(\frac{\partial^2 V}{\partial q_i \partial q_j} \right)_{\mathbf{q}_0} \quad (5.5)$$

whereas Karplus and Kushick [234] suggested

$$F_{ij}^{\text{QHA}} = k_B T [\sigma^{-1}]_{ij} \quad \text{with} \quad (5.6)$$

$$\sigma_{ij} = \langle (q_i - \langle q_i \rangle) (q_j - \langle q_j \rangle) \rangle$$

for the quasiharmonic approximation. In the latter, σ is the covariance matrix of the internal coordinates and its averages $\langle \dots \rangle$ are taken over an ensemble at temperature T , readily obtainable from molecular dynamics (MD) or Monte Carlo (MC) simulations around the equilibrium geometry.

With this definition of the F-matrix, the configurational distribution of the molecule becomes a $3M_N - 6$ dimensional Gaussian and the integral in Eq. (5.2) can be analytically solved (after integration over the momenta and the assumption that \mathbf{G} does not change with \mathbf{q}):

$$Z_{\text{vib}} = \left(\frac{\sqrt{2\pi k_B T}}{h} \right)^{3M_N - 6} \sqrt{\det(\mathbf{G}^{-1} |_{\mathbf{q}_0})} \sqrt{(2\pi)^{3M_N - 6} \det(\sigma)} \quad (5.7)$$

This is the key assumption in the QHA. The Gaussian (given by the equilibrium value \mathbf{q}_0 and the covariance matrix σ) represents the configurational distribution, determined on the full anharmonic surface and considering pairwise correlation of coordinates, while remaining in the harmonic formalism.

As suggested by Karplus and Kushick and pointed out by recent reviews [252, 253] on the QHA, internal coordinates, such as bond-angle-torsion (BAT) coordinates, yield considerably more accurate results than the respective Cartesian coordinates, which are commonly used by standard software packages. Despite this clear suggestion, we explicitly choose for normal coordinates instead of BAT coordinates, for reasons explained in the next section.

5.3.2 Implementation and special considerations

As pointed out in the literature [254], only a valid set of coordinates can be used to uniquely define Wilson's G-matrix. This is tricky, since in the $(\text{ACN})_2\text{H}^+$ cluster, all but the methyl hydrogen atoms are positioned on the z -axis. Consequently, the definition of a set of BAT coordinates is rather impractical, since torsion angles will eventually lead to singularities at the equilibrium geometry. Instead, we decided to use normal coordinates Q_i as a special kind of internal coordinates. Those are readily obtained from ab initio normal mode analysis (NMA) at the equilibrium geometry. GAUSSIAN 16 provides the normal coordinates in form of derivatives $\partial x_k / \partial Q_i$, scaled with the reduced mass μ_i of the respective mode, such that the sum of the squares of the Cartesian displacements is 1 [96]. With this definition, the inverse of Wilson's G-matrix can be calculated as follows

$$[G^{-1}]_{ij} = \sum_{k=1}^{3M_N} \frac{m_k}{\sqrt{\mu_i \mu_j}} \frac{\partial x_k}{\partial Q_i} \frac{\partial x_k}{\partial Q_j} \quad (5.8)$$

According to Eq. (5.8) the matrix is equal to the unity matrix \mathbf{I} , i.e., $\mathbf{G}^{-1} = \mathbf{I}$, which compiles with the definition of mass-weighted normal coordinates. That is, the square

5 Thermochemistry from MD Simulations

of their temporal derivatives directly yields the kinetic energy (right side of the following equation):

$$2\mathcal{E}_{\text{kin}} = \sum_{i=1}^{3M_N-6} \sum_{j=1}^{3M_N-6} [G^{-1}]_{ij} \frac{\partial Q_i}{\partial t} \frac{\partial Q_j}{\partial t} = \sum_{i=1}^{3M_N-6} \left(\frac{\partial Q_i}{\partial t} \right)^2 \quad (5.9)$$

This is only equal to the double sum, which is the kinetic energy for any given set of internal coordinates, when $\mathbf{G}^{-1} = \mathbf{I}$.

At this point only the covariance matrix σ is left for computing the vibrational partition function based on Eq. (5.7). For that, the MD trajectory calculations require the conversion of mass-weighted Cartesian coordinates to normal coordinates at each time step t . This is simply implemented by the transposed, orthonormal transformation matrix between Cartesian and normal coordinates (given by the derivatives $\partial x_k / \partial Q_j$)

$$Q_i(t) = \sum_{k=1}^{3M_N} \frac{\partial Q_i}{\partial x_k} \sqrt{m_k} (x_k(t) - x_k(0)) \quad (5.10)$$

However, since normal coordinates are just a special linear combination of Cartesian coordinates, they suffer from some of the same problems. Overall translation and rotation would erroneously contribute to the normal coordinates although they do not belong to the $3M_N - 6$ internal degrees of freedom. In addition, internal rotations (as do any large amplitude motion) severely change the position of atoms such that their overlap with the normal coordinate is incorrectly diminished. Hence, these motions have to be separated out from the MD trajectory before performing the transformation according to Eq. (5.10).

One can separate out the overall translation and rotation by shifting and rotating the coordinate system to the Eckart frame [97]. For each geometry in the MD trajectory the center of mass (CoM) is set to zero and three Eckart vectors \mathbf{F}_1 , \mathbf{F}_2 and \mathbf{F}_3 are determined to rotate the coordinate system to the Eckart frame $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ [255]. Finally, the Cartesian displacements $\boldsymbol{\rho} = \mathbf{x} - \mathbf{x}(0)$ are then converted into displacements in the Eckart frame $\boldsymbol{\rho}^{(E)}$ [256]. For a detailed description of this transformation the reader is referred to [Appendix A](#) of this chapter.

As pointed out by Sayvetz [98], internal rotations have to be separated out alike the overall translation and rotation. For any arbitrary internal rotor, the Eckart-Sayvetz conditions are given by Kirtman [257–259]. After defining the internal rotor and the rotational axis, we rotate the two fragments (internal rotor and framework) by an angle α and β , respectively, such that the Eckart-Sayvetz condition is fulfilled. A numerical minimization procedure determines those two angles. For a detailed formulation of the condition and the determination of the angles see [Appendix B](#) of this chapter. Note that the internal rotation correction is done prior to the Eckart frame transformation since the internal rotation would eventually lead to an erroneous rotation of the Eckart frame.

For M_{IR} internal rotations, the number of vibrations and hence the number of normal coordinates reduces to $3M_N - 6 - M_{\text{IR}}$. Thus, the dimensionality of the matrices

in Eq. (5.7) changes and the powers have to be reduced as well.

This reduced vibrational partition function is multiplied by the partition function of the internal rotations, which in case of negligible barriers are calculated according to

$$Z_{\text{IR}} = \frac{1}{\sigma_{\text{IR}}} \sqrt{\frac{8\pi^3 I_{\text{IR}} k_B T}{h^2}} \quad (5.11)$$

with σ_{IR} the symmetry number of the internal rotor (e.g., 3 for methyl group rotations) and I_{IR} the reduced moment of inertia of the rotor around the internal rotational axis. For symmetric tops attached to a framework, I_{IR} can be calculated with

$$I_{\text{IR}} = I' \left(1 - \sum_{\lambda \in \{a,b,c\}} \frac{\alpha_i^2 I'}{I_{\lambda\lambda}} \right) \quad (5.12)$$

In this equation, $I' = \sum m_{l'} d_{l'}^2$ represents the moment of inertia of the top around the axis, with the sum running over all rotating atoms l' with masses $m_{l'}$ and their distances $d_{l'}$ with respect to the rotational axis. The sum in Eq. (5.12) runs over the three principal axes a , b , c with the principal moments of inertia $I_{\lambda\lambda}$ of the entire molecule (cf. Subsection 2.2.2) and α_i the direction cosines of the rotational axis onto the principal axes. For non-symmetric tops, the formulas are more complicated but reported in the literature [260–262].

Having separated out those non-vibrational motions the transformation according to Eq. (5.10) can now be performed with the new Cartesian displacements $\boldsymbol{\rho}^{(E)}$. From the MD trajectories in normal coordinates, the covariance matrix and its determinant are readily calculated and the (reduced) vibrational partition function can be obtained by means of Eq. (5.7). As pointed out by Rojas et al. [236], a correction can be made to the configurational partition function Z_c^{QHA} , which is the last term in Eq. (5.7). By calculating the QHA F-matrix from the obtained covariance matrix via Eq. (5.6) and using $V^{\text{QHA}} = 1/2 \mathbf{Q}^T \mathbf{F}^{\text{QHA}} \mathbf{Q}$ (see Eq. (5.3)), the correction can be calculated from the difference of the QHA potential V^{QHA} and the "real" potential energy V obtained from the MD trajectory according to

$$\frac{Z_c^{\text{QHA}}}{Z_c} = \left\langle e^{+\beta(V - V^{\text{QHA}})} \right\rangle \quad (5.13)$$

Z_c is the "real" configurational partition function and the expression $\langle \dots \rangle$ averages over every single time step of the MD simulation. The inverse of this number, termed f_{corr} , can then be multiplied with Z_{vib} .

Finally, we simply multiply the vibrational, rotational (obtained by standard approximations), translational and, if necessary, internal rotational partition functions to obtain the total partition function Z_{tot} , from which the Gibbs enthalpy corrections G_{corr} is readily calculated with

$$G_{\text{corr}} = -k_B T \ln(Z_{\text{tot}}) \quad (5.14)$$

5 Thermochemistry from MD Simulations

It should be mentioned that the ZPE of the molecule never occurs in this formalism since MD simulations assume classical behavior of the nuclei and also the partition functions are determined from classical statistical mechanics.

Once G_{corr} and the electronic energy \mathcal{E}_0 for all reactants and products are calculated, the Gibbs enthalpies of reaction, e.g., for a cluster association reaction, are obtained according to

$$\Delta_R G = \sum_{\text{products}} (\mathcal{E}_0 + G_{\text{corr}}) - \sum_{\text{reactants}} (\mathcal{E}_0 + G_{\text{corr}}) \quad (5.15)$$

5.3.3 Electronic structure and MD simulations

All calculations were carried out with the GAUSSIAN 16 program package [226]. Optimization and NMA as well as PES scans were conducted with the density functional theory applying the B3LYP functional [86, 87] with GD3(BJ) dispersion correction [92] using the def2-TZVPP basis set [75]. The empirical dispersion correction was applied since long range contributions are important in non-covalent bonded systems. Integration grids and convergence criteria were manually set to more grid points and tighter thresholds, respectively. For the MD simulations we decided upon ab initio molecular dynamics (AIMD) for a better description of the bonding situation, instead of using simple force fields. The implemented Atom Centered Density Matrix Propagation (ADMP) [115–117] routine in GAUSSIAN 16 uses a Car-Parrinello approach [114], which is much faster than solving the full electronic Schrödinger equation at each time step. To further speed up the calculations, the smaller def2-SVP basis set was used. 15000×0.2 fs time steps accounted for a total simulation time of 3 ps. The total initial kinetic energy of the molecule in its equilibrium position was randomly assigned to the nuclei with $(3M_N - 6)k_B T$ so that on average each mode would have a total energy of $k_B T$ as should be the case for a thermalized harmonic oscillator. For each system at least three trajectories were calculated with different initial kinetic energy assignments to account for the lack of ergodic mixing in a single MD trajectory.

5.4 Results and Discussion

5.4.1 $(\text{ACN})_2\text{H}^+$

Geometry optimization and NMA Geometry optimization and NMA for the acetonitrile dimer were carried out at the B3LYP-GD3BJ/def2-TZVPP level of theory. As apparent from Fig. 5.1a the equilibrium structure belongs to the D_{3d} point group with the proton (hereafter labeled H^+) as the inversion center symmetrically lying between the two nitrogen atoms. All but the methyl hydrogen atoms are located on the same axis and the methyl groups are found in a staggered conformation. Table 5.1 summarizes all the geometrical parameters of the equilibrium structure. Early theoretical investigations by Deakyne *et al.* [171] suggested an asymmetric equilibrium structure of the dimer to be energetically slightly more favorable than the symmetric

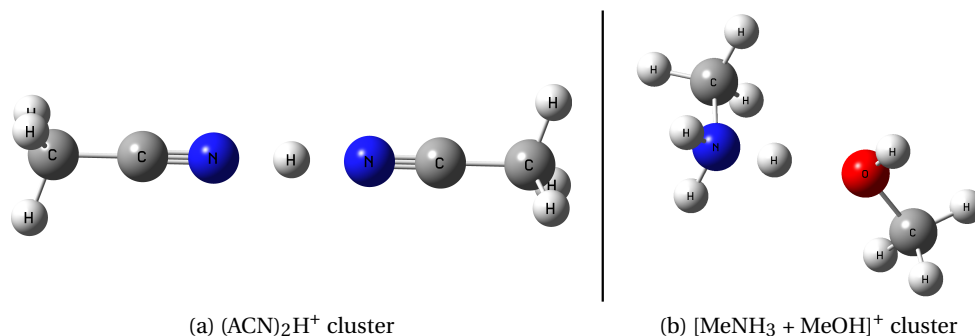


Figure 5.1: B3LYP-GD3BJ/def2-TZVPP optimized geometries of the $(\text{ACN})_2\text{H}^+$ and $[\text{MeNH}_3 + \text{MeOH}]^+$ clusters. While the protonated ACN dimer is highly symmetric (D_{3d} point group), the amine cluster shows no symmetry elements. In case of the $(\text{ACN})_2\text{H}^+$ cluster the H^+ forms an inversion center with equal distance to both binding nitrogen atoms, whereas in case of the amine-methanol cluster the binding proton is closer to the nitrogen atom (1.061 43 Å) than to the oxygen atom (1.618 39 Å). This is expected according to the general difference in basicity between those two functional groups.

Table 5.1: Optimized geometrical parameters for the $(\text{ACN})_2\text{H}^+$ cluster (D_{3d} point group).

Parameter	$R(\text{H}^+, \text{N})$	$R(\text{N}, \text{C})$	$R(\text{C}, \text{C})$	$R(\text{C}, \text{H})$	$A(\text{H}, \text{C}, \text{H})$
Value / Å or °	1.262 02	1.142 36	1.443 62	1.090 25	109.463

one. However, as already shown in a previous publication [144], we cannot confirm their results.

The normal modes, their frequencies and irreducible representations can be found in Table 5.5. As mentioned before, the lowest mode describes an internal rotation of the methyl hydrogens. Because of the large distance, there is basically no interaction and the barrier height is negligible. The negative frequency found in the GAUSSIAN 16 result is very small and rather deemed to be zero, which reasons the assumption of a free rotor. Two also very important modes are Q_6 and Q_7 describing the asymmetric and symmetric stretching modes of the two $\text{N}-\text{H}^+$ distances, respectively. In the first case, the H^+ moves back and forth between two fixed ACN molecules (a_{2u} symmetry) and in the second case the two ACN move relative to the fixed H^+ (a_{1g} symmetry). Fig. 5.2 shows a 2D PES scan (in C_{3v} symmetry) with the two $\text{N}-\text{H}^+$ distances varied and all other parameters allowed to relax. The two modes are indicated by the dashed red lines whereas the solid red line represents the minimum energy paths of the dissociation towards $(\text{ACN})\text{H}^+ + \text{ACN}$. A closer look at this PES clearly illustrates

the problems with the harmonic approximation. First, the potential along Q_6 is very poorly described by a parabola, but rather by a quartic slope [144]. Such a shallow potential drives exceptionally large amplitudes, which means that the asymmetric stretch does in fact not follow the straight line but will essentially bend along the valley. This is interpreted as a coupling of the modes Q_6 and Q_7 . Second, while $Q_6 = 0$ is a minimum for $Q_7 = 0$, it actually is a maximum for $Q_7 > 0$ (towards top right) and the potential along Q_6 becomes a symmetric double well potential (see Fig. 5.3b). Consequently, the calculated harmonic value of 201.4 cm^{-1} for Q_6 is not reliable due to the strong crosstalk from Q_7 . Experimentally a more trustworthy value for Q_6 might well be accessible with Infrared-Multiphoton Dissociation (IRMPD) spectroscopy, since the large change of dipole moment indicates very strong infra-red absorption for this mode. GAUSSIAN 16 calculations predict an extraordinarily high integrated absorption cross section on the order of 7800 km mol^{-1} .

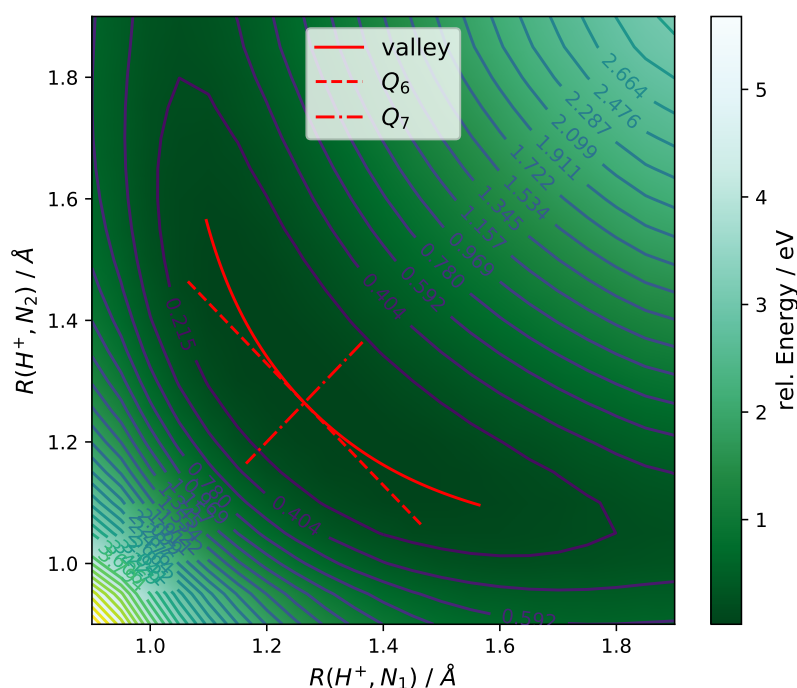


Figure 5.2: Relaxed potential energy surface scan of the two $\text{N}-\text{H}^+$ distances for the $(\text{ACN})_2\text{H}^+$ cluster. Q_6 and Q_7 are the normal modes describing the asymmetric and symmetric stretching modes of the two $\text{N}-\text{H}^+$ distances, respectively. The solid red line shows the lowest energy path towards the two dissociation channels. The molecules geometry at every point on this surface belongs to the C_{3v} point group, whereas the diagonal (along Q_7) belongs to the D_{3d} point group.

