

**CHARGE MODELS TO REPRODUCE
TOPOGRAPHY OF
MOLECULAR ELECTROSTATIC POTENTIAL**

**A Thesis submitted to Goa University for
the award of the Degree of
DOCTOR OF PHILOSOPHY**

**in
CHEMISTRY**

By

Mr. JOHNROSS V. ALBUQUERQUE

GOA UNIVERSITY

Taleigao Plateau, Goa

2020

**CHARGE MODELS TO REPRODUCE
TOPOGRAPHY OF
MOLECULAR ELECTROSTATIC POTENTIAL**

**A Thesis submitted to Goa University for
the award of the Degree of
DOCTOR OF PHILOSOPHY
in
CHEMISTRY**

**By
Mr. JOHNROSS V. ALBUQUERQUE**

**Research Guide
Dr. RAJENDRA N. SHIRSAT**

**GOA UNIVERSITY
Taleigao Plateau, Goa**

2020

SCHOOL OF CHEMICAL SCIENCES

CERTIFICATE

This is to certify that the thesis entitled '**CHARGE MODELS TO REPRODUCE TOPOGRAPHY OF MOLECULAR ELECTROSTATIC POTENTIAL**' submitted by Mr. **JOHNROSS V. ALBUQUERQUE**, is a record of research work carried out by the candidate during the period of study under my supervision and that it has not previously formed the basis for the award of any degree or diploma or other similar titles.

Goa University
January 2020

Dr. Rajendra N. Shirsat
Research Guide
School of Chemical Sciences
Goa University

DECLARATION

I hereby declare that the work embodied in this thesis entitled '**CHARGE MODELS TO REPRODUCE TOPOGRAPHY OF MOLECULAR ELECTROSTATIC POTENTIAL**' is the result of investigations carried out by myself under the guidance of **Dr. RAJENDRA N. SHIRSAT** at the School of Chemical Sciences, Goa University and that the same has not previously formed the basis for the award of any other degree/diploma or other similar titles.

Goa University
January 2020

Mr. Johnross V. Albuquerque
Ph.D. Student
School of Chemical Sciences
Goa University

ACKNOWLEDGEMENT

I am and always will be grateful to my Ph.D. supervisor Professor Rajendra N. Shirsat, School of Chemical Sciences, Goa University, who has supported and guided me throughout my research work. It would have been impossible to complete this task if not for his patience and valuable discussions during times of difficulties. I am also thankful to Professor V. M. S. Verenkar who took up the role of subject expert.

I also take this opportunity to thank Professor Shridhar R. Gadre, Interdisciplinary School of Scientific Computing, Savitribai Phule Pune University, Pune for his support, encouragement words and cheerful discussions which aided in completing this research work.

I wish to thank Dr. Subhash S. Phingale, Department of Chemistry, Savitribai Phule Pune University, Pune for his hospitality on my visit to Savitribai Phule Pune University during June 2014. I thank his two Ph.D. students Ms. Pallavi Rathod and Ms. Anuja Ware for their valuable time spent in training me to use computational chemistry softwares like UNIVIS-2000 and INDPROP, which I have used to carry out my research work. My gratitude also goes out to Dr. C. H. Suresh, Research scientist at the CSIR-National Institute for Interdisciplinary Science and Technology, Trivandrum, India for providing the computational power to carry out some calculations on coordination compounds.

I am thankful to Prof. V. S. Nadkarni, Dean of School of Chemical Sciences, formally known as Department of Chemistry and to the former Heads of this Department namely, Prof. B. R. Srinivasan and Prof. S. G. Tilve for their valuable support, not forgetting other faculty members namely Prof. A. V. Salkar, Prof. J. B. Fernandes, Dr. S. N. Dhuri, Dr. Pranay Morajkar, Dr. Bhanudas Naik, Dr. Vivekanand Gobre, Ms. Siddhali Girkar and the non-teaching staff of School of Chemical Sciences.

My gratitude also goes out to Professor R. V. Pai, Department of Physics, Goa University for allowing me to attend some lectures and practical sessions on FORTRAN programming language during the period between August – September 2014.

Financial aid was provided to me by the University Grants Commission (UGC), New Delhi in the form of UGC-BSR Junior and Senior Fellowship for a period of five years. I am therefore grateful to UGC for the same. My gratitude goes out to my Ph.D. colleagues at the School of Chemical Sciences during the period 2014 – 2019.

Words are short to express my feelings of gratitude towards my parents, Mr. Dominic Felix Albuquerque and Mrs. Augusta Maria Josephine Albuquerque for their unconditional love, support and sacrifices undertaken by them to give me the best possible life.

Johnross V. Albuquerque

Dedicated to my beloved parents

TABLE OF CONTENT

Abbreviations	i
List of publications in peer-reviewed journals	ii
Conferences, symposiums and workshops attended	iii
CHAPTER I. Introduction	1 – 39
I.1 Selective and brief history behind the development of Schrödinger's Wave Equation	2
I.2 Time-Independent Schrödinger Equation	7
I.3 Solutions to the time-independent Schrödinger Equation	7
I.3.1 Hydrogen-like atom	7
I.3.2 Many-particle systems	9
Hartree Product and Hartree approximation	10
Slater Determinants	12
Hartree-Fock method	12
Hartree-Fock Roothaan Equations	15
Basis sets	15
Hartree-Fock energy and Density Matrix	17
I.3.3 Møller-Plesset Perturbation Theory	18
I.3.4 Density Functional Theory	19
I.4 Importance of Theoretical Chemistry	22
I.5 Molecular Electrostatic Potential (MESP)	23
I.5.1 Theoretical evaluation and visualization	23
I.5.2 Tool for predicting susceptibility towards electrophiles	24
I.5.3 Topographical features of MESP	26
I.5.4 Predictor tool for nucleophilic attacks	27
I.5.5 MESP-based Charge Models	28
I.6 Scope of the present work	30
References	33
CHAPTER II. Topography of Molecular Electrostatic Potential	40 – 61
II.1 Introduction	41
II.2 Significance of MESP topography	43

II.3	Analytical approach of locating and characterizing critical points of MESP scalar field	44
II.4	Numerical approach of locating guess points for MESP topography illustrated using pyrazole and imidazole molecules	55
II.5	General remarks and conclusion	58
	References	60

CHAPTER III. Development of Topography-driven Gaussian Charge

	Models	62 – 97
III.1	Introduction	63
III.2	Methodology to develop Gaussian charge models (GCMs)	65
III.2.1	Ammonia molecule	65
III.2.2	Water and hydrogen sulphide molecules	72
III.2.3	Formaldehyde molecule	76
III.2.4	Ethene and acetylene molecules	78
III.2.5	Benzene molecule	82
III.3	Scope for improving MESP and Electron Density values at critical points other than MESP minima	84
III.4	General remarks on GCMs	85
III.5	Additional topography-based Gaussian models	88
III.5.1	Methanol molecule	88
III.5.2	Pyrazine molecule	91
III.6	Future prospects	94
III.7	Concluding remarks	95
	References	96

CHAPTER IV. Investigation of potential energy surfaces via. Gaussian charge models

		98 – 125
IV.1	Introduction	99
IV.2	Computation of interaction energy of a binary complex	101
IV.2.1	Wave function approach	101
IV.2.2	Using GCMs	103
IV.3	Exploration of potential energy surfaces via. GCMs	104
IV.3.1	Water dimer	104

IV.3.2	Ammonia dimer	110
IV.3.3	Ammonia-water complex	115
IV.3.4	Acetylene dimer and acetylene-water complex	116
IV.3.5	Benzene dimer and benzene related systems	117
IV.4	Concluding remarks	121
	References	123

CHAPTER V. Application of quantum mechanical methods in reaction

	mechanisms and forensics	126 – 151
V.1	Cobalt thiocyanate: a useful reagent in forensic	127
V.1.1	Detection of some nitrogen containing compounds via. cobalt thiocyanate and TLC	128
V.1.2	Theoretical structure elucidation of some nitrogen containing compounds with cobalt thiocyanate	129
V.2	Theoretical investigation on 3-aryl-6-ethoxycarbonyl-4-hydroxy-2 <i>H</i> -pyran-2-ones	139
V.3	Concluding remarks	146
	References	148

APPENDIX

152 – 184

A1	Solution to the particle and wave equations	153
A2	Schrödinger's time-dependent equation	154
A3	Schrödinger's time-independent equation	156
A4	Hamiltonian operator for many-particle system	157
A5	Drawback of Hartree Product	158
A6	Slater determinant and antisymmetry principle	159
A7	Cartesian co-ordinates of acetone and acetone-related non-covalent complexes	160
A8	Numerical method to locate critical points of a one-dimensional function via. interpolation	162
A9	The three-dimensional Lagrange interpolation method	168
A10	Guess points for topographical features of MESP scalar field	169
A10.1	Ammonia molecule	169
A10.2	Water molecule	170

A10.3	Hydrogen sulphide	171
A10.4	Formaldehyde molecule	171
A10.5	Ethene molecule	172
A10.6	Benzene molecule	174
A11	Potential due to <i>s</i> -type Gaussians	177
A12	Geometries of some cobalt thiocyanate complexes: comparison between experimental and quantum mechanical structures	180
A13	Certificates of participation/poster presentations of various conferences, symposiums and workshops attended during research period January 2014-December 2019	181
	First page of research articles published during research period January 2014 – December 2019	182

ABBREVIATIONS

3D	Three-dimensional
BSSE	Basis Set Superposition Error
CGTO	Contracted Gaussian-type orbital
CHELP	Charges from electrostatic potentials
CHELPG	Charges from electrostatic potentials using grid
CP	Critical point
DFT	Density Functional Theory
ESP	Electrostatic potential
EPIC	Electrostatic Potential for Intermolecular Complexation
GTO	Gaussian-type orbital
GCM	Gaussian charge model
GP	Guess point
HF	Hartree-Fock
IE	Interaction energy
LP	Lone pair
MESP	Molecular Electrostatic Potential
MK	Merz-Kollman
Molecular orbital	MO
MP2	Second order Møller-Plesset Perturbation theory
NR	Newton Raphson
STO	Slater-type orbital
TLC	Thin layer chromatography
PCM	Point charge model
PES	Potential energy surface
QM	Quantum mechanical
WF	Wave function
vdW	van der Waals

List of publications in peer-reviewed journals

Research articles published:

1. J.V. Albuquerque and R.N. Shirsat, Prelude to molecular dynamics: Topography-driven Gaussian charge models, *Int. J. Quantum Chem.*, DOI: 10.1002/qua.25835.
2. S.R. Kote, J.V. Albuquerque, R.N. Shirsat, R.V. Phadke, J.T. Kohapare and S.S. Dhoble, TLC Detection and Theoretical Structure Elucidation of Nitrogen Containing Compounds with Cobalt Thiocyanate, *Anal. Chem. Lett.*, 9 (2019) 453.

Conferences, symposiums and workshops attended

Conferences:

1. Presented a poster titled ‘Charge Models to reproduce Topography of Molecular Electrostatic Potential of few molecules’ at 2nd National Conference on “New Frontiers in Chemistry – From Fundamentals to Applications – II” organized by Department of Chemistry, BITS Pilani KK Birla Goa Campus on 28th – 29th January 2017 (Poster no. 33).
2. Presented a poster entitled ‘Theoretically proposed structures of Cobalt Thiocyanate with nitrogen containing compounds’ at the Eighth Asian Pacific Conference on Theoretical and Computational Chemistry held at IIT Bombay, Mumbai on 15th – 17th December 2017 (Poster no. P133).

Symposiums

1. Attended a symposium on ‘Chemistry with Computers’ at IICT Hyderabad organized by IIIT Hyderabad and IICT Hyderabad on 18th – 19th January 2014.
2. Presented as a poster titled ‘Prelude to Molecular Dynamics – II: Guidelines for Employing Topography-driven Gaussian Charge Models’ at the 16th Theoretical Chemistry Symposium (TCS – 2019) organized by Department of Chemistry BITS Pilani, Pilani Campus on 13 – 16th February 2019 (Poster no. P177).

Workshops:

1. Attended State level workshop on ‘Quantum Chemistry/Computational Chemistry’ organized by Post Graduate Department of Chemistry, St. Xavier’s College, Mapusa-Goa in collaboration with Indian National Science Academy, New Delhi, on 14th – 16th December 2014.
2. Participated in the Winter School on ‘Computational Chemistry’ at the UGC-Networking Resource Centre, School of Chemistry, University of Hyderabad, Hyderabad during 8th – 24th December 2015.
3. Attended State level workshop on ‘Conceptual Quantum Chemistry’ organized by Post Graduate Department of Chemistry, Ponda Education Society’s Ravi S. Naik College of Arts & Science in collaboration with Indian National Science Academy, New Delhi, on 9th – 11th February 2018.

CHAPTER I

Introduction

	Table of Contents	Page
I.1	Selective and brief history behind the development of Schrödinger's Wave Equation	2
I.2	Time-Independent Schrödinger Equation	7
I.3	Solutions to the time-independent Schrödinger equation	7
I.3.1	The Hydrogen-like atom	7
I.3.2	Many-particle systems	9
	The Hartree Product and Hartree approximation	10
	Slater Determinants	12
	Hartree-Fock method	12
	Hartree-Fock Roothaan Equations	15
	Basis sets	15
	Hartree-Fock energy and Density Matrix	17
I.3.3	Møller-Plesset Perturbation Theory	18
I.3.4	Density Functional Theory	19
I.4	Importance of Theoretical Chemistry	22
I.5	Molecular Electrostatic Potential	23
I.5.1	Theoretical evaluation and visualization	23
I.5.2	Tool for predicting susceptibility towards electrophiles	24
I.5.3	Topographical features of MESP	26
I.5.4	Predictor tool for nucleophilic attacks	27
I.5.5	MESP based Charge Models	28
I.6	Scope of the present work	30
	References	33

I.1 Selective and brief history behind the development of Schrödinger's Wave Equation

The nature of light: whether 'light is a particle or a wave', puzzled scientists for quite some time in human history. Sir Isacc Newton (1643–1727) from his experiments concluded that light was made up of tiny particles. Based on this concept, he stated that the 'bending of light' phenomenon occurring when light travels from air to glass is the consequence of attraction between light particles and glass molecules [1]. This led to the derivation of the so-called Snell's Law, familiar to everyone since high school days. With respect to the extent of bending observed for different colours Newton stated that violet light particles possessed more mass – leading to greater attraction/more bending compared to red light. About a 100 years later, Thomas Young (1773–1829) in 1802 proved that light was made up of waves [2].

James Clerk Maxwell in the 1860s while working with electric (\mathbf{E}) and magnetic (\mathbf{B}) fields, concluded that \mathbf{E} and \mathbf{B} travel in the form of waves [3] (The mathematical equations for \mathbf{E} and \mathbf{B} derived by Maxwell out of which only the former is presented in Eq. (I.1), corresponds to a wave equation. This can be well understood by confining \mathbf{E} to a single dimension – say along the x direction. Eq. (I.1) therefore simplifies to Eq. (I.2), which is analogous to an equation (Eq. (I.3)) of a wave $\psi(x, t)$ travelling in the same direction with speed v). In fact both \mathbf{E} and \mathbf{B} are perpendicular to the propagation direction of the wave.

$$\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (\text{I.1})$$

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (\text{I.2})$$

$$\frac{\partial^2 \psi(x, t)}{\partial x^2} = \frac{1}{v^2} \frac{\partial^2 \psi(x, t)}{\partial t^2} \quad (\text{I.3})$$

The symbol ∇^2 is called the Laplacian operator, defined as $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ while μ_0 (permeability of free space) and ϵ_0 (permittivity of free space) are constants with values $4\pi \times 10^{-7} \text{ Js}^2\text{C}^{-2}\text{m}^{-1}$ and $8.854 \times 10^{-12}\text{C}^2\text{N}^{-1}\text{m}^{-2}$ respectively. These waves were found to travel at $v = 2.998 \times 10^8 \text{ ms}^{-1}$ (value obtained by equating $1/v^2$ to $\mu_0\epsilon_0$) which is the speed of light (c), leading to the conclusion that light is an electromagnetic wave. Thus all electromagnetic waves travelling at speed c are associated with a wavelength λ and frequency ν such that: $\lambda\nu = c$.

In the 1890s, physicists were quite interested in understanding the spectrum of black body radiation (plot of intensity I versus wavelength, illustrated in Figure I.1).

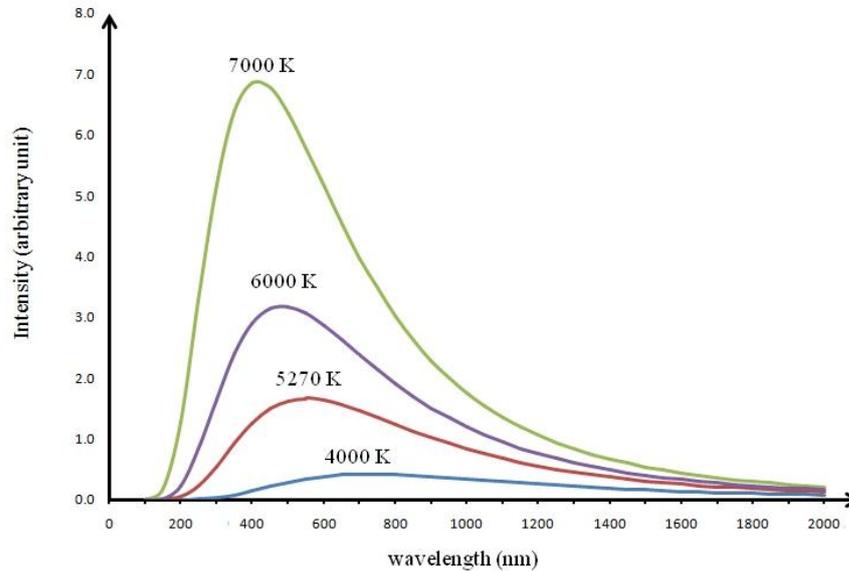


Figure I.1. Spectrum of blackbody radiation observed at various temperatures.

The maximum intensity of light emitted by the blackbody (example: heated metal) was found to increase at higher temperatures, accompanied by decrease in wavelength. The mathematical models proposed by physicists viz. Wein [4]:

$I = \frac{2hc^2}{\lambda^5} \exp(-hc/\lambda k_B T)$ and Baron J. W. S. Rayleigh [5]: $I = \frac{2c}{\lambda^4} k_B T$ (where h is some constant defined in the upcoming section, k_B and T being the Boltzman's constant and temperature in Kelvin) could replicate the spectrum at shorter/longer wavelengths but could not account for the decrease in intensity at very short wavelengths.

German physicist Max Planck (1858–1947) in 1900 accounted for the observed shape of the blackbody radiation spectrum by proposing that the energy of electromagnetic waves is '*quantized*' (though this concept seemed to be quite bizarre at that time) i.e. energy of emitted radiation due to oscillations of electrons in the blackbody must be proportional to integral multiples of frequency via. relation: $E = nh\nu$ where n is a integer number and h is the Planck's constant (6.634×10^{-34} Js). The spectral curves (in Figure I.1) showing maximum at different temperatures were found to obey the mathematical relation: $I = \left(\frac{2hc^2}{\lambda^5}\right) \left(\frac{1}{\exp\left(\frac{hc}{\lambda k_B T}\right) - 1}\right)$ from which Planck obtained the value of h [6-7].

About 5 years later, Albert Einstein (1879–1955) explained the photoelectric effect (ejection of electrons from metal surface on exposure to light) using the idea of quantization of energy [8]. He stated that light composed of particle-like entities called photons, whose energy was related to the frequency via. relation: $E = h\nu$ (and not multiple integrals of the same as proposed by Planck). Such energy when absorbed by an electron in metal, part of the energy would be utilized to overcome the forces responsible for its binding to the metal while the rest would appear as the kinetic energy of ejected electron. Thus the increase in frequency of light would increase the kinetic energy of emitted electrons. Einstein also stated that photons (behaving like waves and particles) possessed no mass but momentum (p) which is related to wavelength via.

equation: $h/\lambda = p$. This was confirmed from the experiment performed by Arthur H. Compton (1892–1962) around 1921 wherein the incident photon when directed towards a free electron led to increase in momentum of the electron [9] (such outcome was possible if light itself had momentum).

Louis de Broglie (1892–1987) in 1923 suggested that electrons like light possessed momentum and wavelength [10]. The Two American physicists, Clinton Davisson (1881–1958) and Lester Germer (1896–1971) experimentally confirmed de Broglie's hypothesis by reflecting electrons from a metal like nickel and observing diffraction effects [11]. German physicist Claus Jönsoon in 1961 performed the double-slit experiment with electrons and observed results similar to light which also confirmed the wave behaviour of electrons [12].

Danish physicist Neils Bohr (1885–1962) in 1913 used the concept of quantization of energy to explain the spectrum of hydrogen atom (which consisted of violet, blue, green and red lines) [13]. The observed colours were the result of emission of light by the electron as it moved from a higher energy level to a lower one, such that the difference of energy between the two levels (ΔE) corresponded to the frequency of emitted radiation i.e. $\Delta E = hv$. This led to the conclusion that the energy levels in the atom must therefore be quantized. Being intrigued by de Broglie's ideas, Austrian physicist Erwin Schrödinger (1887–1961) decided to find the equation [14] for electrons which is briefly described in the upcoming section.

Beginning with the equation of electric field (Eq. (I.4) which is Eq. (I.2) where $\mu_0 \epsilon_0$ is replaced with $1/c^2$), the solution found to satisfy the same is of the form:

$e^{\frac{2\pi i}{\lambda}(x - ct)}$ where i is $\sqrt{-1}$. One can easily verify this by replacing E with this function and checking the math.

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} = \left(\frac{1}{c^2}\right) \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (1.4)$$

Schrödinger stated that the solution chosen must not only satisfy the wave equation i.e. Eq. (I.4) but also the particle equation: $E = pc$. The above mentioned solution is found to satisfy both equations as demonstrated in APPENDIX, section A1.

Relating to electrons, the solution $\Psi(x, t)$ to the corresponding wave equation (which at this point is unknown) is suppose have the form: $e^{\frac{2\pi i}{\lambda}(x - ct)}$. The expression for E is known: it is the sum of kinetic energy (KE) and potential energy (V) of the system under consideration. In terms of mass (m) of the particle (in general) travelling at speed v , the KE (given by $\frac{1}{2}mv^2$) is written in terms of momentum to generate Eq. (I.5), which on multiplying with $\Psi(x, t)$ on both sides results in Eq. (I.6). Its mathematical transformation to the final time-dependent Schrödinger equation (Eq. (I.7)) for one-particle, one-dimensional system is presented in APPENDIX, section A2.

$$E = \frac{p^2}{2m} + V \quad (I.5)$$

$$E \Psi(x, t) = \frac{p^2}{2m} \Psi(x, t) + V \Psi(x, t) \quad (I.6)$$

$$\left(-\frac{\hbar}{i}\right) \frac{\partial \Psi(x, t)}{\partial t} = \left(-\frac{\hbar^2}{2m}\right) \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V \Psi(x, t) \quad (I.7)$$

Taking the wave equation and the corresponding solution of light as a guide, it was possible to obtain the equation for the electron. The task now remains to solve the above equation keeping in mind certain aspects: the wave function (WF) Ψ along with its first and second derivatives are expected to be finite, the quantity $|\Psi|^2$ known as the

probability density [15] also needs to be single valued and approach to zero at larger distances away from the atom.

I.2 Time-Independent Schrödinger Equation

The conversion of (one-dimensional) time-dependent Schrödinger equation (Eq. (I.7)) to a more ‘easy-to-handle’ form involves first by restricting V as a function of variable x , followed by writing $\Psi(x, t)$ as a product of two functions: $\psi(x)$ and $f(t)$ wherein the former and latter are dependent on x and t respectively. After performing a series of mathematical steps (described in APPENDIX, section A3), one can write the time-independent Schrödinger Equation as Eq. (I.8).

$$\left[\left(-\frac{\hbar^2}{2m} \right) \frac{d^2}{dx^2} + V(x) \right] \psi(x) = E \psi(x) \quad (\text{I.8})$$

The entity in square bracket in Eq. (I.8) is the Hamiltonian operator designated as \hat{H} . With this substitution the time-independent Schrödinger equation becomes an eigenvalue problem (Eq. (I.9)). Such equation is called as eigenvalue equation where $\psi(x)$ is the eigenfunction of \hat{H} with eigenvalue E .

$$\hat{H} \psi(x) = E \psi(x) \quad (\text{I.9})$$

I.3 Solutions to the time-independent Schrödinger Equation

I.3.1 Hydrogen-like atom

This section briefly addresses the solution to the Schrödinger equation for the hydrogen-like atom (Eq. (I.10)); r is the distance separating the nucleus and electron; m_e , e and Z being the mass of electron, charge of electron and nuclear charge of hydrogen-like atom which not only accounts for the observed spectral lines but further

supports the concept of ‘quantization of energy levels’ proposed by earlier quantum chemists.

$$\left(-\frac{\hbar^2}{2m_e} \nabla^2 - \frac{Ze^2}{4\pi\epsilon_0 r} \right) \psi = E \psi \quad (\text{I.10})$$

The idea to write $\psi(x, y, z)$ as a product of 3 independent functions: $P(x)$, $Q(y)$ and $S(z)$ is impossible due to the second term of Eq. (I.10). Changing from Cartesian to spherical polar coordinates $\{r, \theta, \phi\}$ transforms the above equation to Eq. (I.11) [wherein the term in square bracket represents ∇^2 - written in terms of spherical polar coordinates], making it separable into two functions: one depending on r $\{R(r)\}$ and the other on (θ, ϕ) $\{Y(\theta, \phi)\}$ viz. Eq. (I.12).

$$\left(-\frac{\hbar^2}{2m_e} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \left(\frac{\partial^2}{\partial \phi^2} \right) \right] - \frac{Ze^2}{4\pi\epsilon_0 r} \right) \psi(r, \theta, \phi) = E \psi(r, \theta, \phi) \quad (\text{I.11})$$

$$\left(-\frac{\hbar^2}{2m_e R(r)} \left(\frac{\partial}{\partial r} \left(r^2 \frac{\partial R(r)}{\partial r} \right) \right) - \frac{Ze^2 r}{4\pi\epsilon_0} - E r^2 \right) + \left(-\frac{\hbar^2}{2m_e Y(\theta, \phi)} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial Y(\theta, \phi)}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \left(\frac{\partial^2 Y(\theta, \phi)}{\partial \phi^2} \right) \right] \right) = 0 \quad (\text{I.12})$$

The exact solution to the separable equations with respect to $R(r)$ and $Y(\theta, \phi)$ are the set of functions known as the associated Laguerre polynomials for the radial equation (first term of Eq. (I.12)) and the spherical harmonics for the angular equation (second term of Eq. (I.12)), described in several quantum chemistry text books [16-17].

The energy expression for the hydrogen atom: $E_n = (-13.6)/n^2$ in electron volt where ‘ n ’ is the principle quantum number, is able to explain the corresponding spectrum (which are a result of electronic transitions from $n = 6, 5, 4$ and 3 levels to $n = 2$, observed in the visible region).

It is worth mentioning that the solution to the radial equation is governed by two quantum numbers n and l (azimuthal quantum number) such that n takes integer values from $1, 2, \dots$ etc. while l takes values from $0, 1, 2, \dots$ upto $n-1$ only. The solution to the angular equation (second term of Eq. (I.12)) gives rise to quantum numbers l and m (magnetic quantum number) such that m takes values of $0, \pm 1, \dots$, upto $\pm l$.

I.3.2 Many-particle systems

The WF ψ for a system consisting several (N number of) nuclei and (M number of) electrons is represented as $\psi(\mathbf{r}; \mathbf{R})$, where \mathbf{r} and \mathbf{R} denote the Cartesian coordinates of electrons and nuclei respectively. Electrons are also associated with the spin co-ordinate – introduced and discussed at a later stage in this chapter. The time-independent Schrödinger equation to be solved (i.e. Eq. (I.13)) contains two unknowns: ψ and E while the expression of the Hamiltonian operator (entity in square bracket whose terms are explained in APPENDIX, section A4) is known.

$$\left[-\frac{\hbar^2}{2} \sum_A^N \frac{\nabla_A^2}{m_A} - \frac{\hbar^2}{2m_e} \sum_{i=1}^M \nabla_i^2 + \frac{e^2}{4\pi\epsilon_0} \sum_A^N \sum_{\substack{B \\ B \neq A}}^A \frac{Z_A Z_B}{R_{AB}} + \frac{e^2}{4\pi\epsilon_0} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^i \frac{1}{r_{ij}} - \frac{e^2}{4\pi\epsilon_0} \sum_A^N \sum_i^M \frac{Z_A}{r_{Ai}} \right] \psi(\mathbf{r}; \mathbf{R}) = E \psi(\mathbf{r}; \mathbf{R}) \quad (\text{I.13})$$

Since the movement of electrons is much faster than the nuclei, it is safe to assume that the latter are stationary, in which case the first term in Eq. (I.13) vanishes. This is known as the Born-Oppenheimer approximation [18]. The third term being a constant (as a result of invoking the Bohn-Oppenheimer approximation) does not affect the WF but only shifts the energy by that much amount, and therefore, can be incorporated at a later stage.

Various constants of different units are found to be present in the Hamiltonian operator: \hbar (like h) has the unit of ‘Joules seconds’; m_A and m_e are in kilograms; the electronic charge e is in ‘Coulombs’ etc. A convenient way to ensure uniformity involves working with *atomic units* where all these quantities are set to unity, giving the Bohr radius a_0 as 1 (0.529 Å) and energy as 1 Hartree (4.3597×10^{-17} J). These implementations leads one to solve the electronic Schrödinger equation (Eq. (I.14)) where the electronic Hamiltonian operator \hat{H}_e (square bracket) operating on $\psi(\mathbf{r}; \mathbf{R})$ (now written as $\psi_e(\mathbf{r})$ as the position of nuclei are fixed parameters and not variables) gives the corresponding electronic energy E_e and the WF back. [All expressions throughout this thesis are expressed in terms of atomic units.] Though the complexity of Eq. (I.13) seems to be reduced by Eq. (I.14), solving the latter is still complicated.

$$\left[-\frac{1}{2} \sum_{i=1}^M \nabla_i^2 - \sum_{A=1}^N \sum_{i=1}^M \frac{Z_A}{r_{Ai}} + \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^i \frac{1}{r_{ij}} \right] \psi_e(\mathbf{r}) = E_e \psi_e(\mathbf{r}) \quad (\text{I.14})$$

Hartree Product and Hartree approximation

Neglecting the inter-electronic repulsion term (which unfortunately is a very drastic approximation) results in \hat{H}_e being dependent on the individual kinetic energy

and nuclear attraction terms of each electron (represented in terms of one-electron operator \hat{h}_i viz. Eq. (I.15) where \hat{h}_i equals $-(\nabla_i^2/2) - \sum_A (Z_A/r_{Ai})$.

$$\left[\sum_i^M \hat{h}_i \right] \psi_e(\mathbf{r}) = E_e \psi_e(\mathbf{r}) \quad (\text{I.15})$$

The WF thus becomes a product of individual functions (or orbitals) $\phi_i(\mathbf{r}_i)$ i.e. Eq. (I.16) which is known as the Hartree Product [19] and Eq. (I.15) hence splits into M differential equations viz. Eq. (I.17) where ε_i denotes the energy associated with corresponding (normalized) orbital ϕ_i .

$$\psi_e(\mathbf{r}) = \phi_1(\mathbf{r}_1) \phi_2(\mathbf{r}_2) \dots \phi_M(\mathbf{r}_M) \quad (\text{I.16})$$

$$\hat{h}_i \phi_i = \varepsilon_i \phi_i \quad (\text{I.17})$$

Introducing a two-electron operator $v(i, j)$ to account the $1/r_{ij}$ terms and assuming that the repulsion felt by the i^{th} electron is due to the charge density of other electrons (i.e. Hartree. approximation: $v(i, j) = \int \frac{|\phi_j(\mathbf{r}_j)|^2}{r_{ij}} d\mathbf{r}_j$), Eq. (I.17) transforms to Eq. (I.18). The key to solve this equation depends on the knowledge of orbitals $\{\phi_i\}$ itself as $v(i, j)$ depends explicitly on $\{\phi_i\}$. One therefore needs to provide a starting guess for $\{\phi_i\}$. Solving the one-electron Schrödinger equations will result in a improved set of $\{\phi_i\}$ which can used to construct a ‘better’ $v(i, j)$ and this process is repeated till there is no substantial difference between $\{\phi_i\}$ obtained in two successive steps. This is known as the *Self-Consistent-Field* (SCF) method [20].

$$\left(\hat{h}_i + \sum_{\substack{j=1 \\ i \neq j}}^M \int \frac{|\phi_j(\mathbf{r}_j)|^2}{r_{ij}} d\mathbf{r}_j \right) \phi_i = \varepsilon_i \phi_i \quad (\text{I.18})$$

Slater Determinants

Recalling the fact that electrons not only have three spatial degrees of freedom but also the spin component, the orbitals $\phi_i(\mathbf{r}_i)$ are transformed to spatial-spin orbitals $\chi_i(\mathbf{x}_i)$ where \mathbf{x}_i equals $\{\mathbf{r}_i, w\}$, w being the spin coordinate. Because electrons are fermions, interchanging coordinates of any two electrons should result in the change in sign of the WF; known as the *antisymmetry principle* (Eq. (I.19)).

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M) = - \Psi(\mathbf{x}_2, \mathbf{x}_1, \dots, \mathbf{x}_M) \quad (\text{I.19})$$

The Hartree product however does not meet this requirement (as demonstrated in APPENDIX, section A5). John C. Slater in 1929 suggested writing the WF in terms of a determinant (Eq. (I.20)), known as the Slater determinant [21].

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M) = \frac{1}{\sqrt{M!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \dots & \chi_M(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \dots & \chi_M(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_M) & \chi_2(\mathbf{x}_M) & \dots & \chi_M(\mathbf{x}_M) \end{vmatrix} \quad (\text{I.20})$$

The term $\frac{1}{\sqrt{M!}}$ is called the normalization constant and the Slater determinant fulfils the antisymmetry principle (as demonstrated in APPENDIX, section A6). More importantly, the determinant is found to vanish if two electrons have same spin thereby satisfying *Pauli principle* which states that no two electrons can have same set of quantum numbers.

Hartree-Fock method

In the Hartree-Fock (HF) approximation [22-23], it is assumed that the WF ψ_e has the form of a single Slater determinant, and the corresponding energy (i.e. the Hartree-Fock energy, E_{HF}) is determined via. Eq. (I.21). Quantities in the first equality

are written in the (bra-ket) Dirac notation (second equality). The third equality holds true since the WF in the Slater determinant form is normalized.

$$E_{\text{HF}} = \frac{\int \psi_e^* \hat{H}_e \psi_e d\tau}{\int \psi_e^* \psi_e d\tau} = \frac{\langle \psi_e | \hat{H}_e | \psi_e \rangle}{\langle \psi_e | \psi_e \rangle} = \langle \psi_e | \hat{H}_e | \psi_e \rangle \quad (\text{I.21})$$

Substituting the Slater determinant in the above equation results to the Hartree-Fock energy expression (Eq. (I.22)) in terms of orbitals $\{\chi_i\}$. Since the orbitals are real, the complex conjugates $\{\chi_i^*\}$ are equivalent to $\{\chi_i\}$ and the asterisk symbols in the below expression can be ignored. [Summation for indices i, j take values from 1 to M .]

$$E_{\text{HF}} = \sum_i \int \chi_i^*(\mathbf{x}_1) \hat{h}_i \chi_i(\mathbf{x}_1) d\mathbf{x}_1 + \left(\frac{1}{2}\right) \sum_i \sum_j \left[\iint \chi_i^*(\mathbf{x}_1) \chi_j^*(\mathbf{x}_2) \frac{1}{r_{12}} \chi_i(\mathbf{x}_1) \chi_j(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \right. \\ \left. - \iint \chi_i^*(\mathbf{x}_1) \chi_j^*(\mathbf{x}_2) \frac{1}{r_{12}} \chi_i(\mathbf{x}_2) \chi_j(\mathbf{x}_1) d\mathbf{x}_1 d\mathbf{x}_2 \right] \quad (\text{I.22})$$

The first integral in above equation will be referred to 'h_i'. The first term in square bracket has some significance: its represents the Coulomb repulsion between two electrons, and is denoted as J_{ij} . [When rearranged, the product $\chi_i^*(\mathbf{x}_1) \chi_i(\mathbf{x}_1)$ corresponds to the probability of electron 1 in i^{th} orbital at \mathbf{x}_1 and likewise similar definition exists for $\chi_j^*(\mathbf{x}_2) \chi_j(\mathbf{x}_2)$. The distance separating electron 1 and 2 is r_{12} and the repulsion between them would have a $1/r_{12}$ term. The rearranged quantity $\chi_i^*(\mathbf{x}_1) \chi_i(\mathbf{x}_1) (1/r_{12}) \chi_j^*(\mathbf{x}_2) \chi_j(\mathbf{x}_2)$ is the quantum mechanical (QM) version of the Coulomb repulsion as the positions of electrons are unknown but their probabilities are known. Integrating over all space and summing all such probabilities yields J_{ij}].

The second term on the other hand has no good physical interpretation and is called the exchange term $\{\mathcal{K}_{ij}\}$ since electrons at positions \mathbf{x}_1 and \mathbf{x}_2 are found to be present in orbitals χ_j and χ_i .

Using a trial WF to approximate the actual WF in Eq. (I.14) results in the energy $\{E_{\text{HF}}\}$ obtained to be greater than the actual value $\{E\}$. One can therefore improve upon E_{HF} and thereby reduce the difference between E_{HF} and E by choosing better set of orbitals $\{\chi_i\}$ via. SCF procedure. Mathematically, this is achieved by setting up a function depending on $\{\chi_i\}$, $\mathcal{L}[\{\chi_i\}]$ defined by Eq. (I.23), subjected to the constraint (second term in Eq. (I.23)) that the orbitals remain orthogonal to each other viz. Eq. (I.24). This is known as Lagrange method of undermined multipliers where ϵ_{ij} are the unknown constants and δ_{ij} is known as cronicker delta. The job is then to minimize \mathcal{L} such that the change $\delta\mathcal{L}$ becomes zero, which leads to Eq. (I.25) where $\hat{F}(\mathbf{x}_1)$ is known as the Fock operator.

$$\mathcal{L}[\{\chi_i\}] = E_{\text{HF}}[\{\chi_i\}] - \sum_{i,j} \epsilon_{ij} (\langle ij \rangle - \delta_{ij}) \quad (\text{I.23})$$

$$\langle ij \rangle = \int \chi_i(\mathbf{x}_1) \chi_j(\mathbf{x}_1) d\mathbf{x}_1 = \delta_{ij} \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases} \quad (\text{I.24})$$

$$\hat{F}(\mathbf{x}_1) \chi_i(\mathbf{x}_1) = \epsilon_i \chi_i(\mathbf{x}_1) \quad (\text{I.25})$$

Though the above final eigenvalue equation appears to be simple, the Fock operator hides the overall complexity: i.e. the construction of \hat{F} depends on \mathcal{J} and \mathcal{K} which in turn depend on the orbitals $\{\chi_i\}$.

Hartree-Fock Roothaan Equations

Roothaan proposed the use of orbitals constructed from a linear combination of some predefined functions/basis functions $\{\tilde{\chi}_\nu\}$ i.e. $\chi_i = \sum_{\nu=1}^k c_{\nu i} \tilde{\chi}_\nu$ leading to Eq. (I.26) [24]. The orbital energies are then obtained by integrating Eq. (I.26) over all space – after multiplying the corresponding complex conjugate term $\tilde{\chi}_\mu^*(\mathbf{x}_1)$ via Eq. (I.27). Since the basis functions are not orthogonal, the right hand side integral is some number (known as overlap integral; represented in matrix notation as $S_{\mu\nu}$). The integral on the left hand side of Eq. (I.27) constitute the elements of the Fock matrix, $F_{\mu\nu}$ and the orbital energies correspond to the elements of the diagonal matrix ε .

$$\hat{F}(\mathbf{x}_1) \sum_{\nu=1}^k c_{\nu i} \tilde{\chi}_\nu(\mathbf{x}_1) = \varepsilon_i \sum_{\nu=1}^k c_{\nu i} \tilde{\chi}_\nu(\mathbf{x}_1) \quad (\text{I.26})$$

$$\sum_{\mu,\nu=1}^k \int \tilde{\chi}_\mu^*(\mathbf{x}_1) \hat{F}(\mathbf{x}_1) \tilde{\chi}_\nu(\mathbf{x}_1) d\mathbf{x}_1 c_{\nu i} = \varepsilon_i \sum_{\mu,\nu=1}^k \int \tilde{\chi}_\mu^*(\mathbf{x}_1) \tilde{\chi}_\nu(\mathbf{x}_1) d\mathbf{x}_1 c_{\nu i} \quad (\text{I.27})$$

$$F_{\mu\nu} c_{\nu i} = \varepsilon_i S_{\mu\nu} c_{\nu i} \quad (\text{I.28})$$

The problem of solving the electronic Schrödinger equation is reduced to a matrix problem (Eq. (I.28)) which can be easily tackled on computers. Known quantities are the elements of Fock and overlap matrices. Chemists prefer the orbitals $\{\chi_i\}$ to correspond to the molecular orbitals of a system under consideration, $c_{\nu i}$ to the molecular orbital (MO) coefficients and $\{\tilde{\chi}_\mu\}$ to represent respective atomic orbitals.

Basis sets

Besides proposing a single Slater determinant to be the solution of Schrödinger's equation for many-particle system, Slater suggested the use of orbitals

(known as Slater orbitals or Slater-type orbitals (STOs)) defined by Eq. (I.29) where the radial function is represented as $(r^{n-1} e^{-\zeta r})$, ζ being the orbital exponent and N_s , the corresponding normalization constant. STOs have the cusp behaviour (non-zero slope) at the nuclei on which they are centered [25]. However the evaluation of multi-center integrals in polyatomic molecules becomes complicated while employing them (as no analytical means exists to solve such two-electron integrals using STOs).

$$\chi_{\text{STO}}(r, \theta, \phi) = N_s Y(\theta, \phi) r^{n-1} e^{-\zeta r} \quad (\text{I.29})$$

The use of Gaussian-type orbitals (GTOs) i.e. Eq. (I.30) where N_g being the normalization constant, non-negative integers $\{a, b, c\}$ describing the shape of orbital along with Gaussian exponent α ; overcomes this problem (since the product of two Gaussians results to a Gaussian and there exist exact analytical expressions for evaluating Gaussian-based integrals) despite their shortcomings: faster decay at larger distances compared to STOs and zero slope at the nucleus on which they are positioned [26 and references within].

$$\chi_{\text{GTO}}(x, y, z) = N_g x^a y^b z^c e^{-\alpha r^2} \quad (\text{I.30})$$

To overcome the cusp weakness, it is a common practice to combine several (n number of) GTOs with a set of (fixed) coefficients into new functions called contracted GTOs (CGTOs)/‘STO- n G’ (which unfortunately also fails to mimic the desired STO).

The STO-3G is known as the minimal basis wherein each orbital of an atom is described by a single CGTO, constructed from 3 GTOs. For example, each atomic orbital of carbon atom would be represented by a single CGTO to yield overall 5 basis functions.

It is more logical and appropriate to differentiate between the inner and valence orbitals in which case the ‘double zeta’ basis set comes into picture wherein each valence orbitals are represented by two Slater orbitals (the valence χ_{2s} orbital in carbon atom is given in Eq. (I.31), where $\chi_{2s}^{\text{STO}}(\mathbf{r}, \zeta_1)$ and $\chi_{2s}^{\text{STO}}(\mathbf{r}, \zeta_2)$ are two Slater orbitals with exponents ζ_1 and ζ_2 ; ‘d’ being the corresponding contraction coefficient.

$$\chi_{2s}(\mathbf{r}) = \chi_{2s}^{\text{STO}}(\mathbf{r}, \zeta_1) + d \chi_{2s}^{\text{STO}}(\mathbf{r}, \zeta_2) \quad (\text{I.31})$$

The simplest double zeta basis set is the 3-21G basis set (also known as split valence) where the inner orbital χ_{1s} (in carbon atom) is represented by a single CGTO constituting 3 GTOs, while the χ_{2s} orbital (an each of the 2p) as per Eq. (I.31) - the first STO $\chi_{2s}^{\text{STO}}(\mathbf{r}, \zeta_1)$ made of 2 GTOs and the second STO $\chi_{2s}^{\text{STO}}(\mathbf{r}, \zeta_2)$ by a single GTO. Describing subsequent classes of basis sets beyond the 3-21G level in itself is an enormous task and therefore the present discussion is concluded at this point.

Hartree-Fock energy and Density Matrix

Based on Eq. (I.25), the orbital energies ε_i are evaluated as per Eq. (I.32). Sum of all orbital energies would yield Eq. (I.33) which is not the true HF energy. The correct equation for E_{HF} with respect to orbital energies would in fact be Eq. (I.34).

$$\varepsilon_i = \int \chi_i(\mathbf{x}_1) \hat{F}(\mathbf{x}_1) \chi_i(\mathbf{x}_1) d\mathbf{x}_1 = h_i + \sum_{j=1}^M (J_{ij} - \mathcal{K}_{ij}) \quad (\text{I.32})$$

$$\sum_{i=1}^M \varepsilon_i = \sum_{i=1}^M h_i + \sum_{i=1}^M \sum_{j=1}^M (J_{ij} - \mathcal{K}_{ij}) \quad (\text{I.33})$$

$$E_{\text{HF}} = \sum_{i=1}^M \varepsilon_i - \left(\frac{1}{2}\right) \sum_{i=1}^M \sum_{j=1}^M (J_{ij} - \mathcal{K}_{ij}) \quad (\text{I.34})$$

The version of HF theory discussed is called the unrestricted HF theory, appropriate for all molecules regardless of the number of electrons and distribution of electron spin. In the restricted Hartree-Fock (RHF) theory, each MO is ‘occupied’ by two electrons and the energy E_{RHF} is given by Eq. (I.35) respectively.

$$E_{\text{RHF}} = 2 \sum_{i=1}^M \varepsilon_i - \sum_{i=1}^M \sum_{j=1}^M (2 J_{ij} - \mathcal{K}_{ij}) \quad (\text{I.35})$$

One important matrix worth discussing is the Density Matrix P. The density, $\rho(\mathbf{r})$, as a function of position in RHF theory is given by Eq. (I.36) and its integral over all space yields M (i.e. number of electrons in given system). In terms of basis functions, Eq. (I.36) transforms to Eq. (I.37) and the elements of P, $P_{\mu\nu} = 2 \sum_{i=1}^{M/2} c_{\mu i}^* c_{\nu i}$. The procedure employed for solving the HF Roothaan equation (Eq. (I.28)) wherein the Fock matrix is re-written in terms of density matrix P is not discussed in this thesis and is available elsewhere [27].

$$\rho(\mathbf{r}) = 2 \sum_{i=1}^{M/2} \chi_i^*(\mathbf{r}) \chi_i(\mathbf{r}) = 2 \sum_{i=1}^{M/2} |\chi_i(\mathbf{r})|^2 \quad (\text{I.36})$$

$$\rho(\mathbf{r}) = \sum_{\mu, \nu=1}^k \left(2 \sum_{i=1}^{M/2} c_{\mu i}^* c_{\nu i} \right) \tilde{\chi}_{\mu}^*(\mathbf{r}) \tilde{\chi}_{\nu}(\mathbf{r}) \quad (\text{I.37})$$

I.3.3 Møller-Plesset Perturbation Theory

The Møller-Plesset Perturbation theory developed by Møller and Plesset in 1934 [28] aimed in correcting the HF energy with respect to the electron correlation. Assuming that $E^{(0)}$ represents the sum of orbital energies obtained from zeroth-order WF (which is the corresponding Hartree-Fock WF) $\psi^{(0)}$ via. $\hat{H}^{(0)}\psi^{(0)} = E^{(0)}\psi^{(0)}$, the

perturbation \hat{V} is given in Eq. (I.38), which is the difference between the true electron-electron repulsion term and its counterpart from HF method.

$$\hat{V} = \hat{H} - \hat{H}^{(0)} = \sum_{i \neq j}^M \frac{1}{r_{ij}} - v(i, j) = \sum_{i \neq j}^M \frac{1}{r_{ij}} - \sum_{i=1}^M \sum_{j=1}^M (J_{ij} - \mathcal{K}_{ij}) \quad (\text{I.38})$$

Applying this perturbation operator, the first-order energy correction $E^{(1)}$ to $E^{(0)}$ results to Eq. (I.39), which when added to $E^{(0)}$ is in fact the E_{HF} itself (Eq. (I.34)). Without further derivation, the second-order energy correction term, $E^{(2)}$ is given in Eq. (I.40) where ‘m’ is the number of occupied orbitals, ‘n’ is the number of orbitals (occupied and virtual) while the term $\langle ij||uv \rangle$ is defined in Eq. (I.41).

$$E^{(1)} = \langle \psi^{(0)} | \hat{V} | \psi^{(0)} \rangle = - \left(\frac{1}{2} \right) \sum_{i=1}^M \sum_{j=1}^M (J_{ij} - \mathcal{K}_{ij}) \quad (\text{I.39})$$

$$E^{(2)} = \frac{1}{4} \sum_{i=1}^m \sum_{j=1}^m \sum_{u=m+1}^n \sum_{v=m+1}^n \frac{|\langle ij||uv \rangle|^2}{(\varepsilon_i + \varepsilon_j - \varepsilon_u - \varepsilon_v)} \quad (\text{I.40})$$

$$\langle ij||uv \rangle = \iint \chi_i^*(\mathbf{r}_1) \chi_j^*(\mathbf{r}_2) \frac{1}{r_{12}} \chi_u(\mathbf{r}_1) \chi_v(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 - \iint \chi_i^*(\mathbf{r}_1) \chi_j^*(\mathbf{r}_2) \frac{1}{r_{12}} \chi_v(\mathbf{r}_1) \chi_u(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (\text{I.41})$$

I.3.4 Density Functional Theory

Density Functional Theory (DFT) is one of the alternatives available to solve the electronic Schrödinger equation. Compared to HF theory, DFT is much cheaper if good approximations are used for the exchange two-electron integral term. Basically DFT works with $\rho(\mathbf{r})$ and not with the WF. The former is dependent on the spatial coordinates of the electron while the latter, has an additional spin component. It was difficult to implement this methodology until Hohenberg and Kohn [29] proved it.

