The oretical and experimental investigations of structure, reactivity and bonding in some organic systems

David M. J. Lilley

THEORETICAL AND EXPERIMENTAL INVESTIGATIONS OF STRUCTURE, REACTIVITY AND BONDING IN SOME ORGANIC SYSTEMS

by

David M. J. Lilley, B.Sc. (Dunelm) M.Sc. (London)

A thesis submitted to the University of Durham for the degree of Doctor of Philosophy.

November 1973

To my parents

MEMORANDUM

The work described in this thesis was carried out in the University of Durham between September 1969 and September 1972. It has not been submitted for any other degree and is the original work of the author except where acknowledged by reference.

Part of the work in this thesis has formed the subject matter of the following publications:

(i) The Electronic Structures of Ethyl Cation and Protonated Ethylene; a Non-empirical LCAO-MO-SCF Investigation.
D. T. Clark and D. M. J. Lilley, J. Chem. Soc. (D) 549 (1970)

(ii) Barriers to Internal Rotation in Carbonium Ions: a Non-empirical LCAO-MO-SCF Investigation of 1- and 2-Fluoroethyl Cations.
D. T. Clark and D. M. J. Lilley, J. Chem. Soc. (D) 603 (1970)

(iii) The Electronic Structures of Bridge-protonated Fluoroethylene and the Interconversion of 1- and 2-Fluoroethyl Cations.
D. T. Clark and D. M. J. Lilley, J. Chem. Soc. (D) 1042 (1970)

(iv) A Non-empirical LCAO-SCF-MO Investigation of Cross Sections through the Potential Energy Surface for the $[C_2H_4Cl]^+$ systems; comparison with the $[C_2H_5]^+$ and $[C_2H_4F]^+$ systems.
D. T. Clark and D. M. J. Lilley, Tetrahedron 29, 845 (1973)

(v) ESCA Studies of a Series of Acetyl Compounds.

D. T. Clark, D. M. J. Lilley and M. Barber. Paper 38, Oxford Conference on Photoionisation and Photoelectron Spectroscopy. Sept. 1970.

(vi) Molecular Core Binding Energies for some Five Membered Ring Heterocycles as determined by X-ray Photoelectron Spectroscopy.

D. T. Clark and D. M. J. Lilley, Chem. Phys. Lett. 9, 234 (1971)

ACKNOWLEDGEMENTS

To my supervisor, Dr. D. T. Clark, I extend my sincere gratitude for his unfailing help and encouragement throughout this work.

I am indebted to the S.R.C. for financial support and provision of equipment.

I should also like to thank the following for their invaluable assistance: My colleagues Dr. D. Kilkast, Mr. D. B. Adams and Mr. I. W. Scanlan for helpful discussion and my tutor, Dr. W. J. Feast for enthusiastic moral support.

D. M. J. Lilley

Durham 1973

ABSTRACT

A theoretical study has been made of some aspects of prototype potential energy surfaces for some simple organic reactions. Addition of prototype electrophiles to simple alkenes has been investigated by means of non-empirical and semi-empirical calculations, within the Hartree-Fock formalism, and the resulting carbonium ions studied. The systems under investigation may be formally considered as being derived from electrophilic addition of H^+ to ethylene, fluoroethylene and chloroethylene, or of X^+ (X=F,Cl) to ethylene, and may thus be represented as $(C_2H_4X)^+$, X=H,F and Cl. For the simplest system, $C_2H_5^+$, two basic structures have been considered, the classical ethylcation and the bridge-protonated ethylene. The energies of these species have been minimised with respect to the C-C bond lengths and also, in the case of the latter ion, with respect to the distance of the bridging H from the CC bond centre. Examination of conformational processes in the classical ion has shown a virtual absence of any barrier to rigid rotation about the cc bond. The calculated relative energies of the species has indicated, subject to limitations imposed by the basis set size and partial geometry optimisation, that in the gas phase the classical ion should be $\sim$5.2k cal $mole^{-1}$ more stable than the bridge protonated ethylene. Furthermore calculations along an idealised reaction coordinate representing transformation between the two spec ies have indicated the absence of an activation barrier thus suggesting the bridged ion to be the transition state for the scrambling of the hydrogen atoms of the ethyl cation. These results have been compared with mass spectrometric data. The approach of a prototype nucleophile (H^-) to ethyl cation has been examined, results suggesting a preferential cis attack. Conformational processes in the 1- and 2- fluoro and chlorethyl cations have been examined. The rotational barrier in the 2- fluoroethyl cation has been shown to be very large

(10.5k cal mole) and, with the exception of the 2- chloroethyl cation, all the barriers for the substituted ethyl cations have been shown to be dominated by attractive terms. In both the fluoro and chloroethyl systems, predicted ordering of stabilities of cations has been 1- haloethyl > bridge-protonated haloethylene > 2- haloethyl, and idealised reaction coordinates have been constructed relating the ions in the fluoro case, the results predict the total absence of any activation barrier in transforming 2- to 1- fluoroethyl cation, whilst, in the analogous chloro case, a small barrier (4.3 k cal $mole^{-1}$) is predicted. Relative thermochemical stabilities of the ions have been computed, and the stabilising/destabilising effects of halogen substitution in these carbonium ions investigated and compared with experimental data.

The halogen bridged 'halonium' ions have been studied, and their total energies minimised with respect to the distance of the halogen atom from the CC bond centre. The calculations have indicated that the fluoronium ion should be of marginally greater stability than the 2-fluoroethyl cation (3.6k cal $mole^{-1}$) and this has been discussed in the light of published nmr studies of the ionisation of 2-halo-3fluoro 2,3-dimethyl butanes in SO_2/SbF_5. Results for the chloronium ion have indicated that this ion should be considerably more stable (15.8k cal $mole^{-1}$) than the corresponding 2- chloroethyl cation.

Electron Spectroscopy for Chemical Applications (ESCA) has been employed for the measurement of core binding energies in three series of closely related molecules

(i) a series of acetyl compounds of general formula

CH_3COX, X=H, CH_3, OH, OCH_3, NH_2, $NHCH_3$, $COCH_3$, CO_2H, CN and $OCOCH_3$.

(ii) a series of five membered ring heterocycles.

(iii) a series of pyrimidine bases and related compounds.

Assignment of core levels has been accomplished in two ways.

(i) Direct correlation of measured binding energies with orbital

energies derived from SCF calculations, i.e. assuming Koopmans' theorem.

(ii) Correlation of shifts in core binding energies with computed electron distributions within the molecule using the charge potential model. In general, assignments based upon the different methods have been found to be in agreement. Furthermore in the case of some members of the pyrimidine series comparison has been possible between charge potential assignments using both ab initio and CNDO/II populations. Agreement between the two sets has been complete.

CONTENTS

INTRODUCTION

This thesis consists of two related yet distinct parts. In the first part (Chapter 2) a theoretical investigation of some aspects of prototype potential energy surfaces for some simple organic reactions are considered. The second part (Chapters 3, 4 and 5) describes an experimental and theoretical study of some aspects of structure and bonding in acetyl compounds (Chapter 3), five membered ring heterocyclic molecules (Chapter 4) and pyrimidines (Chapter 5) investigated by ESCA (Electron Spectroscopy for Chemical Analysis). In Chapter 1 is detailed background theory for both parts of the thesis.

1. Introduction to Part 1

A Theoretical Investigation of aspects of the potential energy surfaces of the systems $(C_2H_4X)^+$ X=H,F and Cl. A relatively detailed examination has been made of prototype electrophilic addition to defins by means of non-empirical quantum mechanical calculations, supplemented where appropriate by semi-empirical calculations. The prototype systems investigated can formally be considered as being derived from electrophilic addition of H^+ to ethylene, fluoroethylene and chloroethylene, or of X^+(X=F,Cl) to ethylene. Computer restrictions have not allowed complete geometry optimisation for species which a priori might be expected to represent energy minima on the potential energy surface. The most thorough geometry optimisations have been performed on bridge-protonated ethylene and ethyl cation for which the computer time necessary was comparatively small. For the halogen substituted species however, with their correspondingly larger basis sets, where possible 'chemical intuition' together with the results for the $C_2H_5^+$ system have been used to utilise as effectively as possible the computer time available for geometry optimisation. For the classical 1- and 2- substituted ethyl cations

conformational processes have been investigated yielding important information on barriers to rotation in simple carbonium ions, the adequacy of the basis sets for describing such processes having been checked in each case by examining conformational processes in the corresponding substituted ethones. Relative energies of bridge-protonated ethylenes, $(C_2H_3XH)^+$ X=F,Cl, and the corresponding classical ions have been computed and prototype reaction coordinates investigated relating the interconversion of classical ions via bridged species. Where possible the results have been compared with experimental evidence.

2. Introduction to Part 2

Investigation by E.S.C.A. of structure and bonding in some series of closely related chemical compounds.

The recent development of Electron Spectroscopy for Chemical Applications, or E.S.C.A., as a method for the determination of molecular core binding energies has provided a valuable tool for obtaining information on structure and bonding in molecules. X-ray photoelectron spectra (Mg $K\alpha_{1,2}$) have been recorded for three series of closely related molecules, all representing important classes of organic molecules with considerable chemical importance namely

(i) a series of acetyl compounds, general formula

CH_3COX, X = H, CH_3, OH, OCH_3, NH_2, $NHCH_3$, $COCH_3$, CO_2H, CN and $OCOCH_3$.