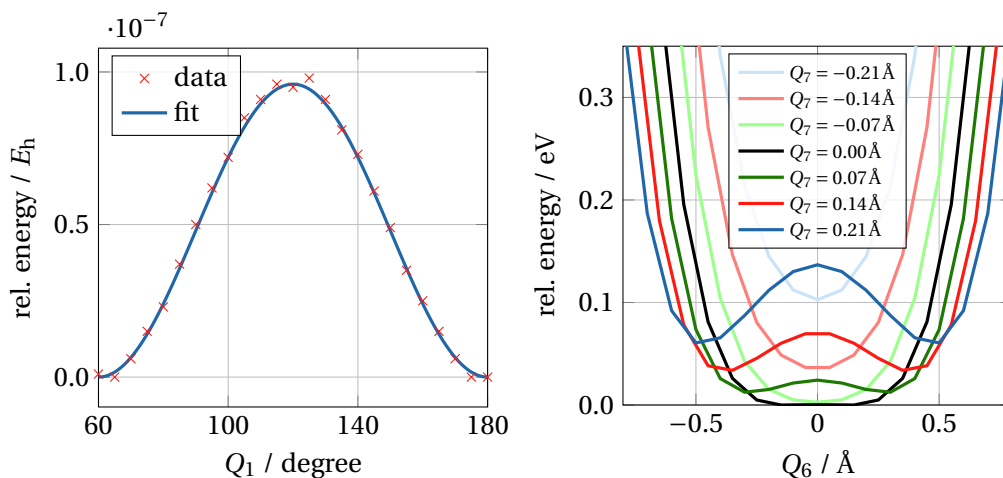


Figure 5.3: PES scans along the internal rotation (Q_1) and the H^+ -movement along the $N-H^+-N$ bond (Q_6 for different positions of Q_7) of the $(ACN)H^+$ cluster. Compared to Fig. 3.8a, the Q_1 scan is performed at a better SCF accuracy. For Q_6 , the double-well character for elongated values of Q_7 becomes apparent (cf. also Fig. 5.2).

Treatment of the vibrations outside the HA The GAUSSIAN 16 software package does allow for vibrational calculations with anharmonicity contributions based on vibrational perturbation theory. However, the implemented VPT2 method gave unreliable results for the investigated cluster structures as well. We performed a partial anharmonic treatment considering only modes Q_2 to Q_{33} since the internal rotation (Q_1) is handled as a free rotor. Also, its negative frequency leads to complications in the anharmonic treatment. One striking result is that the fundamental frequency of Q_6 is calculated to be 8639 cm^{-1} , which is surely not realistic. As stated above, VPT2 can handle only small anharmonic contributions, a quartic potential, however, largely differs from a parabola. Second, the cubic and quartic force constants obtained by numerical differentiation, are often not consistent. We observed values for F_{ijk} and F_{jik} differing by up to 200% though both should in fact be equal. Similar significant deviations were observed for cubic and quartic force constants of degenerate modes although symmetry reasons enforce the exact same values. Above all, these results led us to the conclusion that force constants from numerical differentiation are not consistent for the studied cluster structures and consequently the VPT2 method was deemed inapplicable.

Before following up with the final QHA approach we investigated the aforementioned PES scan method for the moving proton. As stated, Q_6 describes the propagation of H^+ between the two ACN molecules along the z -axis and Q_{18} / Q_{19} are the two degenerate modes (e_u symmetry) with the proton moving along the x - and y -axis, i.e., orthogonal to the molecular axis. In all three modes, the two ACN molecules

virtually do not move. In that way, we calculated a 3D PES of the propagating proton with fixed ACN molecules inside a $41 \times 0.02 \text{ \AA}$ grid in z - and a $61 \times 0.02 \text{ \AA}$ grid in x - and y -direction. Fortunately, symmetry reasons allowed to reduce the calculation effort to one eighth of the cube. The derived potential was handed over to the NuSol program together with the required mass input of 1 amu. The 3D nuclear wavefunction was then calculated with the DVR algorithm [263] and we obtained fundamental transition frequencies of 1141 cm^{-1} for the z -direction and 1157 cm^{-1} for the two orthogonal directions. The vibrational modes Q_{18} and Q_{19} with virtually quadratic progression of the potentials are close to the harmonic value of 1259 cm^{-1} . In contrast, the frequency for Q_6 significantly deviates from the 201.4 cm^{-1} as obtained in the harmonic case. However, as pointed out before, the validity of the latter value seems very questionable since the harmonic treatment entirely neglects the impact of Q_7 on this mode. A final, interesting observation results from the fact that the fundamental frequencies of Q_{18} , Q_{19} and Q_6 are very similar, almost degenerate. This leads to a spatial distribution alike the 1-electron wavefunctions of the hydrogen atom, which indeed can be expected, since in both cases a single particle moves in a nearly centrosymmetric potential. The respective potential can be found in Fig. 5.4.

QHA results Most promising to tackle the vibrational partition functions of the clusters was the QHA method. Herein, the overall translational and rotational partition functions were calculated from standard approximations from quantum mechanics as found in the GAUSSIAN 16 thermochemistry output [96]. For the two fragments ACN, $(\text{ACN})\text{H}^+$ as well as for the proton bound cluster $(\text{ACN})_2\text{H}^+$ we conducted geometry optimization, NMA and three MD simulations with different starting conditions. For the two fragments only the Cartesian coordinates along the MD trajectories were Eckart frame corrected, whereas internal rotation was additionally separated out for the cluster (prior to the Eckart frame correction) according to the numerical approach described in Appendix B. Since both fragments rotate with respect to the fixed coordinate system, the correction was applied for both fragments to prevent an overall rotation of the Eckart frame. Principally the Eckart treatment is designed to manage such situations, however, we noticed severe numerical instabilities for the following reason: When the molecule rotates 60° , the Eckart frame jumps back to its original position due to the C_3 -symmetry of the molecule. At this point small hydrogen vibrations, independent from the rotation, cause numerous jumps around the 60° mark, which leads to virtually random movement and thus large errors in the transformation to the normal coordinates. Accordingly, it is easier to explicitly avoid the overall rotation by correcting the internal rotation of both fragments. The methyl hydrogen atoms were chosen to determine the center of mass (CoM) of the internal rotor and the axis of rotation was chosen to be the vector drawn from the CoM of the entire molecule to the CoM of the internal rotor $\mathbf{R}' - \mathbf{R}$. Since in the equilibrium geometry the rotor is balanced, σ' was set to zero (see Appendix B).

To demonstrate the Eckart-Sayvetz condition, Fig. 5.5 illustrates the result of a sin-

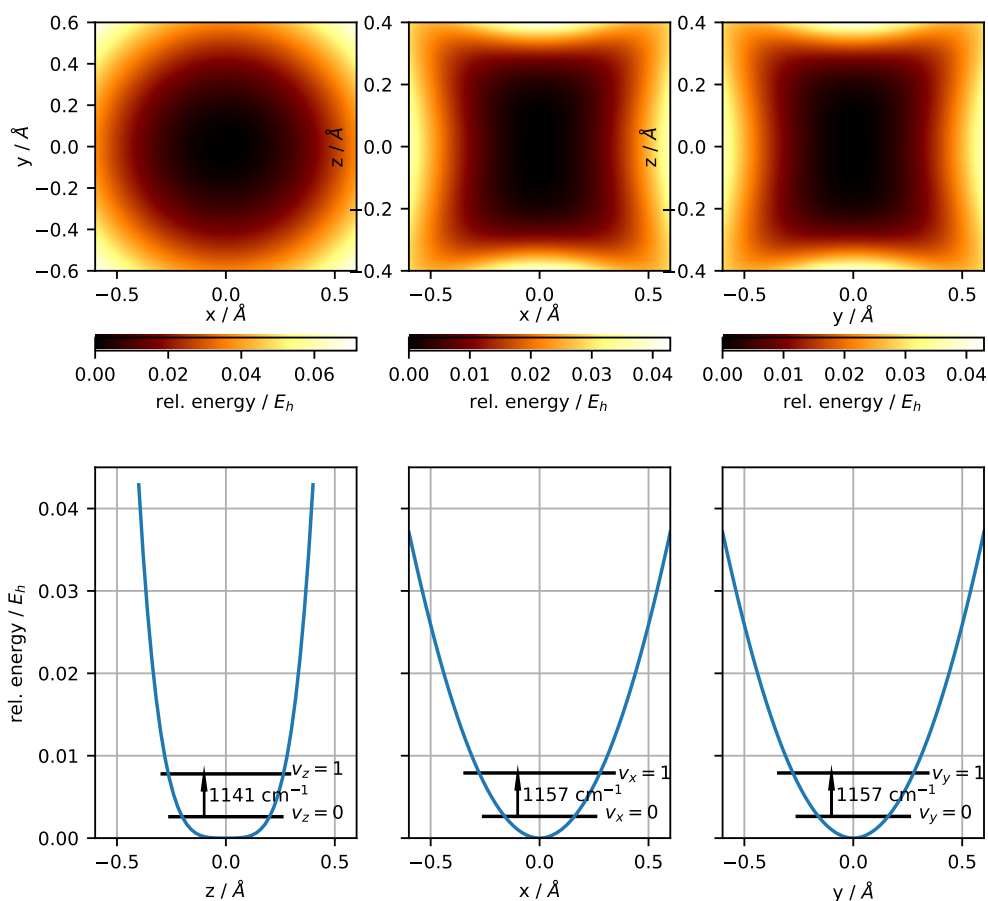


Figure 5.4: 3D PES scans for the H^+ movement between the two stationary ACN molecules. The z -axis connects the two nitrogen atoms while the x - and y -axes are perpendicular to the molecular axis. The H^+ lies in the origin of the coordinate system. Since the x - and y -axes are identical under rotation, the xy potential is circular symmetric and the xz and yz potentials are identical. 1D cuts through the potential are also depicted where the other two coordinates are held at zero. The fundamental frequencies of each mode are shown as well.

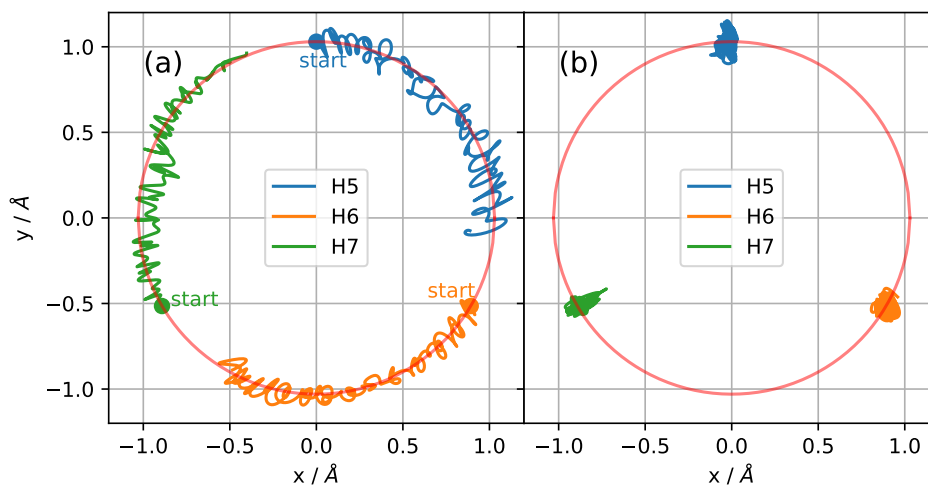


Figure 5.5: $\mathbf{l}_{\alpha'}$ vectors of three methyl hydrogen atoms projected onto the xy -plane (a) before and (b) after the internal rotation correction. The starting points are indicated, the trajectory is shown for 2000 time steps. The clockwise rotation of the methyl group is corrected while retaining all other hydrogen vibrational modes.

gle MD simulation, with the $\mathbf{l}_{\alpha'}$ vectors of three methyl hydrogen atoms projected onto the xy -plane without (Fig. 5.5a) and with (Fig. 5.5b) the correction. The two diagrams clearly show that the clockwise rotation of the methyl group is correctly removed while the other vibrations are retained. Note that a simple rotation of the atoms, e.g., with the H5 positioned on the y -axis, will not have the same effect since its vibration would impact the motion of the two other atoms. Only a correction with the Eckart-Sayvetz condition retains the original motion orthogonal to the rotation.

Subsequent to the internal rotor correction, we used the Eckart frame treatment for each time step to then calculate the normal coordinates from the Cartesian displacements according to Eq. (5.10). Then the covariance matrices σ of all three MD trajectories were averaged and used to calculate the (reduced) vibrational partition function via Eq. (5.7) for the cluster as well as for the fragments. Table 5.2 lists the results from the QHA treatment. The internal rotation partition function of the cluster was calculated from Eq. (5.11) with a reduced moment of inertia of $I_{\text{IR}} = 1.602 \text{ amu}\text{\AA}^2$ and a symmetry number of $\sigma_{\text{IR}} = 3$. There are a couple of interesting points to discuss: (1) In view of the fact that the covariance matrix is the key for the entire QHA, its determinant should be assessed with respect to consistency. In particular its convergence over the MD simulations conducted should be confirmed, since the matrix depends on the averaged fluctuation of the normal modes around their equilibrium positions. (2) The validity of the QHA with assumed Gaussian distributions for each normal mode should be proven.

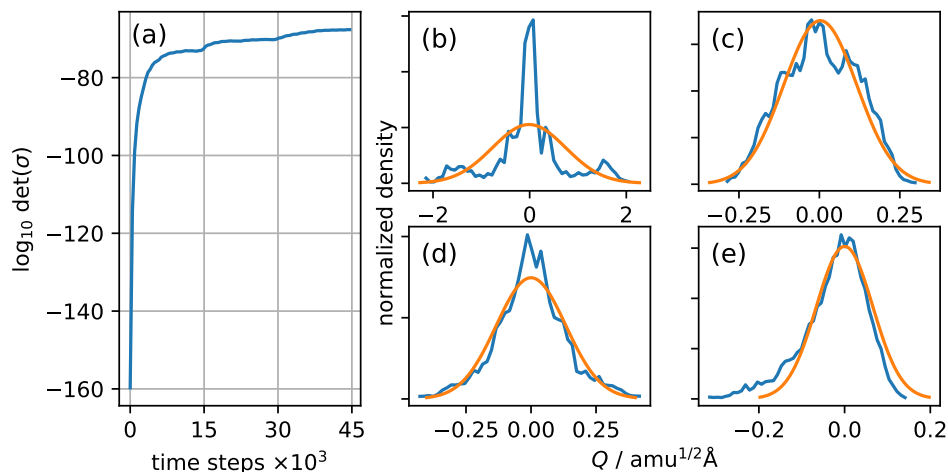


Figure 5.6: Diagnostic plots for the QHA for the ACN system. (a) The convergence of the determinant of the covariance matrix σ and (b)-(e) examples for the MD and derived QHA configurational distributions along selected modes ((b) Q_2 , (c) Q_6 , (d) Q_{18} , (e) Q_{29}) for the $(\text{ACN})_2\text{H}^+$ cluster.

Fig. 5.6a shows the convergence of the covariance matrix for the ACN dimer in dependence of the number of time steps. The determinant is shown in a logarithmic scale and thus the values are linearly proportional to the calculated Gibbs enthalpy correction. The plot clearly shows that a single MD simulation with 15 000 time steps is not enough for the determinant to converge, due to lacking ergodicity. Any new MD simulation with different starting conditions significantly adds to $\det(\sigma)$ and the slope rapidly levels off. Eventually, the determinant sufficiently converges within three MD simulations, though more time steps would certainly increase the accuracy.

Fig. 5.6b-e show the configurational distributions for four selected modes and their respective Gaussian fittings. Their standard deviations were defined by $\sqrt{\sigma_{ii}}$ from the QHA results. In most cases this type of fitting sufficiently describes the real distribution, as exemplary shown for Q_6 and Q_{18} . Deviations from an ideal Gaussian function are observed for Q_{29} and Q_2 . The latter forms two smaller satellites with a distance of $\pm 1.5 \text{ amu}^{1/2}\text{\AA}$ around a relatively broad main peak with $\pm 0.5 \text{ amu}^{1/2}\text{\AA}$ base width. This multimodal character is combined to a broader unimodal configurational distribution. As a result, the QHA predicts a Gaussian function for Q_2 with a rather large standard deviation of $0.757 \text{ amu}^{1/2}\text{\AA}$. The persistence of the two satellites with increasing sampling time is still in question. In case of actual distribution density, though, the QHA will take them into account, whereas the HA reflects the narrow distribution only. Q_{29} is the other mode that needs to be looked at in terms of the Gaussian representation. The apparent asymmetry in the MD distribution cannot be exactly traced by a Gaussian function. Instead, the QHA distribution accounts for

this asymmetry by broadening. This leads to a slower descent towards positive values than the actual distribution, but the left flank is matched nicely for a large portion of the distribution. In conclusion, the QHA accounts for anharmonic behavior by simply broadening the Gaussian functions, which fully complies with the concept of an ensemble average.

As shown in Table 5.2 we noticed a quite large deviation for the correction factor f_{corr} from Eq. (5.13) between the two fragments and the cluster. A value closer to unity relates to a better description of the configurational partition function by the QHA. For the two fragments the treatment is probably sufficient since f_{corr} is relatively large. For the cluster, however, a two orders of magnitude smaller value for f_{corr} is obtained. This translates to a considerable correction and indicates that the QHA description of the PES is a challenging task due to the very flexible system and consequently largely anharmonic surface. Obviously, this correction will significantly increase the accuracy of our treatment.

Additional attention should be paid to the large deviation between the QHA and the harmonic vibrational partition functions in Table 5.2. The HA partition function results from quantum mechanics and accordingly allows for the right description of quantum effects such as the ZPE and the negligence of modes with $h\nu \gg k_B T$. In stark contrast, the QHA partition function is based on classical assumptions, which allow uniformly distributed energy among all degrees of freedom. Consequently, even high frequency modes have an average energy of $k_B T$. We tested the temperature dependence of the QHA partition function for ammonia. For small temperatures the partition function deviated orders of magnitude from the harmonic value, whereas they matched for large temperatures, i.e., when quantum effects become negligible and $h\nu \ll k_B T$ applies for every mode. However, the calculation of the Gibbs reaction enthalpy, which is simply based on the differences of G_{corr} , compensates this effect to a certain extent, even for intermediate temperatures. This is readily explained, since the high frequency modes are retained upon cluster formation, their contribution to G_{corr} should finally cancel out. The newly formed modes in association of the fragments, however, are usually of low frequency and $h\nu \ll k_B T$ applies. Hence, a classical description of these new modes is eligible since quantum effects are rather small.