A functional is a function of a function, thus the density functional (mathematically represented as $E[\rho(\mathbf{r})]$) is a function of $\rho(\mathbf{r})$ which yields E . The energy functional $E[\rho]$ consists of three terms: kinetic energy density functional $T[\rho]$, electron-nuclear attraction functional $E_{eN}[\rho]$ and electron-electron repulsion density functional $E_{ee}[\rho]$. Out of these components, $E_{eN}[\rho]$ can be written viz. Eq. (I.42). One can also easily guess what will be the form of $E_{ee}[\rho]$ - repulsion between the densities of electron 1 ($\rho(\mathbf{r})$) and 2 ($\rho(\mathbf{r}')$) i.e. Eq. (I.43), which in a way represents the Coulomb term in HF theory and therefore the symbol ‘ \mathcal{J} ’ has been used.

$$E_{eN}[\rho] = - \sum_A^{\text{Nuclei}} \int \frac{Z_A \rho(\mathbf{r})}{|\mathbf{R}_A - \mathbf{r}|} d\mathbf{r} \quad (\text{I.42})$$

$$\mathcal{J}[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \quad (\text{I.43})$$

The \mathcal{K}_{ij} integrals in HF theory ensured that the antisymmetry principle was obeyed. To enforce the same in current methodology, the suggestions by Kohn and Sham [30] are implemented, and therefore it is often known as Kohn Sham Density Functional Theory (KS-DFT). They suggested the use $\{\hat{T}_e\}$ expression from the HF theory to calculate $T[\rho]$ (i.e. Eq. (I.44)) using orbitals $\{\phi_i\}$. The density $\rho(\mathbf{r})$ and orbitals $\{\phi_i\}$ are related viz. $\rho(\mathbf{r}) = \sum_{i=1}^M |\phi_i(\mathbf{r})|^2$.

$$T[\rho] = \sum_{i=1}^M \left\langle \phi_i \left| -\frac{1}{2} \nabla_i^2 \right| \phi_i \right\rangle \quad (\text{I.44})$$

The energy expression ($E[\rho]$ represented as $E_{\text{DFT}}[\rho]$) is thus given in Eq. (I.45) where the last term $E_{\text{XC}}[\rho]$ is known as the exchange-correlation functional which takes into account the fact that electrons are antisymmetric and the electron correction term (describing the inadequacies of the HF theory with respect to the

electron repulsion term). Though the exact equation for $E_{XC}[\rho]$ is unknown, there is a need to ‘cook-up’ some approximations for this term.

$$E_{\text{DFT}}[\rho] = T[\rho] + E_{\text{eN}}[\rho] + J[\rho] + E_{\text{XC}}[\rho] \quad (\text{I.45})$$

With the use of orbitals $\{\phi_i\}$ in Eq. (I.44), Eq. (I.45) can be represented in terms of one-electron Kohn Sham operator $\{\hat{h}_{\text{KS}}\}$ operating on $\{\phi_i\}$ to yield orbital energies $\{\epsilon_i\}$ and back the corresponding orbitals (i.e. Eq. (I.46)).

$$\hat{h}_{\text{KS}} \phi_i = \epsilon_i \phi_i \quad (\text{I.46})$$

The operator \hat{h}_{KS} (defined in Eq. (I.47)) contains the kinetic energy component (first term in curly bracket) and the potential energy term, the latter consisting of the attraction towards all nuclei (second term) and the coulomb repulsion term (third). The fourth term (known as the exchange correlation potential) in Eq. (I.47) is related to the exchange functional, it is the derivative of $E_{\text{XC}}(\rho)$ with respect to the density.

$$\left\{ -\frac{1}{2} \nabla_i^2 - \sum_{\text{nuclei}} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} + \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + V_{\text{XC}}(\mathbf{r}) \right\} \phi_i = \epsilon_i \phi_i \quad (\text{I.47})$$

By introducing basis functions $\{\phi_i = \sum_{\mu=1}^k c_{\mu i} \tilde{\phi}_i\}$, the KS-DFT equations become a nonlinear eigenvalue problem (Eq. (I.48)) which is similar to its HF counterpart.

$$\mathbf{F}^{\text{KS}} \mathbf{C} = \epsilon \mathbf{S} \mathbf{C} \quad (\text{I.48})$$

The KS orbitals $\{\phi_i\}$ allow the computation of electron density as pointed in earlier discussion. There is no ‘molecular WF’ defined within this framework. The

descriptions of various approximations made for the exchange correlation functional shall be refrained in this thesis.

I.4 Importance of Theoretical Chemistry

Theoretical Chemistry is a branch of Chemistry that attempts to figure out the mathematical equations behind the chemical phenomena under study. If one desires to study motion of electrons, then the equation (i.e. Schrödinger equation) is well-known. But solving this equation is quite hard. Handling this task (figuring out good approximations to solve Schrödinger equation) is the role of Electronic Structure Theory.

Knowledge about the motion of electrons is necessary to compute several properties, like electronic spectra (wherein the electrons get hit by photons and jump to some excited states; this cannot be modelled without prior information about what the electrons are doing in the atom/molecule), polarizability – how electrons response to applied electric fields, NMR chemical shifts – how electrons respond to applied magnetic fields etc.

At a less expensive cost, it is possible to predict geometries of molecules and corresponding vibrational frequencies if the atoms and bonds connecting them are imagined to be spheres attached by strings, in which case the electronic motion can be ignored. However such methods fail in studies involving transfer of atoms from one molecule to another since it is important to know whether such movement involves electron transfer or not. With respect to reactivity of a molecule, the Molecular Electrostatic Potential (MESP) is a versatile tool in detecting sites susceptible towards electrophilic and nucleophilic attacks, illustrated in the upcoming section with examples.

I.5 Molecular Electrostatic Potential (MESP)

I.5.1 Theoretical evaluation and visualization

For a molecule containing ‘N’ number of atoms with nuclear charges $\{Z_i\}$ situated at $\{\mathbf{R}_i\}$, MESP at point \mathbf{r} in atomic units is defined by Eq. (I.49) where $\rho(\mathbf{r}')$ corresponds to the molecular electron density function discussed earlier. In terms of density matrix elements $P_{\mu\nu}$ and (N_b number of) basis functions $\{\tilde{\chi}_\mu(\mathbf{r}'), \tilde{\chi}_\nu(\mathbf{r}')\}$, Eq. (I.49) is re-written as Eq. (I.50) and all that is required to compute $V(\mathbf{r})$ involves evaluating the integral $\langle \tilde{\chi}_\mu(\mathbf{r}') | \frac{1}{|\mathbf{r} - \mathbf{r}'|} | \tilde{\chi}_\nu(\mathbf{r}') \rangle$.

$$V(\mathbf{r}) = \sum_{i=1}^N \frac{Z_i}{|\mathbf{r} - \mathbf{R}_i|} - \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (\text{I.49})$$

$$V(\mathbf{r}) = \sum_{i=1}^N \frac{Z_i}{|\mathbf{r} - \mathbf{R}_i|} - \sum_{\mu=1}^{N_b} \sum_{\nu=1}^{N_b} P_{\mu\nu} \langle \tilde{\chi}_\mu(\mathbf{r}') | \frac{1}{|\mathbf{r} - \mathbf{r}'|} | \tilde{\chi}_\nu(\mathbf{r}') \rangle \quad (\text{I.50})$$

The interacting basis functions $\tilde{\chi}_\mu(\mathbf{r}')$ and $\tilde{\chi}_\nu(\mathbf{r}')$ are of several types. The s -type basis function is constructed from several s -type ($a = b = c = 0$ in Eq. (I.30)) GTOs. There are three possibilities for a p -type basis function: p_x , p_y and p_z depending upon which two values of $\{a, b, c\}$ of GTOs in corresponding contraction are zero. The integral in Eq. (I.50) therefore needs to be evaluated over all possible combination of GTOs present in $\tilde{\chi}_\mu(\mathbf{r}')$ and $\tilde{\chi}_\nu(\mathbf{r}')$. The works of Gadre [27] and Ramos [31] highlight the formulae involved in the computation of $V(\mathbf{r})$ due to two (different type of) interacting Gaussians.

The MESP, being a three-dimensional (3D) quantity, can be visualized with the aid of a computer. This can be achieved by viewing the scalar field in planes with

the help of contour lines or pixel plots (the latter wherein points are given colours as per MESP values on the grid). Such planes along with the ball-and-stick molecular model give a ‘better’ view of the scalar field. The 3D views of MESP may be obtained using isovalued surfaces or by plotting the scalar field on some predefined surface on which $\rho(\mathbf{r})$ is constant using suitably chosen colour codes (more details in section I.5.4). Several computer programs are available for computation and/or visualization of MESP such as SPARTAN [32], HyperChem [33], UNIVIS-2000 [34] etc.

I.5.2 Tool for predicting susceptibility towards electrophiles

Depending upon which term in Eq. (I.49) is dominant at \mathbf{r} , $V(\mathbf{r})$ takes positive, zero or negative values. A positive unit test charge would therefore be repelled in those regions where $V(\mathbf{r})$ is positive and be attracted towards negative MESP regions. Thus MESP is a useful tool in identifying reactive sites in molecules.

Figure I.2a portrays the MESP contour plot of acetone molecule in the molecular (YZ) plane at HF/6-31G(d,p) level of theory. The $V(\mathbf{r})$ in acetone comprises of both: positive and negative regions wherein the latter is in the proximity of oxygen atom, containing two equivalent minima (V_{\min}) in the molecular plane. Both V_{\min} are situated at a distance of approximately 1.21 Å away from oxygen atom, forming an angle of 129.2° with the carbonyl group. Any incoming electron deficient species would be attracted towards these sites. This is however possible provided the electronic distribution around acetone remains static with respect to the incoming species (which in reality always gets perturbed).

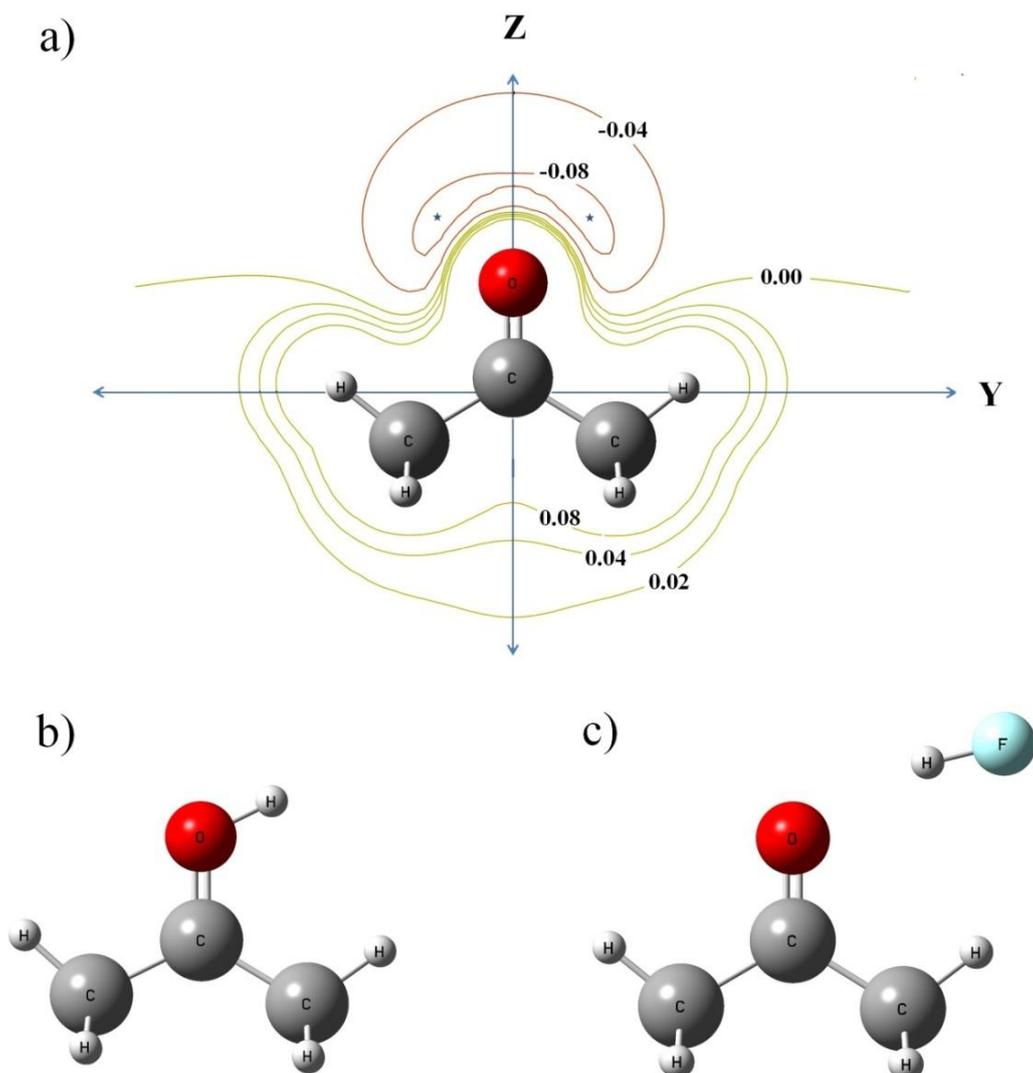


Figure I.2. a) MESP contour plot of acetone molecule at HF/6-31G(d,p) level of theory in the molecular (ZY) plane. The location of two equivalent V_{\min} are indicated by ‘★’ symbol and $V(\mathbf{r})$ value at these sites is about -0.0922 a.u.; optimized structures of b) acetone-proton and c) acetone-hydrogen fluoride complexes at HF/6-31G(d,p) level. Optimized geometry co-ordinates of all structures are provided in APPENDIX, section A7.

To get an idea about perturbation effect, the interaction between acetone and proton is taken for the purpose of illustration. The optimized structure of acetone-proton complex is displayed in Figure I.2b, wherein the proton is situated at a distance of approximately 0.95 Å and inclined by angle of 116.3° with respect to the carbonyl group. Though the proton lies in the molecular plane as predicted by V_{\min} , the O-H bond

length is found to be shorter by 0.26 Å while the H-O-C angle shows 12.9° deviation from expected value (of 129.2°).

Besides a proton, acetone is capable of interacting with hydrogen donor species like hydrogen fluoride (*HF*). The charge distribution in *HF* molecule can crudely be approximated by a positive charge at the hydrogen site and a negative charge on fluorine. It is therefore quite obvious that the hydrogen end of *HF* would be attracted to the negative MESP region in acetone. The HF/6-31G(d,p) geometry of acetone-hydrogen fluoride complex is shown in Figure I.2c. As expected, *HF* is found to lie in the molecular plane, inclined by angle of 123.9° (deviation from expected V_{\min} -O-C angle is about 5°). From the above two cases it is seen that the electrostatic potential does a better job in predicting the directionality of the weaker interaction with hydrogen fluoride than the strong interaction with proton [35]. The presence of two V_{\min} around oxygen in acetone is in agreement with the concept that a sp^2 hybridized oxygen possessing two lone pairs (LPs). Similar conclusions are also drawn from the isopotential maps of ammonia and water molecules available in the works of Scrocco and Tomasi [36] and presented/discussed at a later stage in this thesis.

I.5.3 Topographical features of MESP

The word *topography* is derived from the Greek words *topos* and *graphia* meaning ‘place’ and ‘mapping’ which means mapping of a geographical place. Such a map describes physical features of a region viz. places of equal heights in the form of contours and noting other features like roads, streams etc.

The most negative sites (i.e. positions of V_{\min}) in MESP scalar field, as seen in earlier example, indicates the presence of LPs (and/or π bonds in nuclear frameworks

containing double bonds [37, 38]). Mathematically speaking, the first order partial derivatives of Eq. (I.49): $\frac{\partial V}{\partial x}$, $\frac{\partial V}{\partial y}$ and $\frac{\partial V}{\partial z}$ vanish at these sites. Such points are called critical points (CPs). The CPs are then further classified based on the eigenvalues obtained from the Hessian matrix (matrix of second order partial derivatives) evaluated at the CPs. A generate CP is the one where at least one of the eigenvalues is found to be zero while a nondegenerate CP (whose eigenvalues are non-zero) is characterized by (R, σ) where the former quantity, R corresponds to the rank (number of non-zero eigenvalues) and the latter ‘ σ ’, is the signature – algebraic sum of the signs of eigenvalues.

The rank of a nondegenerate CP of MESP scalar field is 3 and based on the signs of eigenvalues, four type of such CPs are possible: $(3, +3)$, $(3, -3)$, $(3, -1)$ and $(3, +1)$. The former two correspond to a minimum and a maximum while the latter two represent saddle points. Based on this classification, V_{\min} corresponds to a $(3, +3)$ CP. The purpose of this section was to briefly introduce the concept of topography though a detailed description of the same – i.e. significance of MESP CPs is discussed at a later stage in this thesis.

I.5.4 Predictor tool for nucleophilic attacks

The use of MESP to predict sites reactive towards nucleophiles however is not as straight forward as it is for electrophiles. This is due to the fact that the positive $V(\mathbf{r})$ region do not have maxima other than at the position of nuclei, which has been proven by Gadre and Pathak [39]. Sjoberg and Politzer [40] demonstrated that such a prediction was possible by plotting $V(\mathbf{r})$ on a molecular surface having $\rho(\mathbf{r})$ value equal to 0.001 a.u. This strategy has been a success in accounting the stability of acetamide over acetyl fluoride in water (the positive potential on such surface in the former case is

less as compared to the latter, especially near the carbonyl carbon), enhancement in reactivity of p-nitrofluorobenzene over fluorobenzene towards SN^2 nucleophilic displacements etc.

I.5.5 MESP based Charge Models

Hydrogen bonding and solvation are important phenomena in biological processes. The former force is responsible for holding two helical chains of nucleotides in deoxyribose nucleic acid together [41] while various types of interactions such as ions with solvents, water with bio-molecules, protein folding etc. are governed by the latter [42]. In such situations, the electrostatic term is the major contributor to the total interaction energies (IEs) [43, 44], specially at larger separation distances [45]. The *ab initio* QM methods cannot be routinely employed for such studies (calculations involving several molecules simultaneously) due to complicated nature of WF [46]. This problem is overcome by the use of point charge models (PCMs). To demonstrate their usefulness, two examples are presented hereafter.

1. The EPIC (Electrostatic Potential for Intermolecular Complexation) model has been used to study hydration patterns of molecules like formaldehyde, methanol, cyclopropane [47], formamide [48] and uracil [49]. The IEs and geometries of complexes (structures of hydrated species) are found to be in good agreement with *ab initio* counterparts. These geometries when used for QM calculations are found to converge faster.
2. The water models developed by H. Yu and W. Gunsteren [50], comprising of positive point charges placed at hydrogen sites and a negative charge along C_2 axis, are able to simulate properties (viz. heat capacity, self-diffusion constant, thermal expansion coefficients etc.) of liquid water at various thermodynamic

states. These properties correlate well with those derived from experiments and QM calculations.

One of the earlier ways to obtain charges on atoms, involves partitioning the electron charge density obtained from QM calculations into atomic populations viz. Mulliken population analysis [51], Löwdin population analysis [52, 53], Natural population analysis [54] etc. Such charges depend on the level of theory chosen – on the choice of density functional and/or basis set. Momany in 1978 obtained the charges $\{q_i\}$ for a molecule containing N number of atoms situated at \mathbf{R}_i by fitting the QM calculated electrostatic potential (ESP) at various points $\{V(\mathbf{r}_j)$, at M number of points} so as to minimize the standard deviation (σ) defined below [55].

$$\sigma = \left[\frac{\sum_{j=1}^M w(\mathbf{r}_j) [V(\mathbf{r}_j) - V^{\text{CM}}(\mathbf{r}_j)]^2}{\sum_{j=1}^M w(\mathbf{r}_j)} \right]^{\frac{1}{2}} \quad (\text{I.51})$$

The term $V(\mathbf{r}_j)$ in Eq. (I.51) corresponds to the ESP value at point \mathbf{r}_j obtained from QM calculation while $V^{\text{CM}}(\mathbf{r}_j)$ is the potential due to point charges computed via. Eq. (I.52) and $w(\mathbf{r}_j)$ is the weight function employed.

$$V^{\text{CM}}(\mathbf{r}_j) = \sum_{i=1}^N \frac{q_i}{|\mathbf{r}_j - \mathbf{R}_i|} \quad (\text{I.52})$$

The work of Momany refined and extended by Cox and Williams (wherein the charges are obtained via. least square fit method involving matrix algebra) [56], was further improved by Singh and Kollman (by varying the scheme of selecting points for ESP fit) [57]. Several research groups have introduced additional constraints into the fitting procedure – reproducing molecular dipole moments [58, 59] with corresponding monopoles and quadrupoles [60]. Most widely used ESP fitting schemes are the Merz-

Kollman (MK) scheme [61], CHELPG (abbreviation for **CH**arges from **EL**ectrostatic **P**otentials using a **G**rid based method) [62] and CHELP [58], available in computational chemistry software packages like Gaussian [63], GAMESS [64] and ORCA [65]. In the MK scheme, ESP fit is done using points lying on 1.2, 1.4, 1.6 and 1.8 times the van der Waals (vdW) surface of molecules. The CHELPG method is based on generating points in a grid enclosing the molecule followed by selection of only those points which lie beyond the vdW surface while in the CHELP methodology, points lying within a distance of 3 Å from the vdW surface are selected. Such ESP fitted charges (also termed as potential-derived atomic charges) are employed in the EPIC model mentioned earlier. Irrespective of the method used to develop PCMs, such models are incapable of producing topographical features of MESP viz. bond CPs, LPs and π bonds due to lack of continuous charge component [66].

I.6 Scope of the present work

This thesis contains discussion in regard to the development of topography-driven Gaussian charge models (GCMs) wherein the nuclei and electron density function are represented by positive point charges and *s*-type spherical Gaussians associated with negative charges. Such models are then employed to investigate the potential energy surfaces of various chemical systems – in predicting geometries of weak non-covalent complexes, corresponding IEs and compare the same with QM counterparts.

A crucial step to develop GCM for any molecule requires the knowledge of various reactive sites in the molecular framework along with other MESP CPs. Gadre and co-workers have developed the computational package INDPROP [67] which locates and characterizes CPs of MESP scalar field – whose detailed description and

working is presented in Chapter II. The software uses guess points (GPs) for locating CPs, MO coefficients and basis set information. This chapter also contains a brief description of a numerical method developed which acts as a source of generating GPs for INDPROP. It has been tested for some molecules and the pros/cons of the same are also discussed.

Chapter III describes in depth the various steps involved in developing GCMs for molecules like ammonia, water, hydrogen sulphide, formaldehyde, ethylene, acetylene, benzene, methanol and pyrazine. Each model in general basically depends upon three parameters: magnitude of positive charges at nuclei sites, location of spherical Gaussians and Gaussian exponents. The acceptability criteria – GCMs expected to reproduce correct number and type of MESP critical points, are found to be faithfully obeyed by the developed models.

Though there are earlier reports on topography-driven GCMS [68-70], the study presented in Chapter IV describes the potential energy surfaces (PESs) of various molecular systems investigated using GCMs. A detailed description of the formulae involved in computing IEs using QM methods and GCMs is also available in this chapter. Besides the GCM work outlined in Chapters III and IV, Chapter V describes some additional studies (briefly highlighted in the upcoming two paragraphs) carried out using QM methods.

In collaboration with Department of Chemistry, Prof. John Barnabas, Post Graduate School of Biological Studies, Ahmednagar College, Ahmednagar, India, the reaction mechanism leading to the formation of 3-aryl-6-ethoxycarbonyl-4-hydroxy-2*H*-pyran-2-ones (general reaction shown in Figure I.3) is investigated.

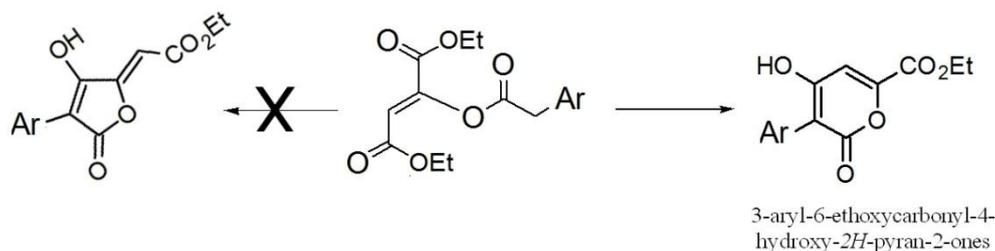


Figure I.3. General reaction leading to the formation of 3-aryl-6-ethoxycarbonyl-4-hydroxy-2H-pyran-2-ones.

Drug molecules (chemical structures displayed in Figure 1.4) form coloured solutions on reaction with cobalt thiocyanate reagent. It is suspected that the colour results due to the formation of coordination type of complexes – due to interaction of nitrogen LPs (belonging to thiocyanate and drug molecules) with cobalt (II) ion. The reactivity of those molecules containing more than one nitrogen atom is investigated using MESP and accordingly the structures of coloured complexes are derived from QM calculations. This study has been carried out in collaboration with the Regional Forensic Science Laboratory from Nagpur, State of Maharashtra, India.

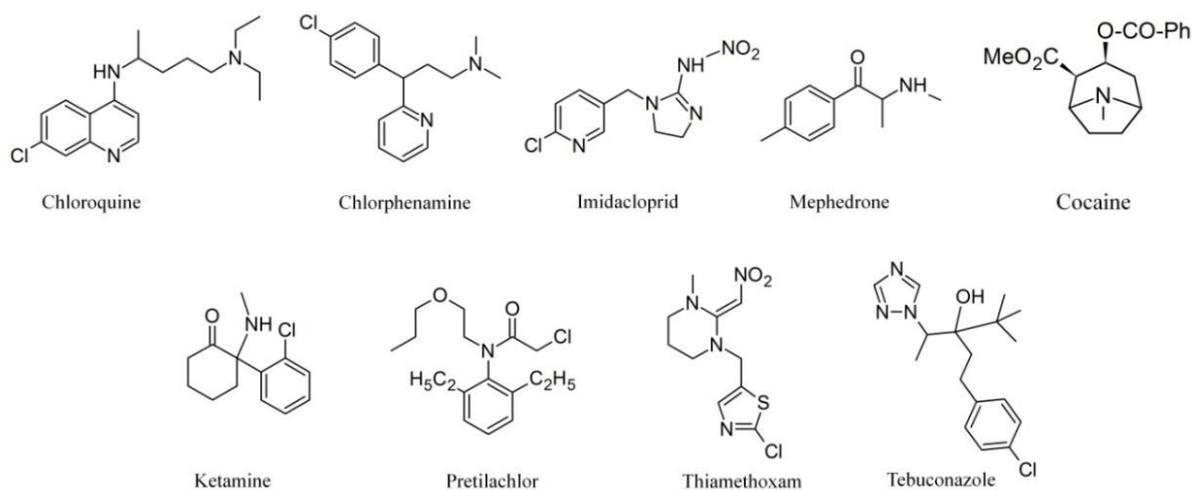


Figure I.4. Structures of nitrogen containing drugs detected by cobalt thiocyanate in combination with TLC technique.

References

1. I. Newton, A Letter of Mr. Isaac Newton, Professor in the Mathematicks in the University of Cambridge; containing his New Theory about Light and Colors: sent by the Author to the Publisher from Cambridge, Febr. 6. 1671/72; in order to be communicated to the R. Society, Phil. Trans. 6, (1671) 3075.
2. T. Young, Experiments and Calculations relative to physical Optics, Phil. Trans., 94 (1804) 1.
3. J.C. Maxwell, A Dynamical Theory of the Electromagnetic Field, Phil. Trans. R. Soc., 155 (1865) 459.
4. W. Wien, Über die Energievertheilung im Emissionsspectrum eines schwarzen Körpers, Ann. d. Physik, 294 (1896) 662.
5. Lord Rayleigh F.R.S., LIII. Remarks upon the law of complete radiation, Phil. Mag., 49 (1900) 539.
6. M. Planck, Über eine Verbesserung der Wein'schen Spectralgleichung, Verh. Dtsch. Phys. Ges. 2 (1900) 202.
7. M. Planck, Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum, Verh. Dtsch. Phys. Ges. 2 (1900) 237.
8. A. Einstein, Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt, Ann. d. Physik, 17 (1905) 132.
9. A.H. Compton, A quantum theory of the scattering of X-rays by light elements, Phys. Rev., 21 (1923) 483.
10. L. de Broglie, A Tentative Theory of Light Quanta, Phil. Mag., 47 (1924) 446.
11. C. Davisson and L.H. Germer, Diffraction of electrons by a crystal of nickel, Phys. Rev., 30 (1927) 705.

12. C. Jönsson, Elektroneninterferenzen an mehreren künstlich hergestellten Feinspalten, *Z. Physik*, 161 (1961) 454.
13. N. Bohr, On the Constitution of Atoms and Molecules, *Phil. Mag.*, 26 (1913) 1.
14. E. Schrödinger, 1. Quantisierung als Eigenwertproblem, *Ann. d. Physik*, 79 (1926) 32.
15. M. Born, Zur Quantenmechanik der Stossvorgänge, *Z. Physik*, 37 (1926) 863.
16. I.N. Levine, *Quantum Chemistry*. 7th ed., Pearson India Education Services Pvt. Ltd., 2016.
17. L. Pauling and E.B. Wilson, Jr., *Introduction to Quantum Mechanics With Applications to Chemistry*, Dover Publications, Inc. New York, 2006.
18. M. Born and J.R. Oppenheimer, Zur Quantentheorie der Molekeln, *Ann. d. Phys.*, 84 (1927) 457.
19. D.R. Hartree, The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods, *Phil. Soc.*, 24 (1928) 89.
20. D.R. Hartree, The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion, *Phil. Soc.*, 24 (1928) 111.
21. J.C. Slater, The Theory of Complex Spectra, *Phys. Rev.*, 34 (1929) 1293.
22. V. Fock, Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems, *Z. Physik*, 61 (1930) 126.
23. J.C. Slater, A Simplification of the Hartree-Fock Method, *Phys. Rev.*, 81 (1951) 385.
24. C.C.J. Roothan, New Developments in Molecular Orbital Theory, *Rev. Mod. Phys.*, 23 (1951) 69.
25. J.C. Slater, Atomic Shielding Constants, *Phys. Rev.*, 36 (1930) 57.

26. I. Shavity, The History and Evolution of Gaussian Basis Sets, *Isr. J. Chem.*, 33 (1993) 357.
27. S.R. Gadre and R.N. Shirsat, *Electrostatics of Atoms and Molecules*, Universities Press (India) Limited, Hyderabad, 2000.
28. C. Møller and M.S. Plesset, Note on an Approximation Treatment for Many-Electron Systems, *Phys. Rev.*, 46 (1934) 618.
29. P. Hohenberg and W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev.*, 136 (1964) 864.
30. W. Kohn and L.J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, *Phys. Rev.*, 140 (1965) 1133.
31. M.J. Ramos and B. Webster, An Analytic Method for the Calculation of Electrostatic Molecular Potentials, *J. Chem. Soc., Faraday Trans. 2*, 79 (1983) 1389.
32. SPARTAN: Wave function, Inc., 18401 Von Karman, Suite 310, Irvine CA 92715, U.S.A.
33. HyperChem: Autodesk, Inc., Scientific Modeling Division, 2320 Marinship Way, Sausalito CA 94965, U.S.A.
34. A.C. Limaye and S.R. Gadre, UNIVIS-2000: An indigenously developed comprehensive visualization package, *Curr. Sci.*, 80 (2001) 1296.
35. T. Brinck, The Use of the Electrostatic Potential for Analysis and Prediction of Intermolecular Interactions, in C. Párkányi (ed.), *Theoretical Organic Chemistry*, Elsevier Science B.V., Amsterdam, The Netherlands, 1998.
36. E. Scrocco and J. Tomasi, The Electrostatic Molecular Potential as a Tool for the Interpretation of Molecular Properties, *New Concepts II*, 42 (1973) 95.

37. A. Kumar, S.R. Gadre, N. Mohan and C.H. Suresh, Lone Pairs: An Electrostatic Viewpoint, *J. Phys. Chem. A*, 118 (2014) 526.
38. S.A. Kulkarni and S.R. Gadre, Electron localization in molecules. A comparative study of scalar fields, *J. Mol. Struct.*, 361 (1996) 83.
39. S.R. Gadre and R.K. Pathak, Non-existence of local maxima in molecular electrostatic potential maps, *Proc. Indian Acad. Sci. (Chem. Sci.)*, 102 (1990) 189.
40. P. Sjoberg and P. Politzer, Use of the Electrostatic Potential at the Molecular Surface to Interpret and Predict Nucleophilic Processes, *J. Phys. Chem.*, 94 (1990) 3959.
41. J.D. Watson and F.H.C. Crick, Molecular structure of Nucleic acids, *Nature*, 171 (1953) 737.
42. P. Ren, J. Chun, D.G. Thomas, M.J. Schnieders, M. Marucho, J. Zhang and N.A. Baker; Biomolecular electrostatics and solvation: a computational perspective, *Q. Rev. Biophys.*, 45 (2012) 427.
43. J. Tomasi, R. Bonaccorsi and R. Cammi, in *Theoretical Models of Chemical Bonding, Part 4*, Z. B. Maksic, Ed., Springer-Verlag, Berlin. (1991).
44. W.A. Sokalski, M. Shibata, R.L. Ornstein and R. Rein, Cumulative Atomic Multipole Moments Complement Any Atomic Charge Model to Obtain More Accurate Electrostatic Properties, *J. Comp. Chem.*, 13 (1992) 883.
45. H. Umeyama and K. Morokuma, The Origin of Hydrogen Bonding. An Energy Decomposition Study, *J. Amer. Chem. Soc.*, 99 (1977) 1316.
46. G.G. Hall and C.M. Smith, Fitting Electron Densities of Molecules, *Int. J. of Quantum Chem.*, 25 (1984) 881.
47. S.R. Gadre and K. Babu, Electrostatic Models for Weak Molecular Complexation, *Resonance*, 4 (1999) 11.

48. S.R. Gadre and A.D. Kulkarni, Molecular electrostatics for exploring hydration patterns of molecules: 2 – Formamide, *Indian J. Chem.*, 39 (2000) 50.
49. S.R. Gadre, K. Babu and A.P. Rendell, Electrostatics for Exploring Hydration Patterns of Molecules. 3. Uracil, *J. Phys. Chem. A*, 104 (2000) 8976.
50. H. Yu and W.F. Gunsteren, Charge-on-spring polarizable water models revisited: From water clusters to liquid water to ice, *J. Chem. Phys.*, 121(2004) 9549.
51. R.S. Mulliken, Electronic Population Analysis on LCAOMO Molecular Wave Functions. 1, *J. Chem. Phys.*, 23 (1955) 1833.
52. P.O. Löwdin, On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals, *J. Chem. Phys.*, 18 (1950) 365.
53. J. Baker, Classical chemical concepts from *ab initio* SCF calculations, *Theor. Chim. Acta*, 68 (1985) 221.
54. A.E. Reed, R.B. Weinstock and F. Weinhold, Natural population analysis, *J. Chem. Phys.*, 83 (1985) 735.
55. F.A. Momany, Determination of Partial Atomic Charges from Ab Initio Molecular Electrostatic Potentials. Application to Formamide, Methanol and Formic Acid, *J. Phys. Chem.*, 82 (1978) 592.
56. S.R. Cox and D.E. Williams, Representation of the Molecular Electrostatic Potential by a Net Atomic Charge Model, *J. Comput. Chem.*, 2 (1981) 304.
57. U.C. Singh and P.A. Kollman, An Approach to Computing Electrostatic Charges for Molecules, *J. Comput. Chem.*, 5 (1984) 129.
58. L.E. Chirlian and M.M. Francl, Atomic Charges Derived from Electrostatic Potentials: A Detailed Study, *J. Comput. Chem.*, 8 (1987) 894.

59. R.J. Woods, M. Khalil, W. Pell, S.H. Moffat and V.H. Smith, Derivation of Net Atomic Charges from Molecular Electrostatic Potentials, *J. Comput. Chem.*, 11 (1990) 297.
60. D.E. Williams, Representation of the Molecular Electrostatic Potential by Atomic Multipole and Bond Dipole Models, *J. Comput. Chem.*, 9 (1988) 745.
61. B.H. Besler, K.M. Merz and P.A. Kollman, Atomic Charges Derived from Semiempirical Methods, *J. Comput. Chem.*, 11 (1990) 431.
62. C.M. Breneman and K.B. Wiberg, Determining Atomic-Centered Monopoles from Molecular Electrostatic Potentials, The Need for High Sampling Density in Formamide Conformational Analysis, *J. Comput. Chem.*, 11 (1990) 361.
63. Gaussian 03, Revision C.01, M.J. Frisch, G.W. Trucks, H.B. Schlegel et al., Gaussian, Inc., Wallingford CT, 2004.
64. M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Ensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, General Atomic and Molecular Electronic Structure System, *J. Comput. Chem.*, 14 (1993) 1347.
65. F. Neese, The ORCA program system, *WIREs Comput. Mol. Sci.*, 2 (2012) 73.
66. G.G. Hall, Point charges and the molecular electrostatic potential, *Int. Rev. Phys. Chem.*, 5 (1986) 115.
67. R.N. Shirsat, S.V. Bapat and S.R. Gadre, Molecular electrostatics. A comprehensive topographical approach, *Chem. Lett. Phys.*, 200 (1992) 373.
68. S.R. Gadre and I.H. Shrivastava, Topography-driven electrostatic charge models for molecules, *Chem. Phys. Lett.*, 204 (1993) 350.
69. I.H. Shrivastava and S.R. Gadre, Molecular Electrostatic Charge Models: A Topographical Approach, *Int. J. Quantum Chem.*, 49 (1994) 397.

70. S.R. Gadre, S.S. Pundlik and I.H. Shrivastava, A “critical” appraisal of electrostatic charge models for molecules, *Prod. Indian Acad. Sci. (Chem. Sci.)*, 106 (1994) 303.

CHAPTER II

Topography of Molecular Electrostatic Potential

	Table of Contents	Page
II.1	Introduction	41
II.2	Significance of MESP topography	43
II.3	Analytical approach of locating and characterizing critical of MESP scalar field	44
II.4	Numerical approach of locating guess points for MESP topography illustrated using pyrazole and imidazole molecules	48
II.5	General remarks and conclusions	58
	References	60

II.1 Introduction

A topographical map of a geographical region describes features like hill tops, valleys, equal altitude points etc. of corresponding landscape. In other words, such a map is a two dimensional representation of a three dimensional (3D) earth surface. Figure II.1 brings out with the help of contours, equal height points of a geographical landscape wherein the two hill peaks P1 and P2 fall in the innermost region of the contour plot while points of equal heights are connected by contour lines. Contours representing heights of 10 and 20 metres engulf both peaks while those of 30 and 40 metres are disjoint. The main advantage of topography is that it provides essential information such as location and values of the function at minimum, maximum, iso-valued points etc.

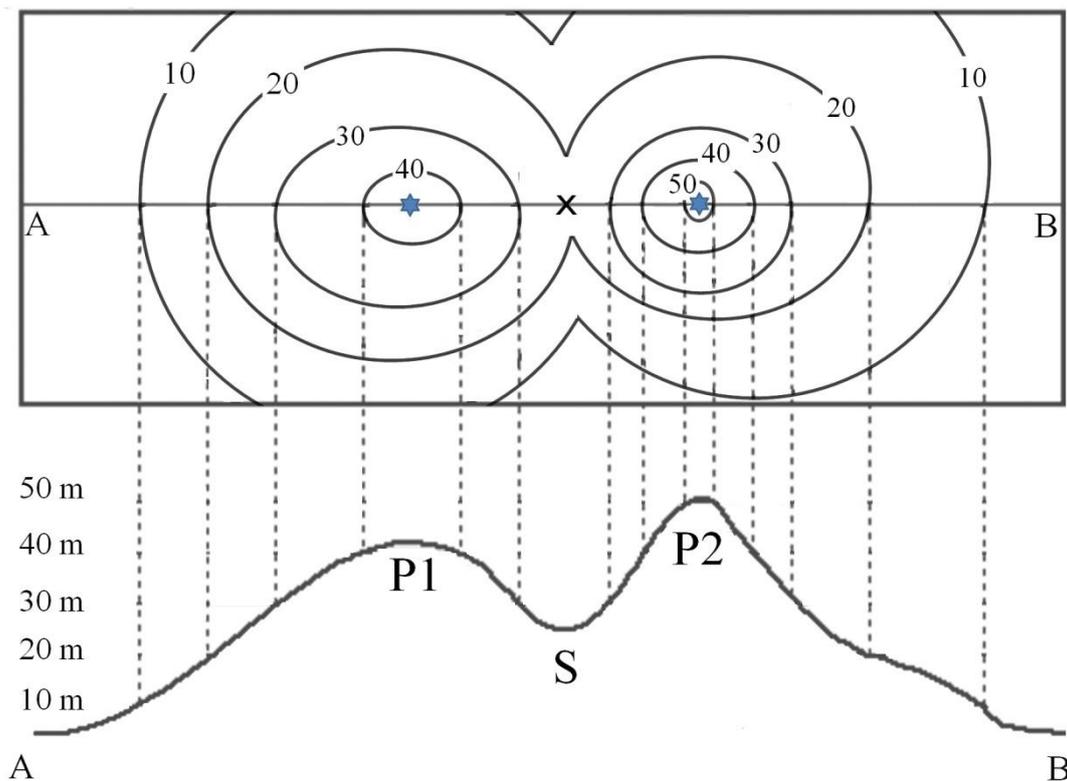


Figure II.1. Planar view of a 3D landscape and the corresponding contour map is shown in rectangle. The numbers indicate height at various levels, ★ symbol represents peak points P1 and P2 while '×' indicates the saddle point S in the contour plot.

The topographical parameters (viz. minima, maxima etc.) of a function can be searched by invoking the well-known conditions involving first and second derivatives, briefly illustrated with the help of a function of one variable $f(x)$ defined by Eq. (II.1). The plot of $f(x)$ from $x = -2$ to $x = 3$ is shown in Figure II.2. [The notations: $f'(x)$ and $f''(x)$ correspond to the first and second derivatives of $f(x)$.]

$$f(x) = \frac{x^4}{4} - \frac{2x^3}{3} - \frac{x^2}{2} + 2x - 2 \quad (\text{II.1})$$

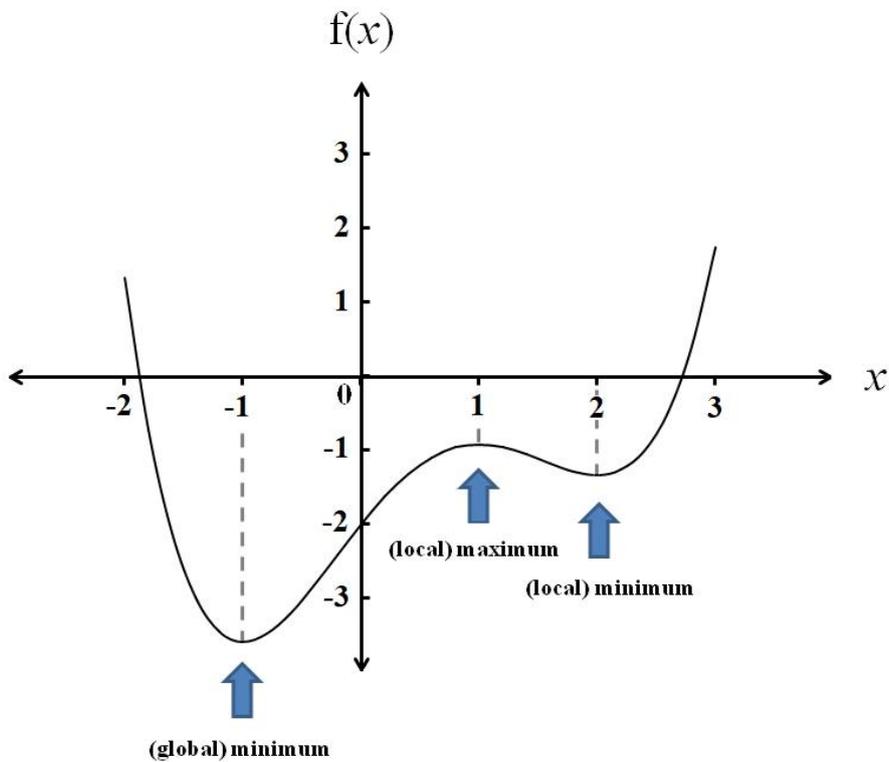


Figure II.2. Plot of $f(x)$ defined in Eq. (II.1) for x in the interval $[-2, 3]$. The function has a global minimum at $x = -1$ with $f(-1) = -3.5833$, a local minimum at $x = 2$ ($f(2) = -1.3333$) and local maximum at $x = 1$ ($f(1) = -0.9167$).

The first derivative of Eq. (II.1), $f'(x)$ is presented in Eq. (II.2). At points $x = -1, 1$ and 2 , Eq. (II.2) becomes zero. Such points are called critical points (CPs). The sign of the second derivative ($f''(x)$) evaluated at the CP is used describe the nature of CP (i.e. if $f''(x) < 0$, the CP is (local) maximum; if $f''(x) > 0$, the CP corresponds to a

(local) minimum and the CP is a saddle if $f''(x)$ equals zero. One can easily check the math and conclude that the CPs $x = -1, 2$ represent minima and there exists a (local) maximum at $x = 1$.

$$f'(x) = x^3 - 2x^2 - x + 2 \quad (\text{II.2})$$

$$f''(x) = 3x^2 - 4x - 1 \quad (\text{II.3})$$

The above results are now generalized for a function of several variables viz. $f(x_1, x_2, \dots, x_n)$. Point P is a CP when the first order partial derivatives $\frac{\partial f}{\partial x_i}$ become zero at P i.e. $\left. \frac{\partial f}{\partial x_i} \right|_P = 0$. The CP is then characterized based on the eigenvalues $(\lambda_1, \lambda_2, \dots, \lambda_n)$ obtained from the Hessian matrix H (i.e. matrix of second order partial derivatives with elements $H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$) evaluated at the CP. If one or more eigenvalues is/are zero, then the CP is termed as a degenerate one, else it is said to be a nondegenerate CP. The nature of the latter is denoted by (rank, signature) where the rank corresponds to the number of non-zero eigenvalues and signature denotes the algebraic sum of the signs of eigenvalues. For 3D scalar fields rank of a nondegenerate CP is 3 and four types of such CPs exist: (3, +3) and (3, -3) which correspond to local minimum and local maximum while (3, +1) and (3, -1) are different types of saddles. The chemical significance of the CPs belonging to MESP scalar field are outlined in the upcoming section.

II.2 Significance of MESP topography

The (3, -1) CP, referred to bond CP is located between each pair of bonded atoms and $V(\mathbf{r})$ value at this site is positive, indicating that the inter-nuclear region is mostly dominated by the nuclear potential (first term of Eq. (I.49)). Its position defines

a physical border point between the bonded atoms and has been related to the covalent radii of respective atoms [1].

The (3, -3) CP corresponds to a local maximum which explicitly represents a nucleus [2]. For a non-nuclear maximum to occur, the three eigenvalues $\{\lambda_1, \lambda_2, \lambda_3\}$ and their sum are expected to be negative quantities. As MESP and electron density functions are related via. Eq. (II.4), the right hand side quantity is always non-negative and therefore $\nabla^2V(\mathbf{r})$ cannot adopt a negative sign. Therefore MESP fails to exhibit non-nuclear maxima [Chapter I, ref. 39].

$$\nabla^2V(\mathbf{r}) = \lambda_1 + \lambda_2 + \lambda_3 = 4\pi\rho(\mathbf{r}) \quad (\text{II.4})$$

The negative valued (3, +3) CPs are those sites where the second term of Eq. (I.49) maximally dominates over the first and correspond to the most reactive sites (lone pairs (LPs) and π bonds) capable of reacting with electrophilic species as illustrated in Chapter I, section I.5.2. A positive valued (3, +3) CP indicates the presence of a cage structure in the molecular framework. Two nearest MESP minima are found to be connected by a negative-valued (3, +1) saddle while a positive-valued (3, +1) CP indicates the presence of a ring structure [3, 4].

II.3 Analytical approach of locating and characterizing critical points of MESP scalar field

Gadre and co-workers have developed the computational chemistry package INDPROP [Chapter I, ref. 67] which aids in locating and characterizing CPs of MESP scalar field. The requirements for INDPROP are: nuclear charges/co-ordinates of atoms and corresponding basis functions, density matrix elements (or occupied molecular orbital (MO) coefficients) and more importantly – guess points (GPs) for MESP

topography. For a given GP, the function CHISQ (sum of the squares of first order partial derivatives of $V(\mathbf{r})$) is minimized using the subroutine STEPIT [5]. When value of CHISQ becomes less than 10^{-13} , the CP located is characterized as per earlier discussion. The STEPIT algorithm finds minimum of a mathematical function without the use of derivatives, provided the function is well-defined in a given domain. The minimization process occurs through cyclic relaxation method: the function CHISQ is minimized with respect to the first variable, keeping others fixed followed by repeating the procedure for all variables. The co-ordinates of initial GPs are supplied to INDPROP by visualizing the scalar field using the graphic visualization package UNIVIS-2000 [Chapter I, ref. 34] and obtaining co-ordinates from appropriate regions (along bonds, reactive sites, etc.). This is demonstrated in the upcoming section using acetone molecule (Figure II.3a) at the Hartree-Fock (HF)/6-31G(d,p) level of theory.