(ii) a series of five membered ring heterocycles

(iii) a series of pyrimidine bases and related compounds.

The experimental measurements are then discussed in terms of two models which provide semi-quantitative interpretations of the data.

(i) The assumption of Koopmans' theorem i.e. a direct correlation of measured binding energies with orbital energies derived from SCF calculation.

(ii) Correlation of shifts in core binding energies with computed electron distributions within the molecule using the charge potential model.

Both these interpretive methods have been used in this work. For the first two series examined, the molecules concerned have been of such size that non-empirical SCF calculations have been possible. Thus a comparison may be made with shifts calculated on the basis of both Koopmans' theorem and the charge potential model. For larger molecules, however, such as the series of pyrimidine bases (Chapter 5), non-empirical calculations were not feasible with the amount of computer time available. In this case, therefore interpretation of core binding energies has been in terms of charge distributions computed within the framework of CNDO/2 and the charge potential model.

CHAPTER I

THEORY OF MOLECULAR ORBITAL APPROACH AND ELECTRON SPECTROSCOPY FOR CHEMICAL APPLICATIONS (ESCA)

I MOLECULAR ORBITAL THEORY

1.1 Introduction

The electronic structure of an atom or molecule is described by a mathematical function Ψ of all the coordinates of the system, including time, called the wave function[1]. This function contains all the information about the properties of the system, and hence the basic problem confronting the quantum chemist is first to derive the wave function of the system in which he is interested, and then to analyse it. Exact treatments are only possible for one electron molecules but, nevertheless, molecular orbital theory provides a sufficiently good approximate description of many electron systems to be of general use.

Theoretical calculations have, in principle, some advantage over experimental measurement. For example any assembly of nuclei and electrons may be investigated irrespective of whether the assembly represents a stable entity or not. Thus it is possible in principle to examine the electronic structure of a transition state. An excellent example of the application of these methods are the studies of the transition state for the geometrical isomenisation of cyclopropane by Salem and co-workers[2,3] and by Goddard[4]. Since however the transition state represents the point of maximum free energy on the lowest energy path from reactants to product it is not directly observable and experimental evidence concerning its electronic structure must therefore be obtained indirectly. The experimentalist is limited to the examination of stable or metastable species with finite lifetimes. Clearly the theoretical method does have its drawbacks, dissadvantages and limitations, and these will be discussed later, but provided these are bourne in mind the method is very powerful and the results obtained are extremely useful.

1.2 Brief Summary of Quantum Mechanics

The starting point for the use of quantum mechanical calculation in chemistry is the Schroedinger equation[5]

$$H\Psi = E\Psi \qquad (1.1)$$

where H is the hamiltonian of the system, or total energy operator, and E is the eigen value corresponding to the total energy of the system. The total wave function Ψ is assumed to be time independent, i.e. separation of the time part has been carried out

$$\Psi(x,y,z,t) = \Psi(x,y,z)\Phi(t) \qquad (1.2)$$

and hence (1.1) is known as the time independent Schroedinger equation. Under these conditions the system is termed a stationary state, the properties of which are obtainable by solution of (1.1). For molecular systems a further stage of separation upon the wavefunction due to Born and Oppenheimes[6] is performed. For a system of nuclear coordinates R and electronic coordinates r, the Hamiltonian is expressed (in atomic units)

$$H(r,R) = -\tfrac{1}{2}\sum_{\mu}\frac{1}{M}\nabla^2_{\mu} - \tfrac{1}{2}\sum_{i}\nabla^2_{i} - \sum_{\mu i}\frac{1}{r_{\mu i}} + \sum\frac{1}{r_{ij}} + \sum\frac{Z_{\mu}Z_{\nu}}{R_{\mu\nu}} \qquad (1.3)$$

If the first term describing the nuclear kinetic energy is separated

$$H(r,R) - Hn(R) = He(r,R) \qquad (1.4)$$

The total wavefunction is assumed separable i.e.

$$\Psi = \Psi_e(r,R)\chi_{ne}(R) \qquad (1.5)$$

and the 'electronic' and 'nuclear' wavefunctions are defined by

$$He(r,R)\Psi_e(r,R) = Ee(R)\Psi_e(r,R) \qquad (1.6)$$

and

$$\big(Hn(R) + Ee(R)\big)\chi_{ne}(R) = E\chi_{ne}(R) \qquad (1.7)$$

Thus the 'electronic' schroedinger equation is solved for fixed configurations of nuclei and the resulting electronic energies Ee form a potential energy surface V(R) and equation (1.7) may be solved to yield a total wavefunction and total energy. The conditions under which the Born-Oppenheimer approximation is valid may be examined. The total

Schroedinger equation may be written

$$\big(He(r,R) + Hn(R)\big)\Psi e(r,R)\chi ne(R) = E\Psi e(r,R)\chi ne(R) \qquad (1.8)$$

Since the only differential terms in He are functions of r,

$$\begin{aligned} He(r,R)\Psi e(r,R)\chi ne(R) &= \chi ne(R)He(r,R)\Psi e(r,R) \\ &= \chi ne(R)Ee(R)\Psi e(r,R) \end{aligned} \qquad (1.9)$$

But Hn is a differential operator of R, and both Ψe and χne are functions of R, hence

$$\begin{aligned} Hn(R)\Psi_e(r,R)\chi_{ne}(R) = -\frac{\hbar^2}{2M}\{&\Psi_e(r,R)\nabla^2_{nuc}\chi_{ne}(R) \\ &+ 2\nabla_{nuc}\Psi_e(r,R)\nabla_{nuc}\chi_{ne}(R) + \chi_{ne}(R)\nabla^2_{nuc}\Psi_e(r,R)\} \end{aligned} \qquad (1.10)$$

Substitution of (1.9) and (1.10) into (1.8) yields (1.7) provided the terms in $\nabla_{nuc}\Psi_e$ and $\nabla^2_{nuc}\Psi_e$ may be neglected i.e. provided that the electronic wavefunction is a slowly varying function of the nuclear coordinates then the Born-Oppenheimer approximation is valid.

For many electron systems the summation representing the operator for interelectronic repulsion, $^1/r_{ij}$, must be included as part of the potential energy in the Schroedinger equation. However, if this could be neglected the wavefunction could be expressed in terms of a summation of products of one electron functions

$$\Psi = \psi_a(1)\psi_b(2) \ldots \psi_k(n) \qquad (1.11)$$

In this form the wave equation would be separable into a set of equations each involving the coordinates of only one electron, the solution of which would give the ψ. However, since the many body problem is not exactly soluble in quantum mechanics (or for that matter in classical mechanics), and since even for simple molecules it is very difficult to deal with functions which explicitly consider inter-electronic repulsion and because of its conceptual simplicity, the orbital approximation is made, in which the wavefunction is expressed in terms of functions each of which depend only on the coordinates of one electron. As will become apparent, within the orbital approximation it is

moderately easy to describe adequately the average repulsion experienced by an electron due to the other electrons. The instantaneous correlation of electronic motions, however, is relatively difficult to incorporate within this framework. The particular virtue of the orbital approximation, other than mathematical simplicity is that it leads to results which are especially acceptable to chemists since they closely resemble their intuitive idea of molecular structure.

Associated with each electron is a spin ($m_s \pm \frac{1}{2}$) and the two possible spin functions are written $\alpha(m_s = \frac{1}{2})$ and $\beta(m_s = -\frac{1}{2})$. The product of the above defined spatial orbital and a spin function is call a spin orbital i.e.

$$\phi i(1) = \psi_i(1)\alpha \text{ or } \psi_i(1)\beta \qquad (1.12)$$

where ψ_i is a function depending only on the space coordinates of electron i. The way in which these spin orbitals are combined to form the total wave function is dictated by the Pauli antisymmetry principle[7] which allows for the fact that electrons are indistinguishable from one another. Thus it states that all acceptable wavefunctions for particles of half integral spin must be anti-symmetric upon permutation of the lables of any two particles. The single product wavefunction

$$\Psi(1,2\ldots2n) = \psi_1(1)\alpha\ \psi_1(2)\beta\ \psi_2(3)\alpha\ \ldots\ldots\ \psi_n(2n)\beta \qquad (1.13)$$

is not antisymmetric but can be made so by writing it in determinantal form[8]

$$\text{i.e. } \Psi(1,2\ldots2n) = \sqrt{\frac{1}{2n!}} \begin{vmatrix} \psi_1(1)\alpha\psi_1(1)\beta\psi_2(1)\alpha\ldots\ldots\ldots\psi_n(1)\beta \\ \psi_1(2)\alpha\psi_1(2)\beta\psi_2(2)\alpha\ldots\ldots\ldots\psi_n(2)\beta \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \psi_1(2n)\alpha\ \ldots\ldots\ldots\ldots\ldots\ \psi_n(2n)\beta \end{vmatrix} \qquad (1.14)$$

ofter called a Slater determinant and abreviated to

$$\begin{aligned} \Psi(1,2\ldots2n) &= |\psi_1\alpha\psi_1\beta\psi_2\alpha\psi_2\beta\ \ldots\ldots\ \psi n\alpha\psi n\beta| \\ &= |\psi_1\overline{\psi}_1\psi_2\overline{\psi}_2\ \ldots\ldots\ \psi n\overline{\psi}n| \end{aligned} \qquad (1.15)$$

which includes only the diagonal elements. Permutation of the lables of any two electrons exchanges two rows of the determinant and hence reverses the sign of the wavefunction, thus ensuring antisymmetry.