The obtained value of $-97.30 \text{ kJ mol}^{-1}$ for the Gibbs reaction enthalpy of the cluster association reaction compares well with the experimental value of $-92.2 \text{ kJ mol}^{-1}$ [54] - significantly better than the harmonic value (with the internal rotor correction already included) of $-108.57 \text{ kJ mol}^{-1}$. This improvement is probably due to mainly two aspects: First, the inclusion of mode coupling effects - in contrast the covariance matrix σ in the harmonic treatment is diagonal - and secondly, mode anharmonicities are considered by broadening of the respective Gaussian distribution. Hence, this provides a more balanced description of the configurational distribution than the pure harmonic results.

Table 5.2: QHA results for the acetonitrile system at 298.15 K: The electronic energy \mathcal{E}_0 , the determinant of the covariance matrix $\det(\boldsymbol{\sigma})$, where $M = 3M_N - 6 - M_{\text{IR}}$ is the number of vibrations and the natural logarithm of all partition functions, i.e. the corrected vibrational partition function $Z_{\text{vib}}^{\text{QHA,corr}} = f_{\text{corr}} Z_{\text{vib}}^{\text{QHA}}$ (the correction factor itself is also given), the harmonic vibrational partition function $Z_{\text{vib}}^{\text{HA}}$ for comparison, the internal rotor partition function Z_{IR} and the overall rotational and translational partition functions Z_{rot} and Z_{trans} , respectively, are given. The resulting Gibbs enthalpy correction G_{corr} and the Gibbs reaction enthalpy $\Delta_R G$ for the cluster association are also reported.

	ACN	(ACN)H ⁺	(ACN) ₂ H ⁺
\mathcal{E}_0 / E_h	-132.816103564	-133.127065632	-265.996251514
$\det(\boldsymbol{\sigma}) / (\text{amu}\text{\AA}^2)^M$	1.1723×10^{-32}	1.2505×10^{-38}	1.4578×10^{-68}
f_{corr}	0.317831	0.600657	0.003921
$\ln(Z_{\text{vib}}^{\text{QHA,corr}})$	-27.042012	-30.567615	-54.663061
$\ln(Z_{\text{vib}}^{\text{HA}})$	-47.534292	-58.670193	-98.827696
$\ln(Z_{\text{IR}})$	-	-	0.963855
$\ln(Z_{\text{rot}})$	7.809736	7.887583	9.947041
$\ln(Z_{\text{trans}})$	16.150450	16.186853	17.208483
G_{corr} / E_h	0.002910	0.006131	0.025063
$\Delta_R G / \text{kJmol}^{-1}$			-97.30

5.4.2 [MeNH₃ + MeOH]⁺

Geometry optimization and NMA As a second example for the QHA approach we investigated the protonated methanol-methylamine cluster [MeNH₃ + MeOH]⁺ as shown in Fig. 5.1b. The closer proximity of the proton towards the nitrogen than to the oxygen atom with distances of 1.06143 Å and 1.61839 Å, respectively, is readily explained by simple basicity considerations. This geometry stands in stark contrast to the perfectly symmetrically ACN dimer. The NMA of the equilibrium geometry shows that many of the fragment modes remain undisturbed despite the cluster formation, which indicates very little coupling between the two fragments. In Fig. 5.7 we assigned each mode of the cluster to the respective mode of the fragments (detailed assignment can be found in Table 5.6). As can be seen, few modes do show notable changes: First of all, the internal rotation around the main bond (C–O or C–N) of each fragment (modes 1) splits into two modes because the overall rotations around these axes becomes an internal motion. That means one rotation of each hetero-atom group is blue shifted for stability reasons which the hydrogen bond brings in, whereas the methyl rotations are clearly red shifted due to the

larger mass of the rotating counterpart. The remaining four additional modes below 500 cm^{-1} belong to a new internal rotation of the fragments against each other (lowest mode), the bending of the almost linear hydrogen bridge in two dimensions and finally the dissociation coordinate, i.e. the elongation of the $\text{H}^+ - \text{O}$ distance. Another significant change is observed for the $\text{N}-\text{H}$ stretch modes. In the separated MeNH_3^+ fragment the three coordinates combine to one symmetric and two asymmetric stretching modes (modes 16-18), whereas the $\text{N}-\text{H}^+$ stretching is completely decoupled from the other two in the clustered structure. The other two eventually combine to a symmetric and asymmetric stretching mode like in H_2O . In principal, the $[\text{MeNH}_3 + \text{MeOH}]^+$ cluster has five internal rotations: Q_1 (rotation of the fragments against each other), Q_2 and Q_5 (rotation of the methyl groups in MeOH and MeNH_3^+ , respectively), Q_7 (rotation of the $-\text{OH}$ group) and Q_8 (rotation of the $-\text{NH}_3^+$ group). The obtained normal frequencies of the latter two are rather high, which suggest a fairly large barrier induced by the hydrogen bond. Also Q_5 shows a significant higher frequency than the lower two modes, which is probably due to the sterical hindrance of the three nitrogen-bonded hydrogen atoms. Consequently, only the two lowest internal rotations need to be corrected in our treatment.

QHA results Consistent with the statement above, no MD simulations showed internal rotations other than the two lowest ones mentioned. Therefore, only these two motions were separated according to the Eckart-Sayvetz condition and all other modes were treated as ordinary oscillators. Also, the internal rotations of the two fragments were treated as ordinary oscillations since we did not observe any rotation in the corresponding MD simulations either. For each of the three molecules we ran up to six MD simulations, with an applied energy of $(3M_N - 6)k_B T$, randomly assigned to the nuclei and with different starting conditions for each simulation. For the cluster, large amplitude motions were separated according to the following order: (1) Correction of the internal rotation of the two fragments rotating against each other. Its rotational axis was more elaborately drawn from the CoM of the entire molecule to the CoM of each fragment, since neither the $\text{O}-\text{H}^+$, the $\text{H}^+ - \text{N}$ nor the $\text{N}-\text{O}$ bond sufficiently encompassed this motion. To prevent an overall rotation, we corrected for both fragments. (2) Correction of the internal methyl group rotation in MeOH with the $\text{C}-\text{O}$ bond as the rotational axis. In this configuration only the hydrogen atoms rotate as a much smaller fragment compared to the residual molecule, which makes it reasonable to correct for the rotation of these hydrogen atoms only. (3) Convert the molecule into the Eckart frame. With the transformation of the Cartesian displacements into normal coordinates, we obtained the covariance matrix for each simulation. Subsequently we averaged the covariance matrices, calculated the correction factor of Eq. (5.13) and determined the (reduced) vibrational partition function of the molecule according to Eq. (5.7). The internal rotational partition functions of the two considered rotations were obtained according to Eq. (5.11) with a reduced moment of inertia, calculated from formulas given in the literature [260–262].

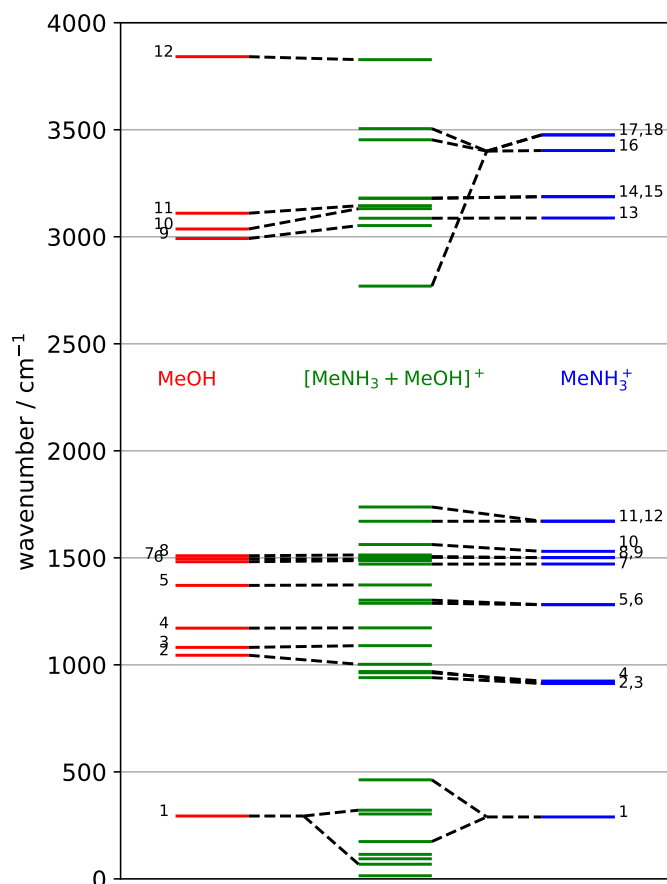


Figure 5.7: Normal frequencies of the two fragments and the $[\text{MeNH}_3 + \text{MeOH}]^+$ cluster. The dashed lines indicate the assignment of the modes, i.e., which mode in the fragments corresponds to which mode in the cluster.

Table 5.3 summarizes the QHA results for the protonated methanol-methylamine cluster. Just as is the case with ACN, the correction factor is much smaller for the cluster than for the fragments. This indicates a significant difference between the QHA and ab-initio potential with a strong anharmonic character. Also, the vibrational partition functions calculated with the QHA approach differ largely from the ones calculated with the standard HA due to neglected quantum effects. Regarding the finally obtained Gibbs reaction enthalpy, the QHA gives a surprisingly inaccurate value compared to the experimental result of $-49.31 \text{ kJ mol}^{-1}$ [184]. In fact, the calculated pure harmonic value of $-59.28 \text{ kJ mol}^{-1}$ is even closer to the literature value, however, still with a significant deviation from the experiment. This outcome is rather unexpected considering the good results for the QHA treatment of the $(\text{ACN})_2\text{H}^+$ cluster.

5 Thermochemistry from MD Simulations

Table 5.3: QHA results for the amine-methanol system at 298.15 K: Listed are the electronic energy \mathcal{E}_0 , the determinant of the covariance matrix $\det(\boldsymbol{\sigma})$, with $M = 3M_N - 6 - M_{\text{IR}}$ as the number of vibrations and the natural logarithm of all partition functions, i.e., the corrected vibrational partition function $Z_{\text{vib}}^{\text{QHA,corr}} = f_{\text{corr}} Z_{\text{vib}}^{\text{QHA}}$ (the correction factor itself is also given), the harmonic vibrational partition function $Z_{\text{vib}}^{\text{HA}}$ for comparison, the internal rotor partition function Z_{IR} and the overall rotational and translational partition functions Z_{rot} and Z_{trans} , respectively, are given. The resulting Gibbs enthalpy correction G_{corr} and the Gibbs reaction enthalpy $\Delta_R G$ for the cluster association are also reported.

	MeOH	MeNH ₃ ⁺	[MeNH ₃ + MeOH] ⁺
\mathcal{E}_0 / E_h	-115.781 709 438	-96.265 936 625 0	-212.082 861 611
$\det(\boldsymbol{\sigma}) / (\text{amu}\text{\AA}^2)^M$	3.8346×10^{-32}	1.5646×10^{-50}	5.9329×10^{-77}
f_{corr}	0.331 338	0.396 865	0.000 512
$\ln(Z_{\text{vib}}^{\text{QHA,corr}})$	-26.417 544	-41.983 030	-64.577 338
$\ln(Z_{\text{vib}}^{\text{HA}})$	-53.822 945	-83.555 063	-136.467 535
$\ln(Z_{\text{IR}})$	-	-	3.117, 1.240
$\ln(Z_{\text{rot}})$	8.053 511	7.356 496	11.499 821
$\ln(Z_{\text{trans}})$	15.778 953	15.780 068	16.819 231
G_{corr} / E_h	0.002 441	0.017 795	0.030 121
$\Delta_R G / \text{kJmol}^{-1}$			-66.50

5.4.3 QHA discussion

Some sources of error should be discussed. It is clear that the real configurational distribution is not a multidimensional Gaussian but can have an asymmetric form along certain modes. Rojas *et al.* [236] developed an extension of the QHA to higher moments to account exactly for these types of shapes. We did not include higher moments but it would be in principle possible. One should also keep in mind that the conducted MD simulation in fact represents a microcanonical (or $NV\mathcal{E}$) ensemble since only one particle at constant energy is considered. To model the canonical (or NVT) ensemble with a partition function at constant temperature, we applied $(3M_N - 6)k_B T$ energy to the molecule randomly distributed among all modes. However, if we calculate the kinetic energy in each mode (each summand of Eq. (5.9) separately), average over the whole simulation time and then determine a temperature $T_i = 2\langle \mathcal{E}_{\text{kin},i} \rangle / k_B$ for each mode Q_i , we find that the mean temperature is not the same for all modes. E.g., for the three MD simulations of (ACN)H⁺, the lowest mean temperature is only 134 K in Q_2 while the highest is 543 K in Q_4 . Though this problem can be addressed by longer simulation time or more MD simulations with different starting conditions, it is worth noting that the mean temperature, averaged

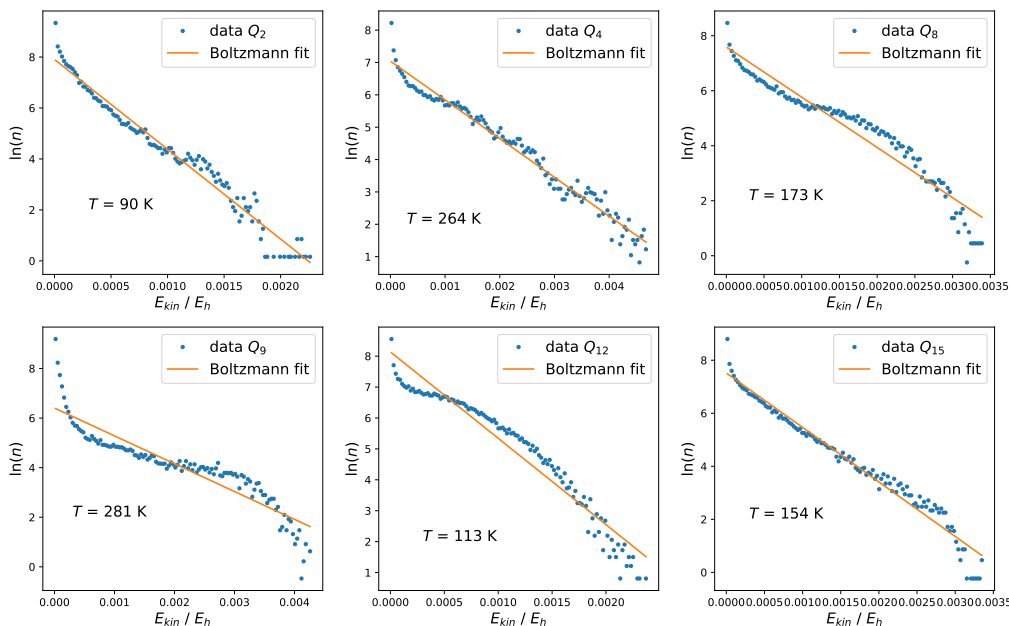


Figure 5.8: Boltzmann plots of \mathcal{E}_{kin} distributions for $(ACN)H^+$. The temperatures depicted are calculated from the slopes of the fitting functions, which should be $-(k_B T)^{-1}$, if the data follow a Boltzmann distribution.

over all modes, is 288 K and thus close to the originally set 298.15 K. Additionally, the distribution of kinetic energy in each mode is an exponential distribution, i.e., higher kinetic energies are less likely and the most probable kinetic energy is zero. This is equivalent to the Boltzmann population distribution of energy levels for a harmonic oscillator in a canonical ensemble according to the term $\exp(-h\nu_i/k_B T)$. Some of these distributions are shown in Fig. 5.8 as Boltzmann plots. Note that most of the latter plots do not exactly follow a straight line as expected for a real Boltzmann distribution, however, they do reflect the exponential dependency of population vs. energy. Since the temperatures are determined from these Boltzmann plots, special attention to the validity of their respective values should be paid. In some cases, those values differ significantly from the above mentioned definition of the temperatures in each mode. Nevertheless, each mode does have different energies according to a Boltzmann distribution, or at least similar to it. This finding supports the capability of our formalism to model canonical ensembles with certain energy distributions in each mode, while fairly correct temperatures are predicted based on the overall energies.

A closer look should as well be paid to low frequency modes with large amplitudes. Therefore the $(ACN)_2H^+$ cluster is a good example with its two degenerated modes Q_2 and Q_3 , which correspond to a bending of the two fragments with the H^+ as the angles vertex. Since the NMA is only valid for small changes in the Cartesian coordi-

nates, large amplitude motions will introduce errors in this formalism, which is the reason for the separation of internal rotations in the first place. The shallow potential for this bending allows the molecule to move atoms far away from the z -axis. In this scenario any other mode with contributions along this axis, e.g. the C–N stretching, will instantly have a smaller overlap with this axis and consequently decrease the value for Q_i since it now stretches along a vector containing z - and x - (or y -) contributions. This problem could be avoided with the use of BAT-coordinates but, as discussed before, this was not applicable in the case of the ACN dimer.

Apparently, our formalism works not as well for the $[\text{MeNH}_3 + \text{MeOH}]^+$ cluster. Quantum effects of the modes are unlikely the cause of significant deviation because all new modes are very low lying and thus classical (cf. Fig. 5.7). All high frequency contributions are very similar to their counterparts in each fragment and consequently annul for the Gibbs reaction enthalpy. The error due to low frequency modes and their large amplitude should be similar to the ACN case and does not intuitively explain the larger deviation. One new type of error could be introduced by the two internal rotations in the following way: The lack of symmetry might cause the equilibrium geometry to slightly deviate in the different wells along the first internal rotation. This speculation is supported by visible bending in the MD simulation since it is not a rotation around the O–H⁺ or N–H⁺ bond but rather a rotation of both fragments around their individual CoM. This bending is essential to retain the H-bond and removing only the rotation but not this bending could introduce an error. Also, the order of sorting the rotations out remains a matter of pure choice. Intuitively we started with the larger motion, however, this could as well introduce some error if the modes are not perfectly orthogonal to each other. Furthermore, the remaining vibrations after the internal rotor correction reveal a quite significant motion of the atoms, although they are not rotating. This can be an indication for an inappropriate correction for this asymmetric cluster geometry, since the same procedure resulted in fixed hydrogen atoms for the symmetric $(\text{ACN})_2\text{H}^+$ example (Fig. 5.5b). Using a partition function derived for a free rotor (Eq. (5.11)), which was a good assumption in the $(\text{ACN})_2\text{H}^+$ case, might also introduce an error having different well depths of the large internal rotation. To minimize errors owing to failed convergence or ergodicity, we ran up to six MD simulations per molecule. Indeed, the configurational distributions along the modes all look very Gaussian and the QHA distribution fits most of them very well. Also, the determinant is very close to convergence with six MD trajectories. Therefore, failed convergence is probably not the reason for this substantial error. A closer look at f_{corr} reveals a value which is by an order of magnitude smaller than for the $(\text{ACN})_2\text{H}^+$ cluster. This could indicate inherent problems with this molecule since the QHA potential seems to largely deviate from the ab-initio one. With eight low lying modes, the amine cluster could just be a bit too flexible to determine accurate values of Q_i via the Cartesian transformation.