Variety of options exists to visualize molecules in UNIVIS-2000 (namely line, stick, tube, space-fill, ball-and-stick models or vdW type of surfaces) out of which the ball-and-stick model view of acetone enclosed in the 3D grid is depicted in Figure II.3b. The '*property settings*' option not only enables the user to obtain basic information about the 3D grid (co-ordinates of initial/final grid points, number of points along the Cartesian axes, etc.) but also aids in visualizing the scalar field in terms of iso-valued surfaces and 3D contours.

Isosurface of the lowest $V(\mathbf{r})$ value i.e. -0.0916 a.u. is hardly noticeable (*lowest* in present context refers to the lowest MESP value computed on the grid), however, two island-like surfaces appear in the vicinity of oxygen atom at a slightly higher $V(\mathbf{r})$ value of about -0.0850 a.u., as depicted in Figure II.3c. Submitting co-ordinates of points lying on these surfaces to INDPROP will definitely lead to exact

locations of MESP minima – corresponding to LPs of sp^2 hybridized oxygen atom. UNIVIS-2000 facilitates extraction of co-ordinates of points lying on isosurfaces. Co-ordinates of two such GPs: GP1/GP2 (locations shown in Figure II.3c) are presented in Table II.1a and the results obtained from INDPROP are displayed in Table II.1b (entries 1 and 2). Their geometric parameters like distances from carbonyl oxygen/angle formed with carbonyl group are mentioned previously in Chapter I, section I.5.2.

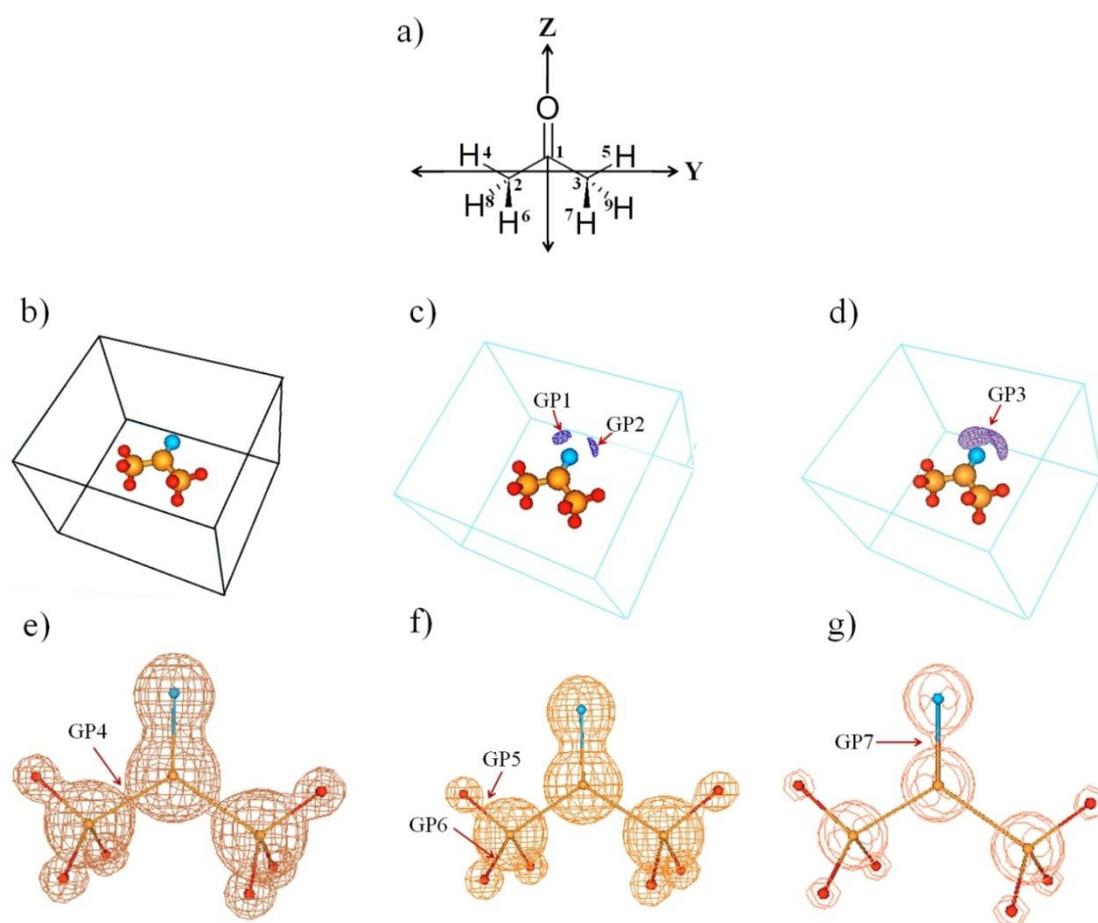


Figure II.3. a) Numbering of carbon/hydrogen atoms in acetone molecule; b) ball-and-stick model view (in UNIVIS-2000) of acetone enclosed in a 3D grid; negative valued MESP isosurfaces of values: c) -0.085 and d) -0.078 a.u. The grid boundary is given less prominence and therefore shown in lighter color; positive valued MESP isosurfaces of values: e) 0.676, f) 0.850 and g) 1.515 a.u. The molecule is viewed in ‘stick model’ mode in figures II.3e, II.3f and II.3g. Grid specification in atomic units: initial co-ordinates (-6.512752, -8.946265, -7.660813) and increment in X, Y and Z direction: 0.333333 a.u. Molecular geometry available in APPENDIX, section A7.

Table II.1a. Co-ordinates of guess points for MESP topography of acetone molecule at HF/6-31G(d,p) level of theory obtained using the visualization package UNIVIS-2000.

Guess points		Co-ordinates			MESP value at GP
Label	for locating	X	Y	Z	
GP1	(3, +3) CP	-0.6545	-1.5900	4.4340	-0.0850
GP2		-0.4650	1.4460	4.2790	-0.0850
GP3	(3, +1) CP	-0.7810	-0.2580	5.0970	-0.0780
for bond saddles					
GP4	between C2-C1	-0.0450	-1.2210	-0.4110	0.6760
GP5	between C2-H4	-0.2960	-3.2300	-0.1010	0.8500
GP6	between C2-H6	1.2150	-2.3640	-1.8750	
GP7	between C1-O	0.1060	-0.0540	1.5710	1.5150

Table II.1b. MESP critical points for acetone molecule at HF/6-31G(d,p) level of theory obtained using GPs from Table II.1a via. INDPROP.

Sr. No.	Type of CP	Location	Co-ordinates			V(r)	$\rho(\mathbf{r})$
			X	Y	Z		
1	(3, +3)	distance: 2.2873 a.u. from oxygen angle: 129.2° with carbonyl group	0.0000	-1.7658	4.0484	-0.0922	0.015
2			0.0000	1.7658	4.0484		
3	(3, +1)	C ₂ axis, distance: 2.4089 a.u. from oxygen	0.0000	0.0000	5.0154	-0.0827	0.009
4	(3, -1)	between bonds C2 – C1 and C3 – C1	0.0000	±1.2193	-0.4052	0.6847	0.271
5		between C2-H4/ C3-H5 bonds	0.0000	±3.4535	-0.3354	0.8552	0.292
6		between C2-H6/H8 and C3-H7/H9 bonds	±1.0546	±2.4878	-1.9233	0.8609	0.285
7		between C1-O bond	0.0000	0.0000	1.4847	1.4086	0.112

Co-ordinates of GP for (3, +1) saddle connecting MESP minima (Table II.1a, entry 3) is obtained from MESP isosurface of value -0.078 a.u. as depicted in Figure II.3d. Its exact location is along the Z-axis at a distance of 2.4089 a.u. (1.2743 Å) from oxygen atom (Table II.1b, entry 3). GPs for bond CPs are obtained by accessing regions where the positive MESP surface between bonded atoms starts to vanish.

The positive isosurface (enclosing the entire molecule) splits between C2-C1 and C3-C1 bonds when $V(\mathbf{r})$ approximately equals 0.676 a.u. as shown in Figure II.3e. Since C2 and C3 atoms (connecting C1) are related via symmetry, it is sufficient to obtain a suitable GP {GP4; co-ordinates in Table II.1a} to locate saddle between C2-C1 bond. The co-ordinates of the corresponding bond CP and its symmetry-related counterpart are displayed in Table II.1b, entry 4. Table II.1a also summarizes the co-ordinates of GPs used to obtain bond saddles located between C2-H4 and C2-H6 bonds (entries 5, 6). The approximate values at which the isosurfaces splits between respective atoms are also mentioned. The exact positions of MESP CPs and their symmetry-related partners are summarized in Table II.1b (entries 5, 6). The last entry in Table II.1b corresponds to the exact location of bond CP between carbon and oxygen along the Z-axis, obtained using respective GP from Table II.1a.

Implementing the above steps to locate CPs of MESP scalar field may turn out to be a tedious task, depending upon several factors: size of molecules under investigation, whether molecules possess symmetry, number of CPs to detect etc. To ease this process and directly obtain GPs from the 3D grid data itself, a numerical method has been developed and tested for few molecules. The basic principle involved is explained with the help of a one-dimensional function in the upcoming section, and the method is extended to the MESP scalar field.

II.4 Numerical approach of locating guess points for MESP topography illustrated using pyrazole and imidazole molecules

Beginning with the supposition that values of some mathematical function are known at 'n+1' points [for given points $\{x_0, x_1, \dots, x_n\}$, the function takes values $\{y_0, y_1, \dots, y_n\}$ such that $x_0 < x_1, \dots, < x_n$], and the expression for the mathematical

function may or may not be known explicitly, it is possible to predict its value at point $\{x_a\}$ lying between x_0 and x_n . This process is called ‘interpolation’ and $\{x_0, x_1, \dots, x_n\}$ are sometimes referred to tabular points. There exist various methods for interpolation depending upon whether any two successive tabular points are equally spaced or not, and discussing all of them is beyond the scope of this thesis, except the Lagrange’s interpolation (LI) method wherein a polynomial $L(x)$ represented by Eq. (II.5) is fitted so as to pass through points $\{y_0, y_1, \dots, y_n\}$ [6]. The terms $L_0(x), L_1(x), \dots, L_n(x)$ or $L_i(x)$ in general, are known as Lagrange’s coefficients (defined by Eq. (II.6)) and they satisfy the property as per Eq. (II.7).

$$L(x) = L_0(x) y_0 + L_1(x) y_1 + \dots + L_n(x) y_n \quad (\text{II.5})$$

$$L_i(x) = \frac{(x - x_0)(x - x_1)\dots(x - x_{i-1})(x - x_{i+1}) \dots(x - x_n)}{(x_i - x_0)(x_i - x_1)\dots(x_i - x_{i-1})(x_i - x_{i+1}) \dots(x_i - x_n)} \quad (\text{II.6})$$

$$L_i(x) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (\text{II.7})$$

To demonstrate the application of this method in detecting CPs of a one-dimensional function, the tabular point data from Table II.2 is employed (tabular points of Eq. (II.1) in the interval $(-2, 3)$, starting from $x = -1.9$ to $x = 2.9$ with step length $(\Delta x$ – difference between two successive tabular points) of 0.2). The search for CP begins in each sub-interval (x_i, x_{i+1}) starting with an initial guess point, which for convenience is taken as the mid-point of the sub-interval under consideration. Rather than interpolating via Eq. (II.5), the fitted-polynomial is constructed using only four tabular points $\{x_{i-1}, x_i, x_{i+1}, x_{i+2}\}$ and respective $\{y_{i-1}, y_i, y_{i+1}, y_{i+2}\}$ values. The expression for $L(x)$ reduces to Eq. (II.8) and corresponding Lagrange’s coefficients are defined by Eqs. (II.9) to (II.12).

Table II.2. Tabular points for Eq. (II.1) for variable x between $x = -2$ and $x = 3$ with $\Delta x = 0.2$ and corresponding $f(x)$ values.

Value of 'i'	x_i	y_i	Value of 'i'	x_i	y_i
0	-1.9000	0.2257	13	0.7000	-1.0136
1	-1.7000	-1.4816	14	0.9000	-0.9270
2	-1.5000	-2.6094	15	1.1000	-0.9263
3	-1.3000	-3.2663	16	1.3000	-0.9956
4	-1.1000	-3.5516	17	1.5000	-1.1094
5	-0.9000	-3.5550	18	1.7000	-1.2323
6	-0.7000	-3.3563	19	1.9000	-1.3196
7	-0.5000	-3.0260	20	2.1000	-1.3170
8	-0.3000	-2.6250	21	2.3000	-1.1603
9	-0.1000	-2.2043	22	2.5000	-0.7760
10	0.1000	-1.8056	23	2.7000	-0.0810
11	0.3000	-1.4610	24	2.9000	1.0177
12	0.5000	-1.1927			

$$L(x) = L_{i-1}(x) y_{i-1} + L_i(x) y_i + L_{i+1}(x) y_{i+1} + L_{i+2}(x) y_{i+2} \quad (\text{II.8})$$

$$L_{i-1}(x) = \frac{(x - x_i)(x - x_{i+1})(x - x_{i+2})}{(x_{i-1} - x_i)(x_{i-1} - x_{i+1})(x_{i-1} - x_{i+2})} \quad (\text{II.9})$$

$$L_i(x) = \frac{(x - x_{i-1})(x - x_{i+1})(x - x_{i+2})}{(x_i - x_{i-1})(x_i - x_{i+1})(x_i - x_{i+2})} \quad (\text{II.10})$$

$$L_{i+1}(x) = \frac{(x - x_{i-1})(x - x_i)(x - x_{i+2})}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)(x_{i+1} - x_{i+2})} \quad (\text{II.11})$$

$$L_{i+2}(x) = \frac{(x - x_{i-1})(x - x_i)(x - x_{i+1})}{(x_{i+2} - x_{i-1})(x_{i+2} - x_i)(x_{i+2} - x_{i+1})} \quad (\text{II.12})$$

At given initial guess, value of $L'(x)$ (defined by Eq. (II.13)) is evaluated. If the computed quantity is less than some predefined tolerance value, then, the CP is characterized by $L''(x)$ expression defined by Eq. (II.14). The Lagrange's coefficients in Eqs. (II.13) and (II.14) will be discussed at a later stage in this thesis.

$$L'(x) = L'_{i-1}(x) y_{i-1} + L'_i(x) y_i + L'_{i+1}(x) y_{i+1} + L'_{i+2}(x) y_{i+2} \quad (\text{II.13})$$

$$L''(x) = L''_{i-1}(x) y_{i-1} + L''_i(x) y_i + L''_{i+1}(x) y_{i+1} + L''_{i+2}(x) y_{i+2} \quad (\text{II.14})$$

If value of $L'(x)$ at initial guess is not less than the defined tolerance value, a new point x_m is obtained via Newton-Raphson (NR) method [7] and the previous step is repeated once again. This iterative procedure is terminated if the updated GP falls outside the sub-interval under consideration or if value of $L'(x)$ becomes less than the tolerance value, or, when the number of NR steps reach some predefined threshold value. A detailed description of the above outlined procedure to obtain CPs using data from Table II.2 is provided in APPENDIX, section A8 along with the definitions of Lagrange's coefficients in Eqs. (II.13) and (II.14).

Extending this methodology to three dimensions, the x , y , z components of point P lie between tabular points (x_i, x_{i+1}) , (y_j, y_{j+1}) and (z_k, z_{k+1}) . In other words, point P is confirmed in a cuboid whose dimensions are Δx , Δy and Δz . The co-ordinates of the corners are as follows: (x_i, y_j, z_k) , (x_i, y_j, z_{k+1}) , (x_i, y_{j+1}, z_k) , (x_i, y_{j+1}, z_{k+1}) , (x_{i+1}, y_j, z_k) , (x_{i+1}, y_j, z_{k+1}) , (x_{i+1}, y_{j+1}, z_k) and $(x_{i+1}, y_{j+1}, z_{k+1})$. Point P is enclosed in a cube if Δx , Δy , Δz are of equal magnitude. The Lagrange polynomial has the form as presented in Eq. (II.15) where the $L_i(x)$, $L_j(y)$ and $L_k(z)$ components are constructed using tabular points $\{x_{i-1}, x_i, x_{i+1}, x_{i+2}\}$, $\{y_{j-1}, y_j, y_{j+1}, y_{j+2}\}$ and $\{z_{k-1}, z_k, z_{k+1}, z_{k+2}\}$ while the term f_{111} (i.e. value of f_{ijk} when i, j, k equals 1) corresponds to the scalar field value at point bearing co-ordinates $(x_{i-1}, y_{j-1}, z_{k-1})$. Terms f_{112} , f_{113} etc. are respective scalar field values at points with co-ordinates (x_{i-1}, y_{j-1}, z_k) , $(x_{i-1}, y_{j-1}, z_{k+1})$ and so on.

$$L(x, y, z) = \sum_{i=1}^4 \sum_{j=1}^4 \sum_{k=1}^4 L_i(x) L_j(y) L_k(z) f_{ijk} \quad (\text{II.15})$$

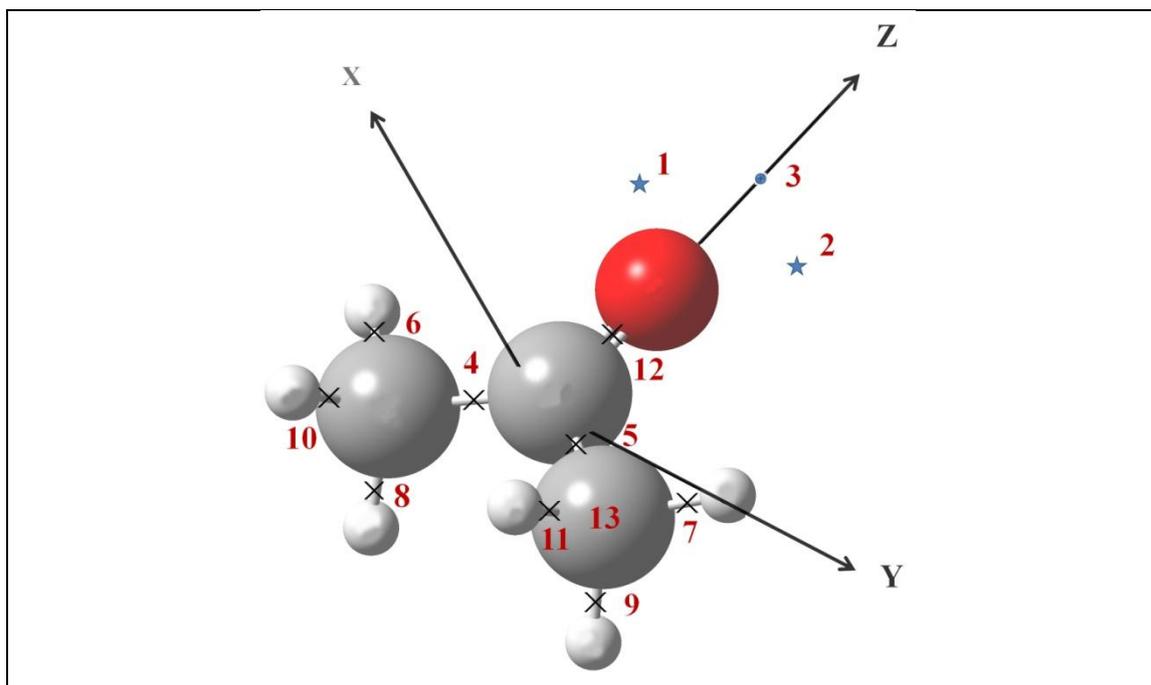
To obtain GPs for MESP topography, scalar field values are generated on the 3D grid via Gaussian cubegen utility. Assuming that there are n_x , n_y , n_z number of tabular points along X, Y and Z directions (and $\Delta x = \Delta y = \Delta z$), there exist a total of ‘ $(n_x-1)(n_y-1)(n_z-1)$ ’ number of cubes to be analyzed. The initial guess point $\{\mathbf{r}_0\}$ is generated at the centre of each cube and Eq. (II.15) is used to interpolate MESP values within the cube. Using the first order partial derivatives and corresponding elements of Hessian matrix derived from Eq. (II.15) (expressions described in APPENDIX, section A9), the initial GP $\{\mathbf{r}_0\}$ is updated to $\{\mathbf{r}_1\}$ using the 3D Newton-Raphson method via Eq. (II.16) [8]. The term $H^{-1}(\mathbf{r}_k)$ corresponds to the inverse of the Hessian matrix evaluated at some reference point $\{\mathbf{r}_k\}$ which in present context is $\{\mathbf{r}_0\}$ and $\nabla L(\mathbf{r}_k)$ represents the first order partial derivatives of Eq. (II.15).

$$\mathbf{r}_{k+1} = \mathbf{r}_k - H^{-1}(\mathbf{r}_k) \cdot \nabla L(\mathbf{r}_k) \quad (\text{II.16})$$

The newly obtained point is once again updated until $\nabla L(\mathbf{r}_k)$ becomes less than 10^{-13} or if the number of NR steps reach some predefined threshold value (in current methodology it is set to 100). If the former condition is satisfied, then the CP is characterized based on the eigenvalues obtained from the Hessian matrix – as per discussion in section II.1. Rather than employing the NR method, value of $\nabla L(\mathbf{r}_k)$ could also be minimized using the subroutine STEPIT.

Since the topographical features of MESP scalar field are already discussed for acetone molecule, the corresponding 3D grid data was investigated and this methodology was found to yield a total of 13 CPs, displayed in Table II.3.

Table II.3. CPs of MESP scalar field obtained using the 3D Lagrange's interpolation method.



Symbols: ★ for GP representing MESP minima, ⊕ for GP representing saddle connecting MESP minima and × for GPs representing bond saddles.

Type of GP		GP label	Co-ordinates			Predicted value of	
Expected	Predicted value of (R, σ)		X	Y	Z	MESP	Electron density
MESP minima	(3, +3)	1	0.0000	-1.7739	4.0344	-0.0922	0.016
		2	0.0000	1.7796	4.0312	-0.0922	0.015
saddle connecting minima	(3, +1)	3	-0.0008	0.0379	5.0336	-0.0827	0.010
C-C bond saddles	(3, -1)	4	0.0019	-1.2101	-0.4194	0.6859	0.256
		5	0.0019	1.2111	-0.4153	0.6859	0.255
C-H bond saddles		6	0.0051	-3.4839	-0.3478	0.8633	0.131
		7	0.0052	3.4783	-0.3712	0.8632	0.170
		8	-1.0768	-2.5098	-1.9130	0.8754	0.096
		9	-1.0776	2.4675	-1.9148	0.8752	0.111
	(3, +1)	10	1.0765	-2.5077	-1.9170	0.8714	0.133
	(3, +1)	11	1.0775	2.4684	-1.9184	0.8716	0.141
C-O bond saddle	(3, -1)	12	0.0017	0.0145	1.4907	1.3492	0.095
----	(3, -3)	13	0.0310	-2.4024	-1.1171	14.5389	-19.689

Based on the locations of GPs, corresponding $V(\mathbf{r})$ and $\rho(\mathbf{r})$ values predicted by the numerical method, the following observations are noted:

- The GPs for MESP minima (CPs numbered 1, 2 in Table II.3) are situated at distances of 2.2773, 2.2796 a.u. (1.2047, 1.2059 Å) from oxygen atom and form angles of 128.8/128.7° with carbonyl group. The expected distance and angle values are 2.2873 a.u. (1.2100 Å) and 129.2°. Thus the deviations between respective quantities are not significant. The MESP and electron density values predicted are also in good agreement with QM results (i.e. entries 1 and 2 in Table II.1b).
- The GP (numbered 3) for saddle connecting MESP minima is situated at a distance of 2.3657 a.u. (1.2515 Å) from oxygen atom and QM counterpart value is about 2.4089 a.u. (1.2743 Å). There is once again a good agreement between predicted MESP and electron density values with those obtained via. WF (Table II.b, entry 3).
- Not much deviation is observed between predicted co-ordinates of bond saddle GPs (numbered 4, 5 in Table II.3) located between carbon-carbon bonds and their QM counterparts (Table II.1b, entry 4). Thus is true even with respect to predicted $V(\mathbf{r})$ and $\rho(\mathbf{r})$ values.
- The above point is also valid for bond saddle GPs numbered 6, 7, 8, 9, 10 and 11. The nature of GPs numbered 10 and 11 predicted by the method however are incorrect.
- Co-ordinates of GP for saddle between carbon-oxygen bond and corresponding $V(\mathbf{r})$, $\rho(\mathbf{r})$ values are reproduced quite well. Unexpectedly, only a single (3, -3) CP (labeled 13 in Table II.3) corresponding to carbon atom, is detected.

This method was initially used by Katan and co-workers to detect CPs of electron density function [9]. In the present context, it is used to analyse topographical features of MESP scalar field – to obtain suitable GPs for INDPROP. Ignoring the predicted (3, -3) CP and errors in signature values at two bond saddles, the GPs when submitted to INDPROP yield the required MESP CPs which are summarized in Table II.1b. The next set of molecules investigated are pyrazole and imidazole (structures in Figure II.4), which have been studied by Scrocco and Tomasi [10].

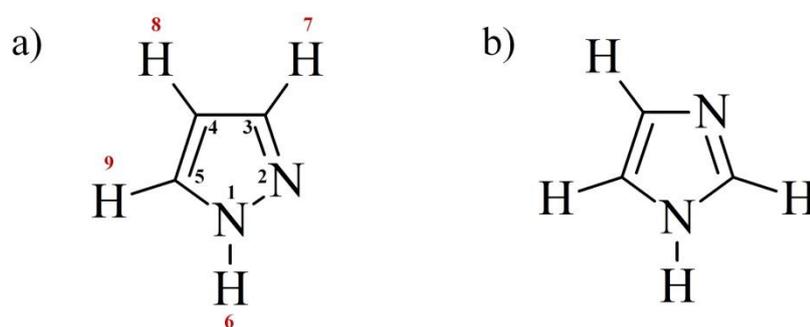
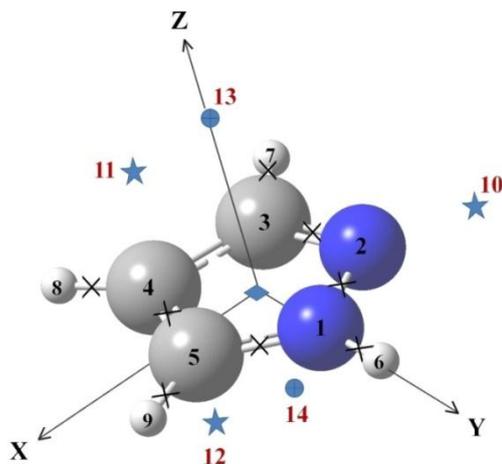


Figure II.4. Structures of a) pyrazole and b) imidazole molecules. Numbering scheme employed for pyrazole ring atoms is applicable for imidazole as well. The hydrogen atoms in the latter are labelled from 6 to 9, in anticlockwise direction starting from N1 as in pyrazole.

The MESP topography of both molecules are expected to exhibit two (3, +3) CPs, each above and below the C4-C5 double bond; a single (3, +3) CP representing the nitrogen LP of respective N2 and N3 atoms (former in pyrazole and latter in imidazole). A connection (negative valued (3, +1) CP) between the MESP minima representing π bond and LP could exist, each above and below the molecular plane. In addition to nine bond CPs and a positive valued (3, +1) CP at the centre of ring (numerical method results displayed in Tables II.5 and II.6).

Table II.4. GPs for MESP topography of pyrazole molecule at HF/6-31G(d,p) level of theory.

GP (label)	(R, σ) predicted	Co-ordinates			V(r)	$\rho(r)$
		X	Y	Z		
nitrogen LP (10)	(3, +3)	-4.2443	1.6504	-0.0001	-0.1086	0.021
		-4.2737	1.6379	0.0000	-0.1086	0.019
π bonds (11 and 12)	(3, +3)	1.0882	-1.3933	3.1683	-0.0322	0.004
		1.1106	-1.3840	-3.1351		
		1.0884	-1.3917	± 3.1564	-0.0322	0.004
saddle connecting LP and π bonds (13 and 14)	(3, +1)	-0.7128	-1.0897	3.4320	-0.0282	0.002
		-0.7123	-1.0904	-3.4342		
		-0.7167	-1.0901	± 3.4428	-0.0282	0.002
ring CP	(3, +1)	-0.0420	-0.0971	0.0011	0.3348	0.065
		-0.0375	-0.0969	0.0000	0.3361	0.055
between N1-N2	(3, -1)	-1.0633	1.3704	0.0025	1.0514	0.361
		-1.0537	1.3584	0.0000	1.0496	0.395
between N2-C3	(3, +1)	-1.7543	-0.4786	0.0021	1.0598	0.536
	(3, -1)	-1.7643	-0.4722	0.0000	1.0780	0.411
between C3-C4	GP not found					
	(3, -1)	-0.0740	-1.7235	0.0000	0.8350	0.309
between C4-C5	(3, +1)	1.6350	-0.6331	0.0021	0.9355	0.409
	(3, -1)	1.6284	-0.6381	0.0000	0.9455	0.337
between C5-N1	(3, -1)	0.9980	1.2642	0.0025	1.0229	0.333
		1.0169	1.2910	0.0000	1.0303	0.370
between N1-H6	(3, -1)	-0.0415	3.2152	0.0024	1.0665	0.731
		-0.0563	3.2327	0.0000	1.1164	0.388
between C3-H7	(3, +1)	-2.3055	-2.5822	0.0039	0.8888	0.193
	(3, -1)	-2.2913	-2.6006	0.0000	0.8797	0.301
between C4-H8	(3, +1)	1.9591	-2.9582	0.0047	0.8918	0.200
	(3, -1)	1.9680	-2.9433	0.0000	0.8839	0.296
between C5-H9	(3, -1)	3.2860	1.0357	0.0009	0.8923	0.390
		3.2604	1.0783	0.0000	0.9134	0.301



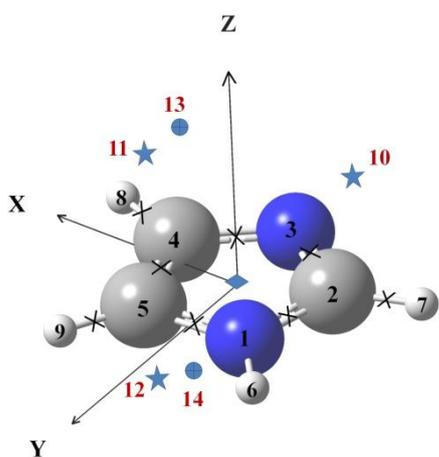
Molecular geometry:
N1 (0, 2.0323, 0); N2 (2.1307, 0.6997, 0)
C3 (-1.4096, -1.6529, 0); C4 (1.2522, -1.8695, 0)
C5 (2.0714, 0.5718, 0); H6 (-0.0869, 3.9037, 0)
H7 (-2.7877, -3.1361, 0); H8 (2.3727, -3.5509, 0)
H9 (3.9328, 1.3628, 0)

Grid specification:
Initial co-ordinates (-7.9223, -8.3822, -6.5127),
Number of tabular points per direction: 51, 49, 40
with increment of 0.333333.

Symbols:
★ for GPs representing MESP minima, ● for saddle connecting MESP minima, ◆ for GP representing ring CP and × for GPs representing bond saddles.

Table II.5. GPs for MESP topography of imidazole molecule at HF/6-31G(d,p) level of theory.

GP (label)		(R, σ) predicted	Co-ordinates			V(r)	$\rho(r)$
			X	Y	Z		
nitrogen LP (10)	(3, +3)	0.1781	-4.6168	-0.0001	-0.1284	0.024	
		0.1775	-4.6193	0.0000	-0.1278	0.020	
π bonds (11 and 12)	(3, +3)	1.9920	0.3949	3.0504	-0.0339	0.005	
		1.9907	0.3926	-3.0538	-0.0339	0.005	
		1.9916	0.3916	± 3.0590	-0.0338	0.005	
saddles (13/14) connecting LP and π bonds	(3, +1)	1.7090	-0.5121	3.3105	-0.0328	0.003	
		1.7092	-0.5097	-3.3131	-0.0328	0.003	
		1.7046	-0.5132	3.3108	-0.0327	0.003	
ring CP	(3, +1)	0.0929	-0.2050	0.0012	0.3308	0.082	
		0.0917	-0.1666	0.0000	0.3338	0.059	
between N1 - C2	(3, +1)	-1.5866	0.3651	0.0022	0.9898	0.516	
	(3, -1)	-1.5845	0.3600	0.0001	1.0127	0.364	
between C2 - N3	(3, -1)	-0.9063	-1.5918	0.0028	1.1221	0.336	
		-0.9028	-1.5882	0.0001	1.1119	0.419	
between N3 - C4	(3, +1)	1.0936	-1.4889	0.0022	0.9114	0.291	
	(3, -1)	1.0725	-1.4724	-0.0001	0.9067	0.352	
between C4 - C5	(3, +1)	1.6661	0.5323	0.0022	0.9581	0.400	
	(3, -1)	1.6471	0.5382	-0.0001	0.9635	0.348	
between C5 - N1	(3, +1)	-0.0159	1.6429	0.0020	0.9419	0.482	
	(3, -1)	-0.0353	1.6492	0.0000	0.9592	0.344	
between N1 - H6	(3, +1)	-2.0537	2.5533	0.0050	1.1070	0.277	
	(3, -1)	-2.0964	2.5108	0.0000	1.1228	0.382	
between C2 - H7	(3, -1)	-3.1524	-1.2954	-0.0051	0.8847	0.477	
		-3.1258	-1.3081	0.0004	0.9051	0.305	
between C4 - H8	(3, -1)	3.3192	-1.0722	0.0014	0.8656	0.374	
		3.3109	-1.0861	-0.0002	0.8745	0.302	
between C5 - H9	(3, +1)	1.8218	2.8988	0.0047	0.9056	0.267	
	(3, -1)	1.8919	2.8855	0.0000	0.9067	0.300	



Molecular geometry:

N1 (0, 2.0323, 0), N2 (2.1307, 0.6997, 0),
 C3 (-1.4096, -1.6529, 0), C4 (1.2522, -1.8695, 0)
 C5 (2.0714, 0.5718, 0), H6 (-0.0869, 3.9037, 0)
 H7 (-2.7877, -3.1361, 0), H8 (2.3727, -3.5509, 0)
 H9 (3.9328, 1.3628, 0)

Grid specification:

Initial co-ordinates (-8.414238, -7.782350, -6.512883)
 Number of tabular points per direction: 52, 49, 40 with increment of 0.333333.

Symbols:

★ for GPs representing MESP minima, ● for saddle connecting MESP minima, ◆ for GP representing ring CP and X for GPs representing bond saddles.

It is observed that the GPs obtained for both molecules mimic different types of MESP CPs. The signature values predicted for GPs in the negative $V(\mathbf{r})$ region are found to be +3 or +1, either representing LPs/ π bonds or saddles connecting MESP minima. In the positive MESP region however, some of the signature values at bond GP sites are found to be incorrect. There has been an instance where the method failed to detect a GP for bond CP. Few GPs with signature values of -3 were also found in both cases (which are not much emphasized upon).

This method has also been employed to study MESP topography of simple molecules like ammonia, water, hydrogen sulphide, formaldehyde, ethane and benzene. These findings will be discussed in the next chapter of this thesis – which deals with the development of topography-based charge models, for which the knowledge of topographical features themselves is essential.

II.5 General remarks and conclusions

To conclude, this chapter has described the means to obtain topographical features of MESP scalar field – by locating points (CPs) where the first order partial derivatives of $V(\mathbf{r})$ vanish, followed by their characterization based on eigenvalues obtained from the Hessian matrix evaluated at these respective points. The reactive sites in molecules such as LPs/ π bonds are indicated by presence of negative valued (3, +3) CPs in the MESP topography, and based on the $V(\mathbf{r})$ values, it is possible to differentiate between them – $V(\mathbf{r})$ values in the former case are more deeper than -0.04 a.u. The positive valued (3, -1) and (3, +1) CPs indicate bonded atoms and ring structures in the molecular framework.

A numerical algorithm for detecting CPs of MESP scalar field topography has been presented and discussed in this chapter. It has been tested for simple molecule like acetone and those possessing rich topographical features viz. pyrazole and imidazole. Almost all of the different types of MESP CPs have been detected by the NR step of the algorithm. These points serve as starters for locating exact topography via computational chemistry INDPROP. Though found to be efficient in locating GPs for MESP topography in negative $V(\mathbf{r})$ regions, it is found to predict incorrect nature at some GPs situated between bonded atoms. There is also the issue of locating all GPs of the type (3, -3). On its own, the method needs some modifications to completely predict GPs of all types of CPs, and also in terms of accuracy of prediction when the results are compared to those obtained from WF. This task remains to be undertaken in the near future.

References

1. S.R. Gadre and P.K. Bhadane, Molecular Electrostatic Potential: Visualization and Topography, *Resonance*, 11 (1999) 14.
2. S.R. Gadre and A. Kuman, Noncovalent Forces, Challenges and Advances in Computational Chemistry and Physics, in S. Scheiner (ed.). Springer International Publishing Switzerland, 2015,
3. M. Leboeuf, A.M. Köster, K. Jug and D.R. Salahub, Topological analysis of the molecular electrostatic potential, *J. Chem. Phys.*, 111 (1999) 4893.
4. S.R. Gadre and A. Kumar, Bonding and Reactivity Patterns from Electrostatic Landscapes of Molecules, *J. Chem. Sci.*, 128 (2016) 1519.
5. STEPIT, QCPE Program 66 written by J.P. Chandler, Quantum Chemistry Program Exchange (QCPE), Department of Chemistry, Indiana University, Bloomington, Indiana. [J.P. Chandler, STEPIT: Finds local minima of a smooth function of several parameters, *Behavioral Science*, 14 (1969) 81].
6. K.A. Hussein, The Lagrange Interpolation Polynomial for Neural Network Learning, *IJCSNS (International Journal of Computer Science and Network Security)*, 11 (2011) 255.
7. J. Verbeke and R. Cools, The Newton-Raphson method, *Int. J. Math. Educ. Sci. Technol.*, 26 (1995) 177.
8. A. Ben-Israel, A Newton-Raphson Method for the Solution of Systems of Equations, *J. Math. Anal. Appl.*, 15 (1966) 243.
9. C. Katan, P. Rabiller, C. Lecomte, M. Guezo and V. Oison M. Souhassou, Numerical computation of critical properties and atomic basins from 3D grid electron densities, *J. Appl. Crystallogr.*, 36 (2003) 65.

10. Chapter I, reference 30; G. Bertheier, R. Bonaccorsi, E. Scrocco and J. Tomasi,
The Electrostatic Molecular Potential for Imidazole, Pyrazole, Oxazole and
Isoxazole, *Theoret. Chim. Acta (Berl.)*, 26 (1972) 101.

CHAPTER III

Development of Topography-driven Gaussian Charge Models

	Table of Contents	Page
III.1	Introduction	63
III.2	Methodology to develop Gaussian charge models (GCM)	65
III.2.1	Ammonia molecule	65
III.2.2	Water and hydrogen sulphide molecules	72
III.2.3	Formaldehyde molecule	76
III.2.4	Ethene and acetylene molecules	78
III.2.5	Benzene molecule	82
III.3	Scope for improving MESP and Electron Density values at critical points other than MESP minima	85
III.4	General remarks on GCMs	87
III.5	Additional topography-based Gaussian models	88
III.5.1	Methanol molecule	88
III.5.2	Pyrazine molecule	91
III.6	Future prospects	94
III.7	Concluding remarks	95
	References	96

Part of this work has been:

- 1) Published in the research article titled ‘Prelude to molecular dynamics: Topography-driven Gaussian charge models’, Int. J. Quantum Chem., DOI:10.1002/qua.25835.
- 2) Presented as a poster ‘Charge Models to reproduce Topography of Molecular Electrostatic Potential of few molecules’ at 2nd National Conference on “New Frontiers in Chemistry – From Fundamentals to Applications – II” organized by Department of Chemistry, BITS Pilani KK Birla Goa Campus on 28 – 29th January 2017 (Poster 33).

III.1 Introduction

The usefulness of point charge models despite their major drawback – unable to reproduce topographical features of Molecular Electrostatic Potential (MESP), has already been illustrated with examples in the introductory chapter of this thesis. Besides such type of charge models (containing only discrete point charges), the upcoming section briefly describes those containing Gaussian charge distributions.

The initial works of Hall and Smith were based on developing approximations for the electron density function using spherical Gaussians [1, 2]. The approximate electron density ρ^* was considered as a sum of product of positive charges $\{q_i\}$ and the associated Gaussian charge distributions $\{G_i\}$. The parameters optimized were the Gaussian exponents and Gaussian positions, subjected to the constraint that sum of charges should equal number of electrons present in the molecule under consideration.

The detailed discussion about the exact method of finding the optimized parameters is avoided in the present context and is available in the cited reference. Taking water as a test case, the Gaussians were placed at the atomic sites while in another set of models, additional Gaussians were placed at lone pair (LP) sites and sometimes between bonded atoms. With the electron density in hand, calculation of MESP was possible via. Eq. (I.49) where $\rho^*(\mathbf{r})$ takes the place of $\rho(\mathbf{r})$. Besides water (H_2O) molecule, ammonia (NH_3) and hydrogen fluoride molecules were also taken as test cases (fitting of electron density followed by calculation of MESP) [3, 4]. In practice however these models were found to increase computational efforts due to increase in the number of Gaussians. Another water model proposed by Hall and Smith [5], comprising of positive charges at nuclei sites and a single diffuse Gaussian

(accompanied with a negative charge) along C_2 axis, was unable to reproduce two negative (3, +3) critical points (CPs) corresponding to two LPs.

Gadre and Shrivastava [Chapter I, ref. 68] initially developed Gaussian charge models (GCMs) to reproduce topographical characteristics of NH_3 and H_2O molecules viz. LPs and a (3, +1) saddle connecting two MESP minima in the latter. They also developed similar topography-driven GCMs for molecules like methane, ethene (C_2H_4), methanol (CH_3OH), hydrogen sulphide (H_2S) and benzene (C_6H_6) using appropriate number of Gaussians by fitting $V(\mathbf{r})$, $\nabla V(\mathbf{r})$ and $\nabla^2 V(\mathbf{r})$ values at negative valued CPs [Chapter I, ref. 69, 70]. Several other research groups have developed Gaussian based models (not related to MESP topography) to tackle issues related to theory and experiments. Two such examples are mentioned in the upcoming section.

The water models developed by Jeon et al. [6], Paricaud et al. [7] and Kiss et al. [8] reproduce/simulate properties (such as diffusion constant, liquid-state interaction energies, radial distribution functions, energy profiles of water dimer, microstructure etc.) of liquid water at ambient temperatures. The thermodynamic/transport properties (viz. enthalpy of vaporization, self-diffusion coefficient, isobaric and isochoric heat capacities etc.) simulated by polarizable and nonpolarizable carbon dioxide GCMs are found to be in good agreement with experimental data [9]. The next section describes the development of topography-driven GCMs for the first set of molecules namely NH_3 , H_2O , H_2S , $HCHO$, C_2H_4 and C_6H_6 . Their molecular geometries have been optimized at the second order Møller-Plesset (MP2) Perturbation theory via. 6-31G(d,p) basis set using the software Gaussian 03 [Chapter I, ref 63]. The corresponding wave functions (WFs), geometries and basis set information were then employed to locate and characterize MESP CPs using the software INDPROP [Chapter I, ref. 67].

III.2 Methodology to develop Gaussian charge models (GCMs)

III.2.1 Ammonia molecule

Table III.1a summarises the optimized geometry of NH₃ molecule, co-ordinates of MESP CPs with corresponding MESP and electron density values. The ammonia GCM comprises of point charges {q_N, q_H}, placed at nitrogen and hydrogen sites. Due to symmetry, charge on all hydrogen atoms equals q_H. From the MESP contour plot displayed in Figure III.1a, a single (3, +3) CP corresponding to LP of electrons is found along the principle axis (Z-axis) at a distance of 2.2859 a.u. (1.2092 Å) from nitrogen atom. A single *s*-type normalized Gaussian (represented as $\varphi(\mathbf{R}_G) = \left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}} e^{-\alpha(\mathbf{r}-\mathbf{R}_G)^2}$ with negative charge q_G is therefore placed along the Z-axis at point \mathbf{R}_G bearing co-ordinates (R_X, R_Y, R_Z) to represent the continuous charge component (components R_X and R_Y vanish on Z-axis). Net charge balance is maintained by taking sum of positive point charges and charge on Gaussian equal to zero (charge on molecule). The potential at point \mathbf{r}_p due to point charges is computed via. Eq. (I.52) and the potential at the same point due to a single Gaussian $\varphi(\mathbf{R}_G)$ is discussed hereafter.

Product of two normalized Gaussians { $\varphi(\mathbf{R}_1)$, $\varphi(\mathbf{R}_2)$ } centred at { \mathbf{R}_1 , \mathbf{R}_2 } with normalization constants and exponents $\left\{\left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}}, \left(\frac{2\beta}{\pi}\right)^{\frac{3}{4}}\right\}$ and { α , β } result to a new Gaussian: $\left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}} \left(\frac{2\beta}{\pi}\right)^{\frac{3}{4}} e^{-\left(\frac{\alpha\beta}{\alpha+\beta}\right)(\mathbf{R}_1 - \mathbf{R}_2)^2} e^{-(\alpha+\beta)\mathbf{R}_{12}^2}$, where $\mathbf{R}_{12} = \frac{\alpha\mathbf{R}_1 + \beta\mathbf{R}_2}{\alpha+\beta}$ is the new Gaussian centre and $(\alpha+\beta)$ is the new exponent. The term $e^{-\left(\frac{\alpha\beta}{\alpha+\beta}\right)(\mathbf{R}_1 - \mathbf{R}_2)^2}$ is known as the pre-exponential factor and $V(\mathbf{r}_p)$ due to product of $\varphi(\mathbf{R}_1)$ with $\varphi(\mathbf{R}_2)$ is given by Eq. (III.1) [10]. The full derivation of Eq. (III.1) is available in APPENDIX, section A11.

Table III.1a. Topographical features of MESP obtained from ammonia WF at the MP2 level of theory employing 6-31G** basis set (refer APPENDIX, section A10, sub-section A10.1 for more details).

Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
	X	Y	Z		
(3, +3)	0.0000	0.0000	2.5075	-0.1365	0.024
(3, -1)	0.0000	0.1284	-0.2579	0.9960	0.361
	± 0.9772	-0.5642	-0.2579		
Molecular geometry: nitrogen at (0, 0, 0.2216) and hydrogen atoms at (0, 1.7652, -0.5170) and (± 1.5287 , -0.8826, -0.5170).					

Table III.1b. Topographical features of MESP obtained from ammonia GCMs. The symbol ‘---’ is used for electron density values less than 0.001 a.u. (in all tables through this chapter).

Model	Parameters	Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
	{ q_N, q_H, α, R_z }		X	Y	Z		
1	0.3232	(3, +3)	0.0000	0.0000	2.5076	-0.1362	0.024
	0.0048	(3, -1)	0.0000	1.5295	-0.4793	0.0978	---
	0.3113		± 1.3245	-0.7647	-0.4793		
1.9402 Charge on Gaussian: -0.3376, Gaussian centre: (0, 0, 1.9402).							
2	2.7336	(3, +3)	0.0000	0.0000	2.5097	-0.1372	0.024
	0.0696	(3, -1)	0.0000	1.4205	-0.4263	0.4146	0.045
	0.2489		± 1.2302	-0.7102	-0.4263		
0.4874 Charge on Gaussian: -2.9424, Gaussian centre: (0, 0, 0.4874).							
3	2.6146	(3, +3)	0.0000	0.0000	2.5075	-0.1366	0.024
	0.0593	(3, -1)	0.0000	1.4352	-0.4332	0.5154	0.041
	0.2479		± 1.2430	-0.7176	-0.4332		
0.5070 Charge on Gaussian: -2.7925, Gaussian centre: (0, 0, 0.5070), $\log(\omega)$ equals negative 4.7486.							
4	0.1665	(3, +3)	0.0000	0.0000	2.5074	-0.1363	0.027
	0.0238	(3, -1)	0.0000	1.2109	-0.4054	0.0994	---
	0.3850		± 1.0487	-0.6055	-0.4054		
2.1876 Charges q_N and q_H are derived from variable q whose value is 0.2379 (the former and latter are 0.7 and 0.1 times q as per Eq. (III.9)). Charge on Gaussian: -0.2379, Gaussian centre: (0, 0, 2.1876) and value of $\log(\omega)$ equals -3.0602.							