Certain mathematical constraints are placed upon the wavefunction. It must be finite, continuous and single valued. Solutions of the Schroedinger equation are required to be normalised i.e.

$$\int \psi_i^2 \partial\tau = 1 \qquad (1.16)$$

which is ensured by multiplying the wavefunction by a normalisation constant such that 1.16 is satisfied. It can also be shown that different eigenfunctions of the same Hamiltonian corresponding to different eigenvalues must be mutually orthogonal, i.e.

$$\int \psi_i \psi_j \partial\tau = 0 \qquad (1.17)$$

and the orthonormalisation can be expressed as

$$\int \psi_i \psi_j \partial\tau = \delta_{ij} \qquad (1.18)$$

where δ_{ij}, the kronecker delta, takes the value unity when i = j and zero for any other case.

From the discussion so far it is apparent that expressing the wavefunction in terms of a Slater determinant of spin orbitals is an approximation and hence a yardstick is required for gauging how close to physical reality is the description of the system provided by the wavefunction. Such a criterion is provided by the variation theorem.

The Variation Method

In practical terms, direct solution of the schroedinger equation is only feasible for one electron systems, and many electron systems are usually solved by the variation method. The variation principle states that[9]

Given any approximate wavefunction satisfying the boundary conditions of the problem, the expectation value of the energy calculated from this function will always be higher than the true ground state energy.

Thus if ψ is an approximation to the exact wave function

$$\int \psi^* H \psi \partial\tau = E \geqslant E^o \qquad (1.19)$$

where E^o is the true energy. Hence the method suggested is to start with a trial wave function containing one or more variational parameters and then to minimise the expectation value of the energy with respect to these parameters. The method generally used for the derivation of suitable trial functions is to take a linear combination of basis functions, known as the basis set, which approaches the perfect wave function (within the Hartree-Fock formalism) as the number of functions tends to infinity. The simplest approach of this kind is the Linear Combination of Atomic Orbitals or LCAO method, which assumes that the electronic distribution in a molecule can be represented approximately as a sum of atomic distributions, and hence the trial functions chosen for the molecular orbital ψ is a linear combination of the appropriate atomic orbitals χ_μ

$$\psi = \sum_\mu C_\mu \chi_\mu \qquad (1.20)$$

and the coeeficients C are used as the variational parameters, though a variational parameter can equally well be placed in the basis functions themselves. The set of atomic orbitals χ_μ is known as the basis set. The use of the variation method is now explained using an LCAO function, first defining the following matrices.

$$\overline{\chi} = (\chi_1 \chi_2 \ \ldots\ \chi^n) \qquad \overline{C}_k = \begin{bmatrix} C_{1k} \\ C_{2k} \\ \cdot \\ \cdot \\ \cdot \\ C_{nk} \end{bmatrix}$$

$$\overline{\Delta} = \overline{\chi}^{\dagger}\,\overline{\chi}$$

$$\overline{H} = \overline{\chi}^{\dagger}\,\overline{\overline{\chi}} \qquad (1.21)$$

where the symbol † refers to the matrix adjoint or complex conjugate transpose.

Thus the energy will be given by

$$\varepsilon = \frac{\overline{C}_k^{\dagger}\overline{H}\,\overline{C}_k}{\overline{C}_k^{\dagger}\overline{\Delta}\,\overline{C}_k} \tag{1.22}$$

and applying the variational principle

$$\delta\varepsilon = 0 = \delta\,\frac{\overline{C}_k^{\dagger}\overline{H}\,\overline{C}_k}{\overline{C}_k^{\dagger}\overline{\Delta}\,\overline{C}_k} = \frac{(\delta\overline{C}_k^{\dagger}\overline{H}\,\overline{C}_k - \varepsilon\delta\overline{C}_k^{\dagger}\overline{\Delta}\,\overline{C}_k)}{\overline{C}_k^{\dagger}\overline{\Delta}\,\overline{C}_k} + \text{complex conjugate} \tag{1.23}$$

i.e. the variation principle requires that

$$\delta\overline{C}_k^{\dagger}(\overline{H}\,\overline{C}_k - \varepsilon\overline{\Delta}\,\overline{C}_k) = 0 \tag{1.24}$$

and for this to be satisfied for a variation $\delta\,C_k$ it is necessary that

$$(\overline{H} - \varepsilon\overline{\Delta})\,\overline{C}_k = 0 \tag{1.25}$$

Hence for a non trivial solution

$$\partial et\ (\overline{H} - \varepsilon\overline{\Delta}) = 0 \tag{1.26}$$

known as the secular equations or secular determinant, and hence one obtains the roots ε_1, ε_2 ... ε_n, and by substitution back one derives the coefficients C.

1.3 Hartree-Fock Selfconsistent Field Theory

The fundamental basis of a selfconsistent field approach derives from the treatment of atomic systems by Hartree[10] and later extended by Fock[11] and Slater[12] to include antisymmetry of wave functions. The method was that of minimising the energy resulting from a single determinantal wave function to yield a set of integrodifferential equations. Thus the Hartree-Fock wave function is the variationally best wave function which can be constructed by assigning each electron to a separate orbital, or function depending only on the coordinates of that electron. For the original atomic calculations the equations could be solved to high accuracy by numerical integration, though for molecules the orbitals ϕ_i are invariably expanded in terms of a set of analytic basis functions. Thus, each molecular orbital ψ_i is expressed

as a linear combination of basis functions ϕ_p

$$\psi_i = \sum_{p=1}^{m} c_{ip}\, \phi_p \qquad (1.27)$$

The ground state of the system (in the closed shell case, with each orbital doubly occupied) is to be wriqten as a single Slater determinant,

$$\Psi = |\psi_1(1)\bar{\psi}_1(2)\psi_2(3)\bar{\psi}_2(4) \ldots\ldots \bar{\psi}n(2n)| \qquad (1.28)$$

where $\bar{\psi}$ represents a spin orbital of spin part β. The non-relativistic Hamiltonian may be expressed (in atomic units) as

$$\mathcal{H} = \sum_{i=1}^{2n} H(i) + \sum_{\substack{i=1 \\ i<j}}^{2n-1} \sum_{j=i+1}^{n} \frac{1}{r_{ij}} \qquad (1.29)$$

and hence one may derive the fundamental expression for the tocal energy of the system

$$E = 2\sum_{i=1}^{n} Hi + \sum_{i=1}^{} \sum_{j=1}^{} (2J_{ij} - K_{ij}) \qquad (1.30)$$

H(i) is the one-electron operator

$$H(i) = -\frac{\nabla_i^2}{2} - \sum_{k=1}^{K} \frac{1}{R_{kj}} \qquad (K \text{ nuclei in the system}) \qquad (1.31)$$

The matrix elements H_i, J_{ij} and K_{ij} are defined by

$$Hi = \langle \psi_i(1) | H(1) | \psi_i(1) \rangle \qquad (1.32)$$

$$Jij = \langle \psi_i(1)\psi_j(2) | \frac{1}{r_{12}} | \psi_i(1)\psi_j(2) \rangle \qquad (1.33)$$

$$Kij = \langle \psi_i(1)\psi_j(2) | \frac{1}{r_{12}} | \psi_j(2)\psi_i(1) \rangle \qquad (1.34)$$

The ψi are solutions to the pseudo-eigenvalue equations

$$F\psi_i = \varepsilon_i \psi_i \qquad (1.35)$$

the Fock operator, F, being given by the expression

$$F = H + \sum_{j=1}^{n} (2J_j - K_j) \qquad (1.36)$$

where J_j and K_j are defined by the equations

$$\langle \psi_i | J_j | \psi_i \rangle = J_{ij} \tag{1.37}$$

$$\langle \psi_i | K_j | \psi_i \rangle = K_{ij} \tag{1.38}$$

The orbital energies are given by

$$\varepsilon_i = H_i + \sum_{j=1}^{n} (2J_{ij} - K_{ij}) \tag{1.39}$$

so that

$$E = \sum_{i=1}^{n} (H_i + \varepsilon_i) \tag{1.40}$$

Equation (1.35) is a pseudo-eigenvalue equation in the sense that both F and the ψ_i are determined by the coefficients c_{ip}. The expectation value for the energy of Ψ is minimised with respect to the coefficients c in the following procedure. The operator satisfying (1.35) can only be found by trial and error and the equations are solved by guessing an initial set of functions, solving the equations for a new set and repeating the process with this new set. This is repeated until the set of functions at each end of a cycle are close enough within the desired degree of accuracy, when self consistency is said to have been achieved.