5.5 Summary and Conclusions

The QHA was applied to small molecular clusters to determine absolute classical vibrational partition functions and to derive thermochemical data from molecular dynamics simulations. For the proton bound acetonitrile dimer, as the first example, normal coordinates from standard ab initio normal mode analysis were used to describe the configurational distribution. It was important to separate large amplitude motions such as the overall translation, overall rotation and internal rotation. This was accounted for in an automated procedure for each time step of the MD simulation, in which the molecules coordinates were converted into the Eckart frame and internal rotations were separated by meeting the Eckart-Sayvetz condition. The partition functions for the separated motions were calculated from standard formula. The covariance matrix of the normal coordinates, derived from the MD simulations, was assessed for convergence, for deviations from the multidimensional Gaussian assumed by the QHA and for the mean temperature in each mode. The resulting Gibbs enthalpy of the association reaction deviated by only 5 % compared to experimental findings, which seemed to prove the applicability of this method to small molecular clusters.

In stark contrast are the QHA results for the second investigated system. The obtained Gibbs enthalpy for the formation of the asymmetric $[\text{MeNH}_3 + \text{MeOH}]^+$ cluster differed by 26 % compared to an experimental value in the literature. A few inherent problems with the application of the QHA were identified, such as the negligence of vibrational quantum effects, the slow convergence regarding ergodic mixing, the flawed Gaussian approximation of asymmetric configurational distributions and errors introduced by large amplitude motions. Nevertheless, it remains unclear why this method worked well in one case and did not in the other. More cluster examples are in preparation to further investigate this approach and additionally we are working on the QHA for the amine-methanol cluster with BAT coordinates.

Although fully automated, our method is computationally expensive because of the ab initio MD simulations additional to the standard NMA (for $(\text{ACN})_2\text{H}^+$, one 15 000 time step long trajectory needed 18 h). However, for the investigated system size, it is still faster than a full VPT2 or VSCF calculation (full VPT2 on $(\text{ACN})_2\text{H}^+$ takes about 20.5 h; VSCF takes multiple days with the same computational resources), in particular since the multiple MD simulations can run in parallel on suitable computer clusters.

5.6 Appendix A - The Eckart Frame

In the equilibrium configuration of a molecule containing M_N atoms define an arbitrary set of three axis given by three normalized and orthogonal vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The vectors from the origin of this coordinate system to each atom α are called \mathbf{r}_α^0 . Let \mathbf{R}^0 be the center of mass (CoM) of the molecule. The so called *static model* is now defined by vectors \mathbf{c}_α^0 drawn from the CoM to each atom [255], thus

$$\mathbf{c}_\alpha^0 = \mathbf{r}_\alpha^0 - \mathbf{R}^0 \quad (5.16)$$

For any instantaneous configuration, i.e., at each time step of the MD simulation, the *dynamic model* can now be used to calculate three Eckart vectors $\mathbf{F}_1, \mathbf{F}_2$ and \mathbf{F}_3 by

$$\mathbf{F}_i = \sum_{\alpha=1}^{M_N} m_\alpha c_{\alpha,i}^0 (\mathbf{r}_\alpha - \mathbf{R}) \quad (5.17)$$

where \mathbf{r}_α and \mathbf{R} are now drawn to the instantaneous position of the atoms. From these, the Gram matrix \mathbf{F} is obtained through

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_1 \cdot \mathbf{F}_1 & \mathbf{F}_1 \cdot \mathbf{F}_2 & \mathbf{F}_1 \cdot \mathbf{F}_3 \\ \mathbf{F}_2 \cdot \mathbf{F}_1 & \mathbf{F}_2 \cdot \mathbf{F}_2 & \mathbf{F}_2 \cdot \mathbf{F}_3 \\ \mathbf{F}_3 \cdot \mathbf{F}_1 & \mathbf{F}_3 \cdot \mathbf{F}_2 & \mathbf{F}_3 \cdot \mathbf{F}_3 \end{pmatrix} \quad (5.18)$$

Via standard matrix algebra a matrix $\mathbf{F}^{-1/2}$ can be obtained for which $\mathbf{F}^{-1/2} \cdot \mathbf{F}^{-1/2} = \mathbf{F}^{-1}$ holds. The Eckart frame, three new normalized and orthogonal vectors to set at the CoM \mathbf{R} , is then calculated from

$$(\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3) = (\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3) \cdot \mathbf{F}^{-1/2} \quad (5.19)$$

Linear algebra basis transformation is then used to bring the vectors drawn from the CoM to the position of the atoms in the equilibrium configuration, i.e., \mathbf{c}_α^0 , from the original to the Eckart orientation:

$$\mathbf{C}_\alpha = \sum_{i=1}^3 c_{\alpha,i}^0 \mathbf{f}_i \quad (5.20)$$

Defining Cartesian displacement vectors $\boldsymbol{\rho}_\alpha$ for each atom α via

$$\boldsymbol{\rho}_\alpha = \mathbf{r}_\alpha - \mathbf{R} - \mathbf{C}_\alpha \quad (5.21)$$

displacements in the Eckart frame $\boldsymbol{\rho}_\alpha^{(E)}$ can be calculated according to [256]:

$$\boldsymbol{\rho}_\alpha^{(E)} = \begin{pmatrix} \boldsymbol{\rho}_\alpha \cdot \mathbf{f}_1 \\ \boldsymbol{\rho}_\alpha \cdot \mathbf{f}_2 \\ \boldsymbol{\rho}_\alpha \cdot \mathbf{f}_3 \end{pmatrix} \quad (5.22)$$

These displacements are now free of overall translation and overall rotation, which might be introduced through numerical error in the MD simulation. Also, the displacements are now small enough that the transformation to normal coordinates can be conducted [256].

5.7 Appendix B - Separating out Internal Rotation

Following Kirtman [259], to separate out internal rotation, we define vectors $\mathbf{c}_{\alpha'}$ drawn from the center of mass (CoM) of the entire molecule \mathbf{R} to each atom of the internal rotor $\mathbf{r}_{\alpha'}$ (all single primed quantities refer to the internal rotor). Further, set the CoM of the internal rotor as \mathbf{R}' and define $\mathbf{d}_{\alpha'} = \mathbf{c}_{\alpha'} - \mathbf{c}_{\alpha'}^0$ the displacements of the atoms from their equilibrium position during the MD simulation. We can now define vectors drawn from the CoM of the internal rotor to the atoms of the rotor:

$$\mathbf{l}_{\alpha'} = \mathbf{c}_{\alpha'} - (\mathbf{R}' - \mathbf{R}) = \mathbf{r}_{\alpha'} - \mathbf{R}' \quad (5.23)$$

If the rotor is unbalanced, set $\boldsymbol{\sigma}'$ as the normal from the axis of internal rotation to \mathbf{R}' , the former being defined as

$$\mathbf{z}' = \boldsymbol{\zeta}' + \lambda \mathbf{e}_{z'} \quad (5.24)$$

where $\boldsymbol{\zeta}'$ is any point on the axis and $\mathbf{e}_{z'}$ a normalized vector parallel to the axis. Then, $\boldsymbol{\sigma}'$ can be calculated from

$$\boldsymbol{\sigma}' = \mathbf{R}' - (\boldsymbol{\zeta}' + (\mathbf{e}_{z'} \cdot (\mathbf{R}' - \boldsymbol{\zeta}')) \mathbf{e}_{z'}) \quad (5.25)$$

The Eckart-Sayvetz condition for the internal rotation now reads [259]

$$0 = \sum_{\alpha'} m_{\alpha'} \left((\boldsymbol{\sigma}'^0 + \mathbf{l}_{\alpha'}^0) \times \mathbf{d}_{\alpha'} \right) \cdot \mathbf{e}_{z'} \quad (5.26)$$

where both $\boldsymbol{\sigma}'$ and $\mathbf{l}_{\alpha'}$ are taken at the equilibrium geometry (superscript "0"), whereas the rotation axis $\mathbf{e}_{z'}$ is updated at every time step since it can change due to other internal motion. At any time step of the MD simulation, this sum is calculated and if it does not yield zero, the atoms of the internal rotor can be rotated around an angle β such that the Eckart-Sayvetz condition is fulfilled. Adjusting a parameter such that another quantity yields zero is a standard minimization procedure and can be performed numerically. With the given notation, rotating all atoms of the internal rotor around \mathbf{z}' reads

$$\tilde{\mathbf{r}}_{\alpha'} = \boldsymbol{\zeta}' + \mathbf{e}_{z'} (\mathbf{e}_{z'} \cdot (\mathbf{r}_{\alpha'} - \boldsymbol{\zeta}')) + \cos \beta (\mathbf{e}_{z'} \times (\mathbf{r}_{\alpha'} - \boldsymbol{\zeta}')) \times \mathbf{e}_{z'} + \sin \beta (\mathbf{e}_{z'} \times (\mathbf{r}_{\alpha'} - \boldsymbol{\zeta}')) \quad (5.27)$$

where $\tilde{\mathbf{r}}_{\alpha'}$ is the new position of the atoms from which the Eckart-Sayvetz condition in Eq. (5.26) is calculated. If the two fragments of the internal rotation are similar in weight, hence, both of them rotate significantly, it is useful to perform this correction to both fragments independently. Otherwise the internal rotation would be transformed to an overall rotation. In principal, this would be taken care of through putting the molecule in the Eckart frame (performed after the internal rotation correction), however, it is numerically more stable to not introduce an overall rotation for reasons explained in the text.

5.8 Supporting Information

Tables 5.4-5.6 show the frequencies (ω) and reduced masses (μ) of the normal modes calculated in this work. For symmetric structures, the irreducible representations are given, for the asymmetric $[\text{MeNH}_3 + \text{MeOH}]^+$ cluster, each mode is assigned to its corresponding mode in the fragments (cf. Fig. 5.7). Experimental data for the neutral fragments, i.e., ACN and MeOH can be found in the literature [264].

Table 5.4: Normal frequencies of ACN, $(\text{ACN})\text{H}^+$, MeOH and MeNH_3^+ calculated at the B3LYP-GD3(BJ)/def2-TZVPP level of theory.

ACN (C_{3v})				$(\text{ACN})\text{H}^+$ (C_{3v})			
No	ω [cm^{-1}]	μ [amu]	Irr. Rep.	No	ω [cm^{-1}]	μ [amu]	Irr. Rep.
1,2	383.8278	3.7968	e	1,2	397.9396	2.6194	e
				3,4	576.2860	1.2340	e
3	931.7197	4.6047	a_1	5	921.6415	4.4269	a_1
4,5	1063.7246	1.4481	e	6,7	1040.9057	1.5187	e
6	1414.8814	1.2317	a_1	8	1392.6331	1.2150	a_1
7,8	1475.2883	1.0381	e	9,10	1426.7417	1.0359	e
9	2367.8959	12.5576	a_1	11	2383.9696	6.4582	a_1
10	3051.4479	1.0335	a_1	12	3029.2283	1.0318	a_1
11,12	3120.3759	1.1020	e	13,14	3109.8293	1.1025	e
				15	3680.6703	1.1425	a_1
MeOH (C_s)				MeNH_3^+ (C_{3v})			
No	ω [cm^{-1}]	μ [amu]	Irr. Rep.	No	ω [cm^{-1}]	μ [amu]	Irr. Rep.
1	293.7752	1.0709	a''	1	289.3908	1.0078	a_2
2	1044.7946	3.2254	a'	2,3	912.9836	1.0535	e
3	1081.5201	1.3491	a'	4	924.3051	5.1137	a_1
4	1171.5037	1.2672	a''	5,6	1281.1962	1.4080	e
5	1371.1656	1.2878	a'	7	1471.0790	1.1406	a_1
6	1481.7174	1.1338	a'	8,9	1501.0055	1.0365	e
7	1495.5962	1.0484	a''	10	1530.2826	1.1625	a_1
8	1509.8121	1.0442	a'	11,12	1671.0430	1.0394	e
9	2991.6967	1.0360	a'	13	3087.6305	1.0296	a_1
10	3036.3218	1.1053	a''	14,15	3187.3223	1.1108	e
11	3110.0860	1.0968	a'	16	3402.8885	1.0306	a_1
12	3841.3936	1.0665	a'	17,18	3476.0606	1.0931	e

Table 5.5: Normal frequencies of the $(\text{ACN})_2\text{H}^+$ cluster calculated at the B3LYP-GD3(BJ)/def2-TZVPP level of theory. The irreducible representations correspond to the point group of the molecule, D_{3d} .

No	ω [cm^{-1}]	μ [amu]	Irred. Rep.
1	-1.4967	1.0078	a_{1u}
2,3	53.3522	3.7362	e_u
4,5	129.9326	5.8025	e_g
6	201.3918	1.0300	a_{2u}
7	355.0749	5.7220	a_{1g}
8,9	406.4703	3.3961	e_u
10,11	418.1781	3.7793	e_g
12	951.9972	4.6350	a_{2u}
13	1000.4961	5.4427	a_{1g}
14,15	1054.7798	1.4970	e_u
16,17	1054.8761	1.4951	e_g
18,19	1258.4921	1.0545	e_u
20	1407.6015	1.2242	a_{1g}
21	1407.8258	1.2162	a_{2u}
22,23	1451.6043	1.0365	e_g
24,25	1451.6476	1.0362	e_u
26	2425.1955	12.2918	a_{2u}
27	2444.9692	12.6540	a_{1g}
28	3050.0515	1.0330	a_{2u}
29	3050.1785	1.0329	a_{1g}
30,31	3128.1794	1.1025	e_g
32,33	3128.1818	1.1025	e_u

5 Thermochemistry from MD Simulations

Table 5.6: Normal frequencies of the $(\text{ACN})_2\text{H}^+$ cluster calculated at the B3LYP-GD3(BJ)/def2-TZVPP level of theory. Each mode is assigned to a mode in the separated fragments if possible. It is also indicated when the original modes red/blue-shifted due to the cluster formation.

No	ω [cm^{-1}]	μ [amu]	Description	MeNH ₃ ⁺		MeOH	
				No	ω	No	ω
1	14.59	2.0785	IR around H ⁺ -bridge				
2	68.53	1.2235	IR of -Me in MeOH			1	293.78
3	93.83	2.3661	N-H ⁺ -O bending				
4	114.61	2.1282	N-H ⁺ -O bending				
5	174.36	1.2737	IR of -Me in MeNH ₃ ⁺	1	289.39		
6	303.15	3.3894	NH ⁺ -O stretch (dissoc. coord.)				
7	321.08	1.1100	IR of -OH in MeOH			1	293.78
8	462.77	1.0923	IR of -NH ₃ ⁺ in MeNH ₃ ⁺	1	289.39		
9	940.20	1.0522	asym twisting in MeNH ₃ ⁺	2	912.98		
10	963.12	5.0211	C-N stretch in MeNH ₃ ⁺	4	924.31		
11	968.74	1.0593	asym wagging in MeNH ₃ ⁺ w/ H ⁺	3	912.98		
12	1002.59	7.6357	C-O stretch in MeOH			2	1044.79
13	1089.65	1.0888	asym wagging in MeOH			3	1081.52
14	1172.74	1.2506	twisting in MeOH			4	1171.71
15	1287.57	1.3838	sym wagging in MeNH ₃ ⁺ w/ H ⁺	5	1281.20		
16	1302.28	1.3865	sym twisting in MeNH ₃ ⁺ w/o H ⁺	6	1281.20		
17	1373.26	1.2881	sym wagging in MeOH			5	1371.17
18	1470.69	1.1472	sym H-C-H bend in MeNH ₃ ⁺	7	1471.08		
19	1486.96	1.1265	sym H-C-H bend in MeOH			6	1481.72
20	1493.55	1.0446	asym1 H-C-H bend in MeOH			7	1495.60
21	1502.30	1.0381	asym1 H-C-H bend in MeNH ₃ ⁺	8	1501.01		
22	1505.80	1.0346	asym2 H-C-H bend in MeNH ₃ ⁺	9	1501.01		
23	1513.38	1.0383	asym2 H-C-H bend in MeOH			8	1509.81
24	1562.16	1.1472	sym H-N-H bend w/ H ⁺	10	1530.28		
25	1670.72	1.0288	asym1 H-N-H bend w/o H ⁺	11	1671.04		
26	1737.50	1.0413	asym2 H-N-H bend (H ⁺ ⊥ N-O)	12	1671.04		
27	2769.50	1.0839	asym1 N-H stretch (H ⁺)	?	3402.89		
28	3052.38	1.0293	sym C-H stretch in MeOH			9	2991.70
29	3086.61	1.0306	sym C-H stretch in MeNH ₃ ⁺	13	3087.63		
30	3131.28	1.1093	asym1 C-H stretch in MeOH			10	3036.32
31	3145.23	1.1069	asym2 C-H stretch in MeOH			11	3110.09
32	3179.72	1.1097	asym1 C-H stretch in MeNH ₃ ⁺	14	3187.32		
33	3180.40	1.1091	asym2 C-H stretch in MeNH ₃ ⁺	15	3187.32		
34	3453.12	1.0530	sym N-H stretch w/o H ⁺	?	3467.06		
35	3505.27	1.0926	asym2 N-H stretch w/o H ⁺	18	3476.06		
36	3827.47	1.0667	O-H stretch			12	3841.39

6 Summary and Outlook

It is the aim of this work to use computational methods to study a number of specific questions arising from MS experiments but also to combine, refine and develop new methods of calculation for more accurate data. Thus, this summary is organized into two sections, addressing the modeling of experimental questions and the development of new computational methods, respectively, followed by a general conclusion.

6.1 Answering Questions from Experiment

In this work, two main questions from experiment were investigated. First, the process of "supercharging" or better "charge retention vs. charge depletion" in gas-phase modified nESI-MS. Second, the effect of dynamic clustering, i.e., the dependence of an ion-cluster distribution on the field induced, effective temperature, in Differential Mobility Spectrometry, DMS.