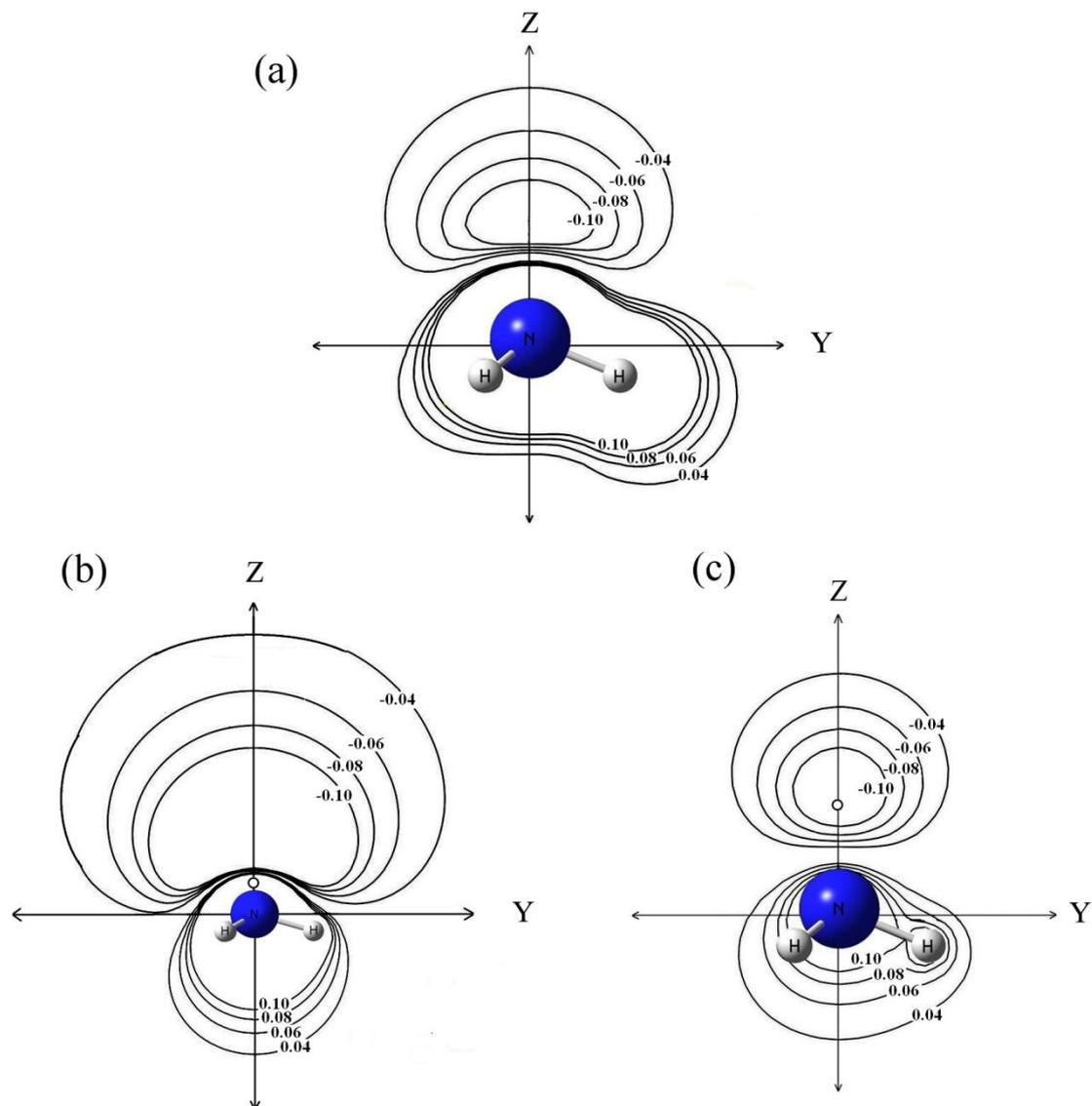


Figure III.1. MESP contour plot for NH₃ molecule in YZ plane obtained from a) MP2/6-31G(d,p) WF; b) GCM numbered 3 and c) GCM numbered 4 from Table III.1b. Position of Gaussian in GCM plots are shown by ‘o’ symbol. For convenience, the ball and stick model is used to represent point charges in GCMs. To save space, Fig. III.1b is resized.

$$V(\mathbf{r}_p) = \left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}} \left(\frac{2\beta}{\pi}\right)^{\frac{3}{4}} e^{-\left(\frac{\alpha\beta}{\alpha+\beta}\right)(\mathbf{R}_1-\mathbf{R}_2)^2} \int_0^1 e^{-(\alpha+\beta)(\mathbf{r}_p-\mathbf{R}_{12})^2 u^2} du \quad (\text{III.1})$$

The integral in above equation belongs to a special class of functions called *incomplete gamma functions* [$F_n(x) = \int_0^1 u^{2n} \exp(-xu^2) du$, for $x \geq 0$] and is represented

as $F_0(t)$ where $t = (\alpha + \beta)(\mathbf{r}_p - \mathbf{R}_{12})^2$ [11]. More details and properties of these functions are available in the cited reference.

A better representation for $V(\mathbf{r}_p)$ due to $\varphi(\mathbf{R}_G)$ is given by the left hand side expression of Eq. (III.2) while the value on the right hand side is obtained from Eq. (III.1) when $\varphi(\mathbf{R}_2)$ equals $\varphi(\mathbf{R}_1)$ (i.e. when both Gaussian centres are the same).

$$q_G \int \varphi(\mathbf{R}_G) \frac{1}{|\mathbf{r} - \mathbf{r}_p|} \varphi(\mathbf{R}_G) d\mathbf{r} = q_G \sqrt{\frac{8\alpha}{\pi}} F_0(t) \quad (\text{III.2})$$

In general, the potential at \mathbf{r}_p due to a GCM comprising of point charges $\{q_i\}$ and single Gaussian $\varphi(\mathbf{R}_G)$, represented as $V^M(\mathbf{r}_p)$ is evaluated via. Eq. (III.3).

$$V^M(\mathbf{r}_p) = \sum_{i=1}^N \frac{q_i}{|\mathbf{r}_p - \mathbf{R}_i|} + q_G \int \varphi(\mathbf{R}_G) \frac{1}{|\mathbf{r} - \mathbf{r}_p|} \varphi(\mathbf{R}_G) d\mathbf{r} \quad (\text{III.3})$$

Since the LP in NH_3 is capable of interacting with electrophilic species, the aim is to find GCM parameters $\{q_N, q_H, \alpha, R_Z\}$ – magnitude of point charges, Gaussian exponent and Gaussian position, which reproduce correct position of (3, +3) CP, corresponding $V(\mathbf{r})$ and $\rho(\mathbf{r})$ values (The quality of the GCM developed is judged on the reproduction of these quantities). This is achieved by making value of objective function ω defined by Eq. (III.4) less than some tolerance value ε via. subroutine STEPIT [Chapter II, ref. 5].

$$\omega = (V^{\text{WF}} - V^{\text{M}})^2 + (\rho^{\text{WF}} - \rho^{\text{M}})^2 + \nabla V^{\text{M}} \quad (\text{III.4})$$

The first and second terms of Eq. (III.4) ensure that the GCM reproduces correct $V(\mathbf{r})$ and $\rho(\mathbf{r})$ values at MESP minimum. [MESP and electron density functions are related by Poisson's equation: $\nabla^2 V^{\text{M}}(\mathbf{r}) = 4\pi\rho^{\text{M}}(\mathbf{r})$.] The third term in the above

equation defined by Eq. (III.5) ensures reproduction of correct position of (3, +3) CP by vanishing first order partial derivatives of Eq. (III.3).

$$\nabla V^M = \left| \frac{\partial V^M}{\partial x} \right| + \left| \frac{\partial V^M}{\partial y} \right| + \left| \frac{\partial V^M}{\partial z} \right| \quad (\text{III.5})$$

Several GCM parameters (also referred to as solutions (S_i)) are found to make value of ω less than $\varepsilon = 10^{-5}$ (tolerance value chosen in present study), namely:

$$S_1 \equiv \{0.0326, 0.0565, 0.3894, 2.3309\},$$

$$S_2 \equiv \{0.3232, 0.0048, 0.3113, 1.9402\},$$

$$S_3 \equiv \{0.3629, 0.0001, 0.3019, 1.8910\},$$

$$S_4 \equiv \{2.7336, 0.0696, 0.2489, 0.4874\} \text{ and others which are not mentioned here.}$$

Magnitude of q_N is expected to be greater than q_H . On this basis, solution S_1 is rejected. Solution S_3 is also rejected as value of q_H is small and meaningless. The topographical features obtained from NH_3 GCMs incorporating S_2 (model numbered 1) and S_4 (model numbered 2) parameters are displayed in Table III.1b. These models produce correct number of MESP CPs with slight deviations in the position of (3, +3) CP and corresponding $V(\mathbf{r})$ value. Nevertheless it is impossible to obtain a unique solution using this strategy.

Perhaps fitting MESP values at more points could generate a unique solution for ω . To test this hypothesis, points lying on a sphere of radius 0.1 a.u. surrounding the (3, +3) CP were generated using polar coordinates (\mathbf{r}, θ, ϕ). Since variables θ and ϕ take values (in radians) from 0 to π and 0 to 2π , a total of 381 points were obtained using increment values of 0.1571 and 0.3142 radians for θ and ϕ . MESP values at these points were computed using the software INDPROP. Further, a weight function $w(\mathbf{r})$ defined by Eq. (III.6) was employed so that MESP value at the minimum

got more weightage compared to the remaining points. For convenience, the MESP minimum is referred to \mathbf{r}_{+3} .

$$w(\mathbf{r}_i) = e^{-|V(\mathbf{r}_i) - V(\mathbf{r}_{+3})|} \quad (\text{III.6})$$

The new objective function ω to be minimized is defined by Eq. (III.7) while the best solution is suppose to make value of ω less than ϵ or in other words, make value of $\log(\omega)$ less than -5.

$$\omega = \left(\sum_{i=1}^{381+\mathbf{r}_{+3}} w(\mathbf{r}_i) (V_i^M - V_i^{WF})^2 \right) + (\rho^{WF} - \rho^M)^2 + \nabla V^M \quad (\text{III.7})$$

Lowest possible value of $\log(\omega)$ was found to be -4.7486 for GCM parameters {2.6146, 0.0593, 0.2479, 0.5070}. This is an acceptable solution as magnitude of q_N is greater than q_H and the corresponding GCM produces correct number and type of CPs (third model in Table III.1b). Another strategy which yields correct magnitude of charges on atoms and reduces the number of GCM parameters to be optimized is presented hereafter.

All point charges are derived from single variable 'q'. Secondly, these charges are assumed to be proportional to nuclear charges of corresponding atoms (due to similarities in the first term of Eq. (III.3) with its QM counterpart i.e. Eq. (I.49)). Thus q_i and Z_i are related via. $q_i = k Z_i$, where k is a proportionality constant. By taking summation on both sides of this equation, value of k is found to be q/Z_{sum} where Z_{sum} represents the sum of nuclear charges in a given molecule. The final relation between point charges and variable q is presented in Eq. (III.8).

$$q_i = \left(\frac{Z_i}{Z_{\text{sum}}} \right) q \quad (\text{III.8})$$

Due to restrictions on point charges (via. above equation), the GCM parameters are obtained by so called constraint-minimization of Eq. (III.7) (for purpose of convenience, this methodology shall be simply referred to as constraint-minimization); and lowest value of $\log(\omega)$ is about -3.0602 for GCM parameters: $\{q, \alpha, R_Z\} \equiv \{0.2379, 0.3850, 2.1876\}$. Using Eq. (III.8) the charges on nitrogen and hydrogen atoms come out to be 0.7 and 0.1 times variable q . The MESP CPs obtained from this model are displayed in Table III.1b (model numbered four).

Irrespective of the strategy used, all NH_3 GCMs displayed in Table III.1b are found to:

1. Produce bond CP between each bonded pair of atoms and a single (3, +3) CP along Z-axis.
2. Reproduce position of (3, +3) CP, corresponding $V(\mathbf{r})$ and $\rho(\mathbf{r})$ values with least errors.

Due to low magnitude of positive charges, the GCMs are unable to reproduce correct MESP values at bond CPs. From the MESP contour plots displayed in Figures III.1b and III.1c, the corresponding GCMs are found to produce positive and negative MESP regions which are in agreement with *ab initio* results (Figure III.1a). A smaller Gaussian exponent value results when the Gaussian centre is situated near the nitrogen atom. This is accompanied by large magnitudes of positive and negative charges as observed in models numbered 2 and 3 compared to 1 and 4 in Table III.1b. The MESP contour lines of model 3 in Fig. III.1b are also found to spread at a larger distance along positive Z-axis direction compared to WF and GCM (model 4) with large exponent value.

To conclude, the third GCM from Table III.1b is chosen to represent NH_3 molecule due to small deviation in MESP value at each bond CP compared to the remaining GCMs. The different strategies (using Eq. (III.4) and Eq. (III.7) with/without Eq. (III.8)) employed in developing the NH_3 GCM need to be tested for the remaining molecules.

III.2.2 Water and hydrogen sulphide molecules

From the MESP contour plot displayed in Figure III.2a, H_2O molecule is found to exhibit two (3, +3) CPs corresponding to LPs in the (YZ) plane perpendicular to the molecular (XY) plane. The co-ordinates of MESP minima and other CPs are displayed in Table III.2a.

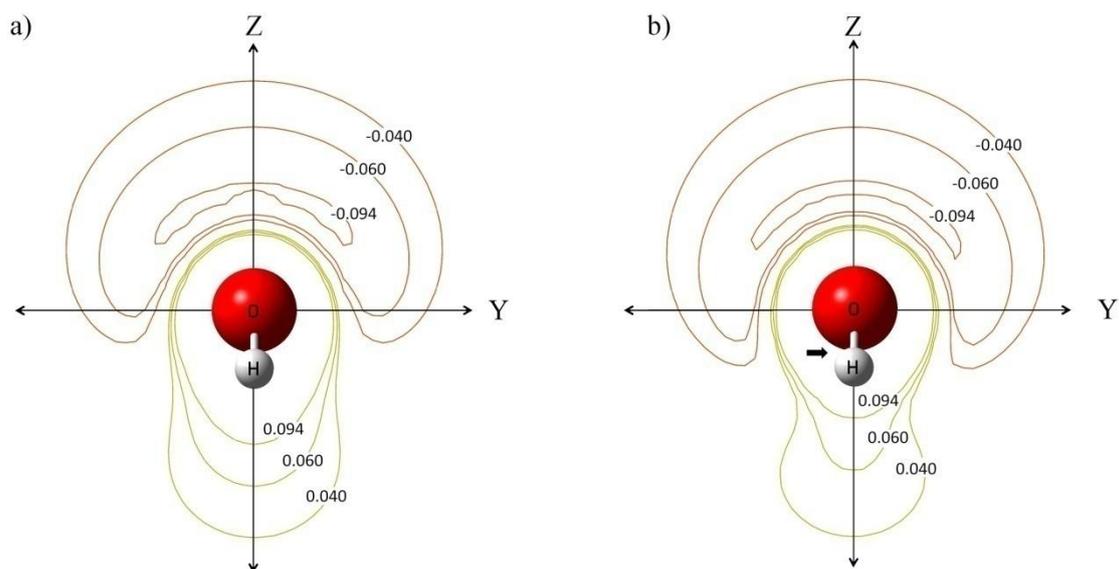


Figure III.2. MESP contour plot for H_2O molecule in YZ plane obtained from a) MP2/6-31G(d,p) WF and b) GCM containing single s -type Gaussian. Since the bonds connecting oxygen and hydrogen atoms are coinciding with Z-axis, the position of Gaussian along Z-axis is indicated by an arrow in Figure III.2b.

Table III.2a. Optimized geometry co-ordinates and topographical features of MESP obtained from water WF at the MP2/6-31G** level of theory (APPENDIX, section A10, sub-section A10.2 provides more details).

Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
	X	Y	Z		
(3, +3)	0.0000	± 1.9791	1.4066	-0.0989	0.017
(3, +1)	0.0000	0.0000	2.5146	-0.0966	0.015
(3, -1)	± 0.9114	0.0000	-0.4996	1.1318	0.423
Molecular geometry: oxygen at (0, 0, 0.2244) and hydrogen atoms at (± 1.4290 , 0, -0.8976).					

Table III.2b. Topographical features of MESP obtained from water GCMs. The first model is developed using two Gaussians in YZ plane and the second by a single Gaussian along Z-axis.

Model	GCM parameters		Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
				X	Y	Z		
1	q_O	0.2924	(3, +3)	0.0000	± 1.9791	1.4066	-0.0987	0.022
	q_H	0.0366						
	α	0.4281	(3, +1)	0.0000	0.0000	2.5834	-0.0405	0.001
	R_Y	1.6762						
	R_Z	1.1001	(3, -1)	± 0.9837	0.0000	-0.6223	0.1531	---
Charge on each Gaussian: -0.1828; Gaussian centres: (0, ± 1.6762 , 1.1001); value of $\log(\omega)$: -4.6014.								
2	q_O	4.6359	(3, +3)	0.0000	± 1.9791	1.4066	-0.0990	0.019
	q_H	0.5795						
	α	0.3137	(3, +1)	0.0000	0.0000	2.4963	-0.0982	0.017
	R_Z	0.1550	(3, -1)	± 0.9597	0.0000	-0.5293	1.0118	0.216
Charge on Gaussian: -5.7949; Gaussian centre: (0, 0, 0.1550); value of $\log(\omega)$ is about -3.6014. Variable q equals 5.7949, $q_O = 0.8q$ and $q_H = 0.1q$.								

It is possible to develop a GCM for H₂O by placing two Gaussians at (0, $\pm R_Y$, R_Z) and finding best set of GCM parameters: $\{q_O, q_H, \alpha, R_Y, R_Z\}$ which reproduce correct positions of MESP minima and corresponding V(**r**), $\rho(\mathbf{r})$ values (though a positive value of R_Y is optimized, the Gaussians are placed at appropriate sites). Due to the presence of symmetry, the two (3, +3) CPs are equidistant from the

oxygen atom and equivalent with respect to MESP and electron density values. It is therefore sufficient to develop the H₂O GCM by fitting $V(\mathbf{r})$, $\nabla V(\mathbf{r})$ and $\rho(\mathbf{r})$ at a single (3, +3) CP. Because of the symmetric nature of MESP minima, the Gaussians in the model bear same magnitude of negative charge $\{q_G$ is obtained by equally dividing the negative sum of positive charges $\}$.

The potential $V(\mathbf{r}_p)$ due to several (K) number of such Gaussians (all with equivalent charge q_G), in general is given by Eq. (III.9). Since interaction of $\varphi(\mathbf{R}_i)$ with $\varphi(\mathbf{R}_j)$ is equivalent to interaction of $\varphi(\mathbf{R}_j)$ with $\varphi(\mathbf{R}_i)$, the constant c_{ij} either takes value of 1 when Gaussian centres \mathbf{R}_i equals \mathbf{R}_j or 2 otherwise.

$$V(\mathbf{r}_p) = \sum_{i=1}^K \sum_{j=1}^i q_G c_{ij} \left[\int \varphi(\mathbf{R}_i) \frac{1}{|\mathbf{r} - \mathbf{r}_p|} \varphi(\mathbf{R}_j) d\mathbf{r} \right] \quad (\text{III.9})$$

The best double-Gaussian model was obtained by using GCM parameters from unconstrained minimization of Eq. (III.4). MESP CPs obtained from the current model are presented in Table III.2b. Despite the fact that $\rho(\mathbf{r})$ value is slightly overestimated, this model is able to reproduce exact positions of MESP minima and corresponding MESP values. It has been possible to develop a simpler model containing only one Gaussian placed along the C₂ axis and dependent on three parameters: $\{q, \alpha, R_Z\}$ by constraint-minimization method. Model parameters and MESP CPs obtained from this GCM are also presented in Table III.2b. The latter GCM is found to be a better representative for H₂O (additionally in terms of $V(\mathbf{r})$ values at (3, +1), (3, -1) CPs and $\rho(\mathbf{r})$ value at (3, +1) CP). It is also observed that the magnitude of point charges at the nuclei sites have become smaller by introducing two Gaussians in the water model. This results to poor $V(\mathbf{r})$ values at bond CPs (as observed in the former

model). Figure III.2b portrays the positive and negative MESP regions obtained from water GCM employing single Gaussian.

LPs in H₂S are located at a distance of 3.3297 a.u. (1.7614 Å) away from sulphur atom. Due to this large distance (compared to LP-oxygen distance of 1.2195 Å in H₂O), the H₂S GCM is developed using two Gaussians. The GCM parameters obtained via. constraint-minimization (model numbered 1) method and from its unrestricted counterpart (i.e. from minimizing Eq. (III.7) without use of Eq. (III.8) – henceforth referred to as unconstraint-minimization) i.e. model numbered 2 in Table III.2b along with corresponding MESP CPs are reported in Tables III.3b. Table III.3a summarizes the MESP CPs obtained from quantum mechanics.

Both models reproduce correct positions of MESP minima and corresponding V(**r**) values, but the latter is found to represent H₂S better – in terms of ρ(**r**) values at (3, +3) CPs, MESP values at (3, -1), (3, +3) CPs and position of saddle connecting minima. The MESP contour plot of H₂S is similar to H₂O, hence not shown.

Table III.3a. Optimized geometry and topographical features of MESP obtained from hydrogen sulphide WF at the MP2/6-31G** level of theory (more insights available in APPENDIX, section A10, sub-section A10.3).

Type of CP	Co-ordinates			V(r)	ρ(r)
	X	Y	Z		
(3, +3)	0.0000	±3.1446	1.3326	-0.0449	0.006
(3, +1)	0.0000	0.0000	4.1283	-0.0190	0.001
(3, -1)	±1.2349	0.0000	-0.9912	0.8229	0.228
Molecular geometry: sulphur along Z-axis at Z = 0.1925 and hydrogen atoms in XZ plane at (±1.8218, 0, -1.5397).					

Table III.3b. MESP topography obtained from hydrogen sulphide GCMs.

Model	Parameters { $q_S, q_H, \alpha, R_X, R_Z$ }	Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
			X	Y	Z		
1	0.2082	(3, +3)	0.0000	± 3.1447	1.3315	-0.0446	0.001
	0.0130						
	0.0662	(3, +1)	0.0000	0.0000	3.4541	-0.0418	---
	1.5532						
	0.3118	(3, -1)	± 1.3032	0.0000	-1.1499	0.0004	0.003
Charge on each Gaussian: -0.11725; Gaussian centres: (0, ± 1.5532 , 0.3118); value of $\log(\omega)$ is -5.1684. Value of q equals 0.2345, $q_S = 0.8889q$ and $q_H = 0.0556q$.							
2	0.0780	(3, +3)	0.0000	± 3.1446	1.3326	-0.0449	0.006
	0.0564						
	0.2596	(3, +1)	0.0000	0.0000	4.0137	-0.0081	---
	2.8651						
	1.1188	(3, -1)	± 0.8555	0.0000	-0.8286	0.0712	---
Charge on each Gaussian: -0.0954; Gaussian centres: (0, ± 2.8651 , 1.1188); $\log(\omega)$ is about -3.2112.							

III.2.3 Formaldehyde molecule

MESP topographical features of HCHO molecule are displayed in Table III.4a. There are two (3, +3) CPs in the molecular (YZ) plane, located at a distance of 2.3192 a.u. (1.2269 Å) from the carbonyl oxygen, each forming an angle of 126.9° with respect to oxygen and carbon atoms. The corresponding GCM is thus developed by placing two *s*-type Gaussians in the YZ plane so as to produce LPs of carbonyl oxygen. As in the case of H₂S molecule, two models for HCHO have been developed using the constraint and unconstraint minimization methods. MESP CPs from both models and corresponding GCM parameters are summarized in Table III.4b.

With respect to positions of MESP minima and corresponding V(**r**) values, both models presented in Table III.4b are in good agreement with WF counterpart, but the electron density value at (3, +3) CPs is better represented by the latter (i.e. the unconstraint parameter model). The V(**r**) values at bond CPs however are comparatively

better in the former (constraint parameter) model and hence this model is chosen to represent HCHO. The MESP contour plots from WF and representative model are displayed in Figures III.3a and III.3b respectively.

Table III.4a. Optimized geometry and topographical features of MESP obtained from formaldehyde WF at the MP2/6-31G** level of theory (refer APPENDIX, section A10, sub-section A10.4 for further details).

Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
	X	Y	Z		
(3, +3)	0.0000	± 1.8535	2.6858	-0.0801	0.014
(3, +1)	0.0000	0.0000	3.7630	-0.0696	0.007
(3, -1)	0.0000	± 1.1236	-1.7253	0.8650	0.239
	0.0000	0.0000	0.1448	1.3339	0.486

Molecule geometry: oxygen and carbon atoms on Z-axis at Z = 1.2918 and Z = -1.0145, hydrogen atoms at (0, ± 1.7589 , -1.5397).

Table III.4b. Topographical features of MESP obtained from formaldehyde GCMs.

GCM parameters { $q_O, q_C, q_H, \alpha, R_Y, R_Z$ }	Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
		X	Y	Z		
0.1365 0.1024 0.0171 0.2567 1.7691 2.1832	(3, +3)	0.0000	± 1.8535	2.6880	-0.0800	0.008
	(3, +1)	0.0000	0.0000	3.3516	-0.0648	0.004
	(3, -1)	0.0000	± 1.2250	-1.8460	0.0700	---
		0.0000	0.0000	0.1855	0.0927	0.001

Charge on each Gaussian: -0.1365; Gaussian centres: (0, ± 1.7691 , 2.1832); $\log(\omega)$ value equals -2.3348, Variable $q = 0.2730$, $q_O = 0.5q$, $q_C = 0.3750q$ and q_H is 0.0625 times variable q .

0.2571 0.0264 0.0053 0.3597 1.5881 2.2492	(3, +3)	0.0000	± 1.8533	2.6857	-0.0802	0.014
	(3, +1)	0.0000	0.0000	3.5073	-0.0524	0.003
	(3, -1)	0.0000	± 1.3156	-1.8420	0.0281	---
		0.0000	0.0000	-0.3517	0.0856	---

Charge on each Gaussian: -0.1471; Gaussian centres: (0, ± 1.5881 , 2.2492); $\log(\omega)$ value equals -4.1017.

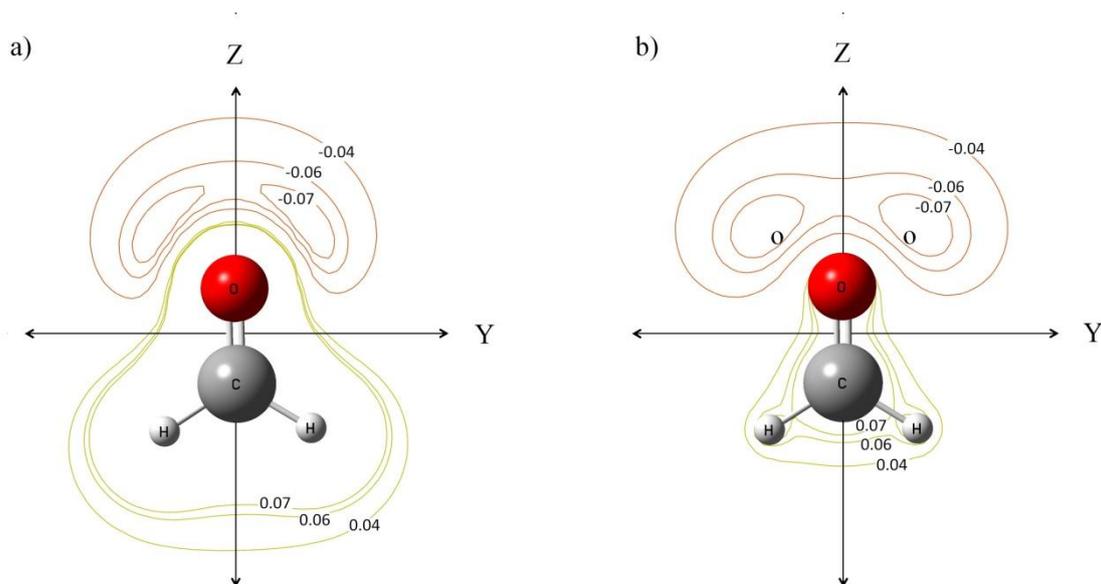


Figure III.3. MESP contour plot of HCHO obtained from a) MP2/6-31G** WF and b) constraint-parameter GCM. The positions of *s*-type Gaussians in GCM plot are shown by ‘o’ symbols.

III.2.4 Ethene and acetylene molecules

In addition to (3, -1) CPs between each C–H and C=C bonds, the MESP topography of ethene molecule is found to exhibit two (3, +3) CPs corresponding to the π bond character present in its molecular framework (Table III.5a). These minima are situated above and below the double bond at a distance of 1.5261 Å from the molecular plane. The corresponding GCMs developed via constraint/unconstraint minimization methods are found to reproduce desired topographical features. Due to insignificant difference between both results, only the former are presented in Table III.5b. The MESP contour plot obtained from this model is found to be in good agreement with its QM counterpart (Figure III.4). Efforts were also made to simplify ethylene GCM by use of single *s*-type Gaussian, placed at the centre of mass of C₂H₄. Irrespective of the method employed to derive GCM parameters, the negative MESP regions (due to (3, +3) CPs) situated above and below the C=C bond were found to be connected by

negative valued (3, +1) CPs as shown in Figure III.5. The model presented in Table III.6 portrays this feature. Compared to the previous model (in Table III.5b), the $V(\mathbf{r})$ and $\rho(\mathbf{r})$ values at bond CPs are found to better.

By representing the electron density function of acetylene (C_2H_2) molecule by single s -type Gaussian, the corresponding charge model is found to replicate positive MESP region near nuclear framework and negative MESP region in the form a ring around the molecular axis (Figure III.6). The C_2H_2 model parameters (obtained from optimal fit of $V(\mathbf{r})$, $\nabla V(\mathbf{r})$ and $\rho(\mathbf{r})$ values at any one degenerate ring CP) are displayed in Table III.7b while MESP topography from WF in Table III.7a

Table III.5a. MESP topographical features of ethene obtained from MP2/6-31G** WF. Section A10, sub-section A10.5 in APPENDIX describes the guess points (GPs) used to obtain corresponding CPs.

Type of CP	Co-ordinates			$V(\mathbf{r})$	$\rho(\mathbf{r})$
	X	Y	Z		
(3, +3)	0.0000	0.0000	± 2.8849	-0.0393	0.007
(3, -1)	0.0000	0.0000	0.0000	0.9835	0.352
	± 1.1123	± 1.9441	0.0000	0.8525	0.068
Molecular geometry: carbon and hydrogen atoms are placed at (0, ± 1.2617 , 0) and (± 1.7419 , ± 2.3293 , 0).					

Table III.5b. MESP topographical features obtained from C_2H_4 GCM.

GCM parameters		Type of CP	Co-ordinates			$V(\mathbf{r})$	$\rho(\mathbf{r})$
			X	Y	Z		
q_C	0.0753	(3, +3)	0.0000	0.0000	± 2.8845	-0.0395	0.007
q_H	0.0065	(3, -1)	0.0000	0.0000	0.0000	0.0548	0.001
α	0.3074		± 1.3172	± 2.0876	0.0000	0.0357	---
R_Z	2.5054	Charge on each Gaussian: -0.0884; Gaussian centres: (0, 0, ± 2.5054); $\log(\omega)$ value is less than -5. Variable q equals 0.2009, $q_C = 0.3750q$ and $q_H = 0.0325q$.					

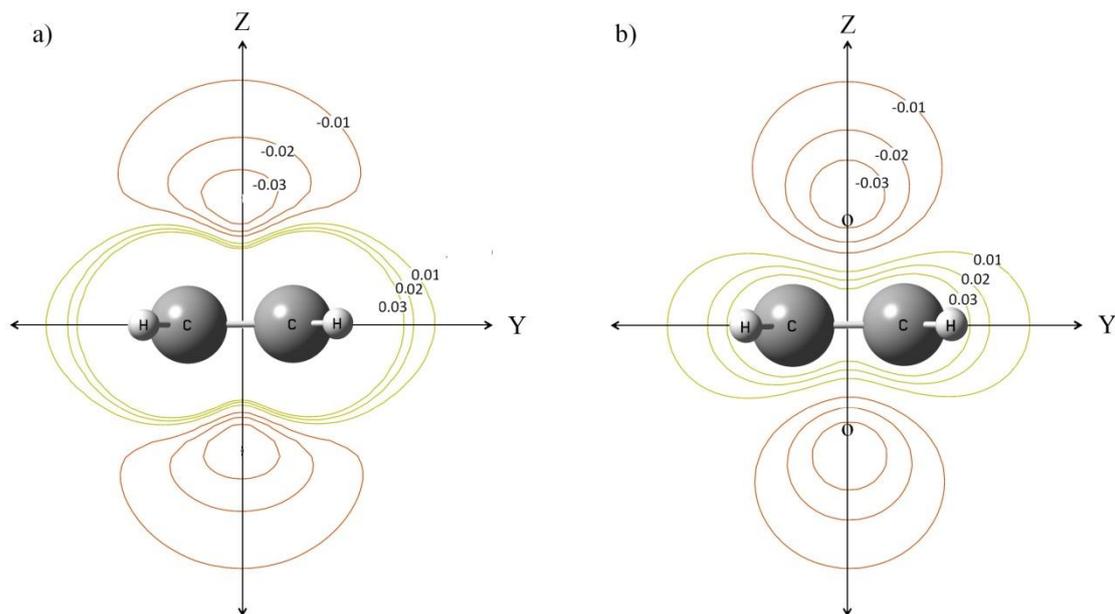


Figure III.4. MESP contour plot of C_2H_4 obtained from a) MP2/6-31G** WF & b) GCM. The symbols 'o' denote positions of *s*-type Gaussians used to represent electron density in ethylene GCM.

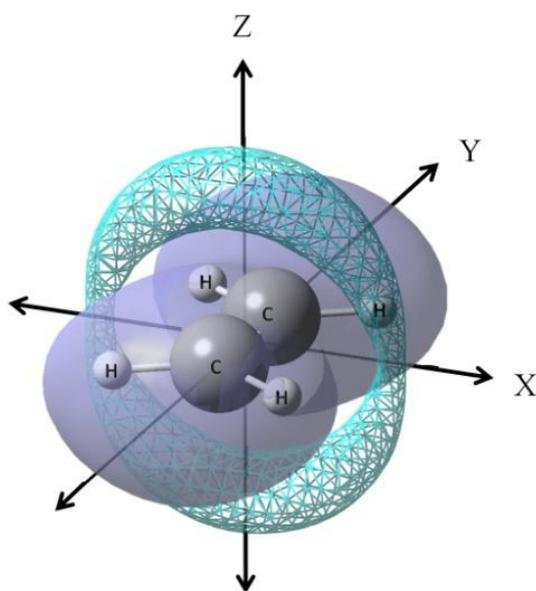


Figure III.5. MESP isosurface plot obtained from C_2H_4 GCM containing single Gaussian. The transparent solid/mesh surfaces (of isovalues ± 0.030) correspond to the positive and negative MESP regions. This plot is generated using GCM parameters from Table III.6.

Table III.6. MESP topographical features of ethene obtained from GCM containing single *s*-type Gaussian.

GCM parameters { q_C, q_H, α }	Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
		X	Y	Z		
0.9466	(3, +3)	0.0000	0.0000	± 2.8847	-0.0408	0.006
0.0502	(3, +1)	± 2.9237	0.0000	0.0000	-0.0318	0.005
0.1175	(3, -1)	0.0000	0.0000	0.0000	0.4241	0.043
		± 1.3362	± 2.1492	0.0000	0.2393	0.010

Charge on Gaussian: -2.0941; Gaussian centres: (0, 0, 0); $\log(\omega) = -3.8469$. The model parameters are derived by unconstraint minimization of Eq. (III.5).

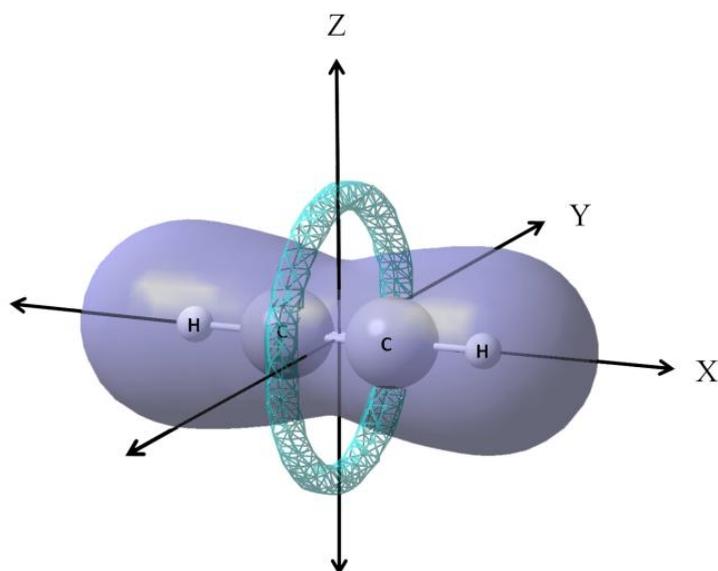


Figure III.6. MESP isosurface plot obtained from C_2H_2 GCM whose parameters are displayed in Table III.7b. The transparent solid/mesh surfaces (of isovalues ± 0.038) correspond to the positive and negative MESP regions.

Table III.7a. MESP topography of acetylene derived from MP2/6-31G** WF.

Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
	X	Y	Z		
(2, 0)	0.0000	1.8080	2.2098	-0.0405	0.007
(3, -1)	0.0000	0.0000	0.0000	1.3331	0.352
	± 2.4310	0.0000	0.0000	0.9472	0.068

Molecular geometry: carbon & hydrogen atoms along X-axis at $X = \pm 1.1514$ and $X = \pm 3.1595$.

Table III.7b. MESP topographical features of C₂H₂ derived from corresponding GCM.

GCM parameters { q_C, q_H, α }	Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
		X	Y	Z		
0.9466	(2, 0)	0.0000	0.0000	± 2.8847	-0.0408	0.006
0.0502	(3, -1)	0.0000	0.0000	0.0000	0.4497	0.039
0.1175		± 2.5103	0.0000	0.0000	0.3580	0.009

Charge on Gaussian: -2.0941; Gaussian centre: (0, 0, 0); $\log(\omega) = -3.8469$. The model parameters are derived by unconstraint minimization of Eq. (III.8).

III.2.5 Benzene molecule

Benzene, the simplest aromatic hydrocarbon is found to exhibit rich topographical features. In addition to twelve bond CPs and one (3, +1) ring CP, six (3, +3) CPs are located above and below the molecular plane of benzene, circling the C₆ axis a distance of 3.4899 a. u. (1.8462 Å) from the centre of mass of C₆H₆. A negative (3, +1) saddle is found to connect two nearest minima. There are negative valued (3, -1) CPs on the C₆ axis at a distance of 3.3322 a. u. (1.7627 Å) from the ring. Co-ordinates of mentioned CPs along with corresponding V(**r**) and $\rho(\mathbf{r})$ are displayed in Table III.8a.

The π cloud in C₆H₆ GCM is represented by a total of twelve *s*-type Gaussians, wherein six Gaussians are placed in (XY) planes sandwiching the aromatic ring. These Gaussians have the same X, Y co-ordinates as MESP minima while the Z co-ordinate is optimized with other GCM parameters to obtain optimal fit at one (3, +3) CP. The best GCM in terms of positions of MESP minima and corresponding V(**r**) values is obtained by unconstraint minimization method and presented in Table III.8b.

Table III.8a. MESP topographical features of benzene obtained from MP2/6-31G** WF (GPs used to obtain CPs are listed in APPENDIX, section A10, subsection A10.6).

Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
	X	Y	Z		
(3, +3)	± 0.6391	± 1.1093	± 3.2466	-0.0332	0.0023
	± 1.2799	0.0000	± 3.2467		
(3, +1)	0.0000	± 1.2058	± 3.2586	-0.0331	0.0021
	± 1.0446	± 0.6015	± 3.2587		
	0.0000	0.0000	0.0000	0.1507	0.0197
(3, -1)	0.0000	0.0000	± 3.3322	-0.0321	0.0010
	± 1.1402	± 1.9748	0.0000	0.8630	0.3209
	± 2.2803	0.0000	0.0000		
	± 3.4170	± 1.9728	0.0000	0.8513	0.2928
	0.0000	± 3.3956	0.0000		

Molecular geometry: carbon atoms at (0, ± 2.6386 , 0) and (± 2.2851 , ± 1.3193 , 0) while hydrogen atoms are situated at (0, ± 4.6846 , 0) and (± 4.0570 , ± 2.3423 , 0).

Table III.8b. MESP topographical features from benzene GCM.

GCM parameters		Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
			X	Y	Z		
q_C	0.0119	(3, +3)	± 0.5665	± 0.9856	± 3.2457	-0.0339	0.0169
			± 1.1365	0.0000	± 3.2457		
q_H	0.0020	(3, +1)	0.0000	± 0.9673	± 3.2580	-0.0312	0.0055
			± 0.8358	± 0.4826	± 3.2582		
α	3.0000	(3, -1)	0.0000	0.0000	± 3.3103	-0.0271	---
			± 2.3746	0.0000	± 3.3103	0.0087	---
± 1.1872	± 2.0564		0.0000				
R_Z	3.2201		± 3.5351	± 2.0410	0.0000	0.0071	---
			0.0000	± 4.0820	0.0000		

Charge on each Gaussian: -0.00695; Gaussian centres: (± 0.6391 , ± 1.1093 , ± 3.2201) and (± 1.2799 , 0, ± 3.2201); $\log(\omega) = -1.6718$.

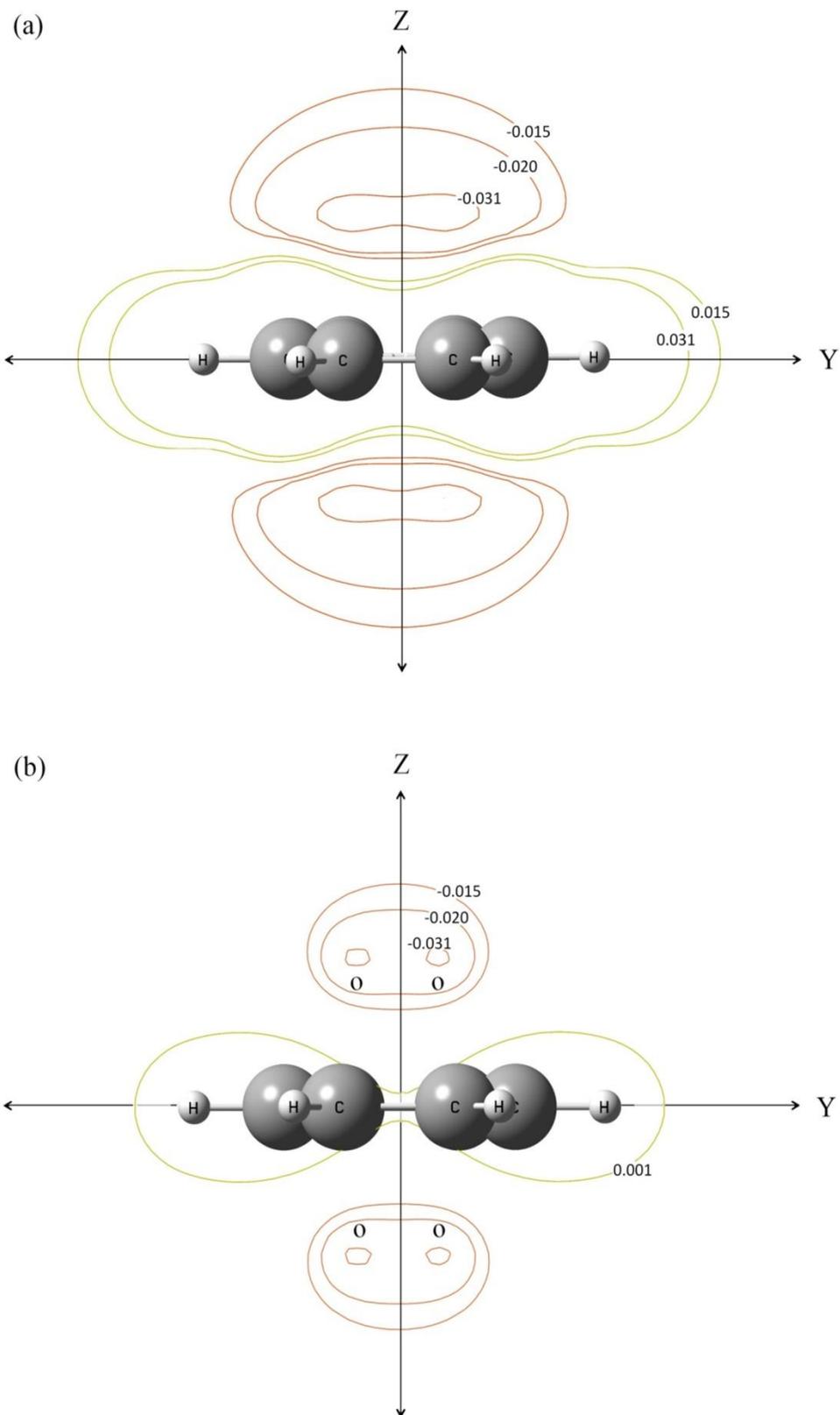


Figure III.7. MESP contour plot of C_6H_6 obtained in the YZ plane using a) WF and b) unconstrained-parameter Gaussian model. The symbols 'o' in GCM plot denote positions of s-type Gaussians.

Electron density values at (3, +3) CPs are found to be overestimated by the model. A larger Gaussian exponent value (typically $\alpha \geq 3$) ensures that the negative MESP regions located above and below the nuclear framework don't intersect/meet each other at the centre of mass (i.e. the MESP value at the ring CP is found to be negative if α takes values less than 3). MESP contour plot (Figure III.7b) obtained from the four-parameter model is found to be in good agreement with its QM counterpart (Figure III.7a).

III.3 Scope for improving MESP and Electron Density values at critical points other than MESP minima

It is observed that the GCMs poorly reproduce $V(\mathbf{r})$ and $\rho(\mathbf{r})$ values at CPs other than MESP minima, primarily at the bond CPs. Though these GCMs produce correct number and type of MESP CPs, the study presented in this section will aid in answering the question: 'Is it possible to develop a GCM which will correctly reproduce MESP and electron density values at CPs other than MESP minima?'

Taking water molecule as a test case, the aim was to find GCM parameters which make value of ω defined by Eq. (III.4) less than ϵ at (3, +3), (3, +1) and (3, -1) CPs. The lowest possible value of ω was found to be 0.3881 for GCM parameters: $\{q_O, q_H, \alpha, R_Z\} = \{1.5760, 0.4104, 0.1513, 0.1406\}$. Since value of ω is greater than ϵ , it is obvious that the current model is not capable of reproducing exact positions of MESP CPs, corresponding MESP and electron density values. [This model was found to produce only three CPs – a single MESP minimum along Z-axis at $Z = 3.1208$ in addition to two bond CPs – which is unacceptable.]

In an attempt to improve the MESP and electron density values at bond CPs, a single Gaussian (with exponent α') was placed at each bond CP site. Eq. (III.5) was then subjected towards minimization by fitting $V(\mathbf{r})$, $\nabla V(\mathbf{r})$ and $\rho(\mathbf{r})$ values at (3, +3), and (3, -1) CPs. Lowest possible value of ω was found to be 0.2219 for parameters: $\{q_O, q_H, \alpha, R_Z, \alpha'\} = \{7.1138, 0.8892, 0.4280, 0.1716, 9000\}$. The charge on Gaussian along C_2 axis is about -8.8922 while the MESP topographical features obtained from this model are displayed in Table III.9.

Table III.9. MESP topography obtained from water GCM consisting additional Gaussians at bond CPs.

CP type		Co-ordinates			V(\mathbf{r})	$\rho(\mathbf{r})$
		X	Y	Z		
Expected	(3, +3)	0.0000	± 1.7385	1.0870	-0.2362	0.056
	(3, +1)	0.0000	0.0000	2.1029	-0.2353	0.052
	(3, -1)	± 0.9910	0.0000	-0.4045	0.9771	0.411
Unexpected	(3, +1)	± 0.8406	0.0000	-0.5960	0.9690	0.417
		0.0000	0.0000	-1.6505	0.0873	0.074
	(3, -1)	0.0000	0.0000	-2.4230	0.0471	0.001

It is seen that the model fails to reproduce correct positions of MESP minima, corresponding $V(\mathbf{r})$ and $\rho(\mathbf{r})$ values though there is improvement in the electron density values at (3, -1) CPs. Additional (unexpected) CPs (as shown in Figure III.8) are found to be present in the MESP topography of water which do not have any significance.

This model is unworthy to represent water molecule. Hence it has not possible to improve the GCM (especially in terms of MESP and electron density values at bond CPs) using the strategies presented in this section.

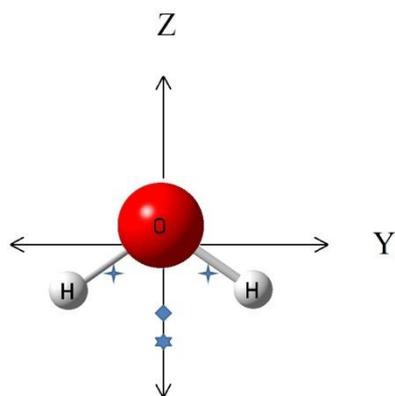


Figure III.8. Position of additional CPs in MESP topography of water obtained from GCM containing additional Gaussian at each bond CP. Symbols ★, ◆ and ★ are used to represent additional (3, +1) CPs with MESP values 0.9690, 0.0873 and (3, -1) CP along Z-axis.

III.4 General remarks on GCMs

The essential topographical features of MESP are found to be restored when charge models are incorporated with *s*-type Gaussians. As the number of Gaussians increase in corresponding model, the magnitude of positive/negative charges located at atomic/Gaussian sites are found to decrease, the Gaussian exponent values become larger and the model produces low MESP values at bond CPs.

The different strategies used to develop GCMs were based on one objective: to obtain a unique and acceptable solution (i.e. GCM parameters wherein different types of atoms are correctly represented by different magnitudes of point charges) for objective function ω . This however has been possible by using Eq. (III.8) or in other words, by using a restraint on point charges.

The models for H₂O, H₂S, HCHO, C₂H₄ and C₆H₆ are developed primarily by fitting position of only one (3, +3) CP along with corresponding MESP and electron density values because these molecules possess symmetry and more importantly, all

MESP minima are equivalent – identical with respect to $V(\mathbf{r})$ and $\rho(\mathbf{r})$ values. Such types of molecules are classified into Group-1 category. The upcoming discussions will further build-up this classification by employing different test molecules.

III.5 Additional topography based Gaussian models

III.5.1 Methanol molecule

The negative MESP region in CH_3OH is due to the presence of two LPs located at a distance of 2.2496 a.u. (1.1904 Å) from oxygen atom which are connected by a (3, +1) saddle as shown in Figure III.9. Value of angles α , β and ϕ are 128, 120.4 and 41.5° . The (3, +1) saddle is located at a distance of 2.2815 a.u. (1.2069 Å) from oxygen atom. The MESP critical points (displayed in Table III.10a) are obtained via computational chemistry package DAMQT [12] using optimized MP2/6-31G** WF. Since MESP and electron density values at (3, +3) CPs are equivalent, methanol appears to fall into group **1** category.