These equations may also be derived from the variation theorem, thus showing that the SCF method does minimise the energy of a system described by a given configurational wave function. It is wished to minimise the energy of Ψ with respect to the coeeficients c_{ip}, such that

$$\langle \psi_i | \psi_j \rangle = \delta_{ij} \tag{1.41}$$

The energy of the system is

$$E = \langle \Psi | \mathcal{H} | \Psi \rangle \tag{1.42}$$

which can be rewritten as in (1.30) where the sums extend over the doubly occupied orbitals of the system (2n electrons). Defining the operators Jj and Kj as in (1.37) and (1.38)

$$E = 2 \sum_{i=1}^{n} \langle \psi_i | H | \psi_j \rangle + \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \psi_i | Gj | \psi_i \rangle \tag{1.43}$$

where $G_j = 2J_j - K_j$

It may be noted that

$$\langle \psi_i | G_j | \psi_i \rangle = G_{ij} = G_{ji} = \langle \psi_j | G_i | \psi_j \rangle \qquad (1.44)$$

A small variation in E will now be considered. The basis set is regarded as fixed and the variation can only be carried out through changes in the coefficients c_{ip}. If the energy is at a true minimum, $\delta E = 0$ for any such variation.

$$\delta E = 2\sum_{i=1}^{n} (\langle \delta\psi_i | H | \psi_i \rangle + \langle \psi_i | H | \delta\psi_i \rangle)$$

$$+ \sum_{i=1}^{n} \sum_{j=1}^{n} (\langle \delta\psi_i | G_j | \psi_i \rangle + \langle \psi_i | G_j | \delta\psi_i \rangle + \langle \psi_i | \delta G_j | \psi_i \rangle) \qquad (1.45)$$

The term involving δG_j may be eliminated since

$$\langle \delta\psi_i | G_j | \psi_i \rangle + \langle \psi_i | G_j | \delta\psi_i \rangle = \langle \psi_j | \delta G_i | \psi_j \rangle \qquad (1.46)$$

thus carrying out the double summation

$$\delta E = 2\sum_{i=1}^{n} (\langle \delta\psi_i | H | \psi_i \rangle + \langle \psi_i | H | \delta\psi_i \rangle$$

$$+ \langle \delta\psi_i | \sum_{j=1}^{n} G_j | \psi_i \rangle + \langle \psi_i | \sum_{j=1}^{n} G_j | \delta\psi_i \rangle) \qquad (1.47)$$

The second and fourth terms are simply the complex conjugates of the first and third, i.e.

$$\langle \psi_i | H | \delta\psi_i \rangle = \langle \delta\psi_i | H | \psi_i \rangle^* \qquad (1.48)$$

since all the operators involved are Hermitian. Therefore

$$\delta E = 2\sum_{i=1}^{n} (\langle \delta\psi_i | H + \sum_{j=1}^{n} G_j | \psi_i \rangle + \langle \delta\psi_i | H + \sum_{j=1}^{n} G_j | \psi_i \rangle^*) \qquad (1.49)$$

and since this is subject to (1.35)

$$\langle \delta\psi_i | \psi_i \rangle + \langle \psi_i | \delta\psi_j \rangle = 0 \qquad (1.50)$$

If an arbitrary set of multipliers ϵ_{ij} are now defined and summation carried out over both i and j

$$-2\sum_{i=1}^{n}\sum_{j=1}^{n}\varepsilon_{ji}\left(\langle\delta\psi_i|\psi_j\rangle + \langle\psi_i|\delta\psi_j\rangle\right) = 0 \qquad (1.51)$$

and making use of the relation

$$\langle\psi_i|\delta\psi_j\rangle = \langle\delta\psi_j|\psi_i\rangle^* \qquad (1.52)$$

(1.51) may be written

$$-2\sum_{i=1}^{n}\sum_{j=1}^{n}\varepsilon_{ji}\langle\delta\psi_i|\psi_j\rangle - 2\sum_{i=1}^{n}\sum_{j=1}^{n}\varepsilon_{ij}\langle\delta\psi_i|\psi_j\rangle^* = 0 \qquad (1.53)$$

According to Lagrange's method of undetermined multipliers, (1.49) and (1.53) are now added

$$\delta E = 0 = 2\sum_{i=1}^{n}\left(\langle\delta\psi_i|H+\sum_{j=1}^{n}G_j|\psi_i\rangle - \sum_{j=1}^{n}\varepsilon_{ji}\langle\delta\psi_i|\psi_j\rangle\right)$$

$$+ 2\sum_{i=1}^{n}\left(\langle\delta\psi_i|H+\sum_{j=1}^{n}G_j|\psi_i\rangle^* - \sum_{j=1}^{n}\varepsilon_{ij}\langle\delta\psi_i|\psi_j\rangle^*\right) \qquad (1.54)$$

Lagrange's method now states that in order for δE to equal zero, each term in the sum must vanish. Furthermore, since the variations are arbitrary, their 'coefficient' must in turn vanish, and hence

$$(H + \sum_{j=1}^{n}G_j)|\psi_i\rangle = \sum_{j=1}^{n}\varepsilon_{ji}|\psi_j\rangle \qquad (1.55)$$

$$(H + \sum_{j=1}^{n}G_j)^*|\psi_i\rangle^* = \sum_{j=1}^{n}\varepsilon_{ij}|\psi_j\rangle^* \qquad (1.56)$$

Taking the complex conjugate of (1.56) and subtracting it from (1.55)

$$\sum_{j=1}^{n}(\varepsilon_{ji} - \varepsilon_{ij}^*)|\psi_j\rangle = 0 \qquad (1.57)$$

Thus $\varepsilon_{ij} = \varepsilon_{ji}^*$ for all i,j and the matrix $\bar{\varepsilon}$ whose elements are ε_{ij} is Hermitian.

Equations (1.55), (1.56) are not quite eigenvalue relations since ψ_i is transformed into a sum over ψ_j. However, the sum can be diagonalised. Considering the following row vector $\bar{\psi}$ whose elements are the ψ_i

$$\bar{\psi} = (\psi_1\psi_2\psi_3 \ldots\ldots \psi_n) \qquad (1.58)$$

Then (1.55), (1.56) may be written in the form

$$\overline{F}\ \overline{\psi} = \overline{\psi}\ \overline{\varepsilon} \tag{1.59}$$

$$\text{or } \overline{F}^{1}\ \overline{\psi}^{1} = \overline{\psi}^{1}\ \overline{\varepsilon}^{1} \tag{1.60}$$

Since $\overline{\varepsilon}^{1}$ is Hermitian, there exists a unitary matrix $\overline{U}$ such that

$$\overline{\varepsilon} = \overline{U}^{\dagger}\ \overline{\varepsilon}^{1}\ \overline{U} \tag{1.61}$$

is a diagonal matrix. Defining $\overline{\psi} = \overline{\psi}^{1}\overline{U}$ it can be proved that

$$\overline{F}^{1}\ \overline{\psi} = \overline{\psi}\ \overline{\varepsilon} \tag{1.62}$$

which is now a true (pseudo) eigenvalue equation with components of the form

$$F^{1}\psi_{i} = \varepsilon_{i}\psi_{i} \tag{1.63}$$

where $\varepsilon_{i} \equiv \varepsilon_{ii}$. However, it is not immediately obvious that this diagonal form still minimises the energy, nor is the meaning of F^{1} entirely well defined. However, it may be noted that the unitary transform $\overline{\psi} = \overline{\psi}^{1}\ \overline{U}$ merely adds certain linear combinations of rows of the Slater determinant Ψ without changing its value; i.e. $\Psi = \psi^{1}$. Furthermore, $F = F^{1}$ $H = H^{1}$ since this does not involve ψ_{i}. It is noted that

$$\sum_{j=i}^{n} G_{j}^{1} = \sum_{j=1}^{n} (2J_{j}^{1} - K_{j}^{1}) = \sum_{j=1}^{n}\sum_{p=1}^{m}\sum_{q=1}^{m} C^{1}{}_{jp}^{*}\, C_{jq}^{1}\, (2\, J_{pq} - K_{pq}) \tag{1.64}$$

where the orbital operators have been expanded in terms of constituent 'basic operators'. The coefficients C^{1}_{jp} are related to the unprimed coefficients C_{jp} by the equation

$$C_{ip} = \sum_{j=1}^{n} C^{1}{}_{jp}\, U_{ji} \tag{1.65}$$

$$\text{or } C^{1}_{ip} = \sum_{j=1}^{n} C_{jp}\, U^{*}_{ij} \tag{1.66}$$

Substituting in (1.64)

$$\sum_{j=1}^{n} G_j^1 = \sum_{j=1}^{n}\sum_{p=1}^{m}\sum_{q=1}^{m}\sum_{k=1}^{n}\sum_{l=1}^{n} c_{kp}^{*}\, c_{lq}\, U_{jk}\, U_{jl}^{*}\, (2J_{pq} - K_{pq}) \quad (1.67)$$

The sum over j introduces a factor of δkl, hence

$$\sum_{j=1}^{n} G_j^1 = \sum_{k=1}^{n}\sum_{p=1}^{m}\sum_{q=1}^{m} c_{kp}^{*}\; c_{kq}^{*}\, (2J_{pq} - K_{pq}) \quad (1.68)$$

The right hand side is, by definition, equal to the total two-electron operator in the unprimed scheme, so

$$\sum_{j=1}^{n} G_j^1 = \sum_{j=1}^{n} G_j \quad (1.69)$$

Hence it has been proved that

$$\bar{F}\,\bar{\psi} = \bar{\varepsilon}\,\bar{\psi} \quad (1.70)$$

or, in component form,

$$F\psi_i = \varepsilon_i \psi_i \quad (1.71)$$

The SCF equations thus follow directly from the variational principle. Also it has been proved that an arbitrary unitary transformation on a Slater determinant alters neither the determinant itself, nor the resultant Fock operator. Thus, if it is wished, the molecular orbital picture may be transformed into an equivalent 'localised orbital' scheme in which the standard orbitals chosen are those most nearly representing atomic functions, or certain 'bond' functions.