Charge Retention/Depletion in nESI-MS Experimental investigations on the charge state distribution of small proteins created by nESI-MS exhibit a pronounced dependence on the modification of the gas phase of the ion source, i.e., adding polar compounds. In particular, using either methanol (MeOH) or acetonitrile (ACN) the main charge state of Substance P (SP) observed is shifted between 2+ and 3+, respectively. Simple considerations regarding the Coulomb repulsion of charges in the gas phase show that further charging of ions in the gas phase is not possible. Hence, the higher charge state can only form in the solution phase and then is either retained or depleted by proton transfer (PT). Whether the charge depletion happens in the solution phase (inside a small droplet), or in the gas phase (via cluster chemistry involving MeOH or ACN), remains unclear. It was hypothesized, why an enriched MeOH atmosphere depletes the higher charge state, while an ACN atmosphere retained it: A protic gas phase modifier as for example MeOH forms hydrogen-bond chains at the protonation site and thus increases its proton affinity until proton transfer becomes thermodynamically favorable. Charge depletion is indeed observed in the presence of protic modifiers (methanol, ethanol, ammonia), while charge retention is observed with aprotic ones (acetonitrile, acetone, diethyl ether).

To further investigate this hypothesis, the PT from a protonated primary amine as model system towards MeOH or ACN clusters is investigated. The geometry and thermochemical stability of the homogeneous $(S)_nH^+$ and mixed $[MeNH_3^+ + (S)_n]^+$ clusters ($S=ACN, MeOH$) are calculated. The most important conclusion drawn from

6 Summary and Outlook

these calculations is that MeOH forms hydrogen-bond chains for the homogeneous as well as the mixed clusters. In particular, the protonated MeOH clusters do already form at the protonated amine site. In contrast, ACN does not offer further binding sites for additional modifier molecules. Any additional ACN molecule binds to the amine directly and not to the already present modifier molecules.

Potential energy surface (PES) scans for the PT reactions from the amine to each of these clusters reveals the important effect of the homogeneous clusters being already formed at the charge site: While the equilibrium cluster geometry always shows the proton residing at the amine (for up to four MeOH added), the PT paths become less steep with increasing cluster size. Although for the calculated systems the PT was always thermodynamically less favorable than evaporation of a single MeOH molecule, the less steeper PT paths lead to enhanced population density in the product channel. Thus, PT is possible under high energy collisions. In stark contrast, ACN showed high energy barriers for the PT reactions due to large geometrical changes necessary for the formation of, e.g., the $(\text{ACN})_2\text{H}^+$. The high energy barrier and low state density in this tight transition state renders PT unlikely.

Since Substance P is observed as multiply charged molecule, the effect of the addition of a second charge on the PT paths is investigated. As model system, double protonated ethylene diamine is used. Although the close proximity of the two charges represents a rather drastic case, the observed trends are useful to draw conclusions. Indeed, very different PT paths are observed: Because of the Coulomb repulsion of the two charges, the PT is energetically favorable for all systems studied. All mixed clusters show a stable minimum, hence, the PT path has to overcome a barrier for deprotonation. However, bound MeOH clusters show decreasing barrier heights with cluster size, while in the ACN case still energetically elevated transition states have to be overcome. This is rationalized by the charge dilution due to the hydrogen-bond chain inside the MeOH clusters, promoting proton transfer.

Thus, the hypothesized model of charge retention/depletion is supported by this theoretical investigation and the deeper insight into the molecular processes occurring are of great benefit to design further experiments. Future work may focus on the actual calculation of rate constants from the determined PES scans via a RRKM theory treatment (cf. Section 2.4). Additionally, performing the PT potential scans for the actual Substance P molecule rather than for model compounds are of interest, to capture the effect of additional charges more realistically. Performing RRKM calculations for the PT at Substance P is the ultimate goal since it enables the estimation of the charge state distribution, which can be directly compared with experiment. Results from these investigations aid to answer the question, whether PT happens inside a small droplet or with gas phase clusters since the rate constants for both processes should be very different and should vary differently with the concentration of the modifier.

Dynamic Clustering in DMS Differential Mobility Spectrometry becomes an increasingly popular tool in analytical chemistry. Through the characteristic difference between the high- and low-field mobility in an electrical field, ions are separated prior to mass analysis. While experimental effects regarding the electrical fields, background gas temperature, and chemical environment used, are well studied, there is no in-depth model yet, which predicts the differential mobility behavior under various experimental conditions. Especially the effect of gas phase modification through the addition of solvent molecules, leading to an increasing mobility with field strength, is difficult to model. While the change of the collision environment certainly affects the collision cross section (CCS), dynamic clustering is also hypothesized to explain the observed effects: At low field strengths, the ions are close to thermal conditions and thus cluster with the solvent molecules. At high field strengths, however, high energy collisions lead to an increased ion temperature, effectively reducing the mean cluster size leading to a higher mobility.

To study this effect from a modeling perspective, the geometry, partition functions and CCS (as a function of temperature) are calculated for a number of cluster structures. Using the Mason-Schamp equation and applying two-temperature theory, the effective temperature of each cluster structure is calculated for a particular field strength. Since the clusters are assumed to be in constant dynamic equilibrium with each other, the ensemble mobility is calculated as a weighted average of the individual cluster mobilities. To determine the weighting factors, which resemble the relative populations of each structure, a modified version of the superposition approximation is used. This modification allows for different temperatures for each cluster, which is necessary due to their different CCSs. The computed cluster distributions show a decreasing mean cluster size with increasing field strength, as suggested by the hypothesis. The fundamental reason for this behavior is the entropy gain upon evaporation of the clusters, eventually overcoming the binding energy at higher ion temperatures. This thermodynamic consideration neglects all kinetic effects as for example rate constants of dissociation.

From the field dependent ensemble mobility, dispersion plots are calculated for a case study and compared to experimental data. The system investigated is a tetramethyl ammonium ion in different environments, i.e., pure N_2 and N_2 seeded with methanol, acetonitrile and acetone. It is shown that the usage of the two-temperature theory, the correct modeling of the temperature dependency of the CCS, and the degree of clustering are all relevant to reproduce the observed dispersion plots. While only semi-quantitative agreement is observed, these findings strongly support the hypothesis that the dynamic clustering mechanism is a major contributor to differential mobility when the gas phase is seeded with polar modifiers.

Future work will focus on the refinement of the presented model. The usage of two-temperature theory is important, however, it is only an approximation and known to yield incorrect mobilities at very high field strengths. Additionally, the thermochemical procedure used to determine the cluster distribution neglects anharmonic effects of the loosely bound clusters, as well as the dynamics of binding and dissociating.

6 Summary and Outlook

Both do most probably contribute significantly to the deviations observed between model and experimental data. Another important contribution to differential mobility could be the change in CCS caused by the presence of modifiers even without any dynamic clustering. The interaction potential between ion and polar neutrals vs. unpolar neutrals is very different and should affect the CCS to a significant degree. Incorporation of these effects in the CCS calculations will increase the accuracy of the model.

6.2 Developing new Calculation Methods

The development of new methods for the calculation of thermochemical data for loosely bound clusters is the second goal of this work since these data are crucial for the understanding of many chemical processes at or close to equilibrium. This is addressed with two different approaches: The development of a new hybrid method incorporating anharmonic contributions at low computational cost, and the usage of the quasi-harmonic approximation (QHA) in normal coordinates with special enforcement of the Eckart-Sayvetz conditions for overall movement and internal rotations.

Hybrid VPT2 Approach Vibrational Perturbation Theory to 2nd order (VPT2) is a well known approach to incorporate anharmonic effects to the partition functions and thus thermochemical stability of molecules. Since its computational cost is rather high, a hybrid model is proposed, where the harmonic contributions are calculated at a high, the anharmonic ones at a low level of theory. Harmonic frequencies and equilibrium rotational constants are determined at higher accuracy, whereas the vibrational anharmonic constants, vibration-rotation coupling matrix and centrifugal distortion constants are calculated with lower accuracy.

For small test molecules, it is shown that similar results are obtained by the hybrid method in comparison to a high level VPT2 treatment while only a fraction of the time was needed. As these results are promising and since this work mainly deals with loosely bound clusters, this new method is applied for these molecules as well. However, the developed formalism shows a number of obstacles when dealing with cluster structures. Large anharmonic effects or internal rotations are known to be troublesome for VPT2 in general, but the hybrid method is even more sensitive to such issues. This is mainly because the equilibrium geometry and Hessian are not the same at the two levels of theory. For such systems the hybrid method yields results with similar accuracy as the standard harmonic approximation but does not show an overall improvement. However, molecules without these troublesome features are accurately studied with this hybrid method, saving a significant amount of computational time compared to the full high-level VPT2 treatment.

Quasi-harmonic approximation in Normal Coordinates A second approach to increase the accuracy of thermochemical data for cluster systems applies the quasi-harmonic approximation (QHA). Within this formalism, molecular dynamics (MD) simulations of a molecule are conducted at a defined temperature around the equilibrium geometry. Within a chosen set of coordinates, the occupation distribution of the configurational space is recorded. Approximating it by a multidimensional Gaussian function, the covariance matrix of the coordinates directly yields the configurational partition function, from which the vibrational partition function is obtained.

Two innovative modifications of this treatment are developed. First, the set of coordinates in which the configurational distribution is represented are chosen to be the normal coordinates of the molecule. They are obtained by a standard normal mode analysis and are generally applicable. In contrast, internal coordinates such as bond lengths, angles or dihedrals are sometimes difficult to define, especially for $(\text{ACN})_2\text{H}^+$, which is of special interest. Second, large amplitude motions such as overall translation, overall rotation or internal rotation (which was troublesome throughout the whole preceding work), are separated out of the configurational space by explicit enforcement of the Eckart-Sayvetz conditions along the MD trajectories.

The thermochemical stability with regard to dissociation for proton bound clusters, i.e., the $(\text{ACN})_2\text{H}^+$ and $[\text{MeNH}_3 + \text{MeOH}]^+$ clusters, motivated by the preceding work, are studied. Multiple MD simulations are conducted for the clusters and fragments and transformed into the respective Eckart-Sayvetz enforced frame. Vibrational partition functions are calculated within the QHA, internal rotations are treated as free rotor, and overall motions (translation and rotation) are treated by standard methods. This formalism worked well for the proton bound ACN dimer. The free enthalpy of dissociation is calculated much closer to the experimental value as compared to results from standard methods, which show large deviations. However, the mixed amine-methanol cluster is not only much more complicated due to more than one internal rotation, its calculated dissociation enthalpy also deviates stronger from experimental values than using standard methods.

The developed formalism clearly needs to be improved. One issue is that the vibrational partition function is treated purely classically. Thus, any quantum effects as for example the vanishing heat capacity for modes with $h\nu_l \gg k_B T$, are not considered. There is a variation of the QHA, where the covariance matrix in Cartesian coordinates is diagonalized to obtain ensemble averaged normal coordinates from the eigenvectors and an analogon to the harmonic frequencies from the eigenvalues. These are then used with the standard formula for the vibrational partition function which do include quantum effects. The developed treatment for the enforcement of the Eckart-Sayvetz conditions is applicable here as well and first test calculations show some promising results. However, some problems still need to be solved regarding the convergence of the covariance matrix, ensuring complete ergodic mixing to cover the complete configurational space. Additionally, the internal rotations of very flexible molecules are still rather troublesome because of their multi-well character and their coupling to other modes.

6.3 General Conclusion

Cluster chemistry has been demonstrated to be very important in API-MS. In particular, clusters formed from the charged analyte with neutral, polar solvents in the gas phase are strongly influencing many areas in MS. For example, their thermochemistry determines the degree of clustering at equilibrium conditions often observed under atmospheric pressure conditions (ion source, IMS/DMS cell). This determines the ion mobility in an electrical field but also the population entering the lower pressure regions inside a mass spectrometer, i.e., mean cluster size, or whether the analyte is charged at all. Under non-equilibrium conditions, often observed at high reduced field strengths (ion transfer, DMS cell), this cluster distribution is exposed to high energy collisions. The molecular dynamics, determined by the shape of the potential energy surface determines the fate of these clusters, e.g., whether a charge is retained at the analyte or is transferred to the solvent molecules. A fundamental understanding of both processes, equilibrium distribution and non-equilibrium dynamics, at various experimental conditions, is crucial for the understanding of the process happening between ionization and detection.

This work clearly demonstrates that Computational Chemistry in conjunction with experimental investigations represents a powerful tool to study these kinds of questions. Not only are thermochemical data, cluster structures or collision cross sections obtainable from calculations, also potential energy pathways for chemical reactions, eventually even their rate constants through RRKM calculations, can be computed. However, since Computational Chemistry is still an evolving field of research, it is important to consider the validity of the used calculation methods. In this work, especially the thermochemical stability of loosely bound clusters is difficult to address, since the standard harmonic treatment to the PES introduces large errors. Thus, efforts are put forward to develop new methods for more accurate predictions. While these show the inherent difficulties when dealing with those types of molecules, some progress was made. There are still many possibilities for further explorations in this field of research.

References

- [1] B. A. Garcia, “What does the future hold for top down mass spectrometry?”, [Journal of the American Society for Mass Spectrometry](#) **21**, 193–202 (2010).
- [2] B. Brügger, “Lipidomics: Analysis of the Lipid Composition of Cells and Subcellular Organelles by Electrospray Ionization Mass Spectrometry”, [Annual Review of Biochemistry](#) **83**, 79–98 (2014).
- [3] A. C. Schrimpe-Rutledge, S. G. Codreanu, S. D. Sherrod, and J. A. McLean, “Untargeted Metabolomics Strategies - Challenges and Emerging Directions”, [Journal of The American Society for Mass Spectrometry](#) **27**, 1897–1905 (2016).
- [4] J. Zaia, “Mass spectrometry of oligosaccharides”, [Mass Spectrometry Reviews](#) **23**, 161–227 (2004).
- [5] M. J. Kailemia, L. R. Ruhaak, C. B. Lebrilla, and I. J. Amster, “Oligosaccharide Analysis by Mass Spectrometry: A Review of Recent Developments”, [Analytical Chemistry](#) **86**, 196–212 (2014).
- [6] N. Zinn, C. Hopf, G. Drewes, and M. Bantscheff, “Mass spectrometry approaches to monitor protein-drug interactions”, [Methods](#) **57**, 430–440 (2012).
- [7] D. Rathore, A. Faustino, J. Schiel, E. Pang, M. Boyne, and S. Rogstad, “The role of mass spectrometry in the characterization of biologic protein products”, [Expert Review of Proteomics](#) **15**, 431–449 (2018).
- [8] E. Niyonsaba, J. M. Manheim, R. Yerabolu, and H. I. Kenttämää, “Recent Advances in Petroleum Analysis by Mass Spectrometry”, [Analytical Chemistry](#) **91**, 156–177 (2019).
- [9] A. T. Lebedev, “Environmental Mass Spectrometry”, [Annual Review of Analytical Chemistry](#) **6**, 163–189 (2013).
- [10] X. Wang, S. Wang, and Z. Cai, “The latest developments and applications of mass spectrometry in food-safety and quality analysis”, [TrAC Trends in Analytical Chemistry](#) **52**, 170–185 (2013).
- [11] T. De Vijlder, D. Valkenburg, F. Lemière, E. P. Romijn, K. Laukens, and F. Cuyckens, “A tutorial in small molecule identification via electrospray ionization-mass spectrometry: The practical art of structural elucidation”, [Mass Spectrometry Reviews](#) **37**, 607–629 (2018).

References

- [12] T. Baer and R. C. Dunbar, "Ion spectroscopy: Where did it come from; where is it now; and where is it going?", *Journal of the American Society for Mass Spectrometry* **21**, 681–693 (2010).
- [13] F. Gunzer, S. Krüger, and J. Grotemeyer, "Photoionization and photofragmentation in mass spectrometry with visible and UV lasers", *Mass Spectrometry Reviews* **38**, 202–217 (2019).
- [14] E. Carrascosa, J. Meyer, and R. Wester, "Imaging the dynamics of ion–molecule reactions", *Chemical Society Reviews* **46**, 7498–7516 (2017).
- [15] N. C. Polfer, "Infrared multiple photon dissociation spectroscopy of trapped ions", *Chemical Society Reviews* **40**, 2211–2221 (2011).
- [16] J. S. Brodbelt, "Photodissociation mass spectrometry: new tools for characterization of biological molecules", *Chemical Society Reviews* **43**, 2757–2783 (2014).
- [17] S. A. McLuckey and M. Mentinova, "Ion/Neutral, Ion/Electron, Ion/Photon, and Ion/Ion Interactions in Tandem Mass Spectrometry: Do We Need Them All? Are They Enough?", *Journal of The American Society for Mass Spectrometry* **22**, 3–12 (2011).
- [18] J. Brodbelt, "Focus on Novel Instrumentation in Mass Spectrometry and Ion Mobility Spectrometry", *Journal of The American Society for Mass Spectrometry* **29**, 211–212 (2018).
- [19] T. Kunsel and G. Phalke, "Mass Spectrometry Market by Technology and Application - Global Opportunity Analysis and Industry Forecast, 2017-2023", Allied Market Research - Diagnostics and Biotech **LI_172639**, 170 (2017).
- [20] J. H. Gross, *Mass Spectrometry*, 3rd ed. (Springer International Publishing AG, Cham, Switzerland, 2017), pp. xxv+968.
- [21] M. L. Gross and R. Caprioli, *Encyclopedia of Mass Spectrometry, Ten-Volume Set* (Elsevier Science & Technology, Amsterdam, Netherlands, 2014), p. 7000.
- [22] E. C. Horning, M. G. Horning, D. I. Carroll, I. Dzidic, and R. N. Stillwell, "New picogram detection system based on a mass spectrometer with an external ionization source at atmospheric pressure", *Analytical Chemistry* **45**, 936–943 (1973).
- [23] D. I. Carroll, I. Dzidic, E. C. Horning, and R. N. Stillwell, "Atmospheric Pressure Ionization Mass Spectrometry", *Applied Spectroscopy Reviews* **17**, 337–406 (1981).
- [24] A. Good, D. A. Durden, and P. Kebarle, "Ion–Molecule Reactions in Pure Nitrogen and Nitrogen Containing Traces of Water at Total Pressures 0.5–4 torr. Kinetics of Clustering Reactions Forming $H^+(H_2O)_n$ ", *The Journal of Chemical Physics* **52**, 212–221 (1970).