There is however a need to differentiate among the hydrogen atoms connecting oxygen and carbon atoms as they are not related via symmetry. [Such situation was absent in GCMs presented earlier.] The group **1** category is hence split into: **1s** and **1a** where ‘s’ and ‘a’ indicate symmetry/asymmetry among different types of atoms. Methanol falls into **1a** group due to the presence of asymmetry among the hydrogen atoms (while NH_3 , H_2O , H_2S , HCHO , C_2H_4 , C_2H_2 and C_6H_6 fall into **1s** category).

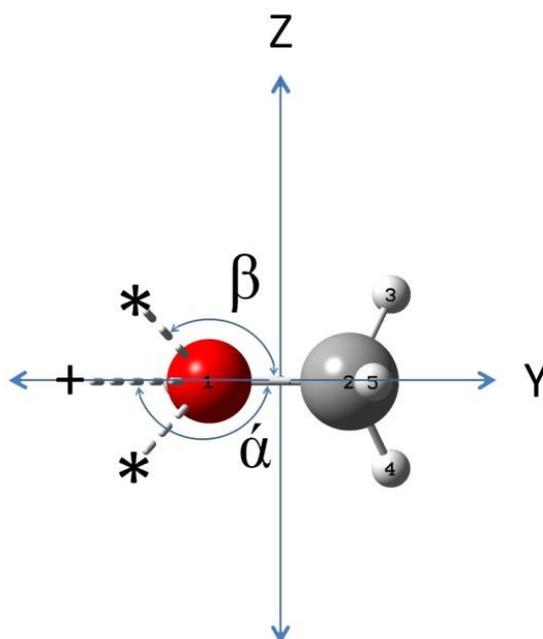


Figure III.9. Location of LPs (denoted by ‘*’ symbols) in methanol molecule and saddle (indicated by ‘+’) connecting them. Oxygen atom (labelled 1), carbon atom (labelled 2) lie on Y axis, methyl hydrogens are labelled from numbers 3 to 5 and the hydroxyl hydrogen (labelled 6) is behind oxygen, ‘trans’ with respect to H5 – hence not observed. The angle formed by the (3, +1) saddle and (3, +3) CP with the Y axis are denoted as α and β . Dihedral angle ϕ is the angle between (3, +3) with plane containing atoms O1, C2 and H5.

The methyl and hydroxy hydrogen charges in the corresponding GCM are represented as $\{q_{\text{H}}^{\text{Me}}, q_{\text{H}}^{\text{OH}}\}$ respectively. Eq. (III.8) cannot differentiate hydrogen atoms belonging to different groups. Hence the methanol GCM depends on 8 parameters: $\{q_{\text{O}}, q_{\text{C}}, q_{\text{H}}^{\text{Me}}, q_{\text{H}}^{\text{OH}}, \alpha, R_{\text{X}}, R_{\text{Y}}, R_{\text{Z}}\}$ which are displayed in Table III.10b along with corresponding MESP CPs. They are found by minimizing Eq. (III.4) at a single (3, +3) CP. The MESP/electron density values at (3, +3) CPs and positions of MESP minima are reproduced with slight deviations from expected values (values of β and ϕ are 121.7 and 37.9°; distance from oxygen 2.1798 a.u. (1.1531 Å)). The (3, +1) saddle is found at a distance of 2.0007 a.u. (1.0584 Å) from oxygen (value of α equals 128.3°).

Table III.10a. MESP topographical features of methanol from MP2/6-31G** WF.

Type of CP	Co-ordinates			V(r)	$\rho(r)$
	X	Y	Z		
(3, +3)	1.5406	-2.5724	± 1.2864	-0.0957	0.017
(3, +1)	1.8873	-2.8366	0.0000	-0.0941	0.015
(3, -1)	-1.0262	-1.7735	0.0000	1.1230	0.423
	1.3478	1.6250	0.0000	0.8395	0.313
	-0.4950	1.7570	± 1.0724	0.8293	0.296
	0.0842	-0.1237	0.0000	0.8013	0.291
Molecular geometry: Oxygen atom situated at (0.0887, -1.4332, 0), carbon atom at (0.0887, 1.2537, 0) while hydrogen atoms of methyl groups at (-0.8247, 2.0392, ± 1.6777) and (2.5069, 1.8369, 0), hydroxy hydrogen at (-1.6495, -1.9722, 0).					

Table III.10b. MESP topographical features from methanol GCM.

GCM parameters		Type of CP	Co-ordinates			V(r)	$\rho(r)$
			X	Y	Z		
q_o	0.0805	(3, +3)	1.5510	-2.5806	± 1.1385	-0.0965	0.022
q_c	0.0509	(3, +1)	1.6598	-2.6722	0.0000	-0.0831	0.009
q_H^{Me}	0.0056						
q_H^{OH}	0.0159	(3, -1)	-0.1259	-0.0743	0.0000	0.0498	---
α	0.6647		-1.1184	-1.6590	0.0000	0.0490	---
R_x	1.4203		-0.6344	1.8481	± 1.2511	0.0376	---
R_y	-2.4638		1.5771	1.7591	0.0000	0.0322	---
R_z	1.2826						
Charge on each Gaussian: -0.0821; Gaussian centres: (1.4203, -2.4638, ± 1.2826); value of $\log(\omega) = -1.8268$.							

III.5.2 Pyrazine molecule

The second aromatic compound, pyrazine –unlike benzene is devoid of π cloud character in its MESP topography but prominently shows negative MESP regions near the nitrogen atoms due to presence of lone pairs. These regions are connected by negative valued (3, +1) saddles located 5.0676 a.u. (2.6808 Å) above and below the molecular plane, as portrayed in Figure III.10 (while the MESP CPs are displayed in Table III.11a). The challenge is to develop a GCM which is capable of reproducing this feature.

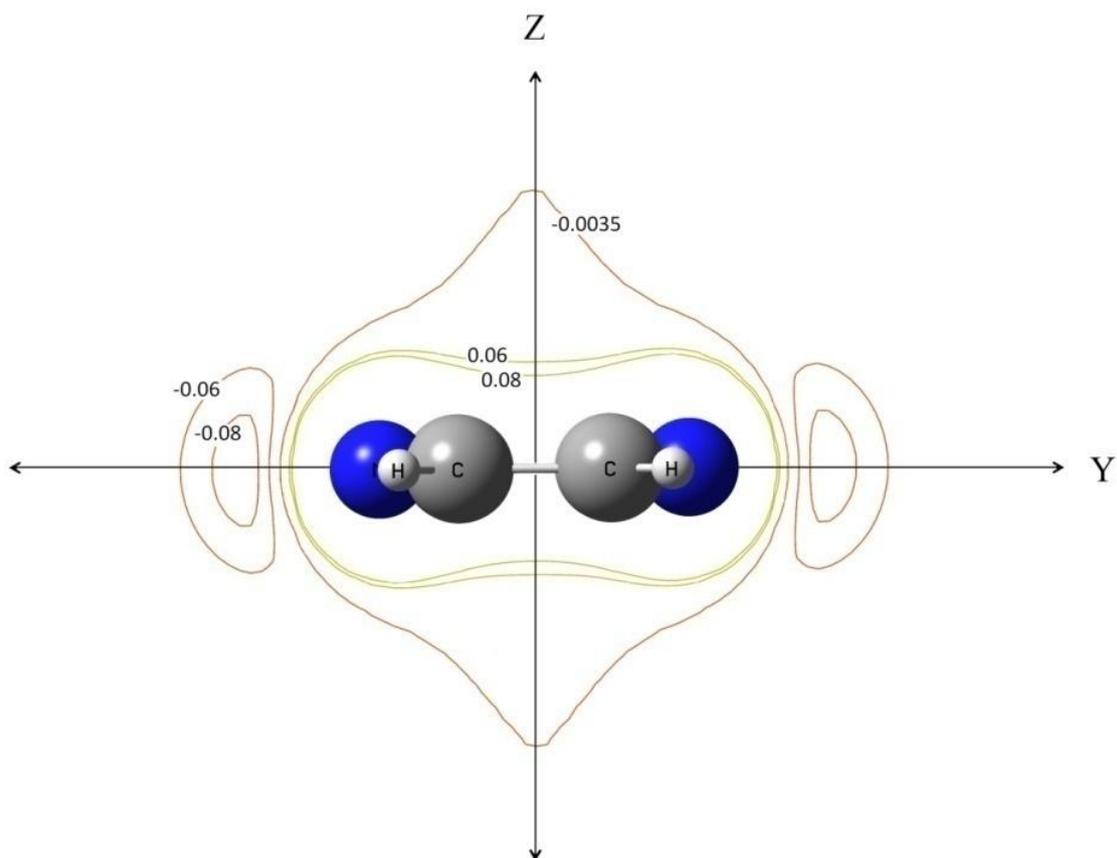


Figure III.10. MESP contour plot of pyrazine molecule in YZ plane obtained from optimized MP2/6-31G** WF.

Table III.11a. MESP topographical features of pyrazine from MP2/6-31G** WF.

Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
	X	Y	Z		
(3, +3)	0.0000	± 5.0353	0.0000	-0.0939	0.019
(3, +1)	0.0000	0.0000	± 5.0676	-0.0035	---
	0.0000	0.0000	0.0000	0.2105	0.023
(3, -1)	± 1.0688	± 2.0028	0.0000	1.0122	0.372
	± 2.1433	0.0000	0.0000	0.9052	0.328
	± 3.2677	± 1.9907	0.0000	0.8743	0.298
Molecular geometry: nitrogen atoms on Y-axis at $Y = \pm 2.6789$, carbons at $(\pm 2.1453, \pm 1.3183, 0)$ and hydrogen atoms at $(\pm 3.9026, \pm 2.3691, 0)$. The software DAMQT is employed for locating MESP topographical features.					

Table III.11b. MESP topographical features from pyrazine GCMs.

Model	GCM parameters { $q_N, q_C, q_H, \alpha, R_Y$ }	Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
			X	Y	Z		
1	0.0466 0.0400 0.0067 0.4033 4.8527	(3, +3)	0.0000	± 5.0353	0.0000	-0.0938	0.018
		(3, +1)	0.0000	0.0000	0.0000	0.0465	---
		(3, -1)	± 2.0145	0.0000	0.0000	0.0598	
			± 0.7853	± 1.6931	0.0000	0.0527	
			± 3.4878	± 2.0119	0.0000	0.0380	
Charge on each Gaussian: -0.1399; Gaussian centres (0, ± 4.8527 , 0); value of $\log(\omega)$ is -8.1819. Value of q equals 0.2798, $q_N = 0.1667q$, $q_C = 0.1429q$ and $q_H = 0.0238q$.							
2	0.2005 0.1802 0.0301 0.0986 3.2242	(3, +3)	± 5.0353	0.0000	0.0000	-0.0935	0.005
		(3, +1)	0.0000	0.0000	± 3.5747	-0.0256	---
			0.0000	0.0000	0.0000	0.0134	0.005
		3, -1)	± 2.2759	0.0000	0.0000	0.1242	0.002
			± 3.4453	± 1.9547	0.0000	0.0852	0.001
± 0.9511	± 1.9387		0.0000	0.0733	0.007		
Charge on each Gaussian: -0.6211; Gaussian centres: (0, ± 3.2242 , 0); $\log(\omega)$ value equals -2.5827.							

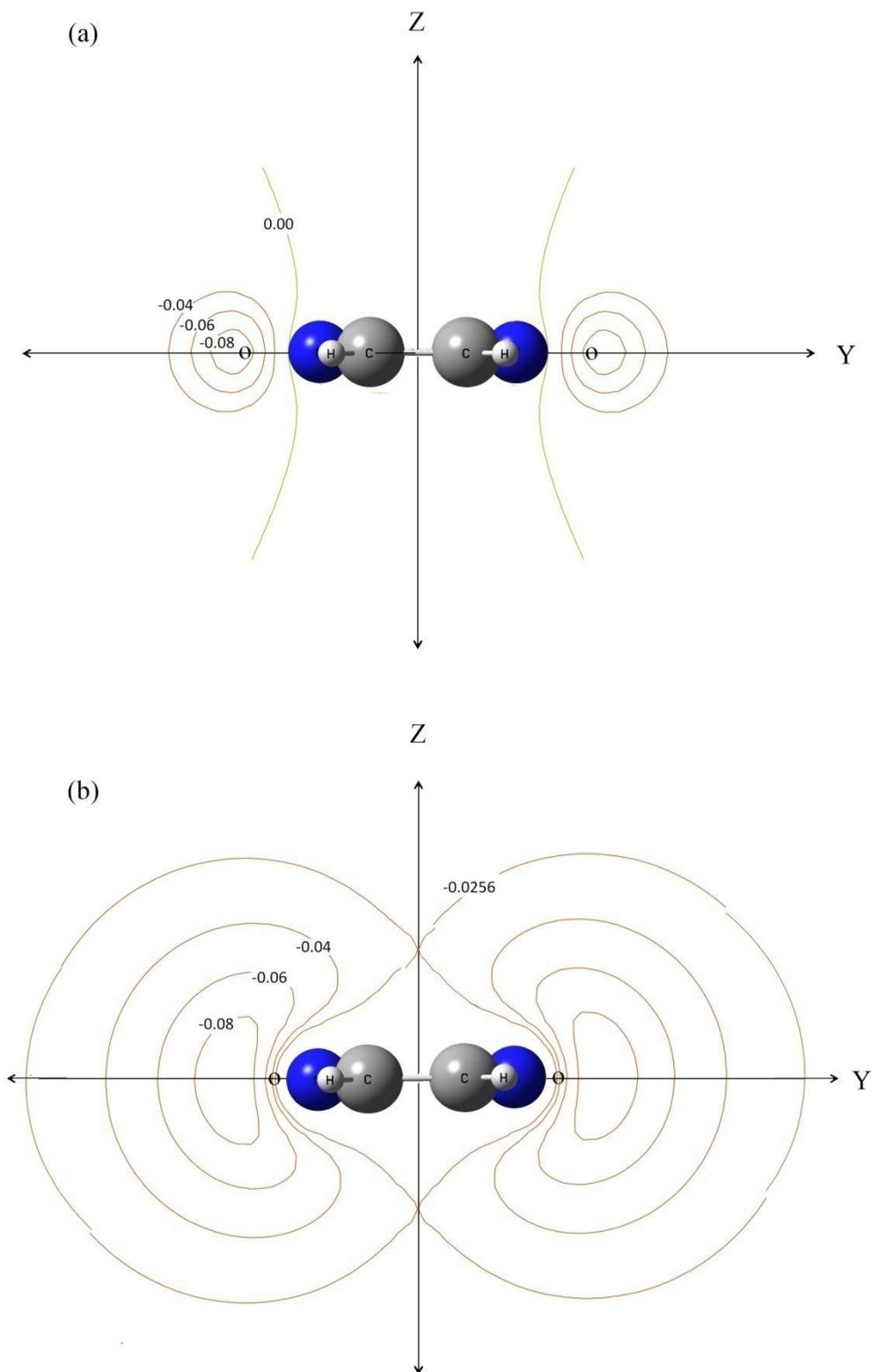


Figure III.11. MESP contour plot in YZ plane obtained from pyrazine GCMs: a) constraint-parameter model with Gaussian exponent value 0.4033 and b) unconstraint-parameter model with Gaussian exponent 0.0986.

Two Gaussians are placed in the pyrazine GCM along Y-axis between $Y = \pm 2.6789$ (nitrogen atoms) and $Y = \pm 5.0353$ (expected MESP minima positions) so as to find best GCM parameters: $\{q_N, q_C, q_H, \alpha, R_Y\}$. Though the first model presented in Table III.11b produces positions of MESP minima, corresponding $V(\mathbf{r})$ and $\rho(\mathbf{r})$ values that are in good agreement with WF counterparts, the two negative regions are found to be disconnected as observed from corresponding MESP contour plot shown in Figure III.11a.

It could be possible to connect the two regions by placing Gaussians closer towards nitrogen atoms, with smaller exponent values (less than 0.4033) and larger magnitudes of positive point charges. With these constraints, the best model (second model in Table III.11b) for pyrazine was successfully developed from unconstrained, partial minimization of Eq. (III.5) i.e. fitting $V(r)$ and $\nabla V(\mathbf{r})$ at (3, +3) and negative valued (3, +1) saddle. Though the electron density value at MESP minima is found to be poorly reproduced, the connection between the negative MESP regions has been restored at a shorter distance of 3.5747 a.u. (1.8910 Å) from the molecular plane by negative (3, +1) saddles (as seen from Figure III.11b).

III. 6 Future prospects

Another molecule which (presumptively) falls in **1a** category is acetone, as the carbon atoms belonging to carbonyl and methyl groups need to be differentiated in the model though the negative MESP region in acetone (as described in the earlier chapter) is similar to formaldehyde. Molecules like oxetane and formic acid possess different types of MESP (3, +3) CPs: the former exhibits LPs and π bonds while the latter possesses LPs belonging to different functional groups (carbonyl and alcohol). They

are hence classified into a separate group – **2**, which is further divided into **2s** and **2a**. The former molecule falls into the **2s** category as the different types of atoms are symmetry related. Formic acid on the other hand does not possess symmetry among different atoms; hence it belongs to the **2a** group. The different strategies presented in this chapter need to be tested for such molecules.

III.7 Concluding remarks

From the study presented in this chapter, it is seen that the GCMs reproduce MESP minima and other essential topographical features. This has been possible since the ‘continuous charge component’ is represented by several *s*-type Gaussians.

It was easier to develop GCMs for NH₃, H₂O, H₂S, HCHO, C₂H₄, C₂H₂ and C₆H₆ molecules as the different types of atoms present in respective molecular frameworks were symmetry related and the respective negative MESP regions were attributed to either LPs or π bonds, which once again were equivalent due to symmetry. As seen in section III.5, methanol contains different non-symmetry related hydrogen atoms which in turn increase the number of GCM parameters. In any case this work needs to be carried forward – especially for molecules possessing different (or a combination of various) functional groups (like amides, acids etc.) which play a key role in biological processes.

References:

1. a) C.M. Smith and G.G. Hall, The approximation of electron densities, *Theor. Chim. Acta.*, 69 (1986) 63.
2. G.G. Hall and C.M. Smith, The electron density of the water molecule, *Theor. Chim. Acta.*, 69 (1986) 71.
3. G.G. Hall and C.M. Smith, Fitting Electron Densities of Molecules, *Int. J. of Quantum Chem.*, 25 (1984) 881.
4. G.G. Hall and K. Tsujinaga, The molecular electrostatic potential of some simple molecules, *Theo. Chim. Acta.*, 69 (1986) 425.
5. G.G. Hall and C.M. Smith, Electrostatic Water Models, *Int. J. Quantum Chem.*, 42 (1992) 1237.
6. J. Jeon, A.E. Lefohn and G.A. Voth, An improved Polarplex water model, *J. Chem. Phys.*, 118 (2003) 7504.
7. P. Paricaud, M. Předota, A.A. Chialvo and P.T. Cummings, From dimer to condensed phases at extreme conditions: Accurate predictions of the properties of water by a Gaussian charge polarizable model, *J. Chem. Phys.*, 121 (2004) 244511.
8. P.T. Kiss and A. Baranyai, A systematic development of a polarizable potential of water, *J. Chem. Phys.*, 138 (2013) 204507.
9. H. Jiang, O.A. Moulton, I.G. Economou and A.Z. Panagiotopoulos, Gaussian-Charge Polarizable and Nonpolarizable Models for CO₂, *J. Phys. Chem. B*, 120 (2016) 984.
10. S.R. Gadre and P.K. Bhadane, *Electrostatics in Chemistry*, 1. Basic Principles, *Resonance*, 4 (1999) 11.

11. T. Helgakerin, *Modern Electronic Structure Theory, Part-2, Chapter 12*, p.809–815, Ed. D.R. Yarkony, World Scientific Publishing Co. Pte. Ltd., Singapore (1995).
12. A. Kumar, S.D. Yeole, S.R. Gadre, R. López, J.F. Rico, G. Ramírez, I. Ema and D. Zorrilla, DAMQT 2.1.0: A New Version of the DAMQT Package Enabled with the Topographical Analysis of Electron Density and Electrostatic Potential in Molecules, *J. Comput. Chem.*, 36 (2015) 2350.

CHAPTER IV

Investigation of Potential Energy Surfaces via. Gaussian Charge Models

	Table of Contents	Page
IV.1	Introduction	99
IV.2	Computation of interaction energy of a binary complex	101
IV.2.1	Wave function approach	101
IV.2.2	Using GCMs	103
IV.3	Exploration of potential surfaces via. GCMs	104
IV.3.1	Water dimer	104
IV.3.2	Ammonia dimer	110
IV.3.3	Ammonia-water complex	115
IV.3.4	Acetylene dimer and acetylene-water complex	116
IV.3.5	Benzene dimer and benzene related systems	117
IV.4	Concluding remarks	121
	References	123

Part of this work has been:

- 1) Published in the research article titled ‘Prelude to molecular dynamics: Topography-driven Gaussian charge models’, Int. J. Quantum Chem., DOI:10.1002/qua.25835.
- 2) Presented as a poster ‘Prelude to Molecular Dynamics – II: Guidelines for Employing Topography-driven Gaussian Charge Models’ at the 16th Theoretical Chemistry Symposium (TCS – 2019) organized by Department of Chemistry BITS Pilani, Pilani Campus on 13 – 16th February 2019 (Poster No. P177).

IV.1 Introduction

Interaction between atoms and molecules either lead to the formation of new molecules or a molecular cluster. Two hydrogen atoms for example, form H₂ molecule wherein the atoms are separated by a distance of $\sim 0.75 \text{ \AA}$ with dissociation energy, D_e of about 103 kcal/mol [1]. The same hydrogen atom when bonded to a *sp*-hybridized carbon is found to have D_e of ~ 126 kcal/mol [2]. Such strong interactions between atoms are referred to ‘covalent interactions’ and they arise due to appreciable overlap of charge distributions in bonded atoms at distances normally less than 2 \AA . Very weak interactions viz. those observed between argon atoms in argon dimer (with D_e of about -1.13 kcal/mol [3]) are referred to ‘non-covalent’ or van der Waals (vdW) interactions. These interactions when recognized by J. D. van der Waals in the last century led to the reformulation of the equation of state for real gases [4].

Strength of interactions between hydrogen and electronegative atoms like oxygen, nitrogen, sulphur or π -charge (associated with ethylene double bonds, aromatic compounds etc.) generally belonging to different molecules, lie between the above mentioned categories. The interacting atoms are separated by distances of $2\text{--}3 \text{ \AA}$ with no substantial overlap between electron distributions, consequently preventing any structural changes. K. Morokuma from his theoretical investigation described five major factors which play a key role in understanding origin of non-covalent interactions [5]. Assuming that two molecules A and B are capable of forming a binary complex A...B, the five components are defined hereafter.

The *Electrostatic* component $\{E_{es}\}$ defined by Eq. (IV.1) is a result of interaction between undistorted charge distributions of molecules A and B. Terms in Eq. (IV.1): γ_A and γ_B are respective charge distributions, inclusive of nuclear and electronic

contributions (\mathbf{r}_1 and \mathbf{r}_2 define positions of electrons, \mathbf{R}_A and \mathbf{R}_B are those of nuclei with nuclear charges Z_A and Z_B). Eq. (IV.1) can be expressed in terms of Molecular Electrostatic Potential (MESP) i.e. Eq. (IV.2) due to either molecule A, $V_A(\mathbf{r}_1, \mathbf{R}_A)$ or B, $V_B(\mathbf{r}_2, \mathbf{R}_B)$; as expressed in Eqs. (IV.3) and (IV.4).

$$E_{ES} = \iint \frac{\gamma_A(\mathbf{r}_1, \mathbf{R}_A) \gamma_B(\mathbf{r}_2, \mathbf{R}_B)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad (IV.1)$$

$$E_{ES} = \int \frac{\gamma_A(\mathbf{r}_1, \mathbf{R}_A) V_B(\mathbf{r}_2, \mathbf{R}_B)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 = \int \frac{\gamma_B(\mathbf{r}_2, \mathbf{R}_B) V_A(\mathbf{r}_1, \mathbf{R}_A)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_2 \quad (IV.2)$$

$$V_A(\mathbf{r}_1, \mathbf{R}_A) = \int \frac{\gamma_A(\mathbf{r}_1, \mathbf{R}_A)}{|\mathbf{r} - \mathbf{r}_1|} d\mathbf{r} = \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} - \int \frac{\rho_A(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|} d\mathbf{r}_1 \quad (IV.3)$$

$$V_B(\mathbf{r}_2, \mathbf{R}_B) = \int \frac{\gamma_B(\mathbf{r}_2, \mathbf{R}_B)}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{r} = \sum_B \frac{Z_B}{|\mathbf{r} - \mathbf{R}_B|} - \int \frac{\rho_B(\mathbf{r}_2)}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{r}_2 \quad (IV.4)$$

The second component i.e. the *polarization* interaction is the effect of distortion (polarization) of electronic distribution of molecule A by B and vice versa. Considering a simple case where a single point charge of magnitude Q is located at \mathbf{r} in the vicinity of molecule A, the polarization energy designated as $P(\mathbf{r})$ by Bonaccorsi *et al.* relies on the self-consistent-field (SCF) calculation of the entire system i.e. molecule A + charge Q [6]. The expression for $P(\mathbf{r})$ is hence given by Eq. (IV.5) where E_A is the energy of isolated molecule A and $E_1(\mathbf{r})$ is the energy obtained after reoptimizing the wave function (WF) of molecule A under the field of point charge Q . The computational cost makes the use of Eq. (IV.5) impractical.

$$P(\mathbf{r}) = E_1(\mathbf{r}) - E_A - V_A(\mathbf{r}_1, \mathbf{R}_A) \quad (IV.5)$$

The *exchange repulsion* {EX} is the interaction caused by exchange of electrons between molecules A and B while *charge transfer* {CT} or electron delocalization interaction is the interaction caused by CT from occupied molecular orbitals (MOs) of A to vacant MOs of B and from occupied MOs of B to vacant MOs of A. The last term, *coupling* is the difference between total interaction energy (IE) of the complex and sum of the above four components E_{ES} , $P(\mathbf{r})$, EX and CT. A detailed explanation on each of these terms is available in the cited reference [5] and therefore avoided in the present context. Based on the decomposition fragments of the total IE, it was possible for Morokuma to study several donor-acceptor type of complexes (viz. water dimer, water-hydrogen fluoride etc.) and concluded that the E_{ES} component plays a major role in complexes possessing linear X-H-Y hydrogen bonds (where X, Y are electronegative atoms like nitrogen, oxygen and fluorine).

Given a pair of molecules capable of forming a weak complex, one would be interested in knowing the strength of interaction and geometry of complex formed, prior to the IE decomposition analysis outlined earlier. The Electrostatic Potential for Intermolecular complexation (EPIC) model has been used to carry out such studies (Chapter I, ref. 47 to 49). Before venturing into the study of potential energy surfaces (PESs) of simple molecular systems using topography-based Gaussian charge models (GCMs) which is the main purpose of this chapter, the upcoming section briefly describes the formulae involved in calculating IEs using WF and GCMs.

IV.2 Computation of interaction energy of binary complex

IV.2.1 Wave function approach

Using quantum mechanical (QM) methods, IE of a binary complex A...B is computed via. Eq. (IV.6), where E_{AB} corresponds to the energy of complex, E_A and E_B

are energies of individual molecules A and B while {AB}, {A} and {B} represent their respective geometries [7].

$$IE = E_{AB}\{AB\} - (E_A\{A\} + E_B\{B\}) \quad (IV.6)$$

At infinite separation distance between molecules A and B in A...B, the IE can be written as a sum of E_A and E_B since overlap between basis functions of A with B (and vice versa) vanish. As intermolecular distance between molecules A and B reach to some minimum value, A...B is found to be extra stabilized as fragment A utilizes extra basis functions from fragment B to describe its electronic distribution and vice versa. The error arising due to this situation is termed as the Basis Set Superposition Error (BSSE) which results in the IE being more negative. If A and B represent the basis sets used for monomers A and B, it is more appropriate to replace $E_A\{A\}$ and $E_B\{B\}$ in Eq. (IV.6) by $E_A^A\{A\}$ and $E_B^B\{B\}$. Similarly $E_{AB}\{AB\}$ is represented by $E_{AB}^{AUB}\{AB\}$ to signify its evaluation using basis sets A and B. Eq. (IV.6) is therefore rewritten as Eq. (IV.7) which needs to be rectified from BSSE.

$$IE = E_{AB}^{AUB}\{AB\} - E_A^A\{A\} - E_B^B\{B\} \quad (IV.7)$$

This is achieved by estimating the amounts (E_A^{BSSE} and E_B^{BSSE}) by which the fragments A and B are stabilised by extra basis functions of B and A, given in Eqs. (IV.8) and (IV.9) respectively.

$$E_A^{BSSE} = E_{AB}^{AUB}\{AB\} - E_A^A\{A\} \quad (IV.8)$$

$$E_B^{BSSE} = E_{AB}^{AUB}\{AB\} - E_B^B\{B\} \quad (IV.9)$$

Subtracting Eqs. (IV.8) and (IV.9) from Eq. (IV.7) results to Eq. (IV.10) where the IE free from BSSE is referred to IE_{corr} . This method was originally proposed by Boys and Bernardi [8] which is known as the counterpoise-correction method.

$$\mathbb{I}E_{\text{corr}} = E_{\text{AB}}^{\text{AUB}}\{\text{AB}\} - (E_{\text{A}}^{\text{AUB}}\{\text{AB}\} + E_{\text{B}}^{\text{AUB}}\{\text{AB}\}) \quad (\text{IV.10})$$

IV.2.2 Using GCMs

The nuclei of molecules A, B in respective GCMs are represented by M_1, N_1 number of point charges: $\{q_{A_i}\}, \{q_{B_j}\}$ situated at $\mathbf{R}_{A_i}, \mathbf{R}_{B_j}$ sites. The intermolecular interaction between them is repulsive and the energy $E_{\text{pc-pc}}$ (due to point charge-point charge interactions) is given by Eq. (IV.11); which is equivalent to the sum of q_{A_i} times the potential at \mathbf{R}_{A_i} sites due to point charges of GCM B (as demonstrated by the second equality in Eq. (IV.11)).

$$E_{\text{pc-pc}} = \sum_{i=1}^{M_1} \sum_{j=1}^{N_1} \frac{q_{A_i} q_{B_j}}{|\mathbf{R}_{A_i} - \mathbf{R}_{B_j}|} = \sum_{i=1}^{M_1} q_{A_i} \left(\sum_{j=1}^{N_1} \frac{q_{B_j}}{|\mathbf{R}_{A_i} - \mathbf{R}_{B_j}|} \right) \quad (\text{IV.11})$$

The electron density functions $\rho(\mathbf{r}_A), \rho(\mathbf{r}_B)$ in respective GCMs are represented by M_2, N_2 number of s -type Gaussians $\phi_k^A(\mathbf{r}_A), \phi_m^B(\mathbf{r}_B)$ situated at $\mathbf{R}_k^A, \mathbf{R}_m^B$ with exponents $\{\alpha_k^A, \alpha_m^B\}$ and charges $\{q_{A_k}^G, q_{B_m}^G\}$. Since interaction between $\rho(\mathbf{r}_A)$ and $\rho(\mathbf{r}_B)$ is also repulsive, the corresponding energy $E_{\text{gc-gc}}$ is given by Eq. (IV.12) and evaluated as per Eq. (IV.13) [9]. The terms C_1, C_2, β (reduced exponent) and \mathbf{R} are defined thereafter.

$$E_{\text{gc-gc}} = \iint \frac{\rho(\mathbf{r}_A)\rho(\mathbf{r}_B)}{|\mathbf{r}_A - \mathbf{r}_B|} d\mathbf{r}_A d\mathbf{r}_B \quad (\text{IV.12})$$

$$\iint \frac{\rho(\mathbf{r}_A)\rho(\mathbf{r}_B)}{|\mathbf{r}_A - \mathbf{r}_B|} d\mathbf{r}_A d\mathbf{r}_B = \sum_{k=1}^{M_2} \sum_{l=1}^{M_2} \sum_{m=1}^{N_2} \sum_{n=1}^{N_2} q_{A_k}^G q_{B_m}^G C_1 C_2 \int_0^1 e^{-\beta \mathbf{R}^2 u^2} du \quad (\text{IV.13})$$

$$C_1 = \left(\frac{2\alpha_k^A}{\pi} \right)^{\frac{3}{4}} \left(\frac{2\alpha_l^A}{\pi} \right)^{\frac{3}{4}} \left(\frac{2\alpha_m^B}{\pi} \right)^{\frac{3}{4}} \left(\frac{2\alpha_n^B}{\pi} \right)^{\frac{3}{4}} \left(\frac{2\pi^{\frac{5}{2}}}{\gamma_1 \gamma_2 (\gamma_1 + \gamma_2)^{0.5}} \right)$$

$$C_2 = \exp \left[\left(\frac{-\alpha_k^A \alpha_l^A}{\alpha_k^A + \alpha_l^A} \right) \mathbf{R}_{kl}^2 + \left(\frac{-\alpha_m^B \alpha_n^B}{\alpha_m^B + \alpha_n^B} \right) \mathbf{R}_{mn}^2 \right]$$

- $\gamma_1 = \alpha_k^A + \alpha_l^A$; $\gamma_2 = \alpha_m^B + \alpha_n^B$; $\beta = \gamma_1 \gamma_2 / (\gamma_1 + \gamma_2)$.
- The summations in Eq. (IV.13) are over four indices: k, l, m and n; analogues to the two-electron repulsion integral in quantum mechanics. The former two indices {k, l} are employed when dealing with spherical Gaussians of molecule A (i.e. k, l take values from 1 to M_2) and likewise the latter two indices {m, n} are employed when dealing with Gaussians of molecule B (m, n take values from 1 to N_2).
- \mathbf{R} is the separation between the two new Gaussian centers \mathbf{R}_{kl} and \mathbf{R}_{mn} formed due to products of Gaussians situated at \mathbf{R}_k^A and \mathbf{R}_m^B with those having centers \mathbf{R}_l^A and \mathbf{R}_n^B respectively.

The attractive interaction energy between point charges and Gaussian charge distributions of respective GCMs (represented as $E_{\text{pc-gc}}$) is given by Eq. (IV.14). The IE calculated by interaction between two GCMs, IE_{GCM} is therefore given by Eq. (IV.15).

$$E_{\text{pc-gc}} = \sum_{i=1}^{M_1} q_{A_i} \int \frac{\rho(\mathbf{r}_B)}{|\mathbf{R}_{A_i} - \mathbf{r}_B|} d\mathbf{r}_B + \sum_{j=1}^{N_1} q_{B_j} \int \frac{\rho(\mathbf{r}_A)}{|\mathbf{R}_{B_j} - \mathbf{r}_A|} d\mathbf{r}_A \quad (\text{IV.14})$$

$$IE_{\text{GCM}} = E_{\text{pc-pc}} + E_{\text{pc-gc}} + E_{\text{gc-gc}} \quad (\text{IV.15})$$

IV.3. Exploration of potential energy surfaces via. GCMs

IV.3.1 Water dimer

A few stationary points (geometries labelled from S1 to S6) along the PES of water dimer $(\text{H}_2\text{O})_2$ with corresponding BSSE uncorrected/corrected IEs are

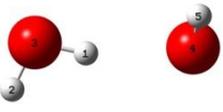
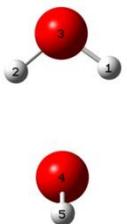
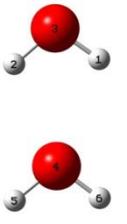
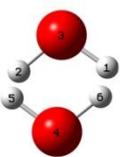
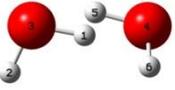
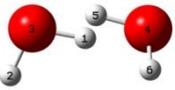
displayed in Table IV.1 respectively. The overall PES exhibits one global minimum (structure S1), wherein one molecule acts as a proton donor while the other is an acceptor.

The study using topography-driven water GCM was performed in three parts: first, investigating whether correct minimum energy structure of $(\text{H}_2\text{O})_2$ was predicted among structures S1 to S6. These studies were performed using water GCM consisting single s -type Gaussian. The IE_{GCM} values calculated via. Eq. (IV.10) for structures S1 to S6 (by appropriately orienting water GCMs as shown in Figure IV.1) are also summarized in Table IV.1. It is seen that the GCMs predicts lower energy for structure S1.

The second study involved monitoring the IE trend as one GCM was translated towards/away from the other in structures S1 to S6. Figure IV.2 depicts the plots of IEs (computed via. WF and GCM approach) as the $R_{\text{O-O}}$ distance is varied from 2.25 to 3.75 Å by step length of 0.25 Å. It is expected that the IE drops (becomes more negative) as the two water molecules approach each other (due to attractive forces operating at larger separation distances) followed by an increase in the IE values (less negative) due to repulsive forces being dominant at shorter distances.

This behaviour is observed when IE values are computed via. WF (curves in Figure IV.2 shown by continuous lines) while the attractive forces are found to dominate in the GCM approach (i.e. IE_{GCM} values are found to become more negative which is depicted by dotted curves in Figure IV.2). The GCMs are unable to replicate IE trend at shorter distances as they comprise of only positive point charges and Gaussian charge distributions associated with negative charges, and energy calculations don't involve orbitals, electron-nuclear attraction/electron-electron repulsion integrals, etc.

Table IV.1. Stationary structures along PES of (H₂O)₂ obtained using second order Møller-Plesset (MP2) Perturbation theory and 6-31G(d,p) basis set. Geometries represented in terms of internal co-ordinates along with respective point group in parenthesis.

Orientation	Internal coordinates (bond lengths in Å while angles and dihedral angles in degrees)	Interaction Energy (kcal/mol)		
		IE	IE _{corr}	IE _{GCM}
S1  (C _s)	2 1 1.5202 3 2 0.9604 1 38.1 4 3 2.9121 2 113.8 1 0.0 5 4 0.9634 3 95.6 2 -127.6 6 5 1.5192 4 38.0 3 -97.2	-7.05	-4.64	-7.33
S2  (C _{2v})	2 1 1.4778 3 2 0.9621 1 39.8 4 3 2.9309 2 50.2 1 0.0 5 4 0.9618 3 127.8 2 -90.0 6 5 1.5204 4 37.8 3 180.0	-4.93	-3.16	-4.68
S3  (C _{2v})	2 1 1.4808 3 2 0.9616 1 39.6 4 3 3.0474 2 50.3 1 0.0 5 4 0.9618 3 128.0 2 0.0 6 5 1.5158 4 38.0 3 180.0	-3.77	-2.43	-3.59
S4  (C _{2h})	2 1 1.4960 3 2 0.9623 1 39.0 4 3 2.9736 2 69.5 1 -62.5 5 4 0.9623 3 69.5 2 -67.8 6 5 1.4960 4 39.0 3 -62.5	-3.69	-0.50	-2.39
S5  (C ₂)	2 1 1.5208 3 2 0.9616 1 37.9 4 3 2.8226 2 118.4 1 -45.3 5 4 0.9648 3 44.2 2 -105.9 6 5 1.5208 4 37.8 3 116.3	-6.24	-2.82	-6.47
S6  (C _i)	2 1 1.5194 3 2 0.9619 1 38.0 4 3 2.8174 2 114.1 1 -45.1 5 4 0.9648 3 43.6 2 -94.7 6 5 1.5194 4 37.9 3 -110.2	-6.93	-3.18	-7.16

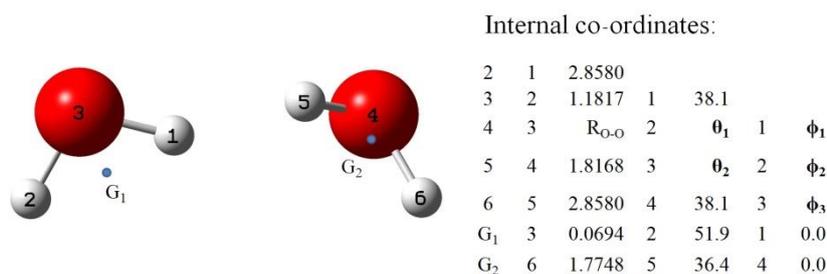


Figure IV.1. Orientation of water GCMs, each comprising of a single *s*-type Gaussian represented by dots and labeled G_1 , G_2 respectively. R_{O-O} is the distance between oxygen atoms while angles $\{\theta_1, \theta_2\}$ and dihedral angles $\{\phi_1, \phi_2, \phi_3\}$ govern orientation of second GCM with respect to the first. Geometries of S1 to S6 are obtained by substituting $\{R_{O-O}, \theta_1, \phi_1, \theta_2, \phi_2, \phi_3\}$ values from corresponding WF structures available in Table IV.1.

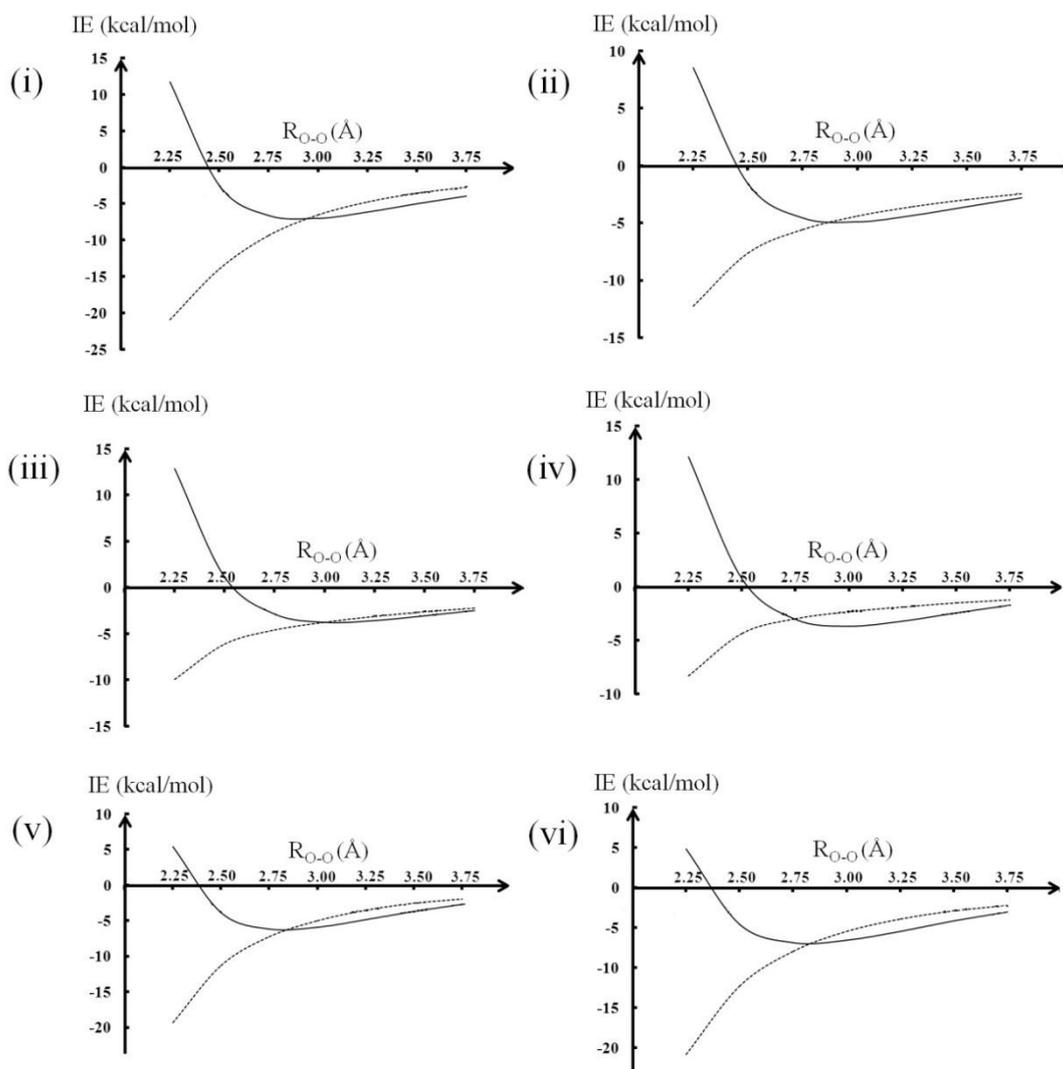


Figure IV.2. Interaction energy as a function of intermolecular oxygen-oxygen distance in structures (i) S1, (ii) S2, (iii) S3, (iv) S4, (v) S5 and (vi) S6. The continuous and dotted curves depict interaction energy values obtained from WF and GCM approach.

To prevent GCMs from coming close towards each other, Gadre and co-workers have used 1.1 and 1.4 Å as the vdW radii for hydrogen (r_H) and oxygen (r_O) atoms while employing the potential derived-atomic charge water model to obtain minimum energy structure for water dimer [Chapter I, ref. 47]. Rather than working with all six structures S1 to S6, for convenience, the present study is undertaken with S1 and S2 geometries only. The GCMs in S1 and S2 orientations are adjusted so that the vdW surfaces of respective models touch each other as shown in Figure IV.3 wherein the R_{O-O} distance in both structures is governed by sum of r_H and r_O values chosen.

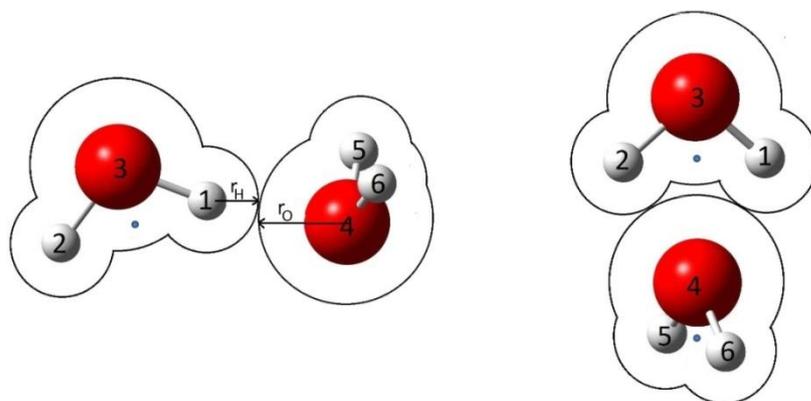


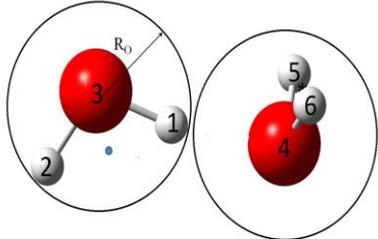
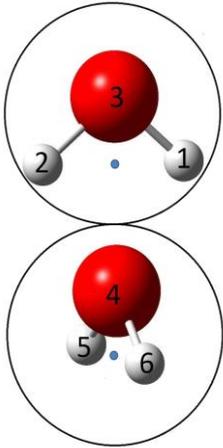
Figure IV.3. Water GCMs in S1 (left) and S2 (right) orientations where respective vdW surfaces touch each other. The dots represent position of s -type Gaussians.

With $r_H = 1.1$ and $r_O = 1.4$ Å, the IE_{GCM} values for S1 and S2 orientations were found to be -3.78 and -4.38 kcal/mol. Thus S2 was found to be more stable than S1. Due to this incorrect prediction, various r_H and r_O values were used in hope of obtaining correct minimum energy structure. The summary of this study is presented in Table IV.2. Despite using different values for vdW radii for oxygen and hydrogen atoms, structure S2 was found to be lower in energy than S1. The R_{O-O} distances in S1 and S2 are 2.91 and 2.93 Å respectively (i.e. almost similar in both). However in earlier study, this distance was found to be larger in S1 orientation which resulted in a wrong prediction.

Table IV.2. Interaction energy (in kcal/mol) for structures S1 and S2 computed using GCMs where the intermolecular R_{O-O} distance (in Å) is dependent on r_H and r_O values (both in Å) respectively.

vdW radius		Sum of r_H and r_O	Geometry of water dimer			
r_H	r_O		S1		S2	
		r_O	R_{O-O}	IE_{GCM}	R_{O-O}	IE_{GCM}
1.0	1.4	2.4	3.3355	-4.25	2.8996	-4.82
0.9	1.4	2.3	3.2358	-4.79	2.7936	-5.33
0.8	1.4	2.2	3.1362	-5.43	2.6880	-5.96
1.0	1.3	2.3	3.2359	-4.79	2.7930	-5.34
0.9	1.3	2.2	3.1362	-5.43	2.6880	-5.96
0.8	1.3	2.1	3.0366	-6.19	2.5900	-6.71
1.0	1.2	2.2	3.1362	-5.43	2.6880	-5.96
0.9	1.2	2.1	3.0366	-6.19	2.5900	-6.71
0.8	1.2	2.0	3.1728	-7.09	2.4750	-7.89

Table IV.3. Interaction energy (kcal/mol) as a function of R_O values (Å) in structures S1 and S2.

Radius R_O	IE_{GCM} values for water dimer geometries	
	S1	S2
		
1.45	-7.45	-4.81
1.46	-7.24	-4.72
1.47	-7.10	-4.64
1.48	-6.85	-4.56
1.49	-6.66	-4.63
1.50	-6.51	-4.40

A spherical surface was therefore employed around each water GCM whose radius (referred to R_O) originated from oxygen atom. With such improvisation, IE_{GCM} values were calculated for S1 and S2 structures with different R_O values ranging from 1.45 to 1.5 Å. These results are shown in Table IV.3. It is seen the GCMs correctly predict minimum energy structure among S1 and S2.