The diagonalised matrix of multipliers, the set εi, has n values and these are known as orbital energies from (1.71)

$$\int \psi_i^{*}\, F\, \psi_i\, \partial\tau = \varepsilon \quad (1.72)$$

It will be shown later that if a singly ionised state of the system is described in terms of the one-electron functions appropriate to the neutral system, then these orbital energies are the energy difference between the neutral and ionised states i.e. correspond to the ionisation energy

$$\text{Ionisation Potential (i)} \simeq -\varepsilon_i \quad (1.73)$$

This was first demonstrated by Koopmans and is widely known as Koopmans' theorem[13]. For closed shell systems, the approximations involved in taking a single determinantal function only, however, must be bourne in mind. Koopmans theorem will be examined in greater depth later.

When the LCAO approximation is incorporated into the Hartree Fock equations the resulting equations are usually known as Roothaan's equations[14].

From (1.71)

$$H^{SCF}\psi_i = \varepsilon_i^{SCF}\psi_i \tag{1.74}$$

and by the LCAO approximation

$$\psi_i = \sum_n c_{in}\chi_n \tag{1.75}$$

thus

$$H^{SCF}\sum_n c_{in}\chi_n = \varepsilon_i^{SCF}\sum_n c_{in}\chi_n \tag{1.76}$$

hence

$$\sum_n c_{in}\int\chi_m H^{SCF}\chi_n dv = \varepsilon_i^{SCF}\sum_n c_{in}\int\chi_m\chi_n dv \tag{1.77}$$

which can be written

$$\sum_n c_{in}\,(H_{mn}^{SCF} - \varepsilon_i^{SCF} S_{mn}) = 0 \tag{1.78}$$

which has a non-trivial solution when

$$\det\left|H_{mn}^{SCF} - \varepsilon_i^{SCF} S_{mn}\right| = 0 \tag{1.79}$$

1.4 Form of the Basis Functions Used

The molecular orbitals for assembly of nuclei and electrons are defined by the expansions

$$\Psi = \sum c_{ai}\psi_i \tag{1.80}$$

$$\psi i = \sum a_{\mu i}\chi_\mu \tag{1.81}$$

and the total set of functions used to describe the system is known as

the basis set. (1.81) is a mathematical description of the linear combination of atomic orbitals procedure and the type of basis functions χ_μ used in non-empirical calculations fall into two distrinct classes, depending upon whether the radial portion of the function is a single exponential ($e^{-\xi r}$) or a gaussian type function ($e^{-\alpha r^2}$).

Single exponential functions. Functions of this type are suggested by the exact solution for one electron atoms, where the wave functions have the form of a product of a function of r (nuclear-electron separation) and spherical harmonics. The function of r may be placed in exponential form

$$\phi(r,\theta,\ ) = Nr^{n-1}\exp\ (-\xi r)\ Y_{lm}\ (\theta,\phi) \qquad (1.82)$$

and Slater[15] proposed a set of rules for the calculation of the orbital exponent ξ for any atom required. The value of the exponent allows the function to vary its radial maximum from the nucleus and is related to the screening effect from the remaining electrons, but it is not usually possible to get good agreement at all distances from the nucleus with a single Slater function per atomic orbital. In their simplest analytical form they are not mutually orthogonal but this is not an important limitation and maybe overcome by the construction of linear combinations.

Gaussian Type functions. The use of these were proposed by Boys[16], and have the form

$$\phi^1(r,\theta,\phi) = N^1 r^{n^1}\ \exp\ (-\alpha r^2)\ Y\ (\theta,\phi) \qquad (1.83)$$

The use of these functions considerably simplifies multicentre integral evaluation, since it can be proved that the product of two gaussian functions centred on two different atoms is another guassian function centred on the line joining the original centres. Thus integrals of the type

$$\int \phi(g)_a \phi(g)_b\ \frac{1}{r_{12}}\ \phi(g)_c \phi(g)\ d\tau \qquad (1.84)$$

are immediately simplified to the form

$$\int \phi(g)_e \frac{1}{r_{12}} \phi(g)_f d\tau \qquad (1.85)$$

Unfortunately, for s type functions, the form of a single gaussian does not closely resemble the form of a true atomic orbital, particularly in the nuclear region where the cusp is lacking,

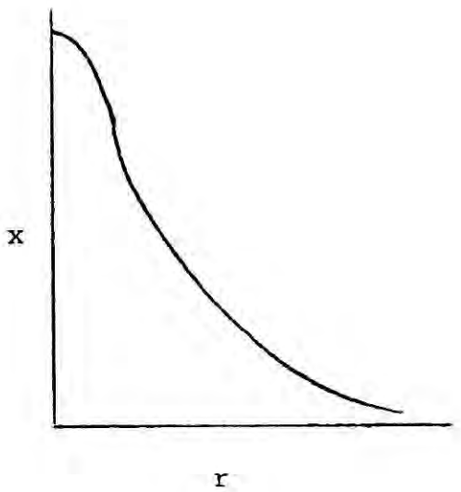


fig. 1.1
Form of a single gaussian function

and also at very large distances. For these reasons it is necessary to use not one gaussian per orbital but several. A minimal basis set is one consisting of the least number of atomic orbitals (of appropriate symmetry) for the atomic ground state, i.e. it includes one function for each occupied atomic orbital with distinct n and l quantum numbers for the component atoms of the molecule. Though calculations using minimal basis sets have had their interpretive uses, it is generally better to use more extended basis sets. However, the number of integrals to be evaluated in a calculation (without regard to the use of symmetry) may be determined from the number of basis functions employed from the formulae

$$\text{Number of 1 electron integrals (p)} = \frac{n(n+1)}{2} \qquad (1.86)$$

Number of 2 electron integrals = $\frac{p(p+1)}{2}$ (1.87)

where n is the number of basis functions used. Thus the number of integrals is roughly proportional to the fourth power of the number of basis functions employed. There are further problems involved with large basis sets. Increased computer time and storage space are required for the construction and diagonalisation of the much larger Fock matrix. Furthermore, as the size of the basis set is increased, so problems with convergence may be encountered in the S.C.F. section of calculations. A further choice must be made as to the nature of the variational freedom.

$$\psi mo = \sum c_{ai}\phi_i \qquad \phi i = \sum a_{\mu i}\chi_\mu \tag{1.88}$$

The exponents of the χ_μ and the coefficients $a_{\mu i}$ may be treated as variational parameters in addition to the MO coefficients (the c_{ai}). However, the greater variational freedom is only achieved at considerable computational expense and in general exponents are fixed (usually having been optimised for the atom) and the coefficients allowed to vary, although calculations employing minimal basis sets with variation in both exponents and coefficients may be employed. Use of exponents optimised for atomic systems is justifiable on the grounds that little distortion will occur of formation of the molecular system. An answer to the problem of the rising number of basis functions has been found to lie in fixing certain coefficients relative to one another i.e. using linear combinations of some gaussians on the same centre e.g.

$$\psi i = \sum_k c_{ik}\chi_k \tag{1.89}$$

$$\chi_k = c_1{}^1G_1 + c_2{}^1G_2 + c_3{}^1G_3 \tag{1.90}$$

In this way the large 'uncontracted' basis set is broken up into a rather smaller basis set of 'contracted' gaussian functions[17]. As explained above the orbital exponents are generally taken from

calculations upon isolated atoms. For example, using a basis set of 10 '1s' and 6 '2p' gaussian functions, Husinaga[18] minimised the energy of the carbon atom (^{3}Pground state) with respect to variation of the exponents and coefficients. The contraction is then carried out by inspection of the exponents-for example, those of large exponent will describe the region close to the nucleus - and then taking linear combinations using the appropriate contraction coefficients. Dunning[19] has recently provided simple criteria for efficient contraction of large basis sets. Of course, in a molecule the orbitals will be distorted relative to free atom orbitals, which can be accommodated by optimising the exponents for molecular environments. This, however, is computationally expensive and has only been investigated for molecules of interest to the organic chemist by Pople et al[20] using minimal STO nG basis sets. From this a standard set of 'molecular' exponents may be obtained. Core orbitals remain effectively unchanged in going from free atom to the colecule, and hence the only exponents worthwhile investigating are those of the valence orbitals.

A good example of the advantages, indeed necessity, of the contraction of primitive gaussian function basis sets may be drawn from the work of Clementi[21] on the hydrogen bonded guanine - cytosine base pair. The calculation of certain aspects of the potential energy surface required 8 days of central processor time on the largest commercial computer available at that time (IBM 360/195) and involved the computation, sorting, retrieving and processing of some 7×10^{10} integrals over gaussian functions. For the uncontracted basis set of 334 functions 2.41×10^9 integrals would have to be processed in the SCF calculation for each point on the potential energy surface. The storage space required would occupy a few hundred magnetic tapes and would be physically impractical. With a contracted basis set of 105 functions, however, the integrals could be stored on two magnetic tapes.

Functions of the type described by (1.82) are usually called cartesian gaussians but this is not the only way in which gaussian type functions may be used. Preuss[22] and Whitten[23] independently proposed an alternative method of gaussian function contraction to avoid the difficulties implicit in integration over the angular part of the basis functions. This method produces the desired angular characteristics by taking a linear combination of simple gaussians centred at different points (for example, the lobes of a p type orbital might be represented by two s type functions centred either side of the atomic nucleus), and is known as the Gaussian - lobe method. For example, a p-type gaussian lobe function can be expressed as

$$g_p^L(\bar{\gamma}) = N_p^L \{\exp(-\alpha(r-R^o\bar{\gamma})^2) - \exp(-\alpha(r+R^o\bar{\gamma})^2)\} \qquad (1.91)$$

where $\bar{\gamma}$ is a unit vector and R^o is a constant defining the distance from the origin or the centre for the two simple gaussians. The apparent disadvantage of using gaussian lobe functions is the extra constants (R^o); however, it can be shown that if R^o is set equal to $C\alpha^{-\frac{1}{2}}$, where α is the gaussian exponent and c is a constant (c~0.03), then for a given set of gaussian functions using the same exponents the results are closely similar whether cartesian or lobe basis sets are employed[24].