- [25] S. Klee, V. Derpmann, W. Wißdorf, S. Klopotoski, H. Kersten, K. J. Brockmann, T. Benter, S. Albrecht, A. P. Bruins, F. Dousty, T. J. Kauppila, R. Kostianen, R. O'Brien, D. B. Robb, and J. A. Syage, "Are clusters important in understanding the mechanisms in atmospheric pressure ionization? Part 1: reagent ion generation and chemical control of ion populations.", *Journal of the American Society for Mass Spectrometry* **25**, 1310–1321 (2014).
- [26] Y. K. Lau, S. Ikuta, and P. Kebarle, "Thermodynamics and kinetics of the gas-phase reactions $\text{H}_3\text{O}^+(\text{H}_2\text{O})_{n-1} + \text{H}_2\text{O} = \text{H}_3\text{O}^+(\text{H}_2\text{O})_n$ ", *Journal of the American Chemical Society* **104**, 1462–1469 (1982).
- [27] I. Dzidic, D. I. Carroll, R. N. Stillwell, and E. C. Horning, "Comparison of positive ions formed in Nickel-63 and corona discharge ion sources using nitrogen, argon, isobutane, ammonia and nitric oxide as reagents in atmospheric pressure ionization mass spectrometry", *Analytical Chemistry* **48**, 1763–1768 (1976).
- [28] R. B. Cody, J. A. Laramée, and H. D. Durst, "Versatile new ion source for the analysis of materials in open air under ambient conditions", *Analytical Chemistry* **77**, 2297–2302 (2005).
- [29] H. Chen, G. Gamez, and R. Zenobi, "What can we learn from ambient ionization techniques?", *Journal of the American Society for Mass Spectrometry* **20**, 1947–1963 (2009).
- [30] J.-C. Wolf, L. Gyr, M. F. Mirabelli, M. Schaer, P. Siegenthaler, and R. Zenobi, "A Radical-Mediated Pathway for the Formation of $[\text{M} + \text{H}]^+$ in Dielectric Barrier Discharge Ionization", *Journal of The American Society for Mass Spectrometry* **27**, 1468–1475 (2016).
- [31] J. A. Syage, "Mechanism of $[\text{M} + \text{H}]^+$ formation in photoionization mass spectrometry", *Journal of the American Society for Mass Spectrometry* **15**, 1521–1533 (2004).
- [32] S. Klee, S. Albrecht, V. Derpmann, H. Kersten, and T. Benter, "Generation of ion-bound solvent clusters as reactant ions in dopant-assisted APPI and APLI.", *Analytical and bioanalytical chemistry* **405**, 6933–6951 (2013).
- [33] G. Nicol, J. Sunner, and P. Kebarle, "Kinetics and thermodynamics of protonation reactions: $\text{H}_3\text{O}^+(\text{H}_2\text{O})_h + \text{B} = \text{BH}^+(\text{H}_2\text{O})_b + (h-b+1)\text{H}_2\text{O}$, where B is a nitrogen, oxygen or carbon base", *International Journal of Mass Spectrometry and Ion Processes* **84**, 135–155 (1988).
- [34] G. I. Taylor, "Disintegration of water drops in an electric field", *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **280**, 383–397 (1964).
- [35] L. T. Cherney, "Structure of Taylor cone-jets: limit of low flow rates", *Journal of Fluid Mechanics* **378**, 167–196 (1999).

References

- [36] J. V. Iribarne and B. A. Thomson, "On the evaporation of small ions from charged droplets", *The Journal of Chemical Physics* **64**, 2287–2294 (1976).
- [37] B. A. Thomson and J. V. Iribarne, "Field induced ion evaporation from liquid surfaces at atmospheric pressure", *The Journal of Chemical Physics* **71**, 4451–4463 (1979).
- [38] M. Dole, L. L. Mack, R. L. Hines, R. C. Mobley, L. D. Ferguson, and M. B. Alice, "Molecular Beams of Macroions", *The Journal of Chemical Physics* **49**, 2240–2249 (1968).
- [39] L. L. Mack, P. Kralik, A. Rheude, and M. Dole, "Molecular Beams of Macroions. II", *The Journal of Chemical Physics* **52**, 4977–4986 (1970).
- [40] J. B. Fenn, "Ion formation from charged droplets: roles of geometry, energy, and time", *Journal of the American Society for Mass Spectrometry* **4**, 524–535 (1993).
- [41] L. Konermann, A. D. Rodriguez, and J. Liu, "On the Formation of Highly Charged Gaseous Ions from Unfolded Proteins by Electrospray Ionization", *Analytical Chemistry* **84**, 6798–6804 (2012).
- [42] E. Ahadi and L. Konermann, "Modeling the Behavior of Coarse-Grained Polymer Chains in Charged Water Droplets: Implications for the Mechanism of Electrospray Ionization", *The Journal of Physical Chemistry B* **116**, 104–112 (2012).
- [43] L. Konermann, E. Ahadi, A. D. Rodriguez, and S. Vahidi, "Unraveling the Mechanism of Electrospray Ionization", *Analytical Chemistry* **85**, 2–9 (2013).
- [44] E. Ahadi and L. Konermann, "Ejection of Solvated Ions from Electrosprayed Methanol/Water Nanodroplets Studied by Molecular Dynamics Simulations", *Journal of the American Chemical Society* **133**, 9354–9363 (2011).
- [45] L. Konermann, R. G. McAllister, and H. Metwally, "Molecular Dynamics Simulations of the Electrospray Process: Formation of NaCl Clusters via the Charged Residue Mechanism", *The Journal of Physical Chemistry B* **118**, 12025–12033 (2014).
- [46] R. G. McAllister, H. Metwally, Y. Sun, and L. Konermann, "Release of Native-like Gaseous Proteins from Electrospray Droplets via the Charged Residue Mechanism: Insights from Molecular Dynamics Simulations", *Journal of the American Chemical Society* **137**, 12667–12676 (2015).
- [47] K. A. Servage, J. A. Silveira, K. L. Fort, and D. H. Russell, "From Solution to Gas Phase: The Implications of Intramolecular Interactions on the Evaporative Dynamics of Substance P During Electrospray Ionization", *The Journal of Physical Chemistry B* **119**, 4693–4698 (2015).

- [48] W. Wissdorf, L. Seifert, V. Derpmann, S. Klee, W. Vautz, and T. Benter, “Monte Carlo Simulation of Ion Trajectories of Reacting Chemical Systems: Mobility of Small Water Clusters in Ion Mobility Spectrometry”, *Journal of The American Society for Mass Spectrometry* **24**, 632–641 (2013).
- [49] N. Krylova, E. Krylov, G. A. Eiceman, and J. A. Stone, “Effect of Moisture on the Field Dependence of Mobility for Gas-Phase Ions of Organophosphorus Compounds at Atmospheric Pressure with Field Asymmetric Ion Mobility Spectrometry”, *The Journal of Physical Chemistry A* **107**, 3648–3654 (2003).
- [50] E. Krylov and E. Nazarov, “Electric field dependence of the ion mobility”, *International Journal of Mass Spectrometry* **285**, 149–156 (2009).
- [51] L. C. Rorrer and R. A. Yost, “Solvent vapor effects in planar high-field asymmetric waveform ion mobility spectrometry: Solvent trends and temperature effects”, *International Journal of Mass Spectrometry* **378**, 336–346 (2015).
- [52] T. Imasaka, D. S. Moore, and T. Vo-Dinh, “Critical assessment: Use of supersonic jet spectrometry for complex mixture analysis (IUPAC Technical Report)”, *Pure and Applied Chemistry* **75**, 975–998 (2003).
- [53] J. M. Purcell, C. L. Hendrickson, R. P. Rodgers, and A. G. Marshall, “Atmospheric pressure photoionization proton transfer for complex organic mixtures investigated by fourier transform ion cyclotron resonance mass spectrometry”, *Journal of the American Society for Mass Spectrometry* **18**, 1682–1689 (2007).
- [54] M. Meot-Ner (Mautner), “The Ionic Hydrogen Bond”, *Chemical Reviews* **105**, 213–284 (2005).
- [55] C. Hardin, T. V. Pogorelov, and Z. Luthey-Schulten, “Ab initio protein structure prediction”, *Current Opinion in Structural Biology* **12**, 176–181 (2002).
- [56] A. Kryshchuk and K. Fidelis, “Protein structure prediction and model quality assessment”, *Drug Discovery Today* **14**, 386–393 (2009).
- [57] S. D. Lam, S. Das, I. Sillitoe, and C. Orengo, “An overview of comparative modelling and resources dedicated to large-scale modelling of genome sequences”, *Acta Crystallographica Section D Structural Biology* **73**, 628–640 (2017).
- [58] M. Tokmina-Lukaszewska, A. Patterson, L. Berry, L. Scott, N. Balasubramanian, and B. Bothner, “The Role of Mass Spectrometry in Structural Studies of Flavin-Based Electron Bifurcating Enzymes”, *Frontiers in Microbiology* **9**, 1–18 (2018).
- [59] F. L. Brancia, M. Stener, and A. Magistrato, “A density functional theory (DFT) study on gas-phase proton transfer reactions of derivatized and underivatized peptide ions generated by matrix-assisted laser desorption ionization”, *Journal of the American Society for Mass Spectrometry* **20**, 1327–1333 (2009).

References

- [60] E. R. Stephens, M. Dumlao, D. Xiao, D. Zhang, and W. A. Donald, “Benzylammonium Thermometer Ions: Internal Energies of Ions Formed by Low Temperature Plasma and Atmospheric Pressure Chemical Ionization”, *Journal of The American Society for Mass Spectrometry* **26**, 2081–2084 (2015).
- [61] Y. Valadbeigi, V. Ilbeigi, B. Michalczuk, M. Sabo, and S. Matejcik, “Effect of Basicity and Structure on the Hydration of Protonated Molecules, Proton-Bound Dimer and Cluster Formation: An Ion Mobility-Time of Flight Mass Spectrometry and Theoretical Study”, *Journal of The American Society for Mass Spectrometry* **30**, 1242–1253 (2019).
- [62] P. B. Armentrout, “The Power of Accurate Energetics (or Thermochemistry: What is it Good for?)”, *Journal of The American Society for Mass Spectrometry* **24**, 173–185 (2013).
- [63] H. Kersten, V. Funcke, M. Lorenz, K. J. Brockmann, T. Benter, and R. O’Brien, “Evidence of neutral radical induced analyte ion transformations in APPI and Near-VUV APLI”, *Journal of the American Society for Mass Spectrometry* **20**, 1868–1880 (2009).
- [64] S. Beuck, T. Schwabe, S. Grimme, N. Schlörer, M. Kamber, W. Schänzer, and M. Thevis, “Unusual mass spectrometric dissociation pathway of protonated isoquinoline-3-carboxamides due to multiple reversible water adduct formation in the gas phase”, *Journal of the American Society for Mass Spectrometry* **20**, 2034–2048 (2009).
- [65] T. J. Kauppila, A. Haack, K. Kroll, H. Kersten, and T. Benter, “Nucleophilic Aromatic Substitution Between Halogenated Benzene Dopants and Nucleophiles in Atmospheric Pressure Photoionization”, *Journal of The American Society for Mass Spectrometry* **27**, 422–431 (2016).
- [66] V. Gabelica, E. De Pauw, and M. Karas, “Influence of the capillary temperature and the source pressure on the internal energy distribution of electrosprayed ions”, *International Journal of Mass Spectrometry* **231**, 189–195 (2004).
- [67] J. E. Carpenter, C. P. McNary, A. Furin, A. E. Sweeney, and P. B. Armentrout, “How Hot are Your Ions Really? A Threshold Collision-Induced Dissociation Study of Substituted Benzylpyridinium ‘Thermometer’ Ions”, *Journal of The American Society for Mass Spectrometry* **28**, 1876–1888 (2017).
- [68] J. L. Campbell, M. Zhu, and W. S. Hopkins, “Ion-molecule clustering in differential mobility spectrometry: Lessons learned from tetraalkylammonium cations and their isomers”, *Journal of the American Society for Mass Spectrometry* **25**, 1583–1591 (2014).

- [69] C. Liu, J. C. Y. Le Blanc, J. Shields, J. S. Janiszewski, C. Ieritano, G. F. Ye, G. F. Hawes, W. S. Hopkins, and J. L. Campbell, "Using differential mobility spectrometry to measure ion solvation: an examination of the roles of solvents and ionic structures in separating quinoline-based drugs", *The Analyst* **140**, 6897–6903 (2015).
- [70] J. B. Foresman and Æ. Frisch, *Exploring Chemistry with Electronic Structure Methods*, 3rd ed. (Gaussian, Inc., Wallingford, CT, USA, 2015), p. 531.
- [71] E. G. Lewars, *Computational Chemistry*, 2nd ed. (Springer Netherlands, Dordrecht, Netherlands, 2011), p. 664.
- [72] J. Reinhold, *Quantentheorie der Moleküle*, 3rd ed. (Vieweg+Teubner, Wiesbaden, Germany, 2006), p. 354.
- [73] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd ed. (Wiley-VCH, Weinheim, Germany, 2001), p. 313.
- [74] P. Atkins and J. de Paula, *Atkins' Physical Chemistry*, 9th ed. (Oxford University Press, Oxford, New York, 2010), p. 972.
- [75] F. Weigend and R. Ahlrichs, "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy", *Physical Chemistry Chemical Physics* **7**, 3297–3305 (2005).
- [76] T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen", *The Journal of Chemical Physics* **90**, 1007–1023 (1989).
- [77] R. A. Kendall, T. H. Dunning, and R. J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions", *The Journal of Chemical Physics* **96**, 6796–6806 (1992).
- [78] D. E. Woon and T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon", *The Journal of Chemical Physics* **98**, 1358–1371 (1993).
- [79] D. E. Woon and T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties", *The Journal of Chemical Physics* **100**, 2975–2988 (1994).
- [80] F. Weigend, F. Furche, and R. Ahlrichs, "Gaussian basis sets of quadruple zeta valence quality for atoms H–Kr", *The Journal of Chemical Physics* **119**, 12753–12762 (2003).
- [81] F. Weigend, "Accurate Coulomb-fitting basis sets for H to Rn", *Physical Chemistry Chemical Physics* **8**, 1057 (2006).
- [82] W. J. Hehre, R. F. Stewart, and J. A. Pople, "Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals", *The Journal of Chemical Physics* **51**, 2657–2664 (1969).

References

- [83] W. J. Hehre, R. Ditchfield, and J. A. Pople, "Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules", *The Journal of Chemical Physics* **56**, 2257–2261 (1972).
- [84] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas", *Physical Review* **136**, B864–B871 (1964).
- [85] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", *Physical Review* **140**, A1133–A1138 (1965).
- [86] A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange", *The Journal of Chemical Physics* **98**, 5648–5652 (1993).
- [87] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", *Physical Review B* **37**, 785–789 (1988).
- [88] S. Grimme, "Semiempirical hybrid density functional with perturbative second-order correlation", *The Journal of Chemical Physics* **124**, 034108 (2006).
- [89] S. Grimme, "Density functional theory with London dispersion corrections", *Wiley Interdisciplinary Reviews: Computational Molecular Science* **1**, 211–228 (2011).
- [90] S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth, "Dispersion-Corrected Mean-Field Electronic Structure Methods", *Chemical Reviews* **116**, 5105–5154 (2016).
- [91] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu", *The Journal of Chemical Physics* **132**, 154104 (2010).
- [92] S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory", *Journal of Computational Chemistry* **32**, 1456–1465 (2011).
- [93] K. Fukui, "The path of chemical reactions - the IRC approach", *Accounts of Chemical Research* **14**, 363–368 (1981).
- [94] S. Maeda, K. Ohno, and K. Morokuma, "Systematic exploration of the mechanism of chemical reactions: the global reaction route mapping (GRRM) strategy using the ADDF and AFIR methods", *Physical Chemistry Chemical Physics* **15**, 3683–3701 (2013).
- [95] E. B. Wilson Jr., J. C. Decius, and P. C. Cross, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra* (Dover Publications, Inc., New York, NY, USA, 1980), p. 388.
- [96] J. W. Ochterski, *Vibrational Analysis in Gaussian*, 2018, <https://gaussian.com/vib/> (visited on 08/15/2019).

- [97] C. Eckart, "Some Studies Concerning Rotating Axes and Polyatomic Molecules", *Physical Review* **47**, 552–558 (1935).
- [98] A. Sayvetz, "The Kinetic Energy of Polyatomic Molecules", *The Journal of Chemical Physics* **7**, 383–389 (1939).
- [99] E. B. Wilson Jr., "A Method of Obtaining the Expanded Secular Equation for the Vibration Frequencies of a Molecule", *The Journal of Chemical Physics* **7**, 1047–1052 (1939).
- [100] E. B. Wilson Jr., "Some Mathematical Methods for the Study of Molecular Vibrations", *The Journal of Chemical Physics* **9**, 76–84 (1941).
- [101] G. N. Lewis and M. Randall, *Thermodynamics*, edited by K. S. Pitzer and L. Brewer, 2nd ed. (McGraw-Hill Inc., New York, NY, USA, 1961), p. 723.
- [102] P. Y. Ayala and H. B. Schlegel, "Identification and treatment of internal rotation in normal mode vibrational analysis", *The Journal of Chemical Physics* **108**, 2314–2325 (1998).
- [103] D. A. Clabo Jr., W. D. Allen, R. B. Remington, Y. Yamaguchi, and H. F. Schaefer, "A systematic study of molecular vibrational anharmonicity and vibration-rotation interaction by self-consistent-field higher-derivative methods. Asymmetric top molecules", *Chemical Physics* **123**, 187–239 (1988).
- [104] V. Barone, "Vibrational zero-point energies and thermodynamic functions beyond the harmonic approximation", *The Journal of Chemical Physics* **120**, 3059–3065 (2004).
- [105] V. Barone, "Anharmonic vibrational properties by a fully automated second-order perturbative approach", *Journal of Chemical Physics* **122**, 014108 (2005).
- [106] J. Bloino, M. Biczysko, and V. Barone, "General Perturbative Approach for Spectroscopy, Thermodynamics, and Kinetics: Methodological Background and Benchmark Studies", *Journal of Chemical Theory and Computation* **8**, 1015–1036 (2012).
- [107] C. C. Marston and G. G. Balint-Kurti, "The Fourier grid Hamiltonian method for bound state eigenvalues and eigenfunctions", *The Journal of Chemical Physics* **91**, 3571–3576 (1989).
- [108] F. Brau and C. Semay, "The Three-Dimensional Fourier Grid Hamiltonian Method", *Journal of Computational Physics* **139**, 127–136 (1998).
- [109] J. Stare and G. G. Balint-Kurti, "Fourier Grid Hamiltonian Method for Solving the Vibrational Schrödinger Equation in Internal Coordinates: Theory and Test Applications", *The Journal of Physical Chemistry A* **107**, 7204–7214 (2003).
- [110] J. Stare, "First-Principle Calculation of Reduced Masses in Vibrational Analysis Using Generalized Internal Coordinates: Some Crucial Aspects and Examples", *Journal of Chemical Information and Modeling* **47**, 840–850 (2007).