Structure S1 was also predicted to have the lowest IE_{GCM} value when water GCMs were subjected towards energy minimization (given initial values of $\{\theta_1, \phi_1, \theta_2, \phi_2, \phi_3\}$), lowest energy structure of $(H_2O)_2$ is found by minimizing Eq. (IV.15) using subroutine STEPIT (Chapter II, ref. 5). The final study involved selecting an appropriate R_O value so that the water GCM could be employed in molecular dynamics studies. From thermal conductivity studies, the interaction energy for water dimer is found to be about -5.44 kcal/mol [10]. For $R_O = 1.5660$ Å, the predicted IE_{GCM} value of -5.45 kcal/mol is found to be in good agreement with experimental value.

IV.3.2 Ammonia dimer

Unlike the methodology used in earlier section, the PES of ammonia dimer $(NH_3)_2$ was investigated in a slightly different manner. The NH_3 GCMs (orientation shown in Figure IV.4) with spherical surfaces of radius $R_N = 1.5$ Å were subjected towards energy minimization. For various initial values of $\{\theta_1, \phi_1, \theta_2, \phi_2, \phi_3\}$, a bifurcated geometry (Figure IV.5) having C_{3v} point group was predicted to possess lowest interaction energy.

Some stationary points (structures S7, S8 and S9) along the PES of ammonia dimer (obtained from QM calculations at MP2/6-31G(d,p) level of theory) are displayed in Table IV.4.

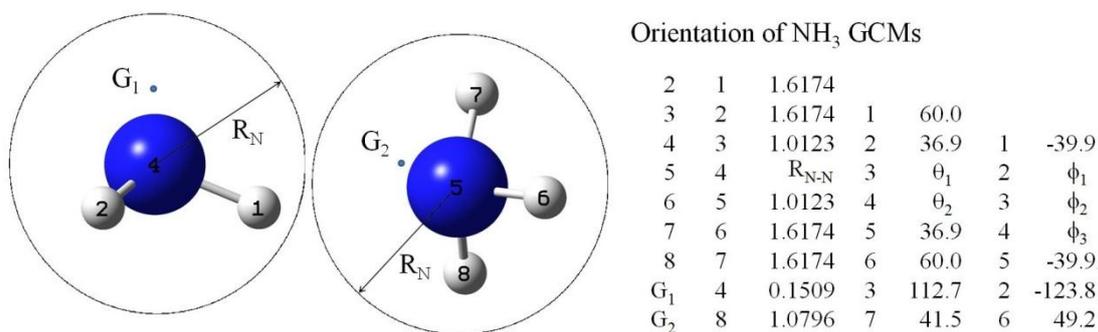


Figure IV.4. Orientation of ammonia GCMs governed by variables $\{\theta_1, \phi_1, \theta_2, \phi_2, \phi_3\}$. Positions of *s*-type Gaussians (with exponent values 0.2479) are represented by dots and labelled G₁ and G₂ respectively. Other GCM parameters are available in Table III.1b.

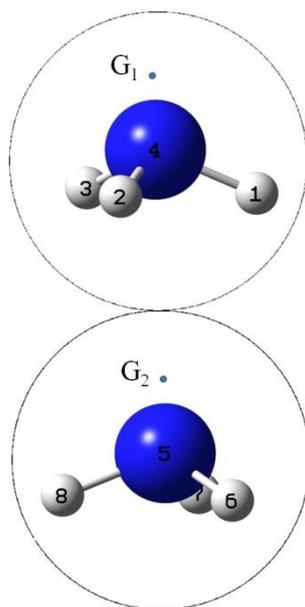
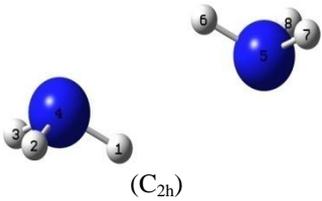
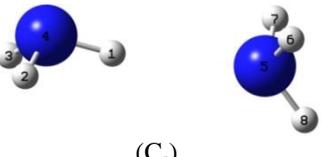
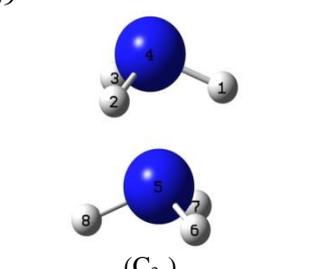


Figure IV.5. Minimum energy structure of (NH₃)₂ predicted from energy minimization using ammonia GCMs with $IE_{\text{GCM}} = -5.55$ kcal/mol and internal coordinates $\{\theta_1, \phi_1, \theta_2, \phi_2, \phi_3\} \equiv \{67.3, 56.2, 112.5, -180.0, -123.8\}$. $R_{\text{N-N}}$ distance is two times the value of R_{N} .

Table IV.4. Stationary structures along PES of $(\text{NH}_3)_2$ at MP2/6-31G(d,p) level of theory. Point group of each orientation given in parenthesis.

Structure	Internal coordinates (bond lengths in Å, angles and dihedral angles in degrees)	Interaction Energy (kcal/mol)		
		IE	IE _{corr}	IE _{GCM}
S7  (C _{2h})	2 1 1.6245 3 2 1.6188 1 60.1 4 3 1.0128 2 37.0 1 -39.3 5 4 3.1357 3 124.1 2 154.2 6 5 1.0140 4 40.6 3 105.1 7 6 1.6245 5 36.7 4 123.5 8 7 1.6186 6 60.1 5 39.6	-4.53	-2.51	-2.51
S8  (C _s)	2 1 1.6231 3 2 1.6120 1 60.2 4 3 1.0133 2 37.3 1 -39.9 5 4 3.2703 3 110.1 2 118.7 6 5 1.0129 4 99.9 3 176.4 7 6 1.6194 5 36.9 4 -103.4 8 7 1.6233 6 60.1 5 -39.4	-4.20	-2.98	-2.27
S9  (C _{3v})	2 1 1.5963 3 2 1.5963 1 60.0 4 3 1.0143 2 38.1 1 -42.6 5 4 3.5546 3 65.3 2 54.1 6 5 1.0128 4 112.8 3 180.0 7 6 1.6175 5 37.0 4 -123.8 8 7 1.6175 6 60.0 5 -40.0	-1.69	-0.49	-3.37

Structure S7 is cyclic where lone pair (LP) of each nitrogen atom is found to interact with single hydrogen atom of the other. Structure S8 is linear with single hydrogen-LP interaction while structure S9 is a bifurcated geometry. From (BSSE uncorrected and corrected) IE values, it is seen that S9 is higher in energy while the minimum energy structure among S7 and S8 varies depending whether BSSE correction is considered or not. It is also seen that the $R_{\text{N-N}}$ distances in S7 to S9 don't vary much, therefore favouring the use of spherical surface around ammonia GCM. In any case, the GCMs are found to make a wrong prediction and there is a need to rectify this problem.

The computational chemistry software DAMQT [Chapter III, ref. 12] employs 5 diffuse functions for nitrogen atom (with respect to 6-31G(d,p) basis set) for

computing $V(\mathbf{r})$ and corresponding first and second order derivatives. Thus in addition to LP and three bond critical points (CPs), positive valued (3, +3) and (3, +1) CPs are found along the C_3 axis. These additional CPs are not obtained when 6 diffuse functions are utilized for nitrogen atom (via. software INDPROP; Chapter I, ref. 67). CPs obtained from both softwares are summarized in Table IV.5. It is seen that there are no significant changes in positions, corresponding MESP/electron density values at LP and bond CP sites derived from both methods. Perhaps the GCM which restores these additional CPs (model parameters presented in Table IV.6) could predict the correct minimum energy structure for $(\text{NH}_3)_2$. The result obtained using this GCM is displayed in Figure IV.6.

QM calculations done at the Hartree-Fock (HF)/MP2 levels of theory employing large basis set like 6-311+G(3d',2p) confirm S10 to be the global minimum structure of $(\text{NH}_3)_2$ [11]. The authors from cited reference have also done several calculations for ammonia dimer at HF and MP2 levels of theory which reveal the following:

- Irrespective of the basis set used, the cyclic C_{2h} structure (S7) was always found to be a stationary point along the PES of ammonia dimer.
- For basis sets like 6-31G, 6-31G(d), 6-31G(d,p) and 6-311G(d,p), the cyclic geometry was found to become a global minimum structure.
- Structure S10 was global minimum when basis sets containing additional polarization functions on nitrogen atom were employed (viz. 6-31+G(d), 6-31+G(d,p), 6-311+G(d,p), 6-31G(d',p'), 6-311+G(2d',p) and 6-311+G(3d',p). Theoretical studies done via. correlation-consistent basis sets confirm this observation [12].

Table IV.5. MESP CPs obtained from computational chemistry softwares INDPROP and DAMQT.

Topography from	Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
		X	Y	Z		
INDPROP	(3, -1)	0.0000	1.1284	-0.2579	0.9660	0.361
		± 0.9772	-0.5642	-0.2579		
	(3, +3)	0.0000	0.0000	2.5075	-0.1365	0.024
DAMQT	(3, -1)	0.0000	1.1284	-0.2579	0.9958	0.361
		± 0.9772	-0.5642	-0.2579		
	(3, +3)	0.0000	0.0000	2.5036	-0.1363	0.024
		0.0000	0.0000	-3.0529	0.0119	0.003
(3, +1)	0.0000	0.0000	-4.2076	0.0113	---	

Geometry of ammonia molecule is available in Chapter III, Table III.1a. Electron density values less than 0.001 a.u. is represented by '---' symbol.

Table IV.6. MESP CPs obtained from NH_3 developed using topographical features from DAMQT software.

GCM parameters		MESP topographical features from corresponding GCM.					
		Type of CP	Co-ordinates			V(r)	$\rho(\mathbf{r})$
			X	Y	Z		
q_N	4.3598	(3, -1)	0.0000	1.2352	-0.3112	0.7960	0.148
q_H	0.3966		± 1.0697	-0.6176	-0.3112		
α	0.2665	(3, +3)	0.0000	0.0000	2.5037	-0.1363	0.024
R_Z	0.2216		0.0000	0.0000	-2.5376	0.0210	0.007
		(3, +1)	0.0000	0.0000	-3.4843	0.0237	---

Gaussian centre: (0, 0, 0.2216), charge on Gaussian: -5.5495. Electron density value less than 0.001 a.u. is represented by '---' symbol. This model was developed by unconstrained minimization of Eq. (III.5).

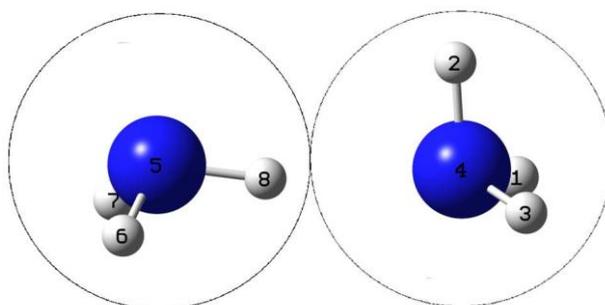
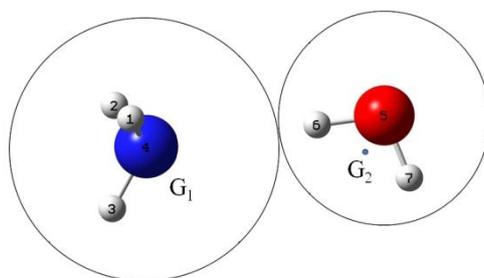


Figure IV.6. Minimum energy structure (S10) of $(\text{NH}_3)_2$ obtained using newly developed NH_3 GCM. The Gaussian centres are situated at nitrogen sites. Internal co-ordinates: $\{\theta_1, \phi_1, \theta_2, \phi_2, \phi_3\} \equiv \{124.7, -90.7, 107.6, -19.4, -114.9\}$.

Tao and Klemperer [13] pointed out that the energy difference between S10 and S7 orientations was not much which in turn made it sensitive towards the basis set chosen. Since the newly developed model could correctly predict the structure of ammonia dimer, the last step in this study is to find a suitable R_N value for ammonia GCM. The S10 structure was subjected towards geometry optimization at MP2 level using one of the correlation-consistent basis set (aug-cc-pVDZ) which resulted in IE and IE_{corr} values of -4.02 and -2.78 kcal/mol. For $R_N = 1.845 \text{ \AA}$, the GCMs predict an IE_{GCM} value of -2.77 kcal/mol which was in good agreement with corresponding QM result.

IV.3.3 Ammonia-water complex

Ammonia and water GCMs could predict minimum energy structure of ammonia-water complex wherein ammonia acts as a proton acceptor while water as hydrogen donor (Figure IV.7). The predicted IE_{GCM} value of -5.45 kcal/mol does not significantly differ from corresponding IE_{corr} value of -6.36 kcal/mol.



Geometry of $\text{NH}_3\text{-H}_2\text{O}$ complex

GCM approach	2	1	1.6174				
	3	2	1.6174	1	60.0		
	4	3	1.0122	2	37.0	1	-39.9
	5	4	3.4110	3	128.0	2	-123.8
	6	5	0.9611	4	0.0	3	-1.9
	7	6	1.5119	5	38.1	4	-178.0
G_1	3	1.0122	2	37.0	1	-39.9	
G_2	7	0.9389	6	36.4	5	0.0	
MP2/6-31G(d,p)	2	1	1.6229				
	3	2	1.6242	1	60.0		
	4	3	1.0127	2	36.7	1	-39.3
	5	4	2.9641	3	122.9	2	-123.3
	6	5	0.9724	4	2.8	3	0.0
	7	6	1.5209	5	37.8	4	180.0

Figure IV.7. Minimum energy structure of ammonia-water complex obtained via. WF and GCMs.

IV.3.4 Acetylene dimer and acetylene-water complex

The simplest GCM of a hydrocarbon molecule (in terms of atoms as well as no. of spherical Gaussians) is that of acetylene molecule. A vdW type of surface, constructed from spheres of radius R_C originating from each carbon atom has been employed around the entire model. The choice of $R_C = 1.9 \text{ \AA}$ is found to be quite satisfactory with respect to the minimum energy structure (T-shape geometry) of acetylene dimer (C_2H_2)₂ obtained using acetylene GCMs (Figure IV.8a) with IE_{GCM} value of about -1.20 kcal/mol (BSSE free value: -1.05 kcal/mol at MP2/6-31G(d,p) level of theory).

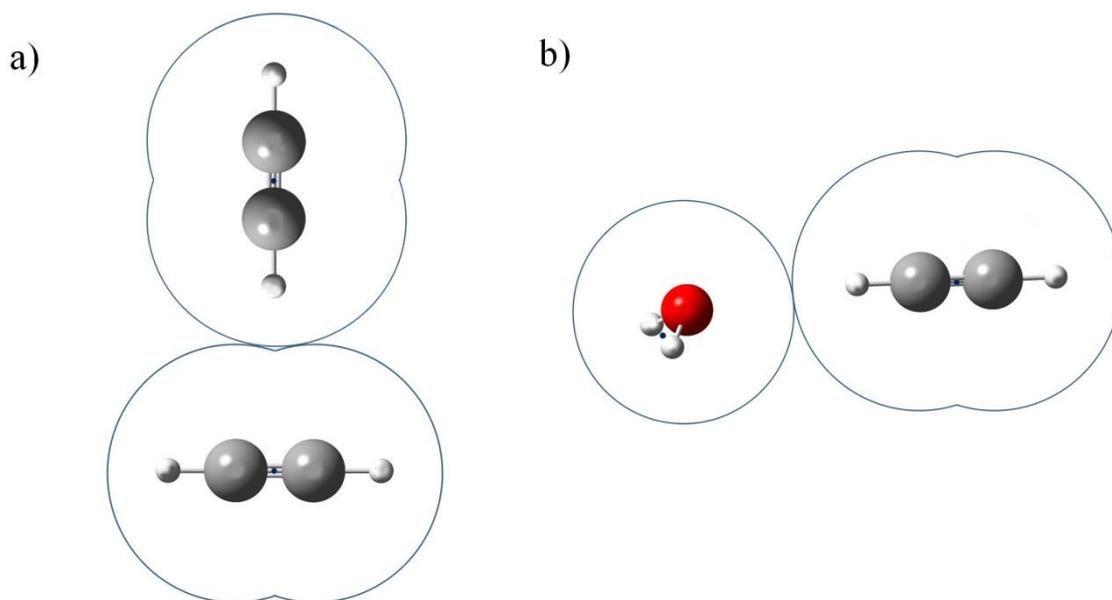


Figure 4.8. Structures of a) acetylene dimer and b) acetylene-water complex predicted by use of acetylene and water GCMs.

Interaction between acetylene and water results to the acetylene-water complex wherein water acts as a proton acceptor (Figure IV.8b) with IE_{GCM} value of -4.64 kcal/mol. At MP2/6-31G(d,p) level of theory, the interaction energy is about -4.46 kcal/mol and -2.43 kcal/mol when BSSE is included.

IV.3.5 Benzene dimer and benzene related systems

Extensive work on aromatic-aromatic interactions has been carried out by Burley and Petsko [14], who defined these interactions as those between phenyl rings whose centroids are separated by a distance of 4.5 – 7.0 Å with energy of formation in the range of -0.6 to -1.3 kcal/mol. They considered them as a separate class of interactions which contributed (to some extent) towards the stability of protein structures (more dominant when protein contains large number of aromatic fragments). These interactions are divided into two categories: π - π stacking (resulting from face-to-face orientation of aromatic rings) and C-H interactions (edge-to-face orientation), which are briefly discussed in the upcoming section.

Based on the electrostatic model proposed by Hunter and Sanders [15], (the kindergarten explanation for) π - π stacking in the simplest system viz. benzene dimer (C_6H_6)₂ is due to repulsion between π electrons of respective rings and, attraction between positively charged benzene ring and π electrons of the other. The columbic interactions between aromatic hydrogen atoms of one ring and π cloud of the other lead to edge-to-face orientation [14]. Early QM calculations on (C_6H_6)₂ predict the T-shaped orientation to be lower in energy than the face-to-face geometry [16, 17]. Hobza and co-workers however found the latter to be more stable than the former [18]. Tsuzuki and co-workers [19] pointed out that the electrostatic interactions favour the T-shape structure while dispersion energy (a ubiquitous weak attractive force) directs the benzene rings in face-to-face geometry.

In the study of PES of (C_6H_6)₂, the benzene GCMs account for the electrostatic component of the total IE. This has been shown from the following investigations.

- Starting from the face-to-face orientation of benzene GCMs (Figure IV.9a; centroids separated by distance of 5.29 Å; spherical surface around each model of radius of 2.645 Å originating from corresponding centre of mass), the energy minimization process predicted the T-shaped geometry (Figure IV.9b) to be highly stable.
- Performing similar calculation with spherical radius of 2.4772 Å (distance from centre of mass to terminal hydrogen of benzene) resulted in a T-shaped structure with C_2 point group (Figure IV.9c).

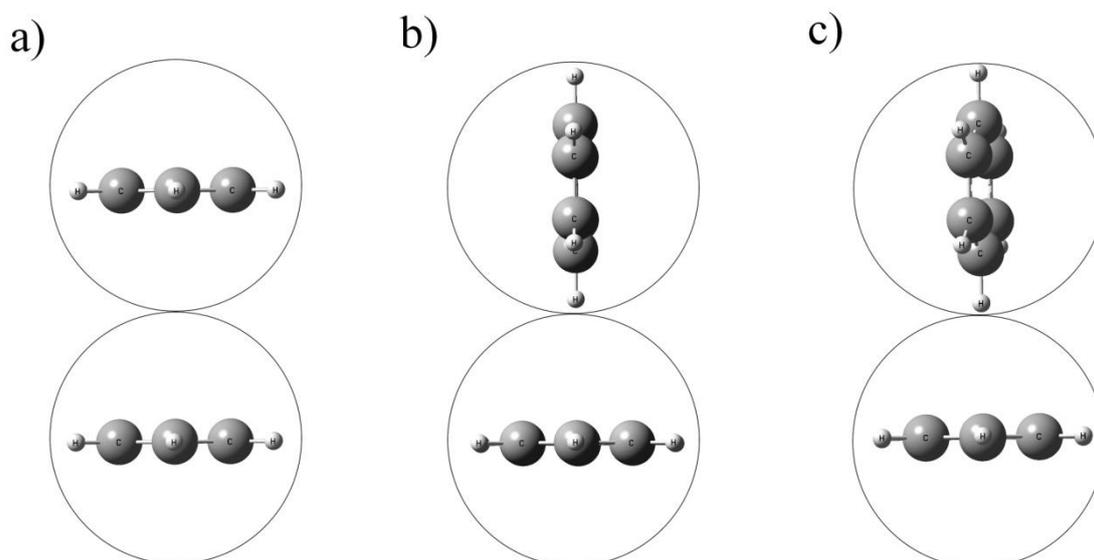


Figure IV.9. Benzene dimer: a) face-to-face orientation, T-shaped orientation with b) C_{2v} and c) C_2 point groups. For convenience, the Gaussians in each model are not shown. The circle represents spherical surface around each GCM.

Table IV.7. Summary of interaction energies calculated for benzene dimer orientations using WF (at MP2/6-31G(d,p) level of theory) and benzene GCMs.

Orientation	Distance between ring centroids (Å)	Interaction energy (kcal/mol)		
		IE	IE _{corr}	IE _{GCM}
Face-to-face	3.7298	-2.36	-0.03	0.56
T-shape (C_{2v})	4.888	-3.46	-1.17	-0.03
T-shape (C_2)	4.8339	-3.46	-1.17	-0.04

At MP2/6-31G(d,p) level of theory, the above mentioned orientations were successfully obtained and the results are summarized in Table IV.7. The benzene GCMs were similarly oriented and IE_{GCM} values calculated are also displayed. It is seen that the GCMs reproduce correct trend in IEs i.e. the face-to-face structure is higher in energy compared to T-shaped geometries.

Unlike the use of spherical surfaces for water and ammonia GCMs, it is more relevant to employ a vdW type of surface for benzene GCM. Such a surface is constructed from spheres (of radius R_C) originating at each carbon atom as shown in Figure IV.10a. An appropriate choice of $R_C = 1.9 \text{ \AA}$ is based on the following observations:

1. Value of R_C needs to be larger than corresponding value of 1.875 \AA selected for ammonia GCM and as far as possible, it is better to avoid multiple values for various carbon containing GCMs (though the latter statement needs to be verified).
2. Gaussian charge distributions in benzene GCM are located at a distance of approximately 1.7 \AA from the plane of the ring. The surface is expected to sufficiently enclose them.
3. It has been observed that for $R_C = 1.77 \text{ \AA}$, the structure predicted to be lower in energy from energy minimization process posses C_s point group (Figure IV.10b) accompanied by further drop in IE_{GCM} value. This is due to the attraction between opposite charges located at the hydrogen and Gaussian sites.

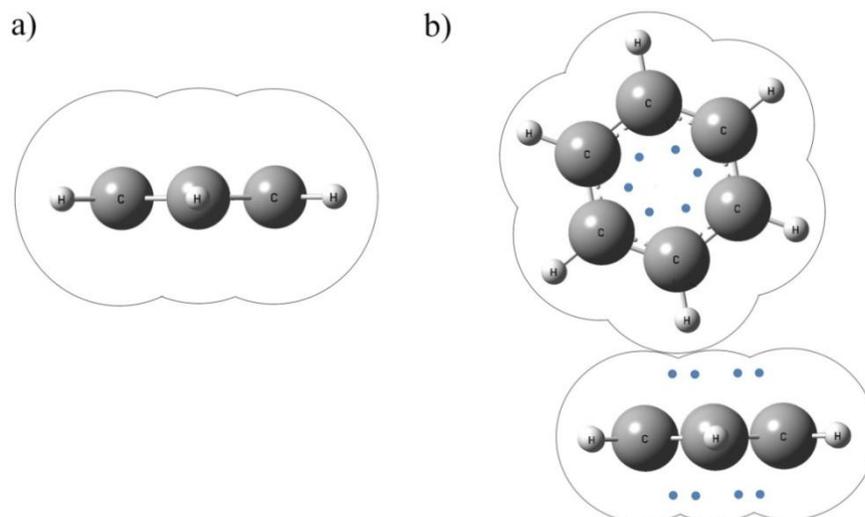


Figure IV.10. a) vdW type of surface around benzene GCM and b) minimum energy structure with C_s point group having IE_{GCM} value of -0.07 kcal/mol. Small spheres represent positions of Gaussian charge distributions in respective benzene models.

The benzene model with R_C value equal to 1.9 \AA was then employed to predict the structures of benzene-water and benzene-ammonia complexes.

The structure of benzene-water complex has been reported by several research groups [20, 21] wherein one of the hydrogen atoms belonging to water molecule interacts with the π cloud of benzene. Energy minimization studies carried out using benzene and water GCMs predict such structure (Figure IV.11a) with IE_{GCM} value of -2.34 kcal/mol which was further confirmed by QM calculations performed at MP2 level of theory employing a slightly larger basis set viz. 6-31+G(d,p) with IE and IE_{corr} values of -4.08 and -2.16 kcal/mol respectively. Ammonia molecule like water also interacts with the π cloud of benzene in similar manner (Figure IV.11b). The IE_{GCM} value is about -1.10 kcal/mol, suggesting that ammonia is weakly bound to the aromatic substrate as compared to water.

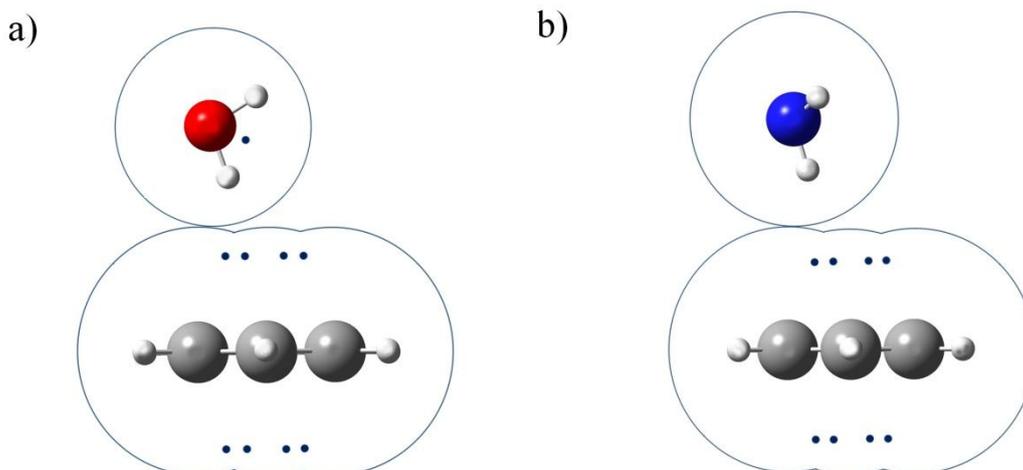


Figure IV.11. Minimum energy structures of a) benzene-water and b) benzene-ammonia complexes predicted by GCMs. The dots represent positions of s-type Gaussians in respective models and in the ammonia GCM, the Gaussian is situated at the nitrogen site.

The geometries of benzene-water, benzene-ammonia, acetylene dimer and acetylene-water complexes predicted by GCMs are in good agreement with earlier theoretical and experimental findings. [21-25].

IV.4 Concluding remarks

From the study carried out on water dimer, it was observed that the GCMs were unable to reproduce expected trend in IEs at shorter distances. It was however possible to obtain correct structure of water dimer by enclosing a spherical surface around each water GCM. The sphere of radius 1.566 Å was sufficient to reproduce interaction energy of water dimer which was in good agreement with experimental findings. Though the global minimum structure of ammonia dimer was basis set dependent, the newly developed NH₃ model (enclosed by a spherical surface of $R_N = 1.845$ Å) could correctly predict its geometry.

The above mentioned models were successful in correctly deducing the structure of ammonia-water complex wherein water acts as a proton donor. In the case of benzene dimer, the benzene GCMs predicted the T-shaped structure be lower in energy with was in good agreement with QM study performed at MP2/6-31G(d,p) level of theory. The vdW surface around C_6H_6 GCM was built using spheres of radius 1.9 Å from each carbon site and the corresponding model was employed to study $C_6H_6-H_2O$ and $C_6H_6-NH_3$ complexes. Compared to benzene-water and benzene-ammonia complexes, the water molecule was found to be strongly bound to benzene than ammonia. In both cases, benzene molecule was found to acts as a proton acceptor.

The acetylene GCM was also found to predict correct structure of its dimer and acetylene-water complex which were in good agreement with earlier reported works. The surface around C_2H_2 GCM was similar to the one used for benzene GCM, with R_C value of 1.9 Å.

From the studies presented in this chapter, it is seen that the GCMs are able to predict correct geometries of some non-covalent type of complexes and the respective IEs don't deviate much from corresponding QM counterparts. The models need to be tested for molecular systems beyond the dimer so as to demonstrate their effectiveness and reliability of their predictions.

References:

1. G. Herzberg and A. Monfils, The Dissociation Energies of the H₂, HD, and D₂ Molecules, *J. Mol. Spectrosc.*, 5 (1960) 482 (and references within).
2. J.A. Montgomery Jr. and G.A. Petersson, On the C-H Bond Dissociation Energy of Acetylene, *Chem. Phys. Lett.*, 168 (1990) 75 (and references within).
3. CRC Handbook of Chemistry and Physics, Ed. R.C. Weast and M.J. Astle, 1982-83, CRC, Florida.
4. J. D. van der Waals, Doctoral Dissertation, Leiden, 1873.
5. K. Morokuma, Why Do Molecules Interact? The origin of Electron Donor-Acceptor Complexes, Hydrogen Bonding, and Proton Affinity, *Accounts of Chemical Research*, (1977) 294.
6. R. Bonaccorsi, E. Scrocco and J. Tomasi, A Representation of the Polarization Term in the Interaction Energy between a Molecule and a Point-Like Charge, *Theoret. Chim. Acta (Berl.)*, 43 (1976) 63.
7. C.C. Pye, R.A. Poirier, D. Yu and P.R. Surján, Ab initio Hartree – Fock calculations of the interaction energy of bimolecular complexes, *J. Mol. Structure (Theochem)*, 307 (1994) 239.
8. S.F. 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 (1970) 553.
9. S. Reine, T. Helgaker and R. Lindh, Multi-electron integrals, *WIREs Comput. Mol. Sci.*, 2 (2012) 290.
10. L.A. Curtiss, D.J. Frurip and M. Blander, Studies of molecular association in H₂O and D₂O vapors by measurement of thermal conductivity, *J. Chem. Phys.*, 71 (1979) 2703.

11. D.M. Hassett, C.J. Marsden and B.J. Smith, The ammonia dimer potential energy surface: resolution of the apparent discrepancy between theory and experiment?, *Chem. Phys. Lettr.*, 183 (1991) 449.
12. J.S. Lee and S.Y. Park, Ab initio study of $(\text{NH}_3)_2$: Accurate structure and energetic, *J. Chem. Phys.*, 112 (2000) 230.
13. F. Tao and W. Klemperer, Ab initio search for the equilibrium structure of the ammonia dimer, *J. Chem. Phys.*, 99 (1993) 5976.
14. S.K. Burley and G.A. Petsko, Aromatic-Aromatic Interaction: A Mechanism of Protein Structure Stabilization, *Science*, 29 (1985) 23.
15. C.A. Hunter and J.K.M. Sanders, The Nature of π - π Interactions, *J. Amer. Chem. Soc.*, 112 (1990) 5525.
16. P. Carsky, H.L. Selzle and E.W. Schlag, Ab initio calculations on the structure of the benzene dimer, *Chem. Phys.*, 125 (1988) 165.
17. P. Hobza, H.L. Selzle and E.W. Schlag, Floppy structure of the benzene dimer: Ab initio calculation on the structure and dipole moment, *J. Chem. Phys.*, 93 (1990) 5893.
18. P. Hobza, H.L. Selzle and E.W. Schlag, New Structure for the Most Stable Isomer of the Benzene Dimer: A Quantum Chemical Study, *J. Phys. Chem.*, 97 (1993) 3937.
19. S. Tzuki, K. Honda, T. Uchimaru, M. Mikami and K. Tanabe, Origin of Attraction and Directionality of the π/π Interaction: Model Chemistry Calculations of Benzene Dimer Interaction, *J. Amer. Chem. Soc.*, 124 (2002) 104.

20. S. Suzuki, P.G. Green, R.E. Bumgarner, S. Dasgupta, W.A. Goddard III and G.A. Blake, Benzene Forms Hydrogen Bonds with Water, *Science*, 257 (1992) 942.
21. J.D. Augspurger, C.E. Dykstra and T.S. Zwier, Hydrogen-Bond Swapping in the Benzene-Water Complex. A Model Study of the Interaction Potential, *J. Phys. Chem.*, 96 (1992) 7252.
22. M. Mons, I. Dimicoli, B. Tardivel, F. Piuzzi, V. Brenner and P. Millié, Energetics of a model NH- π interaction: the gas phase benzene-NH₃ complex, *Phys. Chem. Chem. Phys.*, 4 (2002) 571.
23. C. Enkvist, Y. Zhang and W. Yang, Density Functional Study of a Weakly Hydrogen-Bonded Benzene-Ammonia Complex: The importance of the Exchange Functional, *Int. J. Quantum Chem.*, 79 (2000) 325.
24. A. Karpfen, The Dimer of Acetylene and the Dimer of Diacetylene: A Floppy and a Very Floppy Molecule, *J. Phys. Chem. A*, 103 (1999) 11431.
25. D. Tzeli, A. Mavridis and S.S. Xantheas, A first principles study of the acetylene-water interaction, *J. Chem. Phys.*, 112 (2000) 6178.

CHAPTER V

Application of quantum mechanical methods in reaction mechanisms and forensics

	Table of Contents	Page
V.1	Cobalt thiocyanate: a useful reagent in forensics	127
V.1.1	Detection of some nitrogen containing compounds via. cobalt thiocyanate and TLC	128
V.1.2	Theoretical structure elucidation of some nitrogen containing compounds with cobalt thiocyanate	129
V.2	Theoretical investigation on 3-aryl-6-ethoxycarbonyl-4-hydroxy-2 <i>H</i> -pyran-2-ones	139
V.3	Concluding remarks	146
	References	148

Part of this work has been:

- 1) Published in the research article titled ‘TLC Detection and Theoretical Structure Elucidation of Nitrogen Containing Compounds with Cobalt Thiocyanate’, *Anal. Chem. Lett.*, 9 (2019) 453.
- 2) Presented as a poster ‘Theoretically proposed structures of Cobalt Thiocyanate with nitrogen containing compounds’ at the Eighth Asian Pacific Conference on Theoretical and Computational Chemistry held at IIT Bombay, Mumbai on 15th – 17th December 2017 (Poster no. P133)

V.1. Cobalt thiocyanate: a useful reagent in forensics

In addition to the topography-based charge model work presented in Chapter III and Chapter IV, QM methods have also been used to understand and account certain experimental observations which were briefly discussed in Chapter I, section I.6.

Cobalt thiocyanate reagent has frequently been used in forensic laboratories for preliminary detection of nitrogen containing substances. Its implementation in forensics and law enforcement originated from the development of Scott test [1] around 1973, which involved detecting cocaine and its hydrochloride salt by means of color reaction (positive identification indicated by blue color). Since then, the Scott test has been modified to improve its specificity towards other nitrogen containing compounds [2, 3]. Kote and co-workers found this reagent sensitive towards neonicotinoid insecticides like imidacloprid and acetamiprid, in association with thin-layer chromatographic (TLC) technique [4]. The Regional Forensic Science Laboratory, Nagpur, India detected large number of cases pertaining to human poisoning by an antimalarial and amebicide drug – chloroquine. A case was received by the laboratory where a 24-year-old female victim (health worker by profession) died by consumption of unknown tablets. Police officers seized 70 tablets of chloroquine phosphate from the crime scene and sent for analysis. As per autopsy findings, probable cause of death was suspected due to suicidal poisoning. Chemical analysis revealed chloroquine in visceral samples (pieces of stomach, intestine, liver, kidneys and lungs). The upcoming section will briefly outline the steps involved in the chemical analysis of chloroquine in addition to several other drugs like chlorphenamine, mephedrone, cocaine, ketamine, pretilachlor, thiamethoxam and tebuconazole.

V.1.1 Detection of some nitrogen containing compounds via. cobalt thiocyanate and TLC

Cobalt thiocyanate reagent was prepared by dissolving 3 g of ammonium thiocyanate and 1 g of cobalt (II) chloride in 20 mL distilled water. About 100 mL of ethyl alcohol was added to approximately 50 g of fine pieces of visceral samples, suspected to contain chloroquine and the resulting solution was kept overnight. After 2 hours from addition of 10 mL of ammonia, the mixture was extracted twice with chloroform. On evaporating the solvent, the obtained residue was re-dissolved in 1 mL chloroform and with the help of fine capillary, spotted on pre-coated TLC plate along with standard (pure chloroquine). The plate was then kept in a developing chamber containing mixture of ethyl alcohol, methyl alcohol and ammonia in 8:1:1 volume ratio. After drying in air, it was sprayed with cobalt thiocyanate reagent. An intense blue colour spot (with retardation factor R_F of 0.6) confirmed the presence of chloroquine.

Other nitrogen containing compounds like chlorphenamine, cocaine, ketamine, mephedrone, pretilachlor, tebuconazole and thiamethoxam were also analyzed by TLC method. The R_F values and solvent systems used for their detection are summarized in Table V.1.

Table V.1. Summary of TLC analysis of some nitrogen containing compounds.

Compound	Solvent system		R_F
	Solvents	Volume ratio	
Chlorphenamine	ethyl alcohol, methyl alcohol and ammonia	8 : 1 : 1	0.57
Cocaine		8 : 1.5 : 0.5	0.84
Ketamine			0.81
Mephedrone			0.63
Pretilachlor	ethyl acetate, hexane	1 : 2	0.50
Tebuconazole	acetone, dichloromethane	1 : 2	0.50
Thiamethoxam	acetone, hexane	2 : 8	0.52

All compounds formed blue coloured spots except ketamine which resulted in purple. The intense coloured solutions could not be subjected to spectroscopic analysis as they turned colourless on dilution. It was also difficult to grow single crystals from respective solutions. Hence the structures of chemical species responsible for the observed blue and purple colours were predicted based on QM calculations, described hereafter.

V.1.2 Theoretical structure elucidation of some nitrogen containing compounds with cobalt thiocyanate

Cobalt complexes exhibit blue colour when ligands bind to the metal centre in a tetrahedral manner. Some of the reported ligands coordinate through nitrogen donor atoms belonging to six-membered aromatic substrates viz. 2-methylpyridine, 2-chloropyridine, 2-bromopyridine [5], 3-methylpyridine [6], 4-aminopyridine [7], 1,2-bis(2-pyridyl)ethylene [8], trans-4-(4'-(N-methyl-N-hydroxyethyl)amino)styrylpyridine [9], quinoline [10], while there are those wherein the nitrogen donor atoms belong to five-membered rings: 3,5-dimethylpyrazole [11], 3,5-dimethyl-1-phenylpyrazole [12], 4-acetyl-3-amino-5-methylpyrazole [13], benzimidazole [14], 2-aminobenzimidazole [15], n-butylimidazole [16] and 2,6-dimethylbenzo-[1,2-d;4,5-d']diimidazole [17]. In all these cases, the thiocyanate ligand is found to bind to the cobalt (II) ion via nitrogen donor atoms.

Based on these observations, the drug molecules mentioned earlier are predicted to form complexes of the type CoX_2L_2 (X = thiocyanate ligand, L = drug moiety) with cobalt thiocyanate. Molecules mephedrone and pretilachlor bind via single nitrogen atoms while the others (chloroquine, chlorphenamine, imidacloprid,

tebuconazole and thiamethoxam) contain more than one nitrogen atoms, in which case, the Molecular Electrostatic Potential (MESP) serves as a means to find the most reactive nitrogen site.

Minimum energy structures of the drug molecules under study were obtained first at the Hartree-Fock (HF) level of theory using STO-3G basis set. These geometries were then subjected towards geometry optimization by employing Density Functional Theory (DFT) using 6-31G(d,p) basis set. The B3LYP (Becke, 3-parameter, Lee–Yang–Parr) functional [18-19] was used in the latter case. All these calculations were performed using the Gaussian 03 software [Chapter I, ref. 63]. The MESP analysis of drug molecules was performed using HF/STO-3G and B3LYP/6-31G(d,p) optimized geometries via. computational Chemistry software DAMQT [Chapter III, ref. 12]. The summary of this investigation (position of nitrogen lone pairs (LPs), distances of LPs from respective nitrogen atoms, corresponding MESP values at LP sites, labelling schemes of some atoms in imidacloprid/tebuconazole molecules) is presented in Table V.2.

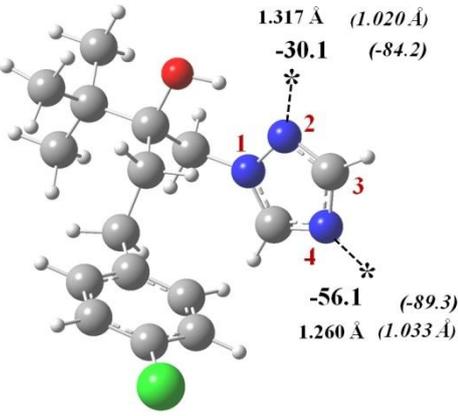
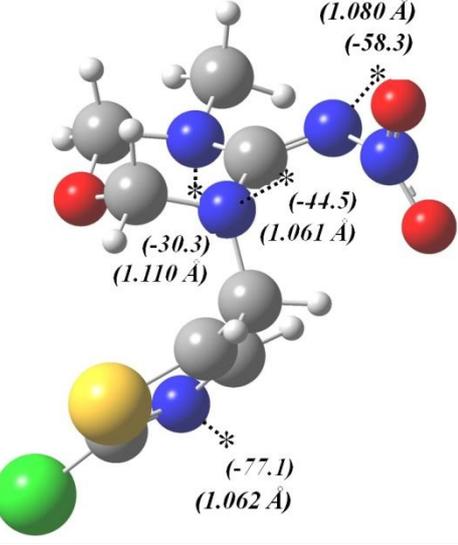
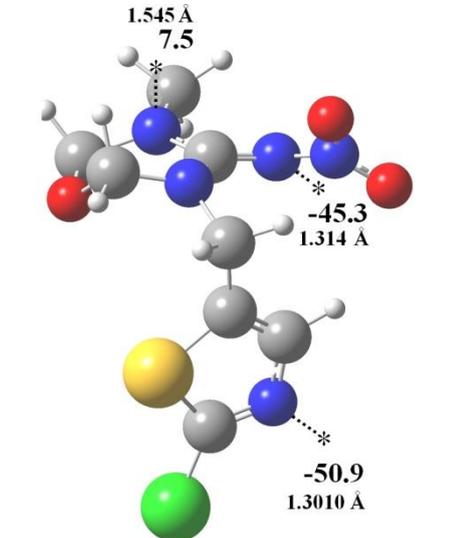
Based on the discussions in ‘Comments’ column in Table V.2, the molecules chloroquine, chlorphenamine and imidacloprid are believed to coordinate to cobalt ion via. LPs belonging to pyridine nitrogen atoms; thiazole – nitrogen LP in the case of thiamethoxam and LP of nitrogen numbered 4 of the 1,2,4-triazole unit in tebuconazole. The tetrahedral geometries of cobalt thiocyanate complexes obtained via. HF and DFT calculations using the LANL2DZ basis set for cobalt atom, STO-3G and 6-31G(d,p) basis sets for non-cobalt atoms (the former employed at HF level of theory and the latter while employing DFT level of theory), are portrayed in Figures V.1 and V.2 respectively. Important parameters like various cobalt – N(ligand) bond lengths and bond angles are displayed in Table V.3.

Table V.2. Nitrogen LP sites (indicated by “*” symbols) in some drug molecules. MESP values (from B3LYP/6-31G(d,p) calculations) at reactive sites are displayed in kcal/mol while distances are given in Å. Values in parenthesis are derived from HF/STO-3G calculations.

Molecule	Locations of reactive nitrogen sites	Comments
Chloroquine	<p> (1.058 Å) (-60.0) * 1.246 Å (1.026 Å) -49.6 (-83.3) * 1.305 Å (1.236 Å) -67.1 (-88.5) * </p>	<p>In the DFT structure, atom numbered 4 is more planar with respect to the quinoline ring, resulting in the absence of N3 LP (due to its participation in conjugation with the π cloud of quinoline ring).</p> <p>This is reflected by increase in dihedral angle value between atoms numbered 1, 2, 3 and 4 from 144° (in HF/STO-3G geometry) to 163.5° in DFT structure.</p>
Chlorphenamine	<p> (1.020 Å) 1.228 Å (-84.2) -59.4 * 1.237 Å (-58.3) * (1.033 Å) (-89.3) </p>	<p>Though HF theory predicts pyridine nitrogen LP to be more reactive, there is no much difference between MESP values (obtained from B3LYP/6-31G(d,p) level of theory) at pyridine and tertiary nitrogen LP sites.</p> <p>Based on the LP distances from respective nitrogen atoms, the former is found to be more accessible than the latter.</p>
Imidacloprid	<p> (1.061 Å) (-42.5) 1.394 Å -6.9 * (-77.0) (1.037 Å) -50.7 * 1.285 Å * -4.8 (-43.6) * 1.350 Å (1.072 Å) (1.049 Å) (-69.9) 1.252 Å -50.7 * </p>	<p>At the B3LYP/6-31G(d,p) level of theory, the reactivity of pyridine nitrogen LP and LP belonging to nitrogen labelled 3 are similar.</p> <p>The MESP value at the former site is deeper than the latter at HF/STO-3G level of theory.</p> <p>Based on the distances of LPs from respective nitrogen atoms, DFT predicts the former to be more accessible while HF predicts the latter.</p>

Continued on next page....

Table V.2 continued...

Molecule	Locations of reactive nitrogen sites	Comments
Tebuconazole	 <p>1.317 Å (1.020 Å) -30.1 (-84.2)</p> <p>-56.1 (-89.3) 1.260 Å (1.033 Å)</p>	<p>The five-membered ring is known as 1,2,4-triazole.</p> <p>MESP value at the LP site belonging to nitrogen numbered 4 is found to be more deeper than at the corresponding N2-LP site.</p>
Thiamethoxam	<p>HF/STO-3G and B3LYP/6-31G(d,p) calculations predicted different conformations of the six-membered ring in thiamethoxam.</p> <p>Chair conformation (at HF/STO-3G level of theory)</p>  <p>(1.080 Å) (-58.3)</p> <p>(-44.5) (1.061 Å)</p> <p>(-30.3) (1.110 Å)</p> <p>(-77.1) (1.062 Å)</p>	<p>Irrespective of the ring conformation, nitrogen LP belonging to the five-membered ring (thiazole unit) is the most reactive site.</p>
	<p>Half-chair conformation (at B3LYP/6-31G(d,p) level of theory)</p>  <p>1.545 Å 7.5</p> <p>-45.3 1.314 Å</p> <p>-50.9 1.3010 Å</p>	

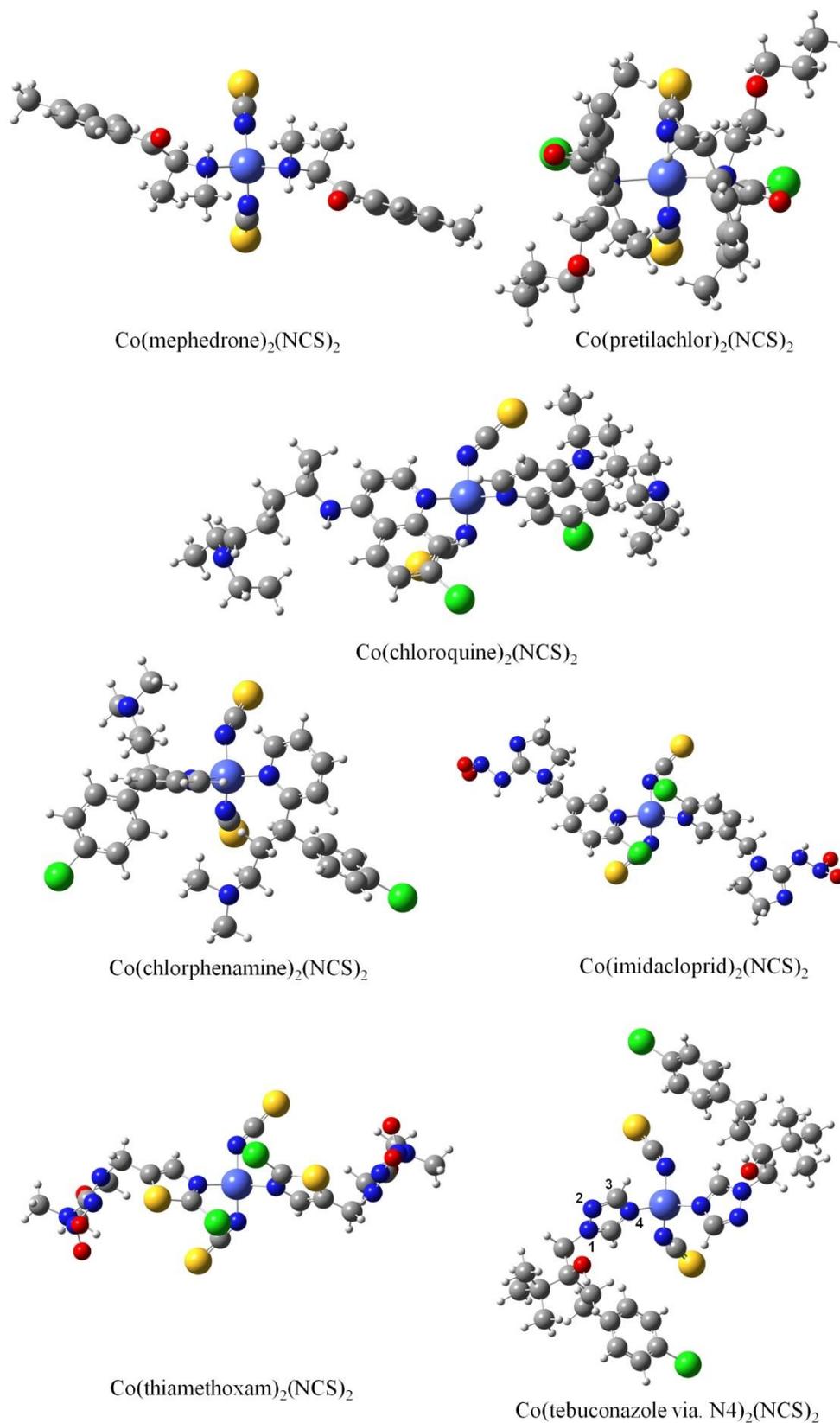


Figure V.1. Minimum energy structures of tetrahedral cobalt thiocyanate complexes of some nitrogen compounds obtained from HF calculations. Coordination of tebuconazole in is via. LP of nitrogen labelled 4 as per numbering scheme of 1,2,4-triazole unit in tebuconazole molecule presented in Table V.2.