A further type of gaussian basis set gives promising results for quite large molecules, known as Floating Spherical Gaussian Orbitals (FSGO). In the cases of cartesian and lobe basis sets, calculations employing these basis sets have generally been performed with the orbitals centred upon atoms in the molecule, thus allowing interpretation of results in terms familiar to the organic chemist. This is not, however, the case with FSGO basis sets, which are defined by

$$\phi_i = N\exp(-(r-R_i)^2/\rho_i^2) \qquad (1.92)$$

where ρ_i is the radius of the orbital i and R_i is its position. For a given basis set, the energy of the molecule is then minimised with

respect to the position of the gaussian orbitals and their radii. The chief advantage of these functions lies in their computational simplicity.

A useful hybrid of the single exponent and gaussian type function approaches may be found in the STO-nG method of Pople et al[20]. This method takes a small linear combination of gaussian functions and fits them by a least squares procedure to a Slater type orbital i.e. each STO is replaced by an atomic orbital ϕ_μ' which is a linear combination of k gaussians (k = 2 - 6). These combinations are obtained for STOs with $\xi = 1$ and then uniformly sealed thus

$$\phi_\mu{}^1(\xi,r) = {}^{3}/_{2}\ \phi_\mu'(1,\xi r) \tag{1.93}$$

where $$\phi_{1s}(1,r) = \sum_k^K \partial_{1s},\ {}_k g_{1s}\ (\alpha_{1k},r) \tag{1.94}$$

$$\phi_{2s}(1,r) = \sum_k^K \partial_{2s},\ {}_k g_{2s}\ (\alpha_{2k},r) \tag{1.95}$$

$$\phi_{2p}(1,r) = \sum_k^K \partial_{2p},\ {}_k g_{2p}\ (\alpha_{2k},r) \tag{1.96}$$

where each g is a standard gaussian type function. The values of ∂ and α are chosen to minimise the integrals

$$\varepsilon_{1s} = \int(\phi_{1s} - \phi_{1s}')^2 \partial\tau \tag{1.97}$$

$$\varepsilon_{2s} + \varepsilon_{2p} = \int(\phi_{2s} - \phi_{2s}')^2 \partial\tau + \int(\phi_{2p} - \phi_{2p}')^2 \partial\tau \tag{1.98}$$

The method is rapidly convergent and the results simulate those based on an STO basis set. The STO-3G basis set is, in particular, economical to use, and can be applied to quite large molecules.

As a variant of this, Pople et al[25] have also used a split valence 4-31G basis set, in which each inner shell is represented by a single basis function taken as a sum of four gaussians and each valence orbital is split into inner and outer parts, described by 3 and 1 gaussian functions respectively. The expansion coefficients and

gaussian exponents are determined by minimising the total calculated energy of the atomic ground state. This method allows for anisotropy in the valence shell.

For atomic systems,[26, 27] where spherical symmetry is present it is possible to solve the Hartree-Fock equations numerically, not requiring an explicit analytical form for the wave functions but merely a tabulation of electronic density with r. In principle this could be extended to molecular systems, and the exact solution produced is equivalent to an LCAO expansion to an infinite number of terms. For a closed shell system the energy corresponding to a perfect single determiental wave function is known as the Hartree-Fock limit, and excludes electron correlation and relativistic effects.

1.5 Limitations upon Hartree-Fock solutions

From the proceeding discussion it will be evident that for closed shell systems it is possible, in principle, to approach the Hartree-Fock limit, providing that a large enough basis set and sufficient computer power are available. In general, in this work the interest has been in relative energies rather than absolute energies. Accurate representation of heats of reaction for the processes of interest are approachable with basis sets which provide total energies some distance from the Hartree-Fock limit. The work of Pople and coworkers[28] on isodesmic processes - i.e. those in which bond types and numbers of electron pairs remain the same in going from reactants to products, has shown that these reactions may be represented with minimal STO-3G basis sets. However there are certain limitations inherent in the Hartree Fock description of molecular electronic structure.

(a) Relativistic Correction. The virial theorem[29, 30] may be written

$$2\langle \mathrm{Top} \rangle_{av} = \langle (\bar{r} \cdot \bar{\nabla} V)_{op} \rangle_{av} \qquad (1.99)$$

and hence an electron in a region of high potential will have a

correspondingly high kinetic energy. Thus the inner shell or core electrons which are close to the nucleus will be subject to relativistic effects, and the heavier the element, the greater the correction necessary. This energy is certainly not negligible and is difficult to estimate but since it remains effectively constant during change of atomic environment, the relative energies calculated will not be affected.

(b) Correlation Energy. This is a rather more serious problem and results from the fact that Hartree Fock theory is based upon a wave function which is a product of one-electron orbitals i.e. each electron experiences an average field provided by the remaining electrons without explicit consideration of the instantaneous correlation of electronic motions. The usual definition of correlation energy is that it is the difference between the experimentally observed energy when it has been corrected for the relativistic term and the calculated energy at the Hartree-Fock limit. Löwdin[31] has defined correlation energy as: the correlation energy for a certain state with respect to a specified Hamiltonian is the difference between the exact eigenvalue of the Hamiltonian and its expectation value in the Hartree-Fock approximation for the state under consideration or, mathematically, as

$$E_{corr} = \langle \mathcal{H} \rangle_{av} \text{ (exact)} - \langle \mathcal{H} \rangle_{av} \text{ (Hartree-Fock)} \qquad (1.100)$$

Thus defined it is known as the non-relativistic correlation energy. The difference expressed in (1.100) results from the fact that an electron is an atom will have instantaneous interactions with all the other electrons in the atom, and particularly[32] with the other electron in the same spatial orbital.

Various methods are available for estimating corrections due to correlation effects.

(i) Within the LCAO-MO formalism:

(a) Use of pair correlation energies

Hollister and Sinanoğlu[33] have proposed the use of atomic

correlation data, known as the pair population method, in which a number of pairs of electrons are assigned to each atomic orbital employed in a minimal basis set SCF wave function for a molecule. The total correlation energy is then evaluated as a sum over all atomic orbitals, of atomic pair correlation energies, ε_{ij}, weighed by the atomic orbital pair populations (i.e. $\frac{1}{2}\varrho i$ where ϱ is the Mulliken atomic orbital charge density). Thus the total correlation energy is given by

$$E_{corr} = \sum_i \varepsilon_{ii} \tfrac{1}{2}\varrho_i + \sum_{i<j} \varepsilon_{ij}\varrho_i\varrho_j \qquad (1.101)$$

where the first term allows for correlation effects between electrons centred upon the same atom, and the second for electrons centred on different atoms.

Snyder and Basch[34] have used this method to try to estimate the correlation corrections required for the calculation of thermochemical data, and some of their results are shown in table 1.1. It can be seen that the effect is quite large. However the effect upon

Table 1.1 Correlation energies for some small molecules.

Molecule	Ecorr (intra-atomic) (au)
H_2	-0.0409
CH_4	-0.2952
C_2H_6	-0.5288
C_2H_4	-0.4772
HCOOH	-0.8235
HCOF	-0.8319

relative energies is considerably smaller, for if this method is applied to the relative energies of protonated ethylene and ethyl cation

ΔEcorr (intra-atomic) (ethylcation-pronated ethylene)	ΔEcorr (inter-atomic)
-0.3 kcal mole^{-1}	0.0 kcal mole^{-1}
+0.7 kcal mole^{-1}*	

*using polarisation functions

(b) The Hartree Fock wave function may be improved by the introduction of configuration interaction (C.I.)[35], that is to allow the calculated Hartree-Fock ground state to mix with other configurations of the same symmetry.

(c) Correlation energies may be estimated by means of a numerical integration of the wave function[36].

(d) The Hartree Fock method may also be improved by the use of a wave function expressed as a linear combination of Slater determinants, rather than the usual single determinantal wave function.

(ii) Other methods are available which are not within the framework of the LCAO formalism:

(e) The generalised valence bond approach[37].

(f) Correlated wave functions have been used. This necessitates dispensing with orbital wave functions and using functions which include interelectronic distance. Such functions[38] have the effect of introducing a coulomb 'hole' for both parallel and anti-parallel spins.