References

- [111] T. Graen and H. Grubmüller, “NuSol - Numerical solver for the 3D stationary nuclear Schrödinger equation”, *Computer Physics Communications* **198**, 169–178 (2016).
- [112] D. L. Bunker, “Classical Trajectory Methods”, in *Methods of computational physics, volume 10: atom and molecular scattering*, edited by B. Adler, S. Fernbach, and M. Rotenberg (Academic Press, New York, London, 1971) Chap. 7, pp. 287–324.
- [113] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, “A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules”, *Journal of the American Chemical Society* **117**, 5179–5197 (1995).
- [114] R. Car and M. Parrinello, “Unified Approach for Molecular Dynamics and Density-Functional Theory”, *Physical Review Letters* **55**, 2471–2474 (1985).
- [115] H. B. Schlegel, J. M. Millam, S. S. Iyengar, G. A. Voth, A. D. Daniels, G. E. Scuseria, and M. J. Frisch, “Ab initio molecular dynamics: Propagating the density matrix with Gaussian orbitals”, *The Journal of Chemical Physics* **114**, 9758–9763 (2001).
- [116] S. S. Iyengar, H. B. Schlegel, J. M. Millam, G. A. Voth, G. E. Scuseria, and M. J. Frisch, “Ab initio molecular dynamics: Propagating the density matrix with Gaussian orbitals. II. Generalizations based on mass-weighting, idempotency, energy conservation and choice of initial conditions”, *The Journal of Chemical Physics* **115**, 10291–10302 (2001).
- [117] H. B. Schlegel, S. S. Iyengar, X. Li, J. M. Millam, G. A. Voth, G. E. Scuseria, and M. J. Frisch, “Ab initio molecular dynamics: Propagating the density matrix with Gaussian orbitals. III. Comparison with Born–Oppenheimer dynamics”, *The Journal of Chemical Physics* **117**, 8694–8704 (2002).
- [118] P. L. Houston, *Chemical Kinetics and Reaction Dynamics*, 1st ed. (Dover Publications, Inc., Mineola, NY, USA, 2006), p. 330.
- [119] D. G. Truhlar and A. D. Isaacson, “Simple perturbation theory estimates of equilibrium constants from force fields”, *The Journal of Chemical Physics* **94**, 357–359 (1991).
- [120] G. Vojta, “Berechnung der Schwingungszustandssumme entarteter anharmonischer Oszillatoren mit Hilfe der Mellinschen Integraltransformation”, *Annalen der Physik* **462**, 397–402 (1961).
- [121] F. Calvo, J. P. K. Doye, and D. J. Wales, “Equilibrium properties of clusters in the harmonic superposition approximation”, *Chemical Physics Letters* **366**, 176–183 (2002).
- [122] D. J. Wales, *Energy Landscapes* (Cambridge University Press, Cambridge, U.K., 2003), p. 681.

- [123] T. V. Bogdan, D. J. Wales, and F. Calvo, “Equilibrium thermodynamics from basin-sampling”, *The Journal of Chemical Physics* **124**, 044102 (2006).
- [124] P. B. Armentrout, K. M. Ervin, and M. T. Rodgers, “Statistical rate theory and kinetic energy-resolved ion chemistry: Theory and applications”, *Journal of Physical Chemistry A* **112**, 10071–10083 (2008).
- [125] H. Eyring, “The Activated Complex in Chemical Reactions”, *The Journal of Chemical Physics* **3**, 107–115 (1935).
- [126] M. G. Evans and M. Polanyi, “Some applications of the transition state method to the calculation of reaction velocities, especially in solution”, *Transactions of the Faraday Society* **31**, 875–894 (1935).
- [127] T. Baer and P. M. Mayer, “Statistical Rice-Ramsperger-Kassel-Marcus quasiequilibrium theory calculations in mass spectrometry”, *Journal of the American Society for Mass Spectrometry* **8**, 103–115 (1997).
- [128] M. T. Rodgers, K. M. Ervin, and P. B. Armentrout, “Statistical modeling of collision-induced dissociation thresholds”, *The Journal of Chemical Physics* **106**, 4499–4508 (1997).
- [129] M. T. Rodgers and P. B. Armentrout, “Statistical modeling of competitive threshold collision-induced dissociation”, *The Journal of Chemical Physics* **109**, 1787–1800 (1998).
- [130] E. Thiele, “Smooth Curve Approximation to the Energy-Level Distribution for Quantum Harmonic Oscillators”, *The Journal of Chemical Physics* **39**, 3258–3262 (1963).
- [131] G. Stock and U. Müller, “Flow of zero-point energy and exploration of phase space in classical simulations of quantum relaxation dynamics”, *Journal of Chemical Physics* **111**, 65–76 (1999).
- [132] E. A. Mason and E. W. McDaniel, *Transport Properties of Ions in Gases* (Wiley-VCH, New York, NY, USA, 1988), p. 560.
- [133] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids*, Corrected (John Wiley & Sons, Inc., New York, NY, USA, 1964), p. 1249.
- [134] Z. Li and H. Wang, “Drag force, diffusion coefficient, and electric mobility of small particles. I. Theory applicable to the free-molecule regime”, *Physical Review E* **68**, 061206 (2003).
- [135] W. E. Siems, L. A. Viehland, and H. H. Hill, “Improved Momentum-Transfer Theory for Ion Mobility. 1. Derivation of the Fundamental Equation”, *Analytical Chemistry* **84**, 9782–9791 (2012).
- [136] A. A. Shvartsburg and M. F. Jarrold, “An exact hard-spheres scattering model for the mobilities of polyatomic ions”, *Chemical Physics Letters* **261**, 86–91 (1996).

References

- [137] M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz, and M. F. Jarrold, “Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential”, *The Journal of Physical Chemistry* **100**, 16082–16086 (1996).
- [138] S. S. Andrews, “Using Rotational Averaging To Calculate the Bulk Response of Isotropic and Anisotropic Samples from Molecular Parameters”, *Journal of Chemical Education* **81**, 877–885 (2004).
- [139] J. W. Lee, K. L. Davidson, M. F. Bush, and H. I. Kim, “Collision cross sections and ion structures: Development of a general calculation method via high-quality ion mobility measurements and theoretical modeling”, *Analyst* **142**, 4289–4298 (2017).
- [140] J. W. Lee, H. H. L. Lee, K. L. Davidson, M. F. Bush, and H. I. Kim, “Structural characterization of small molecular ions by ion mobility mass spectrometry in nitrogen drift gas: improving the accuracy of trajectory method calculations”, *The Analyst* **143**, 1786–1796 (2018).
- [141] C. Ieritano, J. Crouse, J. L. Campbell, and W. S. Hopkins, “A parallelized molecular collision cross section package with optimized accuracy and efficiency”, *The Analyst* **144**, 1660–1670 (2019).
- [142] R. Lai, E. D. Dodds, and H. Li, “Molecular dynamics simulation of ion mobility in gases”, *The Journal of Chemical Physics* **148**, 064109 (2018).
- [143] M. Thinius, C. Polaczek, M. Langner, S. Bräkling, A. Haack, H. Kersten, and T. Benter, “Charge Retention/Charge Depletion in ESI-MS: Experimental Evidence”, *Journal of The American Society for Mass Spectrometry*, DOI: 10.1021/jasms.9b00044 (2020).
- [144] A. Haack, C. Polaczek, M. Tsolakis, M. Thinius, H. Kersten, and T. Benter, “Charge Retention/Charge Depletion in ESI-MS: Theoretical Rationale”, *Journal of The American Society for Mass Spectrometry*, DOI: 10.1021/jasms.9b00045 (2020).
- [145] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse, “Electrospray ionization for mass spectrometry of large biomolecules”, *Science* **246**, 64–71 (1989).
- [146] M. Mann, C. K. Meng, and J. B. Fenn, “Interpreting mass spectra of multiply charged ions”, *Analytical Chemistry* **61**, 1702–1708 (1989).
- [147] X. Han, M. Jin, K. Breuker, and F. W. McLafferty, “Extending Top-Down Mass Spectrometry to Proteins with Masses Greater Than 200 Kilodaltons”, *Science* **314**, 109–112 (2006).
- [148] A. T. Iavarone and E. R. Williams, “Mechanism of Charging and Supercharging Molecules in Electrospray Ionization”, *Journal of the American Chemical Society* **125**, 2319–2327 (2003).

- [149] S. H. Lomeli, I. X. Peng, S. Yin, R. R. Ogorzalek Loo, and J. A. Loo, “New reagents for increasing ESI multiple charging of proteins and protein complexes”, *Journal of the American Society for Mass Spectrometry* **21**, 127–131 (2010).
- [150] H. Metwally, R. G. McAllister, V. Popa, and L. Konermann, “Mechanism of Protein Supercharging by Sulfolane and m -Nitrobenzyl Alcohol: Molecular Dynamics Simulations of the Electrospray Process”, *Analytical Chemistry* **88**, 5345–5354 (2016).
- [151] H. Hinneburg, K. Stavenhagen, U. Schweiger-Hufnagel, S. Pengelley, W. Jabs, P. H. Seeberger, D. V. Silva, M. Wuhrer, and D. Kolarich, “The Art of Destruction: Optimizing Collision Energies in Quadrupole-Time of Flight (Q-TOF) Instruments for Glycopeptide-Based Glycoproteomics”, *Journal of the American Society for Mass Spectrometry* **27**, 507–519 (2016).
- [152] G. S. M. Kammeijer, I. Kohler, B. C. Jansen, P. J. Hensbergen, O. A. Mayboroda, D. Falck, and M. Wuhrer, “Dopant Enriched Nitrogen Gas Combined with Sheathless Capillary Electrophoresis–Electrospray Ionization–Mass Spectrometry for Improved Sensitivity and Repeatability in Glycopeptide Analysis”, *Analytical Chemistry* **88**, 5849–5856 (2016).
- [153] A. Kharlamova and S. A. McLuckey, “Negative Electrospray Droplet Exposure to Gaseous Bases for the Manipulation of Protein Charge State Distributions”, *Analytical Chemistry* **83**, 431–437 (2011).
- [154] J. A. Silveira, K. L. Fort, D. Kim, K. A. Servage, N. A. Pierson, D. E. Clemmer, and D. H. Russell, “From Solution to the Gas Phase: Stepwise Dehydration and Kinetic Trapping of Substance P Reveals the Origin of Peptide Conformations”, *Journal of the American Chemical Society* **135**, 19147–19153 (2013).
- [155] K. L. Fort, J. A. Silveira, N. A. Pierson, K. A. Servage, D. E. Clemmer, and D. H. Russell, “From Solution to the Gas Phase: Factors That Influence Kinetic Trapping of Substance P in the Gas Phase”, *The Journal of Physical Chemistry B* **118**, 14336–14344 (2014).
- [156] S. A. McLuckey, G. J. Van Berkel, and G. L. Glish, “Reactions of dimethylamine with multiply charged ions of cytochrome c”, *Journal of the American Chemical Society* **112**, 5668–5670 (1990).
- [157] B. E. Winger, K. J. Light-Wahl, and R. D. Smith, “Gas-phase proton transfer reactions involving multiply charged cytochrome c ions and water under thermal conditions”, *Journal of the American Society for Mass Spectrometry* **3**, 624–630 (1992).
- [158] E. R. Williams, “Proton Transfer Reactivity of Large Multiply Charged Ions”, *Journal of Mass Spectrometry* **31**, 831–842 (1996).
- [159] K. B. Shelimov and M. F. Jarrold, “Conformations, Unfolding, and Refolding of Apomyoglobin in Vacuum: An Activation Barrier for Gas-Phase Protein Folding”, *Journal of the American Chemical Society* **119**, 2987–2994 (1997).

References

- [160] M. H. Amad, N. B. Cech, G. S. Jackson, and C. G. Enke, "Importance of gas-phase proton affinities in determining the electrospray ionization response for analytes and solvents", *Journal of Mass Spectrometry* **35**, 784–789 (2000).
- [161] A. Moser, K. Range, and D. M. York, "Accurate Proton Affinity and Gas-Phase Basicity Values for Molecules Important in Biocatalysis", *The Journal of Physical Chemistry B* **114**, 13911–13921 (2010).
- [162] E. P. L. Hunter and S. G. Lias, "Proton Affinity Evaluation", in *Nist chemistry webbook, nist standard reference database number 69*, edited by P. J. Linstrom and W. G. Mallard (Gaithersburg MD, 20899).
- [163] A. T. Iavarone, J. C. Jurchen, and E. R. Williams, "Effects of solvent on the maximum charge state and charge state distribution of protein ions produced by electrospray ionization", *Journal of the American Society for Mass Spectrometry* **11**, 976–985 (2000).
- [164] J. C. DeMuth, J. Bu, and S. A. McLuckey, "Electrospray droplet exposure to polar vapors: Delayed desolvation of protein complexes", *Rapid Communications in Mass Spectrometry* **29**, 973–981 (2015).
- [165] E. P. Grimsrud and P. Kebarle, "Gas phase ion equilibriums studies of the hydrogen ion by methanol, dimethyl ether, and water. Effect of hydrogen bonding", *Journal of the American Chemical Society* **95**, 7939–7943 (1973).
- [166] R. Knochenmuss, O. Cheshnovsky, and S. Leutwyler, "Proton transfer reactions in neutral gas-phase clusters: 1-Naphthol with H₂O, D₂O, CH₃OH, NH₃ and piperidine", *Chemical Physics Letters* **144**, 317–323 (1988).
- [167] K. Hirao, M. Sano, and S. Yamabe, "Theoretical study on the gas-phase solvation of the proton by methanol and dimethyl ether", *Chemical Physics Letters* **87**, 181–185 (1982).
- [168] J. J. Fifen, M. Nsangou, Z. Dhaouadi, O. Motapon, and N.-E. Jaidane, "Solvation energies of the proton in methanol", *Journal of Chemical Theory and Computation* **9**, 1173–1181 (2013).
- [169] J. J. Fifen, M. Nsangou, Z. Dhaouadi, O. Motapon, and N.-E. Jaidane, "Structures of protonated methanol clusters and temperature effects", *The Journal of Chemical Physics* **138**, 184301 (2013).
- [170] M. Meot-Ner, "Solvation of the proton by hydrogen cyanide and acetonitrile. Condensation of hydrogen cyanide with ions in the gas phase", *Journal of the American Chemical Society* **100**, 4694–4699 (1978).
- [171] C. A. Deakyne, M. Meot-Ner (Mautner), C. L. Campbell, M. G. Hughes, and S. P. Murphy, "Multicomponent cluster ions. I. The proton solvated by CH₃CN/H₂O", *The Journal of Chemical Physics* **84**, 4958–4969 (1986).
- [172] C. La Manna, "Intermolecular potential of the acetonitrile dimer obtained from ab initio calculations", *Chemical Physics Letters* **103**, 55–58 (1983).

- [173] J. M. Mestdagh, A. Binet, and O. Sublemontier, "Solvation shells of the proton surrounded by acetonitrile, ethanol, and water molecules", *The Journal of Physical Chemistry* **93**, 8300–8303 (1989).
- [174] S. T. Graul and R. R. Squires, "Energy-resolved collision-induced dissociation of proton-bound cluster ions as a structural probe: the acetonitrile-water system", *International Journal of Mass Spectrometry and Ion Processes* **94**, 41–61 (1989).
- [175] M. Samy El-Shall, S. Olafsdottir, M. Meot-Ner (Mautner), and L. Wayne Sieck, "Energy effects on cluster ion distributions: beam expansion and thermochemical studies on mixed clusters of methanol and acetonitrile", *Chemical Physics Letters* **185**, 193–198 (1991).
- [176] X. Zhang and A. Castleman, "Kinetic study of protonated methanol clusters with acetonitrile", *International Journal of Mass Spectrometry and Ion Processes* **149-150**, 521–527 (1995).
- [177] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, *Gaussian 09, Revision C.01* (Gaussian, Inc., Wallingford, CT, 2009).
- [178] S. Grimme, "Semiempirical GGA-type density functional constructed with a long-range dispersion correction", *Journal of Computational Chemistry* **27**, 1787–1799 (2006).
- [179] J. W. Ochterski, *Thermochemistry in Gaussian*, 2000, http://www.gaussian.com/g_whitepap/thermo.htm (visited on 12/12/2019).
- [180] M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, "On the Quantum Nature of the Shared Proton in Hydrogen Bonds", *Science* **275**, 817–820 (1997).
- [181] J. M. Headrick, E. G. Diken, R. S. Walter, N. I. Hammer, R. A. Christie, J. Cui, E. M. Myshakin, M. A. Duncan, M. A. Johnson, and K. D. Jordan, "Spectral Signatures of Hydrated Proton Vibrations in Water Clusters", *Science* **308**, 1765–1769 (2005).