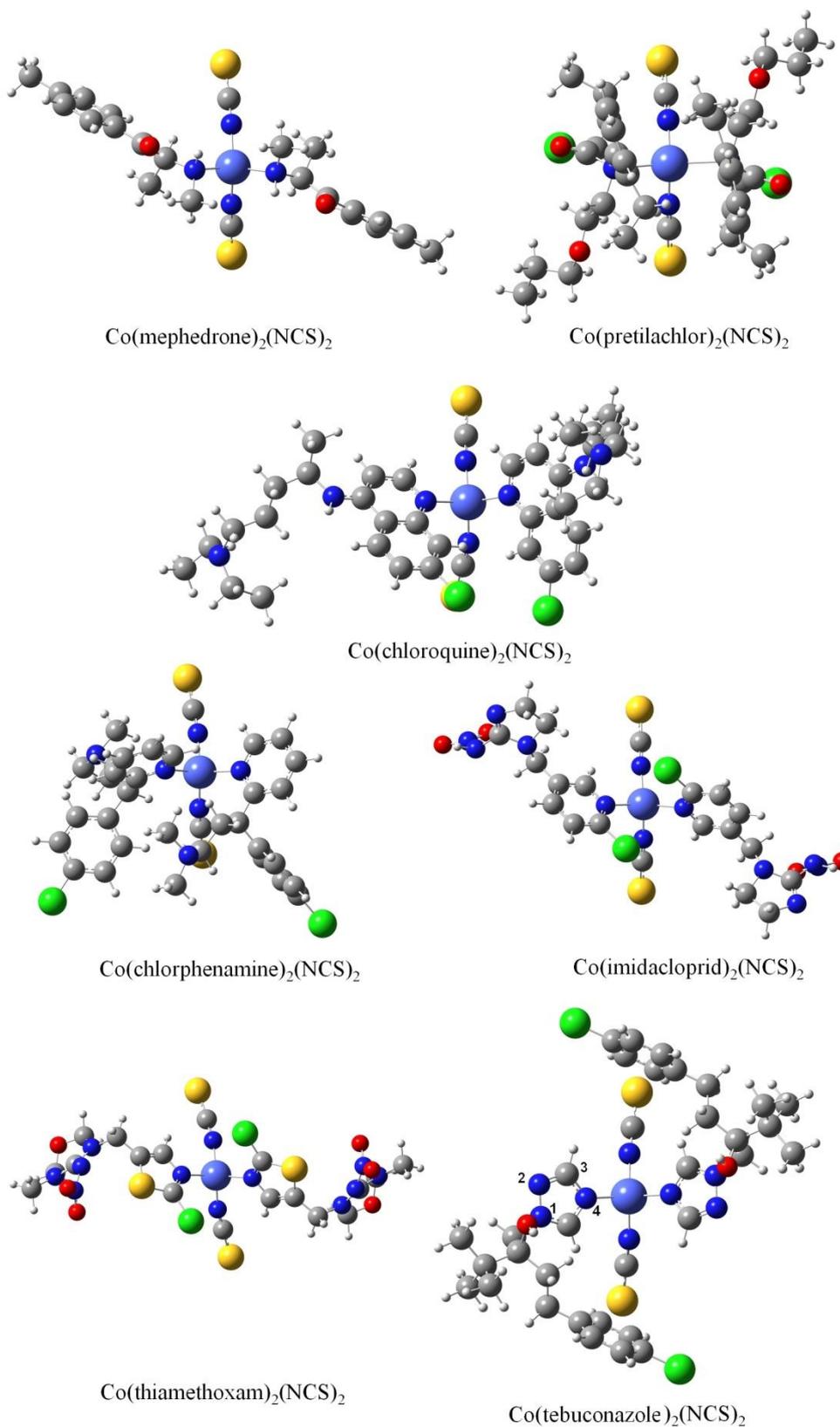
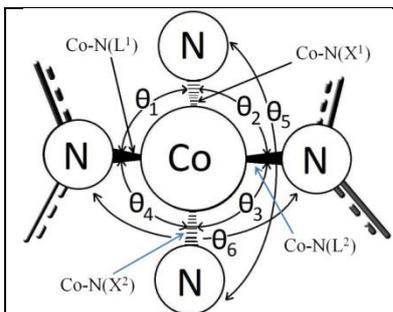


Figure V.2. Minimum energy structures of tetrahedral cobalt thiocyanate complexes of some nitrogen compounds obtained from DFT calculations. Coordination of tebuconazole is via LP of nitrogen labelled 4 as per numbering scheme of 1,2,4-triazole unit in tebuconazole molecule presented in Table V.2.

Table V.3. Various bond lengths and bond angles of ligands with cobalt ion in tetrahedral cobalt thiocyanate complexes obtained from QM calculations.

Optimized geometry from		Bond lengths (Å)		Bond angles (degrees)					
		Co-N(L ¹)/ Co-N(L ²)	Co-N(X ¹)/ Co-N(X ²)	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
Co(mephedrone)₂(NCS)₂									
HF/ STO-3G	2.156 2.156	1.915 1.915	103.8	103.5	107.6	97.7	127.0	106.4	
B3LYP/ 6-31G(d, p)	2.117 2.117	1.935 1.935	103.0	106.8	103.0	106.8	126.6	110.3	
Co(pretilachlor)₂(NCS)₂									
HF/ STO-3G	2.335 2.335	1.917 1.917	99.0	95.7	99.0	95.7	125.0	147.8	
B3LYP/ 6-31G(d, p)	2.439 2.354	1.929 1.924	98.7	94.0	99.3	93.8	126.2	151.3	
Co(chloroquine)₂(NCS)₂									
HF/ STO-3G	2.110 2.112	1.944 1.933	103.8	102.2	113.2	104.3	123.1	110.0	
B3LYP/ 6-31G(d, p)	2.111 2.120	1.957 1.931	101.0	100.2	112.8	106.3	124.6	110.8	
Co(chlorphenamine)₂(NCS)₂									
HF/ STO-3G	2.166 2.186	1.920 1.909	104.4	101.1	113.1	109.9	123.2	111.7	
B3LYP/ 6-31G(d, p)	2.127 2.146	1.934 1.936	101.4	103.1	108.0	113.5	122.9	106.4	
Co(imidacloprid)₂(NCS)₂									
HF/ STO-3G	2.132 2.128	1.903 1.924	102.2	104.6	100.0	102.5	129.0	120.7	
B3LYP/ 6-31G(d, p)	2.141 2.141	1.923 1.923	100.4	106.6	100.4	106.6	126.9	117.3	
Co(thiamethoxam)₂(NCS)₂									
HF/ STO-3G	2.116 2.116	1.917 1.917	103.8	103.5	107.6	97.7	147.1	124.1	
B3LYP/ 6-31G(d, p)	2.120 2.120	1.924 1.924	100.1	109.6	100.1	109.6	126.2	111.2	
Co(tebuconazole – via. N4)₂(NCS)₂									
HF/ STO-3G	2.084 2.084	1.933 1.933	109.9	101.2	109.2	101.2	123.8	110.8	
B3LYP/ 6-31G(d, p)	2.439 2.354	1.929 1.924	104.4	99.6	104.4	99.7	134.1	115.6	
Co(tebuconazole – via. N2)₂(NCS)₂									
HF/ STO-3G	2.158 2.158	1.924 1.925	112.3	111.5	111.9	111.0	116.5	91.0	
B3LYP/ 6-31G(d, p)	2.109 2.109	1.931 1.931	110.4	99.2	110.4	99.2	128.7	108.1	



Thiocyanate ligands are along the vertical axis (going into the plane of paper, depicted by dotted wedges) and organic moieties along the horizontal (coming out of the plane of paper, depicted by solid wedges). Cobalt – thiocyanate bonds are denoted as Co – N(X), and since there are two such bonds, Co – N(X¹), Co – N(X²) notations are used. The same applies to cobalt – organic ligand bonds i.e. Co – N(L): Co – N(L¹), Co – N(L²). Angle between the thiocyanate ligands is θ_5 , while that between organic ligands is θ_6 . Angles between thiocyanate and organic ligands are denoted by $\theta_1, \theta_2, \theta_3$ and θ_4 .

Irrespective of the choice of basis set/level of theory employed, the Co-N(X) type of bond lengths are found to be shorter than the Co-N(L) ones, and largest deviation (of about $> 9^\circ$) from expected tetrahedral value is observed for θ_5 . As mentioned earlier, crystal structures of several pyridine related ligands are deposited in the Cambridge Crystallographic Database. Since chloroquine, chlorphenamine and imidacloprid possess pyridine nucleus in their molecular frameworks, the cobalt environment in the newly proposed complexes was compared with reported data (summarized in Table V.4).

Table V.4. List of ligands forming tetrahedral complexes with cobalt thiocyanate and various parameters describing the cobalt environment in them.

Structures of compounds numbered 1 to 7: 1) 2-methylpyridine, 2) 2-chloropyridine, 3) 2-bromopyridine, 4) 3-methylpyridine, 5) 1,2-bis(2-pyridyl) ethylene, 6) trans-4-(4'-N(N-methyl-N-hydroxyethyl)amino styryl pyridine and 7) quinoline.									
Thiocyanate complex of ligand	CCDC No.	Bond lengths (Å)		Bond angles (degrees)					
		Co-N(L ¹), Co-N(L ²)	Co-N(X ¹), Co-N(X ²)	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
1	910141	2.048 2.042	1.939 1.954	113.9	104.9	111.2	105.2	109.8	111.7
*2 (α)	910145	2.050 2.044	1.932 1.937	115.4	103.3	113.2	105.3	109.0	110.7
*2 (β)	934076	2.035 2.035	1.932 1.932	104.3	110.4	104.3	110.4	107.3	119.4
3	910143	2.041 2.041	1.941 1.941	103.5	110.3	103.5	110.3	106.9	121.4
4	832182	2.021 2.021	1.937 1.933	106.6	106.6	107.8	107.8	119.5	107.9
5	935552	2.031 2.031	1.945 1.945	110.4	109.0	110.4	109.0	105.2	112.4
6	1194974	2.044 2.029	1.940 1.934	103.2	114.2	111.6	114.5	104.5	108.3
7	737174	2.015 2.015	1.939 1.939	109.6	104.0	109.6	104.0	118.7	110.9

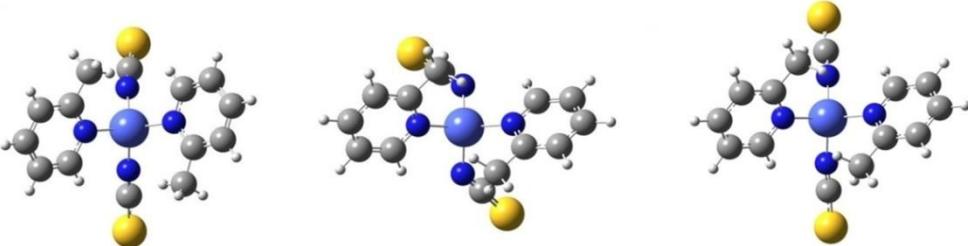
*Cobalt thiocyanate complex of 2-chloropyridine is found to crystallize in two forms - the α form crystallizes in $P2_12_12_1$ while the β form in $Pbcn$ space groups respectively. The former is denoted at 2 (α) and the latter 2 (β).

The Co-N(L¹)/Co-N(L²) and Co-N(X¹)/Co-N(X²) bond lengths lie within the range of 2.010 – 2.050 and 1.932 – 1.954 Å respectively. Angles θ_1 to θ_5 are found to deviate at the most by $\pm 6^\circ$ from expected tetrahedral value. Value of θ_6 is found to be greater than 119° when bulky substituents are present at *ortho*-positions in pyridine (observed in 2-bromo and 2 (β) complexes). The corresponding QM values are slightly on the higher side for Co-N(L) bonds and slightly shorter for Co-N(X) type of bonds. The value of θ_5 is also found to be less than 120° . These minor deviations are due to the fact that calculations are performed on single monomeric units, not including the presence of other such units in the neighbourhood. To confirm this hypothesis, geometry optimizations were performed on monomeric crystal structure geometries of Co(2-methylpyridine)₂(NCS)₂, Co(2-chloropyridine)₂(NCS)₂ and Co(quinoline)₂(NCS)₂ complexes. Details about cobalt-ligand bond lengths and angles of the former are presented in Table V.5, and those of the latter two are available in APPENDIX, section A11. It is seen that all QM geometries have slightly larger Co-N(L) and slightly shorter Co-N(X) bonds with value of θ_5 greater than 118° ; trends which match with newly proposed complexes of chloroquine, chlorphenamine and imidacloprid.

From the study carried out by Darby and Vallario [20], tetrahedral complexes of several pyridine derivatives were found to be intact, with no change in coordination number from 4 to 6 when pyridine had a substituent at its *ortho*-position. Pyridine nucleus in chlorphenamine and imidacloprid molecules possesses this feature, thereby contributing to the stabilities of newly proposed complexes in solutions. This phenomenon has also been observed in smaller aromatic systems like imidazole and its derivatives [21]. The thiazole ring in thiamethoxam coordinates to the cobalt ion via nitrogen LP whose adjacent carbon possesses a halogen substituent. The same can be

observed if triazole ring in tebuconazole coordinates through LP of nitrogen numbered 2 (refer Table V.2 for numbering scheme of 1,2,4-triazole unit).

Table V.5. Comparison between cobalt environment in experimental and theoretical structures of Co(2-methylpyridine)₂(NCS)₂.



Monomeric units (left to right) from: crystal structure [4], HF/STO-3G and DFT calculations.

Geometry from	Bond lengths (Å)		Angles (degrees)					
	Co-N(L ¹)/ Co-N(L ²)	Co-N(X ¹)/ Co-N(x ²)	θ ₁	θ ₂	θ ₃	θ ₄	θ ₅	θ ₆
Crystal structure	2.048 2.042	1.939 1.954	113.9	104.9	111.2	105.2	109.8	111.7
HF/STO-3G calculation	2.123 2.123	1.935 1.935	105.9	104.2	105.9	104.3	122.5	114.6
B3LYP/6-31G** calculation	2.124 2.124	1.936 1.936	107.0	102.1	107.1	102.4	125.3	113.2

The corresponding minimum energy structure of the coordination compound and bond lengths/bond angles of ligands with cobalt centre are displayed in Figure V.3 and Table V.3 (entry 8) respectively. Based on DFT calculations, there is no much difference between the coordination compounds of tebuconazole occurring through N2 and N4 LPs (refer Table V.2 for numbering scheme of 1,2,4-triazole unit in tebuconazole molecule; energy difference is just about 1 kcal/mol, giving rise to another probable structure).

The purple colour observed for ketamine on reaction with cobalt thiocyanate is believed to occur due to the formation of an octahedral type of complex wherein ketamine behaves as a bidentate ligand. This theoretical study however is yet to be undertaken.

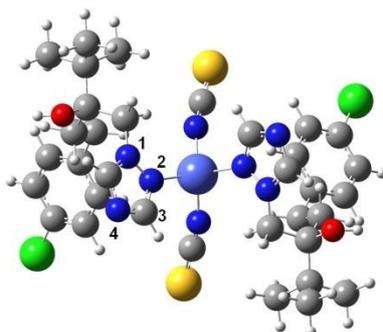


Figure V.3. Coordination compound of cobalt with tebuconazole via. LP of nitrogen labelled 2 of 1,2,4-triazole unit obtained from HF calculation. The DFT structure is similar, hence not displayed.

V.2. Theoretical investigation on 3-aryl-6-ethoxycarbonyl-4-hydroxy-2H-pyran-2-ones

Ring formations are important and commonly occurring processes in organic chemistry. Some of the synthetic routes used to construct rings involve intramolecular cyclizations (either by enolate/radical species) [22-23], pericyclic reactions [24] and cycloaromatization processes [25]. The formation of pyran-2-one framework in ‘3-aryl-6-ethoxycarbonyl-4-hydroxy-2H-pyran-2-ones’ class of compounds takes place starting from a triester compound, which on treatment with a weak base like triethylamine cyclizes (through enolate species) to form the required products (reaction mechanism shown in Figure V.4). Though there are two processes possible: the first – formation of pyran-2-one structure and, the second – formation of furan-5-one moiety, the former route undertaken by the enolate species must therefore be lower ‘energy-demanding’ thereby accounting experimental findings [26]. This study is undertaken in the upcoming section via. DFT calculations. The enolate species formed after abstraction of acidic hydrogen can exist in two possible structures: the ‘E’ form (where ‘E’ stands for *Entgegen*; this terminology is used when bulky groups are across carbon – carbon double bond) and the ‘Z’ form (‘Z’ for *Zusammen*, which means *together*) [27].

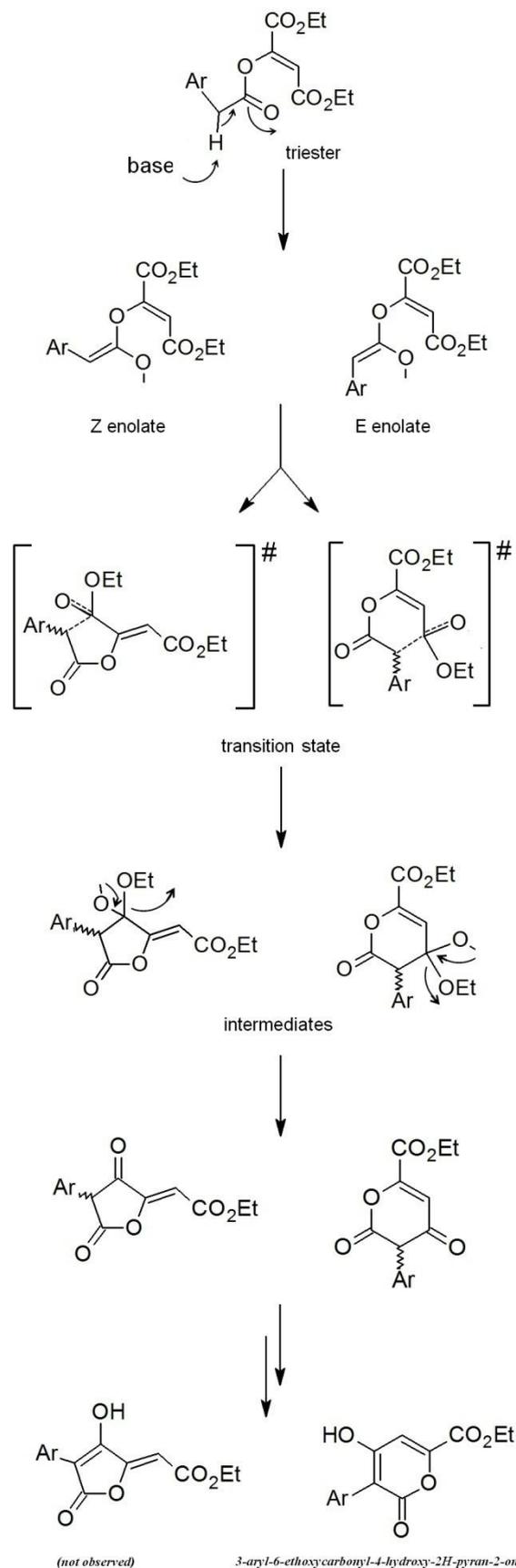
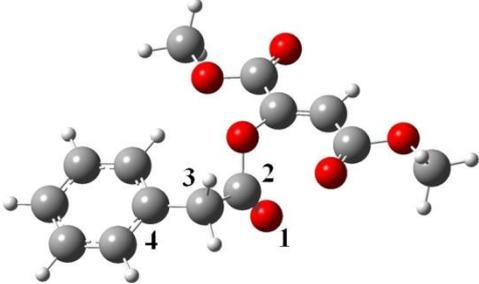
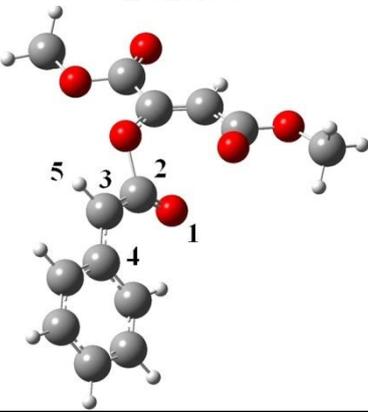
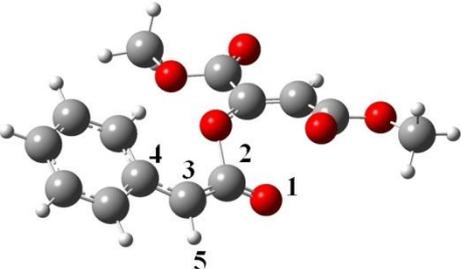
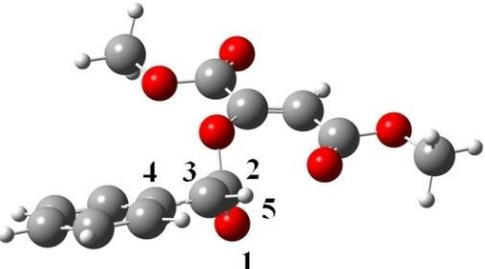


Figure V.4. Reaction mechanism leading to the formation of 3-aryl-6-ethoxycarbonyl-4-hydroxy-2H-pyran-2-one and its ‘experimentally not observed’ isomer.

Table V.6. Optimized geometries and some important structural parameters of triester compound, E/Z enolates and transition state connecting E and Z species. Dihedral angle between atoms numbered i, j, k and l is designated as D(i, j, k, l).

Structure	Geometric parameters
<p>triester with methy groups</p> 	$D(1, 2, 3, 4) = 93.8^\circ$
<p>E enolate</p> 	$D(1, 2, 3, 4) = 0.1^\circ$ $D(1, 2, 3, 5) = 179.0^\circ$ Expected values: $D(1, 2, 3, 4) = 0.0^\circ$ $D(1, 2, 3, 5) = 180.0^\circ$
<p>Z enolate</p> 	$D(1, 2, 3, 4) = 172.7^\circ$ $D(1, 2, 3, 5) = 5.2^\circ$ Expected values: $D(1, 2, 3, 4) = 180.0^\circ$ $D(1, 2, 3, 5) = 0.0^\circ$
<p>E-Z transition state</p> 	$D(1, 2, 3, 4) = 86.9^\circ$ $D(1, 2, 3, 5) = -91.4^\circ$ Expected values: $D(1, 2, 3, 4) = 90.0^\circ$ $D(1, 2, 3, 5) = -90.0^\circ$

In present context, bulky groups correspond to the phenyl ring and oxygen bonded to an alkene unit bearing two ester groups. In the first step of the investigation, minimum energy structure of the starting triester compound was determined. To save computational time, the ethyl groups in corresponding ester units were replaced with methyl, the aryl group was a simple phenyl ring and calculations were performed at the B3LYP/3-21G level of theory.

Optimized geometry of the triester compound is displayed in Table V.6, along with those of E, Z enolates. Both enolate species are interconvertible, i.e. the phenyl group must undergo an out-of-plane rotation to form either of them. [The plane under consideration contains atoms numbered from 1 to 3 in the E enolate structure displayed in Table V.6. Atoms 4 and 5 are found to be almost planar (based on corresponding dihedral values) while in the Z enolate, slight deviations are observed for dihedral angles D(1, 2, 3, 4) and D(1, 2, 3, 5) from expected values. In the transition state geometry, these atoms are found to be perpendicular.] The amount of energy required for E enolate to form the Z species is about 58.4 kcal/mol (or 54.6 kcal/mol for the reverse process).

Energy profile diagram for the cyclization process of E/Z enolates is displayed in Figure V.5a. The transition state structures by which E/Z enolates form corresponding pyran ring skeletons are labelled 6E-TS and 6Z-TS respectively. Corresponding 5E-TS and 5Z-TS geometries represent transition state structures leading to the formation of furan rings. The intermediates formed are labelled 6EI, 6ZI, 5EI and 5ZI. Based on the energy values displayed in Figure V.5a, the E enolate is found to favour the formation of pyran-2-one structure while the Z enolate favours the isomeric furan-5-one ring. The reaction therefore is found to proceed via E enolate species.

Transition state geometries of 6E-TS and 5E-TS are shown in Figure V.5b. The phenyl ring in both is found to be in equatorial position (in 6Z-TS and 5Z-TS geometries, the phenyl group is in axial orientation).

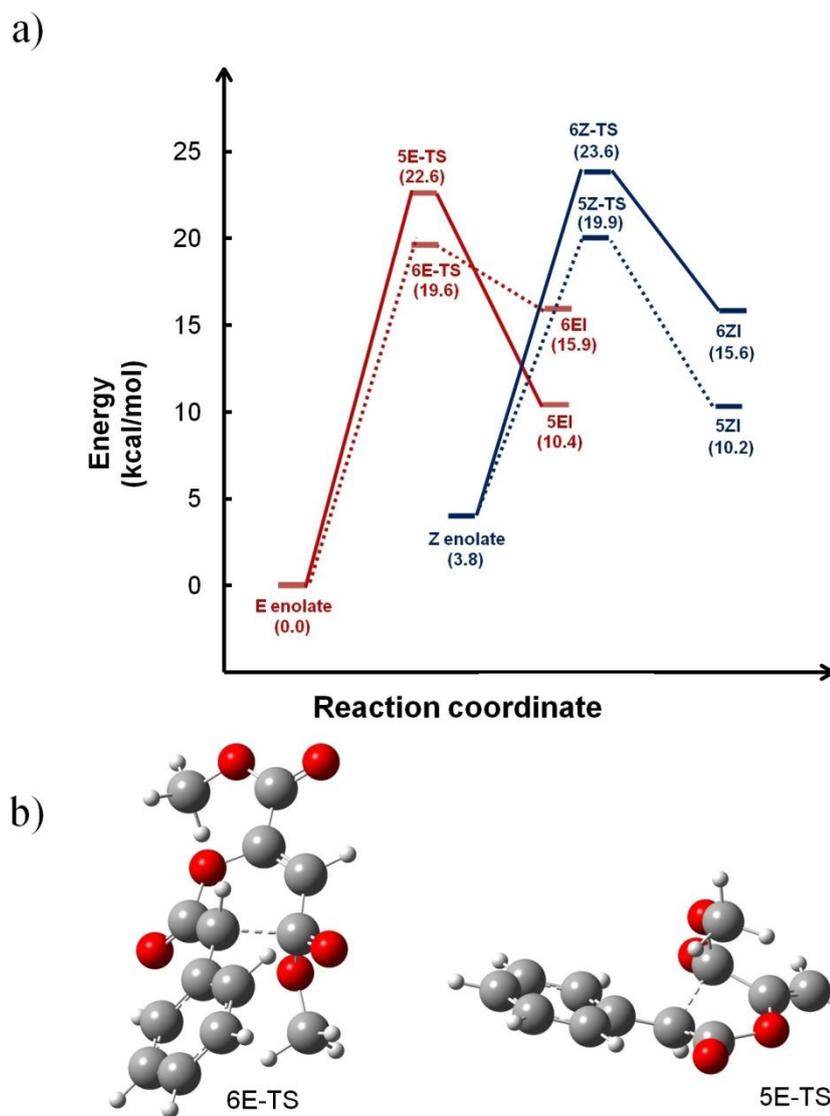
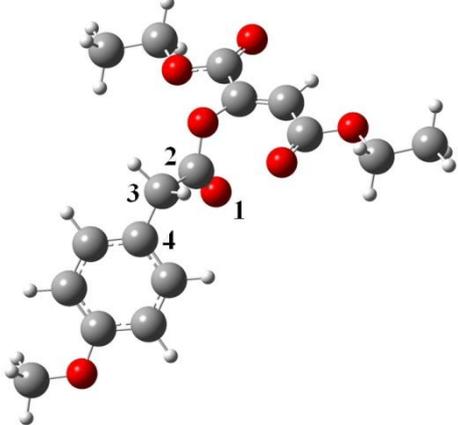
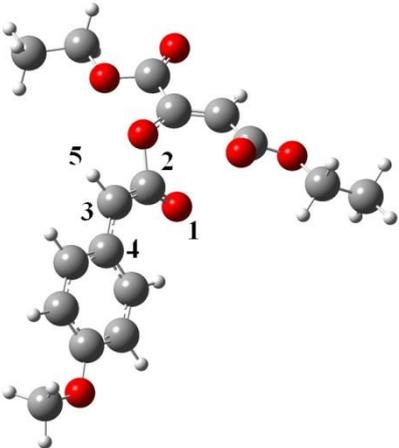
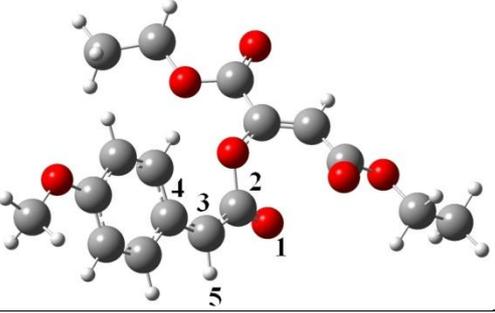
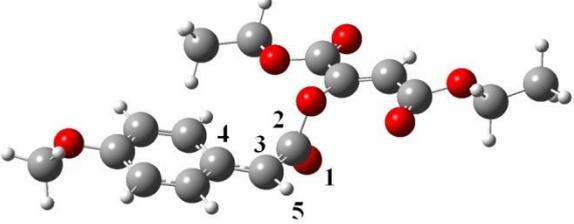


Figure V.5. a) Energy profile diagram for cyclization pathways available for E/Z enolates (lower energy ones denoted by dotted lines, and those higher are shown by continuous lines) and b) transition state structures of 6E-TS and 5E-TS wherein the phenyl group is in equatorial position. Similar geometries are observed for 6Z-TS and 5Z-TS, the only difference being that the phenyl group is in axial position.

Table V.7. Optimized geometries and some important structural parameters of triester compound, E/Z enolates and transition state connecting E and Z species.

Structure	Geometric parameters
triester with ethyl and <i>para</i> -methoxy groups 	$D(1, 2, 3, 4) = 18.1^\circ$
E enolate 	$D(1, 2, 3, 4) = 0.4^\circ$ $D(1, 2, 3, 5) = 179.0^\circ$ Expected values: $D(1, 2, 3, 4) = 0.0^\circ$ $D(1, 2, 3, 5) = 180.0^\circ$
Z enolate 	$D(1, 2, 3, 4) = 175.8^\circ$ $D(1, 2, 3, 5) = 2.1^\circ$ Expected values: $D(1, 2, 3, 4) = 180.0^\circ$ $D(1, 2, 3, 5) = 0.0^\circ$
E-Z transition state 	$D(1, 2, 3, 4) = 99.1^\circ$ $D(1, 2, 3, 5) = 98.3^\circ$ Expected values: $D(1, 2, 3, 4) = 90.0^\circ$ $D(1, 2, 3, 5) = 90.0^\circ$

In the second phase of investigation, the basis set is upgraded to 6-31G(d,p), ethyl ester groups are considered and the phenyl ring contains methoxy group at *para*-position. The lowest energy conformer of the triester compound, structures of corresponding E/Z enolates and the transition state connecting them are shown in Table V.7. The geometries appear to be similar to those presented in Table V.6; while the energy profile diagram is shown in Figure V.6. It is seen that:

- The Z enolate is higher in energy than the E species by ~ 3.9 kcal/mol.
- The E enolate favors formation of 6-membered ring structure while the Z enolate shows opposite trend.
- The energy difference between the 6-membered and 5-membered transition states in the former case is quite substantial (~ 4.3 kcal/mol) as compared to the latter case (~ 1.3 kcal/mol). This accounts for the reaction selectivity.
- The interconversion process from E to Z species requires energy of about 26 – 27 kcal/mol which is higher than the energy needed to form the pyran ring.

Thus the overall reaction is found to proceed via E enolate. The investigation carried out using basic model of the triester compound (with simple phenyl ring and methyl esters) also leads to the same conclusion. [The transition state geometries of 6E'-TS and 5E'TS are similar to those shown for 6E-TS and 5E-TS in Figure 5.5b and therefore are not presented in this chapter. Similar structural relations exist between the transition state geometries 6Z'-TS and 5Z'TS with 6Z-TS and 5Z-TS.]

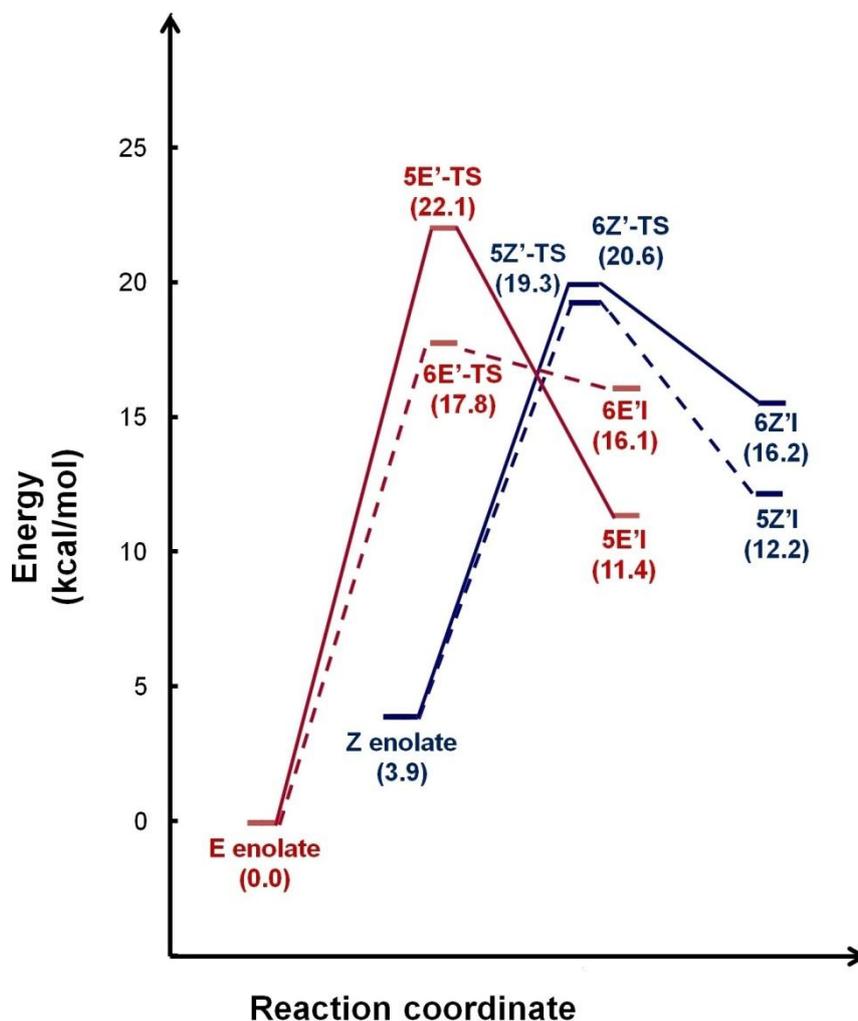


Figure V.6. Energy profile diagram of for E/Z enolates bearing methoxy group the on phenyl ring. Lower energy pathways are shown by dotted lines. The transition states are denoted by 6E'-TS, 5E'-TS, 6Z'-TS and 5Z'-TS. Intermediates formed from respective transition states are 6E'-I, 5E'-I, 6Z'-I and 5Z'-I.

V.3 Concluding remarks

To conclude, the blue colour formed by reaction of cobalt thiocyanate with chloroquine, chlorphenamine, imidacloprid, mephedrone, pretilachlor, thiamethoxam and tebuconazole is attributed to formation of tetrahedral CoL_2X_2 complex. The reactive sites of drug molecules containing several nitrogen atoms have been explored using MESP and based on the most negative nitrogen LP sites, minimum

energy structures of corresponding cobalt complexes have successfully been obtained. Ketamine however is found to form purple colour, which is attributed to $\text{Co}(\text{ketamine})_2(\text{NCS})_2$ complex wherein ketamine binds in bidentate fashion. Looking at the structure of mephedrone, it is also capable of behaving as a bidentate ligand, however it functions as a monodentate in present case. It would be interesting to investigate the reasons for such preferences.

With respect to the formation of 3-aryl-6-ethoxycarbonyl-4-hydroxy-2*H*-pyran-2-ones, the corresponding E enolate species is found to form the pyran scaffold which eventually leads to the experimentally observed product.

References:

1. L.J. Scott Jr., Specific field test for cocaine, *Microgram*, 6 (1973) 179.
2. J.A. Morris, Modified Cobalt Thiocyanate Presumptive Color Test for Ketamine Hydrochloride, *J. Forensic. Sci.*, 52 (2007) 84.
3. Y. Tsumura, T. Mitome and S. Kimoto, False positives and false negatives with a cocaine-specific field test and modification of test protocol to reduce false decision, *Forensic Science International*, 155 (2005) 158 and reference within.
4. S.R. Kote, M.R. Bhandarkar, N.L. Chutke, Use of Cobalt Thiocyanate for Thin-Layer Chromatographic Detection of Neonicotinoid Insecticide Imidacloprid, *J. Planar Chromatogr.*, 28 (2015) 352.
5. S. Wöhlert, I. Jess, U. Englert and C. Näther, Synthesis and crystal structures of Zn(II) and Co(II) coordination compounds with ortho substituted pyridine ligands: two structure types and polymorphism in the region of their coexistence, *CrystEngComm*, 15 (2013) 5326.
6. J. Boeckmann, B. Reimer and C. Näther, Synthesis, Crystal Structures, and Spectroscopic and Thermal Properties of New Cobalt Thiocyanato Coordination Compounds Based on 3-Methylpyridine as a Neutral Coligand, *Z. Naturforsch.*, 66 (2011) 819.
7. H. Sugiyama, A. Sekine and H. Uekusa, Crystal Structure of Bis(4-aminopyridine)bis(isothiocyanato)cobalt(II), *X-Ray Struct. Anal. Online*, 31 (2015) 27.
8. S. Wohlert, I. Jess and C. Nather, [Co(NCS)₂(1,2-bis(2-pyridyl)ethylene)]_n: A new 1D coordination polymer that shows a metamagnetic transition, *Polyhedron*, 63 (2013) 21.

9. L. Kong, W. Li, X. Li, W. Geng, F. Hao, J. Wu, H. Zhou, J. Yang, Y. Tian and B. Jin, Four divalent metal thiocyanate coordination compounds containing a rigid functional pyridine ligand, *Polyhedron*, 29 (2010) 1575.
10. A. Mirčeva, Structure of Cobalt Diquinoline Diisothiocyanate, *Acta Cryst. C*, 46 (1990) 1001.
11. X. Chen, S. Liu, X. You, H. Fun, K. Chinnakali and I.A. Razak, Bis(3,5-dimethylpyrazole-N2)bis(thiocyano-N)cobalt(II) monohydrate, *Acta Cryst. C*, 55 (1999) 22.
12. A.M.G.D. Rodrigues, Y.P. Mascarenhas and M.M. Rodrigues, Bis(3,5-dimethyl-1-phenylpyrazole)bis(isothiocyanato)cobalt(II), *Acta Cryst. B*, 36 (1980) 159.
13. V.M. Leovac, Z.D. Tomic, K.M. Szecsenyi, L.S. Jovanovic and M.D. Joksovic, Transition metal complexes with pyrazole based ligands, Part 27, Structural and thermal characterization of cobalt(II) halide and pseudohalide complexes with 4-acetyl-3-amino-5-methylpyrazole, *J. Serb. Chem. Soc.*, 72 (2007) 1281.
14. Z. Zhongqiang, D. Di, Z. Jingshuo, W. Lili, Z. Qinlei, X. Yanqing and H. Rudan, Synthesis, Structures and Properties of Two Complexes Based on Benzimidazole or Benzothiazole Ligand, *Chem. Res. Chin. Univ.*, 30 (2014) 185.
15. H. Zhao, X. Wang, E. Yang and X. Zhao, Synthesis and Crystal Structure of a Mixed-Ligand Cobalt(II) Complex with 2-Aminobenzimidazole, *J. Chem. Crystallogr.*, 38 (2008) 625.
16. R. Zhuang, F. Jian, K. Wang, Synthesis, Characterization and performance of two new Co(II) complexes as electrocatalyst of hydrogen peroxide and trichloroacetic acid, *J. Mol. Struct.*, 938 (2009) 254.

17. C. Cai, Y. Tian, Y. Li and X. You, The mononuclear cobalt(II) complex $\text{Co}^{\text{II}}(\text{DMBDIZ})_2(\text{NCS})_2$, where DMBDIZ is 2,6-dimethylbenzo-[1,2-d:4,5-d']diimidazole, *Acta Cryst. C*, 58 (2002) 459.
18. C. Lee, W. Yang and R.G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Phys. Rev. B*, 37 (1998) 785.
19. A.D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, *Phys. Rev. A*, 38 (1988) 3098.
20. W.L. Darby and L.M. Vallarino, Stereochemical Trends for Complexes of Cobalt(II) Halides and Thiocyanate with A Series of Mono-, Di- and Tri-methylpyridines, *Inorganica Chim. Acta*, 48 (1981) 215.
21. A. Maslejova, S. Uhrinova, J. Mrozinski, B. Zurowska, M.C. Munoz and M. Julve, Study of the mutual influence of ligands in cobalt(II) complexes containing thiocyanate and imidazole derivatives, *Inorganica Chim. Acta*, 255 (1997) 343.
22. J.E. Baldwin and M.J. Lusch, Rules for Ring Closure: Application to Intramolecular Aldol Condensations in Polyketonic Substrates, *Tetrahedron*, 38 (1982) 2939.
23. A.L.J. Beckwith, C.J. Easton, A.K. Serelis, Some Guidelines for Radical Reactions, *J.C.S. Chem. Comm.*, (1980) 482.
24. E.M. Greer and C.V. Cosgriff, Reaction mechanisms: pericyclic reactions, *Annu. Rep. Prog. Chem., Sect. B: Org. Chem.*, 109 (2013) 328.
25. R.K. Mohamed, P.W. Peterson, I.V. Alabugin, Concerted Reactions That Produce Diradicals and Zwitterions: Electronic, Steric, Conformational, and Kinetic Control of Cycloaromatization Processes, *Chem. Rev.*, 113 (2013) 7089.

26. G.R. Dhage, S.R. Thopate, S.N. Ramteke and P.P. Kulkarni, One-pot synthesis and evaluation of novel 3- aryl-6-ethoxycarbonyl-4-hydroxy-2H-pyran-2-one as a potent cytotoxic agent, RSC Adv., 4 (2014) 56870.
27. D. Nasipuri, Stereochemistry of Organic Compounds: Principles and Applications, New Age International (P) Limited, Publishers, Delhi, 2010.

APPENDIX

Table of Contents		Page
A1	Solution to the particle and wave equations	153
A2	Schrödinger's time-dependent equation	154
A3	Schrödinger's time-independent equation	156
A4	Hamiltonian operator for many-particle system	157
A5	Drawback of Hartree Product	158
A6	Slater determinant and antisymmetry principle	159
A7	Cartesian co-ordinates of acetone and acetone-related non-covalent complexes	160
A8	Numerical method to locate critical of a one-dimensional function via. interpolation	162
A9	The three-dimensional Lagrange interpolation method	168
A10	Guess points for topographical features of MESP scalar field	169
A10.1	Ammonia molecule	169
A10.2	Water molecule	170
A10.3	Hydrogen sulphide molecule	171
A10.4	Formaldehyde molecule	171
A10.5	Ethene molecule	172
A10.6	Benzene molecule	174
A11	Potential due to <i>s</i> -type Gaussians	177
A12	Geometries of some cobalt thiocyanate complexes: comparison between experimental and quantum mechanical structures	180
A13	Certificates of participation/poster presentations of various conferences, workshops and symposiums attended during research period January 2014 – December 2019	181
A14	First page of research articles published during research period January 2014 – December 2019	182

A1. Solution to the particle and wave equations

The function $f(x) = e^{\frac{2\pi i}{\lambda}(x-ct)}$ is found to satisfy the wave equation (represented by Eq. (A1)). Substituting $f(x)$ in place of E yields Eq. (A2), and cancelling all quantities except λ and c results in Eq. (A3).

$$\frac{\partial^2 \mathbf{E}}{\partial x^2} = \left(\frac{1}{c^2}\right) \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (\text{A1})$$

$$e^{\frac{2\pi i}{\lambda}(x-ct)} \left(\frac{2\pi}{\lambda}\right)^2 i^2 = \left(\frac{1}{c^2}\right) e^{\frac{2\pi i}{\lambda}(x-ct)} \left(\frac{2\pi}{\lambda}\right)^2 i^2 c^2 \quad (\text{A2})$$

$$\left(\frac{1}{\lambda}\right)^2 = \left(\frac{1}{c^2}\right) \left(\frac{1}{\lambda}\right)^2 c^2 \quad (\text{A3})$$

Replacing (the numerator term) c by λv in the right hand side of Eq. (A3) gives Eq. (A4) followed by Eq. (A5) on multiplication with h^2 .

$$\left(\frac{1}{\lambda}\right)^2 = \left(\frac{1}{c^2}\right) v^2 \quad (\text{A4})$$

$$\left(\frac{h}{\lambda}\right)^2 = \frac{h^2 v^2}{c^2} \quad (\text{A5})$$

Since p equals h/λ , the left hand side term of Eq. (A4) is equivalent to p^2 . The numerator in the right hand side of the above equation can be replaced by E^2 as $E = hv$. Based on these results, Eq. (A5) comes out to be $p^2 = E^2/c^2$. It is therefore shown that $f(x) = e^{\frac{2\pi i}{\lambda}(x-ct)}$ satisfies the particle equation.

A2. Schrödinger's time-dependent equation

The wave equation for electrons is unknown but the solution $\Psi(x, t)$ has the form $\Psi(x, t) = e^{\frac{2\pi i}{\lambda}(x - ct)}$. Double differentiation of $\Psi(x, t)$ with respect to variable x gives Eq. (A6) which on multiplication by h^2 (on both sides) and rearrangements yields Eq. (A7).

$$\frac{\partial^2 \Psi(x, t)}{\partial x^2} = \frac{i^2 (2\pi)^2}{\lambda^2} \Psi(x, t) = -\frac{(2\pi)^2}{\lambda^2} \Psi(x, t) \quad (\text{A6})$$

$$\left(-\frac{h^2}{(2\pi)^2}\right) \frac{\partial^2 \Psi(x, t)}{\partial x^2} = \left(\frac{h^2}{\lambda^2}\right) \Psi(x, t) = p^2 \Psi(x, t) \quad (\text{A7})$$

Differentiating $\Psi(x, t)$ with respect to variable t results in Eq. (A8) which on multiplication by h gives Eq. (A9).

$$\frac{\partial \Psi(x, t)}{\partial t} = -\frac{i2\pi c}{\lambda} \Psi(x, t) = -2\pi i v \Psi(x, t) \quad (\text{A8})$$

$$\left(-\frac{h}{2\pi i}\right) \frac{\partial \Psi(x, t)}{\partial t} = h v \Psi(x, t) = E \Psi(x, t) \quad (\text{A9})$$

Since the constant $h/2\pi$ is found to occur frequently in above equations, it is given its own unique symbol \hbar (h-bar) and with this substitution, equations (A7) and (A9) become (A10) and (A11) respectively.

$$(-\hbar^2) \frac{\partial^2 \Psi(x, t)}{\partial x^2} = p^2 \Psi(x, t) \quad (\text{A10})$$

$$\left(-\frac{\hbar}{i}\right) \frac{\partial \Psi(x, t)}{\partial t} = E \Psi(x, t) \quad (\text{A11})$$

Substituting the left hand side expressions for $E \Psi(x, t)$ and $p^2 \Psi(x, t)$ from Eqs. (A10) and (A11) in Eq. (A12) results in the time-dependent Schrödinger's equation i.e. Eq. (A13).

$$E \Psi(x, t) = \frac{p^2}{2m} \Psi(x, t) + V \Psi(x, t) \quad (\text{A12})$$

$$\left(-\frac{\hbar}{i}\right) \frac{\partial \Psi(x, t)}{\partial t} = \left(-\frac{\hbar^2}{2m}\right) \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V \Psi(x, t) \quad (\text{A13})$$

A3. Schrödinger's time-independent equation

Making the substitution $\psi(x)f(t)$ for $\Psi(x, t)$ in Eq. (A13) and dividing throughout by $\psi(x)f(t)$ yields Eq. (A14), whose left hand side term is independent of variable x while the right hand side is independent of variable t .

$$\left(-\frac{\hbar}{i}\right)\frac{1}{f(t)}\frac{df(t)}{dt} = \left(-\frac{\hbar^2}{2m}\right)\frac{1}{\psi(x)}\frac{d^2\psi(x)}{dx^2} + V(x) \quad (\text{A14})$$

Since the left and right sides in above equation are independent of x and t respectively, it must therefore be a constant (chosen to be) E . Equating E to the left hand side of Eq. (A14), followed by integration leads to Eq. (A15), which can be rewritten as $f(t) = e^{\text{constant}} e^{-\frac{iEt}{\hbar}}$. The term e^{constant} being a scalar can be included in $\psi(x)$ as it is multiplied to $f(t)$ and therefore be omitted to yield Eq. (A16).

$$\ln f(t) = -\frac{iEt}{\hbar} + \text{constant} \quad (\text{A15})$$

$$f(t) = e^{-\frac{iEt}{\hbar}} \quad (\text{A16})$$

Equating E to right hand side of Eq. (A14) followed by rearranging the term $\psi(x)$ results to Eq. (A17) which corresponds to the time-independent equation for a particle moving in one-dimension with mass m .

$$\left(-\frac{\hbar^2}{2m}\right)\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \quad (\text{A17})$$

A4. Hamiltonian operator for many-particle system

The various terms in the Hamiltonian operator \hat{H} are written in Eq. (A18). The kinetic energy of all nuclei $\{\hat{T}_N\}$ and electrons $\{\hat{T}_e\}$ are represented by equations (A19) and (A20) where m_A corresponds to mass of atom A in the former equation.

$$\hat{H} = \hat{T}_N + \hat{T}_e + V_{NN} + V_{EE} + V_{EN} \quad (\text{A18})$$

$$\hat{T}_N = -\frac{\hbar^2}{2} \sum_A^N \frac{\nabla_A^2}{m_A} \quad (\text{A19})$$

$$\hat{T}_e = -\frac{\hbar^2}{2m_e} \sum_{i=1}^M \nabla_i^2 \quad (\text{A20})$$

The third term in Eq. (A18) represents the nuclear-nuclear repulsion between atoms, defined by Eq. (A21) where R_{AB} is the distance separating atoms A and B. The fourth term, V_{EE} (defined by Eq. (A22)) corresponds to the electron-electron repulsions where r_{ij} is the distance between the i^{th} and j^{th} electron. The last quantity V_{EN} accounts for the attraction between electrons and nuclei where r_{Ai} is the distance separating the interacting species (i.e. Eq. (A23)).

$$V_{NN} = \frac{e^2}{4\pi\epsilon_0} \sum_A^N \sum_{\substack{B \\ B \neq A}}^N \frac{Z_A Z_B}{R_{AB}} \quad (\text{A21})$$

$$V_{EE} = \frac{e^2}{4\pi\epsilon_0} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^i \frac{1}{r_{ij}} \quad (\text{A22})$$

$$V_{EN} = -\frac{e^2}{4\pi\epsilon_0} \sum_A^N \sum_i^M \frac{Z_A}{r_{Ai}} \quad (\text{A23})$$

A5. Drawback of Hartree Product

The drawback of Hartree Product is demonstrated by considering a simple case of two electrons. The corresponding wave function (WF) is given in Eq. (A24). Interchanging \mathbf{x}_1 with \mathbf{x}_2 and vice-versa results in Eq. (A25), which is different from the expected result i.e. Eq. (A26).