The single most important defect of the Hartree-Fock treatment is the inability to describe bond breaking. Fig. 1.2 shows the energy

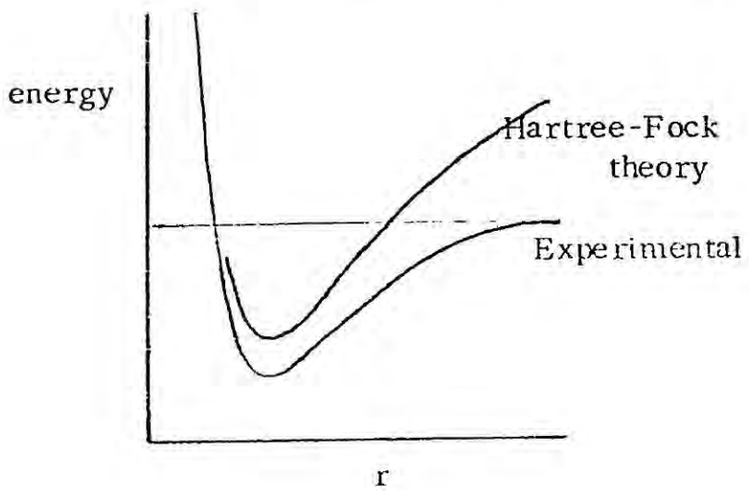


Fig 1.2 Behaviour of Hartree Fock energy and experimental energy as functions of internuclear distance

of the H_2 molecule as obtained from Hartree-Fock calculations compared with the exact non-relativistic energies. Expansion of the single-determinantal wavefunction corresponding to the ground state for H_2 shows that, at large internuclear distances, instead of arriving at a dissociation limit $2\dot{H}$, the Hartree-Fock wavefunction behaves incorrectly. This arises from spurious ionic terms. It is found that the Hartree-Fock energy curve and the exact energy curve are parallel to one another up to ~2.5 a.u., a distance corresponding to twice the HH bond length, the correlation energy remaining fairly constant in this region. At large internuclear distances the correlation energy increases rather rapidly and it is in this region that the Hartree-Fock method goes seriously astray. Thus bond stretching up to a certain point will be described adequately by Hartree-Fock theory, i.e. the solution will remain parallel to the exact solution, but in the extreme case of atomisation of a molecule, the correlation errors will rise and hence heats of atomisation are poorly described within this framework. However, the parallel behaviour of the curves at normal bond distances does mean that one of the most readily calculable properties of a molecule, even with a poor basis set, is its geometry. Furthermore, in the absence of near degeneracy of states, and provided that no differences in electron pairing are introduced during a molecular change, then relative energies will be predicted with accuracy.

1.6 Symmetry Considerations

If the result of a certain set of operations upon a figure produces an equivalent configuration, identical in all respects with the initial configuration, then the figure is said to possess symmetry, and the set of operations are known as symmetry operations which define the 'symmetry group' or 'group' for the figure[39,40]. Some of the possible symmetry operations are:

E Identity operation, leaving all points where they were

Cn Rotation about axis by $2\pi/n$ degrees

σ(ij) Reflection through plane defined by axes i and j

i Inversion through a centre of symmetry

Sn Rotation of $2\pi/n$ about an axis followed by reflection through a plane perpendicular to it.

The set of symmetry operations possessed by a molecule is known as the molecular point group, and this group is seen to be a mathematical 'group' as defined by the following relations:

(i) An identity operation E exists such that ER=RE=R where R is a member of the set

(ii) The product of any two operations in the set is another element in the set

(iii) Multiplication is associative i.e.

$$P(QR) = (PQ)R \tag{1.102}$$

(iv) There exists for every element R of the set, a reciprocal element S defined as

$$RS = SR = E \tag{1.103}$$

From the second point in this definition it is clear that a 'multiplication table' can be drawn up for a group showing the result of multiplication of any element of the group by itself or any other.

The operations can be divided into certain 'classes' defined as: two elements P and Q belong to a class if

$$X^{-1}PX \text{ and } X^{-1}QX \tag{1.104}$$

yield either P or Q for any X of the group.

The next task in group theory is to associate with every symmetry operation of the group a known mathematical quantity having a one to one correspondence with each symmetry operation and also obeying the four requirements of a group. This constitutes a 'representation' of the group. Clearly the association of 1 with every operation will fulfill

the definition of a representation. Others will usually be found by associating 1 with some operations and -1 with others. Another will arise from associating with each symmetry operation, a matrix containing information on the coordinates for any point in the molecular before and after the symmetry operation. In general any set of matrices which multiply together in such a way as to satisfy the group multiplication table may be called a representation of the group. The case of associating 1 or -1 with each operation is a special case where the matrices are of order 1.

A representation is called reducible if it is possible to find a matrix α such that by performing $\alpha^{-1} R \alpha$ on any one of the elements of the representation, a new matrix R^1 is obtained being more diagonal than R. If this is not possible then the representation is said to be irreducible.

The 'character' of a representation is the sum of the diagonal elements of one matrix of the irreducible representation, and a 'character table' of the groups is built up to, tabulating the characters for each operation in each representation. For example for the ammonia molecule

	E	(a)	(b)	(c)	$C_3(Z)$	$C_3^1(Z)$
Γ_1	1	1	1	1	1	1
Γ_2	1	-1	-1	-1	1	1
Γ_3	2	0	0	0	-1	-1

In a point group containing several operations it is not necessary to always consider the full symmetry of the group and it is possible to form smaller sub-groups. The formal way to accomplish this or the reverse process is to form a 'direct product'. Thus if there are two groups G_1, containing operations R_i (i=1,2...n) and, for G_2, R_j (j=1,2...m), if all the operation R_i and R_j commute, then the new group containing all the products $R_i R_j = R_j R_i$ is called the direct product group. This

operation may also be performed between irreducible representations of the same group, for example in the ammonia group

$$\Gamma_1 \times \Gamma_1;\ \Gamma_2 \times \Gamma_2 = \Gamma_1;\ \Gamma_3 \times \Gamma_3 = \Gamma_1 + \Gamma_2 + \Gamma_3;\ \Gamma_2 \times \Gamma_1 = \Gamma_2;$$
$$\Gamma_3 \times \Gamma_1 = \Gamma_3;\ \Gamma_2 \times \Gamma_3 = \Gamma_3 \qquad (1.104)$$

The use of group theory in quantum mechanics is extensive and some of the uses will be considered.

(a) If operation R belongs to the symmetry group of a molecular system, then the Schroedinger equation for the system may be written

$$H\psi_1 = E_1\psi_1 \qquad (1.105)$$

but also

$$HR\psi_1 = E_1R\psi_1 \qquad (1.106)$$

(b) Molecular orbitals transform as one of the irreducible representations of the group which provides a useful notation for electronic configuration. Thus for ammonia the configuration is

$$(1\Gamma_1)^2\ (2\Gamma_1)^2\ (3\Gamma_1)^2\ (1\Gamma_3)^4$$

(c) If the direct product of any integral over molecular orbitals does not contain the Γ_1 representation (i.e. is not totally symmetric), then its value is zero.

(d) Construction of Symmetry Adapted Functions. This involves taking the basis set of functions which represent the symmetry of the component atoms of the molecule, and transforming to a basis set with symmetry characteristics of the molecule as a whole. The general method of generation of symmetry adapted functions is to take one function ϕ of the basis set and operate on it with each of the operations of the group. The result of each operation is multiplied by the appropriate character χ_t of the irreducible representation and the results summed i.e.

$$\psi_t = \sum_R \chi_t(R)(R\phi) \qquad (1.107)$$

where ψ_t is the symmetry adapted function which transforms like the irreducible rep^n labelled t.

1.7 Ab Initio and Semi-empirical LCAO-MO-SCF calculations

In (1.3) the Hartree-Fock equations of the form

$$\overline{F}\,\overline{\psi} = \overline{\varepsilon}\,\overline{\psi} \tag{1.108}$$

were derived, where the Fock operator could be written in the form

$$F_{ij} = H^{o}_{ij} + \sum_{k}\sum_{l} P_{kl}\left(\langle ij|kl - \tfrac{1}{2}\, jk|il\rangle\right) \tag{1.109}$$

where

$$P_{kl} = 2\sum_{m}^{occ} C^{*}_{mk}\, C_{ml} \quad (1.10), \quad H^{o}_{ij} = \int \chi_i(\mu)\left(-\tfrac{1}{2}\nabla^2_{\mu} - \sum_{A}\frac{ZA}{r_{A\mu}}\right)\chi_j(\mu)\,d\tau \tag{1.111}$$

$$\langle ij|kl\rangle = \iint \chi_i(\mu)\chi_k(\nu)\frac{1}{r_{\mu\nu}}\chi_j(\mu)\chi_l(\nu)\,\partial\tau_{\mu}\,\partial\tau_{\mu} \tag{1.112}$$

and thus where P_{kl} is the total electron population in the overlap region between atomic orbitals k and 1. The solution of the secular equations derived from (1.108), i.e.

$$|F_{ij} - ES_{ij}| = 0 \tag{1.113}$$

requires the evaluation of the constituent matrix terms, F_{ij}. The solution of Roothaans equations must be an iterative one in which an initial trial set of coefficients is used to evaluate the Fock matrix elements and hence solve the secular determinant. This solution leads to a better approximation to the wave function i.e. an improved set of P_{kl}. The processes are repeated a number of times, each cycle obtaining an improved set of Pkl until sufficient consistency is obtained between successive iterations. Three main criteria for self consistency are used.

(i) Convergence of total energy to within specific limits viz 10^{-6}au between successive iterations.

(ii) Convergence of energy components i.e. one and two electron energies.

(iii) Comparison of coefficients.

Method (i) is used by POLYATOM and method (ii) by IBMOL

Clearly the main obstacles to the solution of the problem arise from the large number of multicentre integrals (i.e. 3 and 4 centre two electron integrals) requiring evaluation, which place large demands upon computer capability and hence on upper limit upon the size of molecular system for which investigation is feasible. For these reasons a number of semi empirical methods have been devised in which use is made of spectroscopic data for parametrisation to reduce the overall computation necessary.

(a) Ab-Initio methods. In ab-initio (non-empirical) Hartree-Fock calculation all integrals of the Fock matrix are evaluated, leading in general to reasonably reliable estimates of total energy and populations but the above disadvantages with respect to the limited size of system which may be reasonably examined.