References

- [182] S. M. Craig, F. S. Menges, C. H. Duong, J. K. Denton, L. R. Madison, A. B. McCoy, and M. A. Johnson, "Hidden role of intermolecular proton transfer in the anomalously diffuse vibrational spectrum of a trapped hydronium ion", *Proceedings of the National Academy of Sciences* **114**, E4706–E4713 (2017).
- [183] D. Bing, T. Hamashima, C.-W. Tsai, A. Fujii, and J.-L. Kuo, "Proton location in $(\text{CH}_3)_3\text{N}-\text{H}^+-(\text{CH}_3\text{OH})_n$: A theoretical and infrared spectroscopic study", *Chemical Physics* **421**, 1–9 (2013).
- [184] M. Meot-Ner (Mautner), "The ionic hydrogen bond and ion solvation. 1. $\text{NH}^+\cdots\text{O}$, $\text{NH}^+\cdots\text{N}$, and $\text{OH}^+\cdots\text{O}$ bonds. Correlations with proton affinity. Deviations due to structural effects", *Journal of the American Chemical Society* **106**, 1257–1264 (1984).
- [185] M. Meot-Ner (Mautner) and L. W. Sieck, "The ionic hydrogen bond and ion solvation. 4. $\text{SH}^+\cdots\text{O}$ and $\text{NH}^+\cdots\text{S}$ bonds. Correlations with proton affinity. Mutual effects of weak and strong ligands in mixed clusters", *The Journal of Physical Chemistry* **89**, 5222–5225 (1985).
- [186] D. Marx, "Proton Transfer 200 Years after von Grotthuss: Insights from Ab Initio Simulations", *ChemPhysChem* **7**, 1848–1870 (2006).
- [187] W. Y. Feng, M. Iraqi, and C. Lifshitz, "Reactions of $\text{H}^+(\text{CH}_3\text{OH})_3$ with a series of base molecules", *The Journal of Physical Chemistry* **97**, 3510–3514 (1993).
- [188] W. Y. Feng, M. Goldenberg, and C. Lifshitz, "Reactions of proton-bound dimers", *Journal of the American Society for Mass Spectrometry* **5**, 695–703 (1994).
- [189] M. Tsolakakis, "Ab-initio-Berechnungen zur Thermochemie protonengebundener Cluster", Bachelor Thesis (Bergische Universität Wuppertal, 2017), p. 50.
- [190] K. A. Peterson, D. Feller, and D. A. Dixon, "Chemical accuracy in ab initio thermochemistry and spectroscopy: current strategies and future challenges", *Theoretical Chemistry Accounts* **131**, 1079 (2012).
- [191] E. P. L. Hunter and S. G. Lias, "Evaluated Gas Phase Basicities and Proton Affinities of Molecules: An Update", *Journal of Physical and Chemical Reference Data* **27**, 413–656 (1998).
- [192] T. I. Williams, J. W. Denault, and R. Cooks, "Proton affinity of deuterated acetonitrile estimated by the kinetic method with full entropy analysis", *International Journal of Mass Spectrometry* **210-211**, 133–146 (2001).
- [193] K. Honma, L. S. Sunderlin, and P. B. Armentrout, "Guided-ion beam studies of the reactions of protonated water clusters, $\text{H}(\text{H}_2\text{O})_n^+$ ($n = 1 - 4$), with acetonitrile", *Journal of Chemical Physics* **99**, 1623–1632 (1993).

- [194] A. Haack, J. Crouse, F.-J. Schlüter, T. Benter, and W. S. Hopkins, “A First Principle Model of Differential Ion Mobility: the Effect of Ion-Solvent Clustering”, *Journal of The American Society for Mass Spectrometry* **30**, 2711–2725 (2019).
- [195] J. Crouse, A. Haack, T. Benter, and W. S. Hopkins, “Understanding non-traditional differential mobility behavior: a case study of the tricarbastannatrane cation, $N(CH_2CH_2CH_2)_3Sn^+$ ”, *Journal of the American Society for Mass Spectrometry*, DOI: 10.1021/jasms.9b00042 (2020).
- [196] F.-J. Schlüter, “Ab-initio-Berechnungen zur Bestimmung differentieller Mobilitäten von geklusterten Tetramethylammoniumionen”, Bachelor-Thesis (Bergische Universität Wuppertal, 2018), p. 49.
- [197] W. Wissdorf, L. Pohler, S. Klee, D. Müller, and T. Benter, “Simulation of Ion Motion at Atmospheric Pressure: Particle Tracing Versus Electrokinetic Flow”, *Journal of The American Society for Mass Spectrometry* **23**, 397–406 (2012).
- [198] W. Wissdorf, M. Lorenz, T. Pöhler, H. Hönen, and T. Benter, “Atmospheric pressure ion source development: experimental validation of simulated ion trajectories within complex flow and electrical fields.”, *Journal of the American Society for Mass Spectrometry* **24**, 1456–66 (2013).
- [199] W. Wissdorf, M. Lorenz, K. Brockmann, and T. Benter, “Systematic Ion Source Parameter Assessment by Automated Determination of the Distribution of Ion Acceptance (DIA) Using APLI”, *Journal of The American Society for Mass Spectrometry* **30**, 1262–1275 (2019).
- [200] G. A. Eiceman and Z. Karpas, *Ion Mobility Spectrometry*, 2nd ed. (CRC Press, Boca Raton, FL, USA, 2005), pp. xvi + 350.
- [201] C. Laphorn, F. Pullen, and B. Z. Chowdhry, “Ion mobility spectrometry-mass spectrometry (IMS-MS) of small molecules: Separating and assigning structures to ions”, *Mass Spectrometry Reviews* **32**, 43–71 (2013).
- [202] V. Gabelica and E. Marklund, “Fundamentals of ion mobility spectrometry”, *Current Opinion in Chemical Biology* **42**, 51–59 (2018).
- [203] L. A. Viehland, “Velocity distribution functions and transport coefficients of atomic ions in atomic gases by a Gram-Charlier approach”, *Chemical Physics* **179**, 71–92 (1994).
- [204] L. A. Viehland and W. F. Siems, “Uniform Moment Theory for Charged Particle Motion in Gases”, *Journal of The American Society for Mass Spectrometry* **23**, 1841–1854 (2012).
- [205] L. A. Viehland, “Uniform Moment Theory for Charged Particle Motion in Gases. 2. Second Approximation”, *Journal of Statistical Physics* **163**, 175–196 (2016).

References

- [206] L. Boltzmann, "Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen", *Sitzungsberichte der kaiserlichen Akademie der Wissenschaften (II)* **66**, 275–370 (1872).
- [207] I. Campuzano, M. F. Bush, C. V. Robinson, C. Beaumont, K. Richardson, H. Kim, and H. I. Kim, "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).
- [208] H. Ouyang, C. Larriba-Andaluz, D. R. Oberreit, and C. J. Hogan, "The Collision Cross Sections of Iodide Salt Cluster Ions in Air via Differential Mobility Analysis-Mass Spectrometry", *Journal of The American Society for Mass Spectrometry* **24**, 1833–1847 (2013).
- [209] T. G. Flick, I. D. G. Campuzano, and M. D. Bartberger, "Structural Resolution of 4-Substituted Proline Diastereomers with Ion Mobility Spectrometry via Alkali Metal Ion Cationization", *Analytical Chemistry* **87**, 3300–3307 (2015).
- [210] C. D. Chouinard, V. W. D. Cruzeiro, C. R. Beekman, A. E. Roitberg, and R. A. Yost, "Investigating Differences in Gas-Phase Conformations of 25-Hydroxyvitamin D3 Sodiated Epimers using Ion Mobility-Mass Spectrometry and Theoretical Modeling", *Journal of The American Society for Mass Spectrometry* **28**, 1497–1505 (2017).
- [211] G. H. Wannier, "Motion of Gaseous Ions in Strong Electric Fields", *Bell System Technical Journal* **32**, 170–254 (1953).
- [212] L. A. Viehland and E. A. Mason, "Gaseous Ion mobility in electric fields of arbitrary strength", *Annals of Physics* **91**, 499–533 (1975).
- [213] L. A. Viehland and E. A. Mason, "Gaseous ion mobility and diffusion in electric fields of arbitrary strength", *Annals of Physics* **110**, 287–328 (1978).
- [214] W. F. Siems, L. A. Viehland, and H. H. Hill, "Correcting the fundamental ion mobility equation for field effects", *The Analyst* **141**, 6396–6407 (2016).
- [215] I. Buryakov, E. Krylov, E. Nazarov, and U. K. Rasulev, "A new method of separation of multi-atomic ions by mobility at atmospheric pressure using a high-frequency amplitude-asymmetric strong electric field", *International Journal of Mass Spectrometry and Ion Processes* **128**, 143–148 (1993).
- [216] A. A. Shvartsburg, *Differential Ion Mobility Spectrometry: Nonlinear Ion Transport and Fundamentals of FAIMS* (CRC Press, Boca Raton, FL, USA, 2009), p. 322.
- [217] B. B. Schneider, T. R. Covey, S. L. Coy, E. V. Krylov, and E. G. Nazarov, "Chemical effects in the separation process of a differential mobility/mass spectrometer system.", *Analytical chemistry* **82**, 1867–1880 (2010).

- [218] B. B. Schneider, T. R. Covey, S. L. Coy, E. V. Krylov, and E. G. Nazarov, "Planar differential mobility spectrometer as a pre-filter for atmospheric pressure ionization mass spectrometry.", *International journal of mass spectrometry* **298**, 45–54 (2010).
- [219] B. B. Schneider, T. R. Covey, S. L. Coy, E. V. Krylov, and E. G. Nazarov, "Control of chemical effects in the separation process of a differential mobility mass spectrometer system.", *European journal of mass spectrometry (Chichester, England)* **16**, 57–71 (2010).
- [220] E. V. Krylov, S. L. Coy, J. Vandermeij, B. B. Schneider, T. R. Covey, and E. G. Nazarov, "Selection and generation of waveforms for differential mobility spectrometry", *The Review of scientific instruments* **81**, 024101 (2010).
- [221] R. W. Purves and R. Guevremont, "Electrospray Ionization High-Field Asymmetric Waveform Ion Mobility Spectrometry-Mass Spectrometry", *Analytical Chemistry* **71**, 2346–2357 (1999).
- [222] D. J. Wales and J. P. K. Doye, "Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms", *The Journal of Physical Chemistry A* **101**, 5111–5116 (1997).
- [223] W. S. Hopkins, R. A. Marta, and T. B. McMahon, "Proton-Bound 3-Cyanophenylalanine Trimethylamine Clusters: Isomer-Specific Fragmentation Pathways and Evidence of Gas-Phase Zwitterions", *The Journal of Physical Chemistry A* **117**, 10714–10718 (2013).
- [224] U. C. Singh and P. A. Kollman, "An approach to computing electrostatic charges for molecules", *Journal of Computational Chemistry* **5**, 129–145 (1984).
- [225] B. H. Besler, K. M. Merz, and P. A. Kollman, "Atomic charges derived from semiempirical methods", *Journal of Computational Chemistry* **11**, 431–439 (1990).
- [226] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, M. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, *Gaussian 16, Revision A.03* (Gaussian, Inc., Wallingford, CT, USA, 2016).

References

- [227] T. A. Halgren, “Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94”, *Journal of Computational Chemistry* **17**, 490–519 (1996).
- [228] T. A. Halgren, “Merck Molecular Force Field. II. MMFF94 van der Waals and Electrostatic Parameters for Intermolecular Interactions”, *J. Comput. Chem.* **17**, 520–552 (1996).
- [229] T. Wyttenbach, G. Helden, J. J. Batka, D. Carlat, and M. T. Bowers, “Effect of the long-range potential on ion mobility measurements”, *Journal of the American Society for Mass Spectrometry* **8**, 275–282 (1997).
- [230] M. A. Biondi and L. M. Chanin, “Blanc’s Law - Ion Mobilities in Helium-Neon Mixtures”, *Physical Review* **122**, 843–847 (1961).
- [231] C. Liu, J. C. Y. Le Blanc, B. B. Schneider, J. Shields, J. J. Federico, H. Zhang, J. G. Stroh, G. W. Kauffman, D. W. Kung, C. Ieritano, E. Shepherdson, M. Verbuyst, L. Melo, M. Hasan, D. Naser, J. S. Janiszewski, W. S. Hopkins, and J. L. Campbell, “Assessing Physicochemical Properties of Drug Molecules via Microsolvation Measurements with Differential Mobility Spectrometry”, *ACS Central Science* **3**, 101–109 (2017).
- [232] N. J. A. Coughlan, C. Liu, M. J. Lecours, J. L. Campbell, and W. S. Hopkins, “Preferential Ion Microsolvation in Mixed-Modifier Environments Observed Using Differential Mobility Spectrometry”, *Journal of The American Society for Mass Spectrometry* **30**, 2222–2227 (2019).
- [233] E. V. Krylov, S. L. Coy, and E. G. Nazarov, “Temperature effects in differential mobility spectrometry”, *International Journal of Mass Spectrometry* **279**, 119–125 (2009).
- [234] M. Karplus and J. N. Kushick, “Method for estimating the configurational entropy of macromolecules”, *Macromolecules* **14**, 325–332 (1981).
- [235] R. M. Levy, M. Karplus, J. Kushick, and D. Perahia, “Evaluation of the configurational entropy for proteins: application to molecular dynamics simulations of an α -helix”, *Macromolecules* **17**, 1370–1374 (1984).
- [236] O. L. Rojas, R. M. Levy, and A. Szabo, “Corrections to the quasiharmonic approximation for evaluating molecular entropies”, *The Journal of Chemical Physics* **85**, 1037–1043 (1986).
- [237] W. S. Hopkins, “Dynamic Clustering and Ion Microsolvation”, in *Advances in ion mobility-mass spectrometry: fundamentals, instrumentation and applications*, edited by W. A. Donald and J. S. Prell, Volume 83 (Elsevier B.V., 2019) Chap. Four, pp. 83–122.
- [238] S. W. C. Walker, A. Anwar, J. M. Psutka, J. Crouse, C. Liu, J. C. Y. Le Blanc, J. Montgomery, G. H. Goetz, J. S. Janiszewski, J. L. Campbell, and W. S. Hopkins, “Determining molecular properties with differential mobility spectrometry and machine learning”, *Nature Communications* **9**, 5096 (2018).

- [239] P. Simidzija, M. J. Lecours, R. A. Marta, V. Steinmetz, T. B. McMahon, E. Fillion, and W. S. Hopkins, "Changes in Tricarbostannatrane Transannular N–Sn Bonding upon Complexation Reveal Lewis Base Donicities", *Inorganic Chemistry* **55**, 9579–9585 (2016).
- [240] S. Boys and F. Bernardi, "The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors", *Molecular Physics* **19**, 553–566 (1970).
- [241] A. Bondi, "van der Waals Volumes and Radii", *The Journal of Physical Chemistry* **68**, 441–451 (1964).
- [242] C. Thierfelder, B. Assadollahzadeh, P. Schwerdtfeger, S. Schäfer, and R. Schäfer, "Relativistic and electron correlation effects in static dipole polarizabilities for the group-14 elements from carbon to element $Z = 114$: Theory and experiment", *Physical Review A* **78**, 052506 (2008).
- [243] A. Haack, T. Benter, and H. Kersten, "Computational analysis of the proton bound acetonitrile dimer, $(\text{ACN})_2\text{H}^+$ ", *Rapid Communications in Mass Spectrometry*, DOI: 10.1002/rcm.8767 (2020).
- [244] E. Horning, D. Carroll, I. Dzidic, K. Haegele, M. Horning, and R. Stillwell, "Liquid chromatograph-mass spectrometer-computer analytical systems", *Journal of Chromatography A* **99**, 13–21 (1974).
- [245] T. J. Kauppila, T. Kuuranne, E. C. Meurer, M. N. Eberlin, T. Kotiaho, and R. Kostianen, "Atmospheric Pressure Photoionization Mass Spectrometry. Ionization Mechanism and the Effect of Solvent on the Ionization of Naphthalenes", *Analytical Chemistry* **74**, 5470–5479 (2002).
- [246] M. L. Gross and R. Caprioli, "Volume 6: Ionization Methods", in *Encyclopedia of mass spectrometry*, 1st ed. (Elsevier, Oxford, 2007), p. 1030.
- [247] M. M. Shahin, "Mass-Spectrometric Studies of Corona Discharges in Air at Atmospheric Pressures", *The Journal of Chemical Physics* **45**, 2600–2605 (1966).
- [248] S. Klee, M. Thinius, K. J. Brockmann, and T. Benter, "Capillary atmospheric pressure chemical ionization using liquid point electrodes", *Rapid Communications in Mass Spectrometry* **28**, 1591–1600 (2014).
- [249] M. L. Huggins, "Hydrogen Bridges in Ice and Liquid Water", *The Journal of Physical Chemistry* **40**, 723–731 (1935).
- [250] J. M. Bowman, "The self-consistent-field approach to polyatomic vibrations", *Accounts of Chemical Research* **19**, 202–208 (1986).
- [251] G. M. Chaban, J. O. Jung, and R. B. Gerber, "Ab initio calculation of anharmonic vibrational states of polyatomic systems: Electronic structure combined with vibrational self-consistent field", *The Journal of Chemical Physics* **111**, 1823–1829 (1999).

References

- [252] C.-E. Chang, W. Chen, and M. K. Gilson, "Evaluating the Accuracy of the Quasiharmonic Approximation", *Journal of Chemical Theory and Computation* **1**, 1017–1028 (2005).
- [253] S. Hikiri, T. Yoshidome, and M. Ikeguchi, "Computational Methods for Configurational Entropy Using Internal and Cartesian Coordinates", *Journal of Chemical Theory and Computation* **12**, 5990–6000 (2016).
- [254] J. H. Frederick and C. Woywod, "General formulation of the vibrational kinetic energy operator in internal bond-angle coordinates", *The Journal of Chemical Physics* **111**, 7255–7271 (1999).
- [255] J. D. Louck and H. W. Galbraith, "Eckart vectors, Eckart frames, and polyatomic molecules", *Reviews of Modern Physics* **48**, 69–106 (1976).
- [256] D. Janežič, M. Praprotnik, and F. Merzel, "Molecular dynamics integration and molecular vibrational theory. I. New symplectic integrators", *The Journal of Chemical Physics* **122**, 174101 (2005).
- [257] B. Kirtman, "Interactions between Ordinary Vibrations and Hindered Internal Rotation. I. Rotational Energies", *The Journal of Chemical Physics* **37**, 2516–2539 (1962).
- [258] B. Kirtman, "Interactions between Ordinary Vibrations and Hindered Internal Rotation. II. Theory of Internal Rotation Fine Structure in Some Perpendicular Bands of Ethane-Type Molecules", *The Journal of Chemical Physics* **41**, 775–788 (1964).
- [259] B. Kirtman, "Dependence of Inertial Constants on Ordinary Vibrational Coordinates in Molecules with an Internal-Rotation Degree of Freedom", *The Journal of Chemical Physics* **49**, 2257–2260 (1968).
- [260] K. S. Pitzer and W. D. Gwinn, "Energy Levels and Thermodynamic Functions for Molecules with Internal Rotation I. Rigid Frame with Attached Tops", *The Journal of Chemical Physics* **10**, 428–440 (1942).
- [261] K. S. Pitzer, "Energy Levels and Thermodynamic Functions for Molecules with Internal Rotation: II. Unsymmetrical Tops Attached to a Rigid Frame", *The Journal of Chemical Physics* **14**, 239–243 (1946).
- [262] J. E. Kilpatrick and K. S. Pitzer, "Energy Levels and Thermodynamic Functions for Molecules with Internal Rotation. III. Compound Rotation", *The Journal of Chemical Physics* **17**, 1064–1075 (1949).
- [263] D. T. Colbert and W. H. Miller, "A novel discrete variable representation for quantum mechanical reactive scattering via the S-matrix Kohn method", *The Journal of Chemical Physics* **96**, 1982–1991 (1992).
- [264] T. Shimanouchi, *Tables of Molecular Vibrational Frequencies Consolidated Volume I* (National Bureau of Standards (U.S.), 1972), p. 164.