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \chi_1(\mathbf{x}_1) \chi_2(\mathbf{x}_2) \quad (\text{A24})$$

$$\psi(\mathbf{x}_2, \mathbf{x}_1) = \chi_1(\mathbf{x}_2) \chi_2(\mathbf{x}_1) \quad (\text{A25})$$

$$\psi(\mathbf{x}_2, \mathbf{x}_1) = -\chi_1(\mathbf{x}_1) \chi_2(\mathbf{x}_2) \quad (\text{A26})$$

A6. Slater determinant and antisymmetry principle

Writing the WF in the form of Slater determinant is found to satisfy the antisymmetry principle. As in section A5, the two-electron case has been considered where the corresponding Slater determinant is given in Eq. (A27).

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) \end{vmatrix} = \frac{1}{\sqrt{2!}} [\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2) - \chi_2(\mathbf{x}_1)\chi_1(\mathbf{x}_2)] \quad (\text{A27})$$

Interchanging \mathbf{x}_1 and \mathbf{x}_2 in Eq. (A27) results to Eq. (A28) wherein the sign of $\psi(\mathbf{x}_1, \mathbf{x}_2)$ is changed.

$$\begin{aligned} \psi(\mathbf{x}_2, \mathbf{x}_1) &= \frac{1}{\sqrt{2!}} [\chi_1(\mathbf{x}_2)\chi_2(\mathbf{x}_1) - \chi_2(\mathbf{x}_2)\chi_1(\mathbf{x}_1)] = \\ &= -\frac{1}{\sqrt{2!}} [-\chi_1(\mathbf{x}_2)\chi_2(\mathbf{x}_1) + \chi_2(\mathbf{x}_2)\chi_1(\mathbf{x}_1)] = -\psi(\mathbf{x}_1, \mathbf{x}_2) \end{aligned} \quad (\text{A28})$$

A7. Cartesian co-ordinates of acetone and acetone-related non-covalent complexes

Molecular geometry of acetone obtained from geometry optimization at the Hartree-Fock (HF) level of theory using 6-31G(d,p) basis set along with co-ordinates of V_{\min} are displayed in Table A1. Geometries of acetone-proton and acetone-hydrogen fluoride complexes at mentioned level of theory/basis set are displayed in Tables A2 and A3 respectively.

Table A1. Optimized geometry of acetone at HF/6-31G(d,p) level of theory. Co-ordinates of V_{\min} also displayed. All values are in atomic units.

Atom/site	X	Y	Z
C	0.0000	0.0000	0.3539
C	0.0000	±2.4335	-1.1481
H	±1.6525	±2.5192	-2.3624
H	0.0000	±4.0305	0.1254
O	0.0000	0.0000	2.6065
V_{\min}	0.0000	±1.7658	4.0484

Table A2. Optimized geometry (in atomic units) of acetone-proton complex at HF/6-31G(d,p) level of theory.

Molecule	Atoms	X	Y	Z
Acetone	C	0.0000	0.1383	0.0000
	C	-2.6147	-0.8600	0.0000
	H	-2.8654	-2.0543	±1.6514
	H	-3.9802	0.6562	0.0000
	C	2.2392	-1.5507	0.0000
	H	2.1429	-2.7680	±1.6505
	H	4.0077	-0.5236	0.0000
	O	0.2206	2.4999	0.0000
Proton	H	1.9056	3.1471	0.0000

Table A3. Optimized geometry (in atomic units) of acetone-hydrogen fluoride complex at HF/6-31G(d,p) level of theory.

Molecule	Atom	X	Y	Z
Acetone	C	0.0000	1.3078	0.0000
	C	0.2703	4.1460	0.0000
	H	-0.6524	4.9402	± 1.6529
	H	2.2482	4.6569	0.0000
	C	-2.6458	0.2413	0.0000
	H	-3.6654	0.9083	± 1.6524
	H	-2.5857	-1.7987	0.0000
	O	1.8308	-0.0313	0.0000
Hydrogen fluoride	H	1.4450	-3.4090	0.0000
	F	0.7928	-5.0073	0.0000

A8. Numerical method to locate CPs of a one-dimensional function via interpolation

This section describes the application of Lagrange interpolation method in detecting critical points (CPs) of a one-dimensional function whose tabular point data is presented in Table A4. For the benefit of readers, the function under investigation is presented in Eq. (A29), which possesses three CPs: global minimum at $x = -1$, local minimum at $x = 2$ and a local maximum at $x = 1$.

$$f(x) = \frac{x^4}{4} - \frac{2x^3}{3} - \frac{x^2}{2} + 2x - 2 \quad (\text{A29})$$

Table A4. Tabular points and corresponding function values (denoted as y_i) of a one-dimensional function between $x = -2$ and $x = 3$ with $\Delta x = 0.2$.

i	x_i	y_i	i	x_i	y_i
0	-1.9000	0.2257	13	0.7000	-1.0136
1	-1.7000	-1.4816	14	0.9000	-0.9270
2	-1.5000	-2.6094	15	1.1000	-0.9263
3	-1.3000	-3.2663	16	1.3000	-0.9956
4	-1.1000	-3.5516	17	1.5000	-1.1094
5	-0.9000	-3.5550	18	1.7000	-1.2323
6	-0.7000	-3.3563	19	1.9000	-1.3196
7	-0.5000	-3.0260	20	2.1000	-1.3170
8	-0.3000	-2.6250	21	2.3000	-1.1603
9	-0.1000	-2.2043	22	2.5000	-0.7760
10	0.1000	-1.8056	23	2.7000	-0.0810
11	0.3000	-1.4610	24	2.9000	1.0177
12	0.5000	-1.1927			

The search for CP is implemented in all sub-intervals (x_i, x_{i+1}) ; the Lagrange polynomial $L(x)$ constructed using tabular points $\{x_{i-1}, x_i, x_{i+1}, x_{i+2}\}$ and corresponding function values $\{y_{i-1}, y_i, y_{i+1}, y_{i+2}\}$ is presented in Eq. (A30). The Lagrange's coefficients (or simply referred to LCs): $L_{i-1}(x)$, $L_i(x)$, $L_{i+1}(x)$ and $L_{i+2}(x)$ are defined in equations (A31) to (A34) respectively.

$$L(x) = L_{i-1}(x) y_{i-1} + L_i(x) y_i + L_{i+1}(x) y_{i+1} + L_{i+2}(x) y_{i+2} \quad (\text{A30})$$

$$L_{i-1}(x) = \frac{(x - x_i)(x - x_{i+1})(x - x_{i+2})}{(x_{i-1} - x_i)(x_{i-1} - x_{i+1})(x_{i-1} - x_{i+2})} \quad (\text{A31})$$

$$L_i(x) = \frac{(x - x_{i-1})(x - x_{i+1})(x - x_{i+2})}{(x_i - x_{i-1})(x_i - x_{i+1})(x_i - x_{i+2})} \quad (\text{A32})$$

$$L_{i+1}(x) = \frac{(x - x_{i-1})(x - x_i)(x - x_{i+2})}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)(x_{i+1} - x_{i+2})} \quad (\text{A33})$$

$$L_{i+2}(x) = \frac{(x - x_{i-1})(x - x_i)(x - x_{i+1})}{(x_{i+2} - x_{i-1})(x_{i+2} - x_i)(x_{i+2} - x_{i+1})} \quad (\text{A34})$$

For the starting sub-interval (-1.9, -1.7), tabular point $\{x_{i-1}\}$ does not exist, therefore Eq. (A30) is modified to Eq. (A35). The corresponding LCs are defined by Eqs. (A36) to (A39).

$$L(x) = L_i(x) y_i + L_{i+1}(x) y_{i+1} + L_{i+2}(x) y_{i+2} + L_{i+3}(x) y_{i+3} \quad (\text{A35})$$

$$L_i(x) = \frac{(x - x_{i+1})(x - x_{i+2})(x - x_{i+3})}{(x_i - x_{i+1})(x_i - x_{i+2})(x_i - x_{i+3})} \quad (\text{A36})$$

$$L_{i+1}(x) = \frac{(x - x_i)(x - x_{i+2})(x - x_{i+3})}{(x_{i+1} - x_i)(x_{i+1} - x_{i+2})(x_{i+1} - x_{i+3})} \quad (\text{A37})$$

$$L_{i+2}(x) = \frac{(x - x_i)(x - x_{i+1})(x - x_{i+3})}{(x_{i+2} - x_{i-1})(x_{i+2} - x_i)(x_{i+2} - x_{i+3})} \quad (\text{A38})$$

$$L_{i+3}(x) = \frac{(x - x_i)(x - x_{i+1})(x - x_{i+2})}{(x_{i+3} - x_i)(x_{i+3} - x_{i+1})(x_{i+3} - x_{i+2})} \quad (\text{A39})$$

With the substitutions $\{x_i, x_{i+1}, x_{i+2}, x_{i+3}\} \equiv \{-1.9, -1.7, -1.5, -1.3\}$ and $\{y_i, y_{i+1}, y_{i+2}, y_{i+3}\} \equiv \{0.2257, -1.4816, -2.6094, -3.2663\}$, the function value for x between (-1.9, -1.7) can be interpolated via. Eq. (A40).

$$\begin{aligned}
L(x) = & (x + 1.7)(x + 1.5)(x + 1.3)(-4.70221) + (x + 1.9)(x + 1.5)(x + 1.3)(-92.6000) + \\
& (x + 1.9)(x + 1.7)(x + 1.3)(163.0875) + (x + 1.9)(x + 1.7)(x + 1.5)(-68.0479)
\end{aligned} \tag{A40}$$

The expression for $L'(x)$ (i.e. derivative for Eq. (A35)) is given in Eq. (A41) and the LC: $L'_i(x)$ is defined by equation (A42). Likewise, one can easily guess the expressions for $L'_{i+1}(x)$, $L'_{i+2}(x)$, $L'_{i+3}(x)$ based on Eqs. (A37), (A38) and (A39).

$$L'(x) = L'_i(x) y_i + L'_{i+1}(x) y_{i+1} + L'_{i+2}(x) y_{i+2} + L'_{i+3}(x) y_{i+3} \tag{A41}$$

$$L'_i(x) = \frac{(x - x_{i+2})(x - x_{i+3}) + (x - x_{i+1})(x - x_{i+2}) + (x - x_{i+1})(x - x_{i+3})}{(x_i - x_{i+1})(x_i - x_{i+2})(x_i - x_{i+3})} \tag{A42}$$

It is a well known fact that if some function of single variable $g(x)$ defined over an interval (a, b) is continuous and changes its signs at points $x = a$ and $x = b$, then there exist a point $x = c$ such that $g(c)$ equals zero. A CP therefore exists in any given interval (x_i, x_{i+1}) if the signs of $L'(x)$ at $x = x_i$ and $x = x_{i+1}$ are opposite. Coming back to the numerical example under discussion, the derivative expression of $L(x)$ in the region $(-1.9, -1.7)$ is given by Eq. (A43).

$$\begin{aligned}
L'(x) = & [(x + 1.5)(x + 1.3) + (x + 1.7)(x + 1.5) + (x + 1.7)(x + 1.3)] (-4.70221) + \\
& [(x + 1.5)(x + 1.3) + (x + 1.9)(x + 1.5) + (x + 1.9)(x + 1.3)] (-92.6000) + \\
& [(x + 1.7)(x + 1.3) + (x + 1.9)(x + 1.7) + (x + 1.9)(x + 1.3)] (163.0875) + \\
& [(x + 1.7)(x + 1.5) + (x + 1.9)(x + 1.7) + (x + 1.9)(x + 1.5)] (-68.0479)
\end{aligned} \tag{A43}$$

At $x = -1.9$ and $x = -1.7$ the value of $L'(x)$ equals -10.1662 and -6.9972 respectively. Since the sign of $L'(x)$ remains unchanged at respective tabular points, this

sub-interval does not contain a CP. Going to the next sub-interval (-1.7, -1.5), the $L(x)$ expression is constructed using tabular points $\{x_{i-1}, x_i, x_{i+1}, x_{i+2}\} \equiv \{-1.9, -1.7, -1.5, -1.3\}$ and respective function values. As tabular point $\{x_{i-1}\}$ exists, $L(x)$ has the form as per Eq. (A30). The values of $L'(x)$ at $x = -1.7$ and $x = -1.5$ are -6.9972 and -4.3712 respectively. Thus this sub-interval is also devoid of a CP. The change in sign of $L'(x)$ is detected in three sub-intervals (-1.1, -0.9), (0.9, 1.1) and (1.9, 2.1). These observations are summarized in Table A5.

Table A5. Sub-intervals (x_i, x_{i+1}) where value of $L'(x)$ computed at $x = x_i$ and $x = x_{i+1}$ via. Lagrange interpolation method is found to change signs. The corresponding data used for interpolation is available in Table A4.

Sub-interval (x_i, x_{i+1})	Value of $L'(x)$ at	
	x_i	x_{i+1}
(-1.1, -0.9)	-0.6552	0.5547
(0.9, 1.1)	0.2050	-0.1847
(1.9, 2.1)	-0.2652	0.3447

This paragraph shall briefly describe the process of locating CP in sub-interval (-1.1, -0.9). An initial guess point (GP) $x_{\text{guess}} = -1.0$ is generated which lies at the centre of (-1.1, -0.9). For chosen tolerance value ϵ of 10^{-13} , $L'(x_{\text{guess}})$ defined in Eq. (A44) equals 0.1409×10^{-6} , and $L'(x_{\text{guess}})$ is greater than ϵ .

$$\begin{aligned}
 L'(x) = & [(x + 0.9)(x + 0.7) + (x + 1.1)(x + 0.9) + (x + 1.1)(x + 0.7)] (68.0479) + \\
 & [(x + 0.9)(x + 0.7) + (x + 1.3)(x + 0.9) + (x + 1.3)(x + 0.7)] (-221.9750) + \\
 & [(x + 1.1)(x + 0.7) + (x + 1.3)(x + 1.1) + (x + 1.3)(x + 0.7)] (222.1875) + \\
 & [(x + 1.1)(x + 0.9) + (x + 1.3)(x + 1.1) + (x + 1.3)(x + 0.9)] (-69.9229)
 \end{aligned}
 \tag{A44}$$

A new guess $x_{\text{guess}} = -0.9999$ is obtained via. Newton-Raphson method and $L'(x_{\text{guess}})$ at this point equals 0.3689×10^{-15} . Since $L'(x_{\text{guess}}) < \varepsilon$, the found CP is characterized by the second order derivative of $L(x)$ – the general form given in Eq. (A45) and the corresponding LCs: $L''_{i-1}(x)$, $L''_i(x)$, $L''_{i+1}(x)$, $L''_{i+2}(x)$ defined in Eqs. (A46) to (A49). The CP $x = -0.9999$ represents a minimum based on the second derivative value computed via. Eq. (A50). The total number of CPs detected for data in Table A4, summarized in Table A6, are in good agreement with expected results.

$$L''(x) = L''_{i-1}(x)y_{i-1} + L''_i(x)y_i + L''_{i+1}(x)y_{i+1} + L''_{i+2}(x)y_{i+2} \quad (\text{A45})$$

$$L''_{i-1}(x) = \frac{2(x - x_i) + 2(x - x_{i+1}) + 2(x - x_{i+2})}{(x_{i-1} - x_i)(x_{i-1} - x_{i+1})(x_{i-1} - x_{i+2})} \quad (\text{A46})$$

$$L''_i(x) = \frac{2(x - x_{i-1}) + 2(x - x_{i+1}) + 2(x - x_{i+2})}{(x_i - x_{i-1})(x_i - x_{i+1})(x_i - x_{i+2})} \quad (\text{A47})$$

$$L''_{i+1}(x) = \frac{2(x - x_{i-1}) + 2(x - x_i) + 2(x - x_{i+2})}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)(x_{i+1} - x_{i+2})} \quad (\text{A48})$$

$$L''_{i+2}(x) = \frac{2(x - x_{i-1}) + 2(x - x_i) + 2(x - x_{i+1})}{(x_{i+2} - x_{i-1})(x_{i+2} - x_i)(x_{i+2} - x_{i+1})} \quad (\text{A49})$$

$$\begin{aligned} L''(x) = & [(x + 1.1) + (x + 0.9) + (x + 0.7)](136.0958) + [(x + 1.3) + (x + 0.9) + \\ & (x + 0.7)](-443.9500) + [(x + 1.3) + (x + 1.1) + (x + 0.7)](444.3750) + \\ & [(x + 1.3) + (x + 1.1) + (x + 0.9)](-139.8458) \end{aligned} \quad (\text{A50})$$

Table A6. Summary of CPs obtained via. Lagrange interpolation method for data available in Table A4.

CP x	Value of		
	$L(x)$	$L'(x)$	$L''(x)$
-0.9999	-3.5836	$< 10^{-13}$	6.0494
1.0001	-0.9169		-1.9486
2.0001	-1.3336		3.0510

The sub-intervals $(2.5, 2.7)$ and $(2.7, 2.9)$ do not have tabular points $\{x_{i+2}\}$ and $\{x_{i+1}\}$ in the latter case. Thus if n_x represents the total number of tabular points in X direction then $L(x)$ in these two cases are constructed using tabular points $\{x_{n_x-3}, x_{n_x-2}, x_{n_x-1}, x_{n_x}\}$ and respective function values.

A9. The three-dimensional Lagrange interpolation method

The Lagrange polynomial used for interpolation in three dimensions is given by Eq. (A51) wherein the LCs: $L_i(x)$, $L_j(y)$ and $L_k(z)$ are constructed using tabular points $\{x_{i-1}, x_i, x_{i+1}, x_{i+2}\}$, $\{y_{j-1}, y_j, y_{j+1}, y_{j+2}\}$ and $\{z_{k-1}, z_k, z_{k+1}, z_{k+2}\}$. The scalar field values at points bearing co-ordinates $(x_{i-1}, y_{j-1}, z_{k-1})$, (x_{i-1}, y_{j-1}, z_k) , $(x_{i-1}, y_{j-1}, z_{k+1})$, $(x_{i-1}, y_{j-1}, z_{k+2})$, (x_{i-1}, y_j, z_{k-1}) , ..., $(x_{i+2}, y_{j+2}, z_{k+2})$ are denoted as f_{111} , f_{112} , f_{113} , f_{114} , f_{121} , ..., f_{444} respectively. The first order partial derivatives of (A51) are given by Eqs. (A52), (A53) and (A54). Constructions of LCs: $L'_i(x)$, $L'_j(y)$ and $L'_k(z)$ are discussed in APPENDIX, section A7.

$$L(x, y, z) = \sum_{i=1}^4 \sum_{j=1}^4 \sum_{k=1}^4 L_i(x) L_j(y) L_k(z) f_{ijk} \quad (\text{A51})$$

$$\frac{\partial L(x, y, z)}{\partial x} = \sum_{i=1}^4 \sum_{j=1}^4 \sum_{k=1}^4 L'_i(x) L_j(y) L_k(z) f_{ijk} \quad (\text{A52})$$

$$\frac{\partial L(x, y, z)}{\partial y} = \sum_{i=1}^4 \sum_{j=1}^4 \sum_{k=1}^4 L_i(x) L'_j(y) L_k(z) f_{ijk} \quad (\text{A53})$$

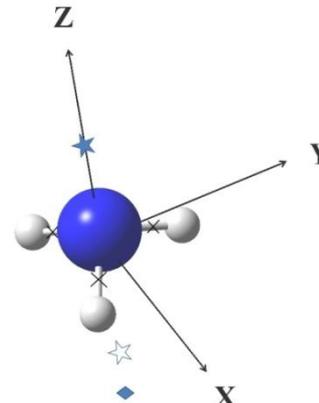
$$\frac{\partial L(x, y, z)}{\partial z} = \sum_{i=1}^4 \sum_{j=1}^4 \sum_{k=1}^4 L_i(x) L_j(y) L'_k(z) f_{ijk} \quad (\text{A54})$$

The terms: $L'_i(x)$, $L'_j(y)$, $L'_k(z)$ in Eqs. (A52), (A53), (A54) are replaced by $L''_i(x)$, $L''_j(y)$, $L''_k(z)$ to yield elements of Hessian matrix $H_{11} = \frac{\partial^2 L}{\partial x^2}$, $H_{22} = \frac{\partial^2 L}{\partial y^2}$ and $H_{33} = \frac{\partial^2 L}{\partial z^2}$. Elements H_{21} , H_{31} and H_{32} correspond to $\frac{\partial^2 L}{\partial y \partial x}$, $\frac{\partial^2 L}{\partial z \partial x}$ and $\frac{\partial^2 L}{\partial z \partial y}$, and the former is obtained by replacing $L(y)$ in Eq. (A52) by $L'_j(y)$. The latter two are derived by replacing the $L(z)$ components in Eqs. (A53) and (A54) with $L'_k(z)$.

A10. Guess points for topographical features of MESP scalar field

A10.1. Ammonia molecule

Table A7. MESP topography for ammonia molecule: comparison between those predicted by numerical method (darkened rows) and those derived from MP2/6-31G(d,p) wave function. Electron density values less than 0.001 a.u. is represented by '---'. The symbols for various types of GPs presented is applicable for Tables A8 to A11.

					Symbols for GPs located at various sites	
					LPs/ π bonds ★	Bond saddles ×
CP	Nature	Co-ordinates			V(r)	$\rho(\mathbf{r})$
		X	Y	Z		
LP	(3, +3)	-0.0005	-0.0005	2.5163	-0.1369	0.027
		0.0000	0.0000	2.5075	-0.1365	0.024
bond CPs	(3, -1)	-0.9813	-0.5719	-0.2338	1.0187	0.160
		0.9766	-0.5682	-0.2280	1.0216	0.160
		± 0.9772	-0.5642	-0.2579	0.9960	0.361
		-0.0053	1.1341	-0.2372	0.9792	0.435
		0.0000	1.1284	-0.2579	0.9960	0.361
Not found	(3, +3)	-0.0007	-0.0004	-3.0774	0.0122	0.002
	(3, +1)	-0.0007	-0.0006	-4.1828	0.0134	---

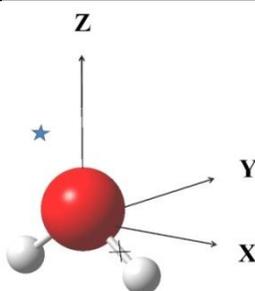
The Lagrange method predicts a total of 7 GPs for ammonia molecule. The 3D grid data is obtained at MP2/6-31G(d,p) level of theory with increment of 0.33333 and 34 tabular points per direction. The starting grid point bears co-ordinates (-5.483229, -5.483229, 5.261677). Co-ordinates of GPs and corresponding results obtained after submission to the computational chemistry software INDPROP are

displayed in Table A7 (presented on last page). Though there is no much difference between the co-ordinates of guess points and corresponding MESP values from the corresponding QM counterparts, the electron density values however are unreliable. The numerical method however predicts correct nature of GPs. No CPs were detected using last two entries however.

A10.2. Water molecule

The numerical method has detected only two GPs for water molecule and the corresponding QM counterparts obtained are displayed in Table A8.

Table A8. MESP topography predicted by three-dimensional Lagrange method for water molecule (darkened rows) and corresponding counterparts derived from MP2/6-31G(d,p) WF.

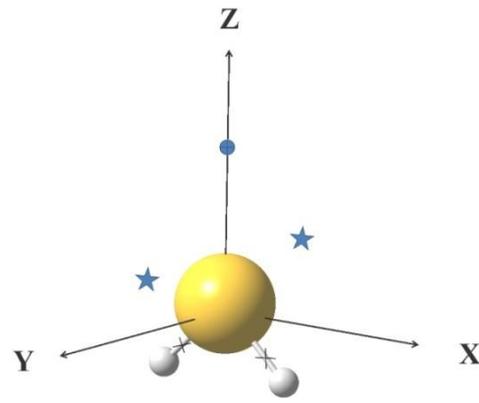
						
CP	Nature	Co-ordinates			V(r)	$\rho(\mathbf{r})$
		X	Y	Z		
lone pair	(3, +3)	0.0034	2.0569	1.3082	-0.0989	0.017
		0.0000	1.9791	1.4066	-0.0989	0.017
bond CP	(3, -1)	0.9262	0.0289	-0.4823	1.1457	0.364
		0.9114	0.0000	-0.4996	1.1318	0.423
Grid specification: Starting point (-4.970736, -4.970736, -4.746573), number of tabular points per direction: 31 with increment of 0.333333.						

The remaining (3, +3) and (3, -1) CPs could easily be obtained using symmetry. The numerical method fails to generate a GP for saddle connecting MESP minima. Thus the software UNIVIS-2000 had to be employed to obtain a suitable GP (steps involved are similar to those explained for acetone molecule in Chapter II, section II.3).

A10.3. Hydrogen sulphide molecule

Unlike the case of water, the numerical method is found to yield a total of 5 GPs: two for (3, +3) and (3, -1) CPs in addition to a saddle (3, +1) saddle (Table A9).

Table A9. MESP topography for hydrogen sulphide molecule: comparison between those predicted by numerical method (darkened rows) and those derived from MP2/6-31G(d,p) WF.



CP	Nature	Co-ordinates			V(r)	$\rho(\mathbf{r})$
		X	Y	Z		
lone pair	(3, +3)	0.0021	-3.1365	1.3307	-0.0450	0.007
		0.0021	3.1596	1.3129	-0.0449	0.007
		0.0000	± 3.1446	1.3326	-0.0449	0.006
saddle connecting MESP minima	(3, +1)	0.0041	0.0019	4.1262	-0.0190	0.001
		0.0000	0.0000	4.1283	-0.0190	0.001
bond CP	(3, -1)	-1.2631	0.0183	-0.9535	0.8256	0.162
		1.2560	0.0232	-0.9748	0.8367	0.084
		± 1.2349	0.0000	-0.9912	0.8229	0.228

Grid specification: Starting point (-8.571798, -8.571798, -8.379329), number of tabular points per direction: 52 with increment of 0.333333.

A10.4. Formaldehyde molecule

A total of six GPs (two for lone pairs belonging to carbonyl oxygen, a saddle connecting them, and three bond CPs) yield required MESP topographical features of formaldehyde molecule (displayed in Table A10).

Table A10. MESP topography for formaldehyde molecule: comparison between those predicted by numerical method (darkened rows) and those derived from MP2/6-31G(d,p) WF.

CP	Nature	Co-ordinates			V(r)	$\rho(\mathbf{r})$
		X	Y	Z		
lone pair	(3, +3)	0.0002	-1.8706	2.6677	-0.0801	0.014
		0.0001	1.8702	2.6667	-0.0801	0.014
		0.0000	± 1.8535	2.6858	-0.0801	0.014
saddle connecting MESP minima	(3, +1)	-0.0006	0.0037	3.7336	-0.0698	0.001
		0.0000	0.0000	3.7630	-0.0696	0.007
bond CPs between C-H atoms	(3, -1)	0.0042	-1.1347	-1.7150	0.8669	0.270
		0.0044	1.1336	-1.7164	0.8663	0.280
		0.0000	± 1.1236	-1.7253	0.8650	0.239
bond CP between C-O atoms	(3, -1)	0.0025	0.0025	0.1133	1.3137	0.695
		0.0000	0.0000	0.1448	1.3339	0.486

Grid specification: Starting point (-6.512752, -6.512752, -7.527282); tabular points in X, Y direction are 40 and 42 in Z direction; increment per direction 0.333333.

A10.5. Ethene molecule

Table A11 portrays the GPs obtained from the 3D Lagrange method for ethene molecule. It has been observed that the numerical method does a good job in predicting correct electron density values at negative valued CPs than the positive ones.

Table A11. MESP topography for ethene molecule: comparison between those predicted by numerical method (darkened rows) and those derived from MP2/6-31G(d,p) WF.

CP	Nature	Co-ordinates			V(r)	$\rho(r)$
		X	Y	Z		
π bond	(3, +3)	0.0003	-0.0038	-2.8946	-0.0393	0.007
		0.0003	-0.0038	2.8932	-0.0393	0.007
		0.0000	0.0000	± 2.8849	-0.0393	0.007
between C-H atoms	(3, -1)	-1.1212	-1.9332	0.0039	0.8524	0.289
		-1.1210	1.9408	0.0037	0.8465	0.283
		1.1197	-1.9348	0.0042	0.8523	0.296
		1.1167	1.9454	0.0040	0.8463	0.293
		± 1.1123	± 1.9441	0.0000	0.9835	0.352
between C-C atoms		0.0019	0.0018	0.0096	0.9575	0.527
		0.0000	0.0000	0.0000	0.9835	0.352

Grid specification: Starting point (-6.512752, -6.512752, -7.774458), number of tabular points in X, Y direction: 40 and 48 in Z direction with increment of 0.333333.

As it shall be discussed in the development of ethene GCM, a single Gaussian placed at the centre of the carbon-carbon double bond produces two MESP minima as required but also connects them by negative valued (3, +1) CPs, thereby generating a ring type of surface encircling the carbon-carbon double bond. This ring feature however is observed in the MESP plot of acetylene, and therefore encouraged to develop a model for acetylene. The Lagrange method was however not used to obtain

suitable GPs for obtaining degenerate ring CPs, however this was achieved using UNIVIS-2000.

A10.6. Benzene molecule

Interestingly, the numerical method is found to predict GPs for all MESP critical points (Table A12). A single bond saddle between carbon – carbon bond was missed but could be obtained due to its symmetry-related counterparts.

Table A12. MESP topography for benzene molecule: comparison between those predicted by numerical method (darkened rows) and those derived from MP2/6-31G(d,p) wave function.

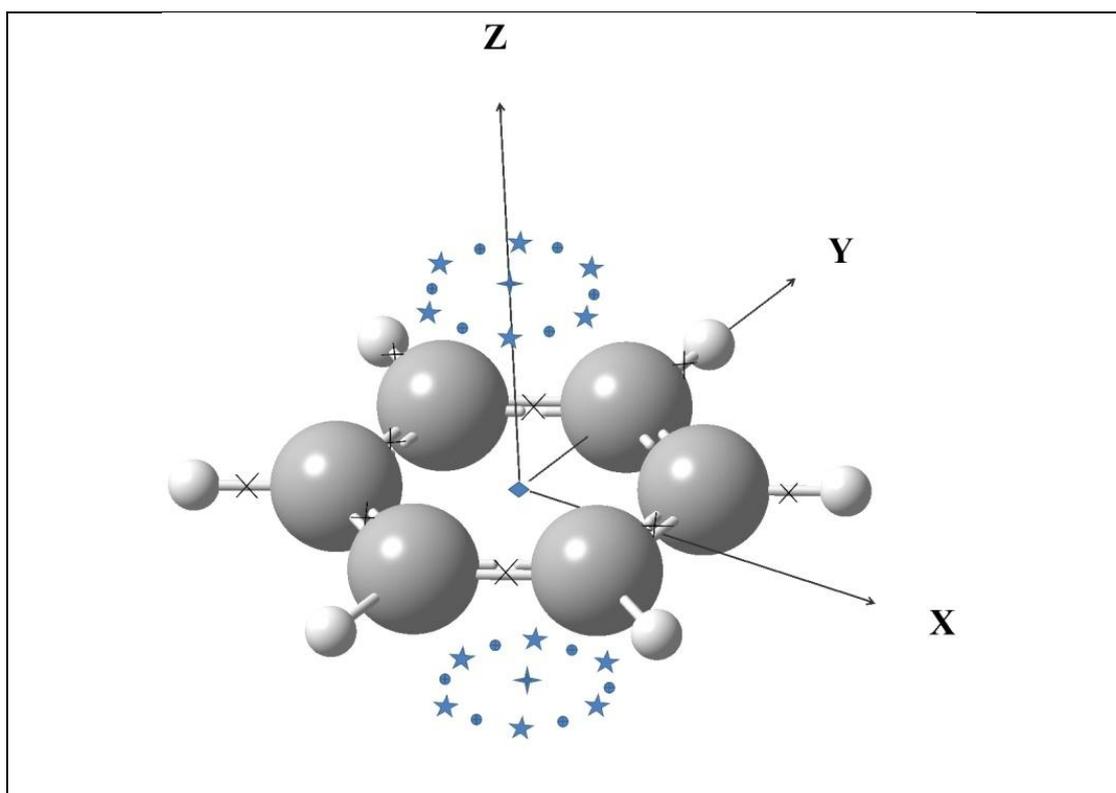


Table A12 continued...

Table A12 continued...

CP	Nature	Co-ordinates			V(r)	$\rho(\mathbf{r})$
		X	Y	Z		
π bonds	(3, +3)	-0.6399	-1.1088	-3.2517	-0.0332	0.0023
		-0.6400	-1.1096	3.2505		
		-0.6399	1.1089	-3.2517		
		-0.6402	1.1097	3.2505		
		0.6401	-1.1086	-3.2518		
		0.6403	-1.1095	3.2505		
		0.6403	1.1087	-3.2518		
		0.6407	1.1095	3.2505		
		± 0.6391	± 1.1093	± 3.2466	-0.0332	0.0023
		-1.2786	0.0000	-3.2521	-0.0332	0.0023
		-1.2791	0.0001	3.2509		
		1.2750	0.0000	-3.2522		
		1.2758	0.0001	3.2510		
		± 1.2799	0.0000	± 3.2467	-0.0332	0.0023
saddles connecting MESP minima	(3, +1)	-0.0269	-1.2022	-3.2631	-0.0331	0.0021
		-0.0267	-1.2030	3.2618		
		-0.0267	1.2009	-3.2630		
		-0.0274	1.2016	3.2617		
		0.0000	± 01.2058	± 3.2586	-0.0331	0.0021
		-1.0477	-0.6042	-3.2633	-0.0331	0.0021
		-1.0480	-0.6051	3.2620		
		-1.0498	0.6016	-3.2633		
		-1.0501	0.6025	3.2620		
		1.0485	-0.6000	-3.2634		
		1.0491	-0.6004	3.2621		
		1.0507	0.5972	-3.2634		
		1.0515	0.5974	3.2621		
		± 1.0446	± 0.6015	± 3.2587	-0.0331	0.0021
ring cp		0.0036	0.0019	0.0009	0.1499	0.02466
		0.0000	0.0000	0.0000	0.1507	0.0197
saddles between C-C atoms	(3, -1)	-1.1212	-1.9332	0.0039	0.8524	0.2890
		-1.1210	1.9408	0.0037	0.8465	0.2830
		1.1197	-1.9348	0.0042	0.8523	0.2960
		1.1167	1.9454	0.0040	0.8463	0.2930
		± 1.1402	± 1.9748	0.0000	0.8630	0.3209
		-2.2774	0.0024	0.0018	0.8410	0.4656
		2.2687	0.0024	0.0018	0.8417	0.4659
		± 2.2803	0.0000	0.0000	0.8630	0.3209
saddles between C-H atoms	(3, -1)	-3.4238	-1.9606	0.0039	0.8486	0.3157
		-3.4229	1.9630	0.0038	0.8471	0.3183
		3.4287	-1.9564	0.0027	0.8479	0.3075
		3.4303	1.9558	0.0024	0.8459	0.3111
		± 3.4170	± 1.9728	0.0000	0.8513	0.2928

Continued on next page

Table A12 continued...

CP	Nature	Co-ordinates			V(r)	$\rho(\mathbf{r})$
		X	Y	Z		
saddles between C-C atoms	(3, -1)	0.0104	3.9689	-0.0045	0.8069	0.3157
		0.0080	-3.9623	-0.0035	0.8042	0.3183
		0.0000	± 3.3956	0.0000	0.8513	0.2928
saddles along Z-axis		-0.0043	-0.0019	-3.3321	-0.0321	0.0010
		-0.0043	-0.0020	3.3312		
		0.0000	0.0000	± 3.3322	-0.0321	0.0010

A11. Potential due to *s*-type Gaussians

Two normalized *s*-type Gaussians $\varphi(\mathbf{R}_A)$ and $\varphi(\mathbf{R}_B)$ are defined by Eqs (A55) and (A56). Gaussian centres \mathbf{R}_A and \mathbf{R}_B bear coordinates (A_x, A_y, A_z) and (B_x, B_y, B_z) respectively and the terms $(2\alpha/\pi)^{3/4}$ and $(2\beta/\pi)^{3/4}$ are normalization constants of respective Gaussians.

$$\varphi(\mathbf{R}_A) = \left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}} e^{-\alpha(\mathbf{r} - \mathbf{R}_A)^2} \quad (\text{A55})$$

$$\varphi(\mathbf{R}_B) = \left(\frac{2\beta}{\pi}\right)^{\frac{3}{4}} e^{-\beta(\mathbf{r} - \mathbf{R}_B)^2} \quad (\text{A56})$$

The potential at point \mathbf{R}_C due to interaction of $\varphi(\mathbf{R}_A)$ with $\varphi(\mathbf{R}_B)$ can be written as $\langle \varphi(\mathbf{R}_A) | \frac{1}{|\mathbf{r} - \mathbf{R}_C|} | \varphi(\mathbf{R}_B) \rangle$, defined by Eq. (A57) [\mathbf{R}_N corresponds to the newly formed Gaussian centre due to products of $\varphi(\mathbf{R}_A)$ with $\varphi(\mathbf{R}_B)$ where ‘ k ’ is the pre-exponential factor given by $e^{-\left(\frac{\alpha\beta}{\alpha+\beta}\right)(\mathbf{R}_A - \mathbf{R}_B)^2}$].

$$\langle \varphi(\mathbf{R}_A) | \frac{1}{|\mathbf{r} - \mathbf{R}_C|} | \varphi(\mathbf{R}_B) \rangle = \left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}} \left(\frac{2\beta}{\pi}\right)^{\frac{3}{4}} k \left[\int \frac{e^{-(\alpha+\beta)(\mathbf{r} - \mathbf{R}_N)^2}}{|\mathbf{r} - \mathbf{R}_C|} d\mathbf{r} \right] \quad (\text{A57})$$

The term $1/|\mathbf{r} - \mathbf{R}_C|$ in Eq. (A57) is replaced by the employing the integral transformation $\frac{1}{|\mathbf{r} - \mathbf{R}_C|} = \int_{-\infty}^{\infty} \frac{e^{-(\mathbf{r} - \mathbf{R}_C)^2 t^2}}{\sqrt{\pi}} dt$, resulting to Eq. (A58) where $k_1 = \left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}} \left(\frac{2\beta}{\pi}\right)^{\frac{3}{4}} k$ and $\gamma = \alpha + \beta$.

$$\langle \varphi(\mathbf{R}_A) | \frac{1}{|\mathbf{r} - \mathbf{R}_C|} | \varphi(\mathbf{R}_B) \rangle = k_1 \left[\int e^{-\gamma(\mathbf{r} - \mathbf{R}_N)^2} d\mathbf{r} \right] \left[\int_{-\infty}^{\infty} \frac{e^{-(\mathbf{r} - \mathbf{R}_C)^2 t^2}}{\sqrt{\pi}} dt \right] \quad (\text{A58})$$

Invoking Gaussian product rule once again for $e^{-\gamma(\mathbf{r} - \mathbf{R}_N)^2}$ and $e^{-(\mathbf{r} - \mathbf{R}_C)^2 t^2}$ in above equation results to a new Gaussian $e^{-(\gamma+t^2)(\mathbf{r} - \mathbf{R}_S)^2}$ with new Gaussian centre \mathbf{R}_S and exponential factor given by $e^{-\left(\frac{\gamma t^2}{\gamma+t^2}\right)(\mathbf{R}_N - \mathbf{R}_C)^2}$, yielding Eq. (A59).

$$\langle \varphi(\mathbf{R}_A) | \frac{1}{|\mathbf{r} - \mathbf{R}_C|} | \varphi(\mathbf{R}_B) \rangle = k_1 \int_{-\infty}^{\infty} \left\{ \int \frac{e^{-(\gamma+t^2)(\mathbf{r} - \mathbf{R}_S)^2}}{\sqrt{\pi}} d\mathbf{r} \right\} e^{-\left(\frac{\gamma t^2}{\gamma+t^2}\right)(\mathbf{R}_N - \mathbf{R}_C)^2} dt \quad (\text{A59})$$

Integrating the quantity in curly brackets over \mathbf{r} yields Eq. (A60), corresponding to one-dimensional integral over variable t. The same is transformed to Eq. (A61) by taking limits of integration from 0 to infinity.

$$\langle \varphi(\mathbf{R}_A) | \frac{1}{|\mathbf{r} - \mathbf{R}_C|} | \varphi(\mathbf{R}_B) \rangle = \frac{k_1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \left(\frac{\pi}{\gamma + t^2} \right)^{\frac{3}{2}} e^{-\left(\frac{\gamma t^2}{\gamma+t^2}\right)(\mathbf{R}_N - \mathbf{R}_C)^2} dt \quad (\text{A60})$$

$$\langle \varphi(\mathbf{R}_A) | \frac{1}{|\mathbf{r} - \mathbf{R}_C|} | \varphi(\mathbf{R}_B) \rangle = \frac{2k_1}{\sqrt{\pi}} \int_0^{\infty} \left(\frac{\pi}{\gamma + t^2} \right)^{\frac{3}{2}} e^{-\left(\frac{\gamma t^2}{\gamma+t^2}\right)(\mathbf{R}_N - \mathbf{R}_C)^2} dt \quad (\text{A61})$$

The next step involves substituting $u^2 = \frac{t^2}{\gamma+t^2}$ which upon rearranging leads to $\gamma t^2 + 1 = u^2$. Differentiating the given expression with respect to variable t, results in $\gamma t^{-3} = u^{-3} \frac{du}{dt}$ which when reorganized becomes $dt = \frac{1}{\gamma} \left(\frac{t^2}{u^2} \right)^{\frac{3}{2}} du$. It can easily be shown that u equals 0 and 1 when t equals 0 and ∞ . Performing these substitutions in Eq. (A61) transforms it to Eq. (A62) and finally to Eq. (A63).

$$\langle \varphi(\mathbf{R}_A) | \frac{1}{|\mathbf{r} - \mathbf{R}_C|} | \varphi(\mathbf{R}_B) \rangle = \frac{2k_1}{\sqrt{\pi}} \int_0^1 \left(\frac{\pi^{\frac{3}{2}}}{\gamma} \right) e^{-\gamma(\mathbf{R}_N - \mathbf{R}_C)^2 u^2} du \quad (\text{A62})$$

$$\langle \varphi(\mathbf{R}_A) | \frac{1}{|\mathbf{r} - \mathbf{R}_C|} | \varphi(\mathbf{R}_B) \rangle = (2k_1) \left(\frac{\pi}{\gamma} \right) \int_0^1 e^{-\gamma(\mathbf{R}_N - \mathbf{R}_C)^2 u^2} du \quad (\text{A63})$$

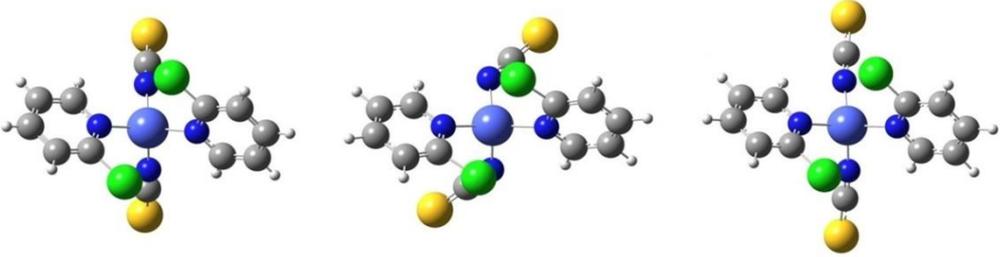
The integral in above equation is represented as $F_0(t)$ where $t = \gamma(\mathbf{R}_N - \mathbf{R}_C)^2$

Substituting for k_1 and γ gives Eq. (A64).

$$\langle \varphi(\mathbf{R}_A) | \frac{1}{|\mathbf{r} - \mathbf{R}_C|} | \varphi(\mathbf{R}_B) \rangle = \left(\frac{2\alpha}{\pi} \right)^{\frac{3}{4}} \left(\frac{2\beta}{\pi} \right)^{\frac{3}{4}} \left[\left(e^{-\left(\frac{\alpha\beta}{\alpha+\beta} \right) (\mathbf{R}_A - \mathbf{R}_B)^2} \right) \left(\frac{2\pi}{\alpha + \beta} \right) F_0(t) \right] \quad (\text{A64})$$

A12. Geometries of some cobalt thiocyanate complexes: comparison between experimental and quantum mechanical structures

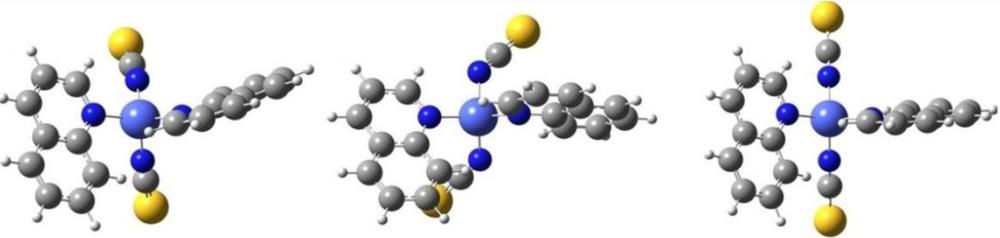
Table A13. Comparison between cobalt environment in experimental and theoretical structures of $\text{Co}(\text{2-chloropyridine})_2(\text{NCS})_2$.



Monomeric units (left to right) from: crystal structure, HF/STO-3G and B3LYP/6-31G(d,p) calculations.

Geometry from	Bond lengths (Å)		Angles (degrees)					
	Co-N(L ¹)/ Co-N(L ²)	Co-N(X ¹)/ Co-N(x ²)	θ ₁	θ ₂	θ ₃	θ ₄	θ ₅	θ ₆
Crystal structure	2.035 2.035	1.932 1.932	104.3	110.4	104.3	110.4	107.3	119.4
HF/STO-3G optimization	2.119 2.119	1.940 1.940	102.1	105.3	102.1	105.3	119.7	123.8
B3LYP/6-31G** optimization	2.137 2.137	1.924 1.924	101.1	107.4	101.1	107.4	124.7	116.2

Table A14. Comparison between cobalt environment in experimental and theoretical structures of $\text{Co}(\text{quinoline})_2(\text{NCS})_2$.



Monomeric units (left to right) from: crystal structure, HF/STO-3G and B3LYP/6-31G(d,p) calculations.

Geometry from	Bond lengths (Å)		Angles (degrees)					
	Co-N(L ¹)/ Co-N(L ²)	Co-N(X ¹)/ Co-N(x ²)	θ ₁	θ ₂	θ ₃	θ ₄	θ ₅	θ ₆
Crystal structure	2.044 2.029	1.940 1.934	103.2	114.2	111.6	114.5	104.5	108.3
HF/STO-3G optimization	2.133 2.128	1.930 1.925	103.2	103.1	111.7	105.7	124.2	107.9
B3LYP/6-31G** optimization	2.134 2.115	1.936 1.925	102.1	106.6	105.9	111.2	124.0	105.7

Tables A13 and A14 summarize the various bond lengths and bond angles between cobalt and ligands in $\text{Co}(\text{2-chloropyridine})_2(\text{NCS})_2$ & $\text{Co}(\text{quinoline})_2(\text{NCS})_2$ complexes obtained from experimental geometries and corresponding quantum mechanical calculations.

A13. Certificates of participation/poster presentations of various conferences, symposiums and workshops attended during research period January 2014 – December 2019

Conferences:



Symposiums:



Workshops:



**First page of research articles published during
research period January 2014 – December 2019**