(b) Semi-empirical all valence electron, Neglect of Diatomic Overlap Method. This method[41] represents the closest semi-empirical approximation to the ab-initio approach and is particularly suitable for simplifying the Hartree-Fock problem owing to the simplicity and adequacy of its approximations, viz:

(i) Only valence electrons are specifically considered, the inner shells being regarded as an unpolarinable core. Thus only atomic orbitals of the same principle quantum number as that of the highest occupied orbital in the isolated atom are included in the basis set.

(ii) Diatomic differential overlap is neglected i.e.

$$S_{ij} = \int \chi_i(\mu)\chi_j(\mu)d\tau = 0 \qquad (1.114)$$

if the orbitals χ_i and χ_j are not on the same atom, and

$$\langle ij|kl\rangle = 0 \qquad (1.115)$$

unless χ_i and χ_j are atomic orbitals of the same atom and χ_k and χ_l are atomic orbitals also localised upon a given atom. The

diatomic differential overlap approximation considerably reduces the initial number of integrals requiring evaluation. All three and four centre integrals are thus set equal to zero and some two centre integrals also. A small atomic orbital basis set is used for integral evaluation.

(c) Complete Neglect of Differential Overlap method. Even using the above approximations investigation of large molecules necessitates the computation of too many integrals and further simplifications are necessary. These are complicated, however, by the requirement for rotational invariance. Thus, whilst the results for two centre integral evaluation in the non-empirical approach are invariant with respect to an orthogonal transformation of the axes this is not in general true for an approximate treatment, and in such approaches will be dependent upon choice of coordinate systems and orbital hybridisation. To ensure the maintainence of rotational invariance further approximations are restricted to either complete neglect of differential overlap or partial neglect of differential overlap methods. In the Complete Neglect of Differential Overlap method [41,42] both one and two centre integrals involving differential overlap are set equal to zero. Writing the electronic interaction integrals $\langle ii|jj\rangle$ as Γ_{AB}, the Fock matrix elements become

$$F_{ii} = H_{ii} + (P_{AA} - \tfrac{1}{2}P_{ii})\,\Gamma_{AA} + \sum_{B\neq A} P_{BB}\,\Gamma_{AB} \qquad (1.116)$$

$$F_{ij} = H_{ij} - \tfrac{1}{2}P_{ij}\,\Gamma_{AB} \qquad (1.117)$$

where χ_i and χ_j are centred on atoms A and B respectively, and P_{ij} are the components of the charge density and bond order matrix

$$P_{ij} = 2\sum_{m}^{occ} a_{mi}\,a_{mj} \qquad (1.118)$$

and P_{AA} is the total charge density on atom A.

$$P_{AA} = \sum_{i}^{A} P_{ii} \qquad (1.119)$$

The core matrix element H_{ii} may be separated into two components - the diagonal matrix elements of χ_i with respect to the one-electron Hamiltonian containing only the core of its own atom (U_{ii}), and the interaction (V_{AB}) of an electron in χ_i on atom A with the cores of other atoms B. Thus

$$H_{ii} = U_{ii} - \sum_{B \neq A} V_{AB} \tag{1.120}$$

$$\text{hence } F_{ii} = U_{ii} + (P_{AA} - \tfrac{1}{2}P_{ii})\,\Gamma_{AA} + \sum_{B \neq A} (P_{BB}\,\Gamma_{AB} - V_{AB}) \tag{1.121}$$

and the total energy may be expressed as the sum of one-and two-atom terms

$$E = \sum_{A} E_A + \sum_{A<B} \sum E_{AB} \tag{1.122}$$

$$\text{where } E_A = \sum_{i}^{A} P_{ii}\,U_{ii} + \tfrac{1}{2} \sum_{i}^{A} \sum_{j}^{A} (P_{ii}\,P_{jj} - \tfrac{1}{2}P_{ij}^2)\,\Gamma_{AA} \tag{1.123}$$

$$E_{AB} = \sum_{i}^{A} \sum_{j}^{B} (2P_{ij}H_{ij} - \tfrac{1}{2}P_{ij}^2\Gamma_{AB}) + (Z_A Z_B \frac{1}{R_{AB}} - P_{AA}V_{AB} - P_{BB}V_{BB} + P_{AA}P_{BB}\Gamma_{AB}) \tag{1.124}$$

where R_{AB} is the internuclear distance separating atoms A and B. The integrals are evaluated with the aid of semi-empirical parametrisation:

(i) One electron integral U_{ii}. An estimate of this integral may be obtained from spectroscopic data i.e.

$$U_{ii} = -I_A - (Z_A - 1) \tag{1.125}$$

where I_A is the ionisation energy of the electron. An improved method takes into account the electron affinity (A_A), obtaining an 'orbital electronegativity'

$$U_{ii} = -\tfrac{1}{2}(I_A + A_A) - (Z_A - \tfrac{1}{2})\Gamma_{AA} \tag{1.126}$$

(ii) One-centre two-electron integral Γ_{AA}. These are calculated as the electrostatic repulsion energy of two electrons in a Slaters orbital irrespective of actual orbital types involved. Thus

$$\Gamma_{AA} = \iint \chi_{S_A}^2(\mu)\,\frac{1}{r_{\mu\nu}}\,\chi_{S_A}^2(\nu)\,\partial\tau_\mu\,\partial\tau_\nu \tag{1.127}$$

(iii) Two centre one-electron integral Hij. The resonance integral is regarded as being directly proportional to the overlap integral S_{ij} between the orbitals χ_i and χ_j centred on A and B respectively

$$H_{ij} = \beta^{o}_{AB} S_{ij} \qquad (1.128)$$

where S later atomic orbitals are used to calculate S_{ij}. To preserve rotational invariance β^{o}_{AB} should be characteristic of χ_i and χ_j but independent of their spatial positions and hence the parameter is averaged for each atom,

$$\beta^{o}_{AB} = \tfrac{1}{2}(\beta_A{}^{o} + \beta_B{}^{o}) \qquad (1.129)$$

where β^{o}_{A} and β^{o}_{B} etc. are chosen empirically to reproduce experimental results or those from ab initio calculation.

(iv) Two-centre two electron integral Γ_{AB}. This integral is the most difficult to evaluate, and is calculated as

$$\Gamma_{AB} = \langle ii|jj\rangle = \iint \chi_{S_A}^{2}(\mu)\frac{1}{r_{\mu\nu}}\chi_{S_B}^{2}(\nu)\partial\tau_\mu\partial\tau_\nu \qquad (1.130)$$

where χ_{S_A} and χ_{S_B} are the Slater s type orbitals for the atoms A and B.

(v) Coulomb Penetration V_{AB}. The effect of the interaction of an electron in χ_i on atom A with the cores of other atoms B are neglected and the coulomb penetration integrals estimated as

$$V_{AB} = Z_B\Gamma_{AB} \qquad (1.131)$$

The original form of CNDO, due to Pople, Santry and Segal, has now been largely superseeded by a reparametrised form CNDO II[68] and this is the programme which has been used in this work. The main improvement was the inclusion of the above penetration integrals in order to reproduce experimental bond lengths with greater accuracy.

(d) Intermediate Neglect of Differential Overlap. INDO[43] is very similar to CNDO/II and the results obtainable are generally as good or better.

The basic approximations are the same but the one-centre Fock matrix elements are given by

$$F_{ii} = U_{ii} + \sum_{jk}^{A} \left(P_{jk} \langle ii|jk\rangle - P_{jk}^{\alpha} \langle ij|ik\rangle\right) + \sum_{B\neq A} (P_{BB} - Z_B)\Gamma_{AB} \quad i \text{ on atom A} \tag{1.132}$$

Thus in CNDO and INDO methods the Hartree-Fock equations are solved after most of the atomic integrals have been eliminated or parameterised from spectroscopic data and the advantages of this approach are evident upon examination of table 1.2 which shows that number of two electron integrals requiring evaluation in a calculation of propane at various levels of sophistication

Integrals	Hartree-Fock using minimal basis set	NDDO	CNDO
1 - centre	368	173	11
2 - centre	6652	568	55
3-4 centre	31206	0	0
total	38226	741	66

Table 1.2 Number of two-electron integrals requiring evaluation for calculations of the propane molecule.

1.8 Computer Programmes for Ab Initio Calculations

Several packages are now available for performing non empirical calculations on molecular systems. The writing and development of these programmes requires many man-years of labour and hence once written they tend to be generally available through organisations such as the Quantum Chemistry Programme Exchange. The programmes are required to possess certain features in their basic design, which are:

(i) The system should be applicable to general molecular systems i.e. not restricted to particular types such as diatomics or linear triatomics.

(ii) They should be largely machine independent and thus capable of implementation on different types of computer. For this reason the programmes must be written in a high level language, which is invariably FORTRAN. In general this requirement is not fulfilled.

(iii) The system should require minimal input i.e. it should be possible to create prior data files to avoid the necessity of inputting large numbers of cards with the possibility of mechancial error. Furthermore, the input should be in as flexible format as possible, such as the NAMELIST format.

(iv) The system should include restart features at any point so that very long calculations may be run in steps.

(v) The system should be as easy to use as possible, and include check facilities to prevent wastage of machine time.

(vi) The system should be capable of running in a multiprogramming environment, which tends to be the mode of operation in university computing laboratories. This will depend upon the size of the programme. The object code produced by a FORTRAN computer is fast in execution but storage is inefficiently used since the arrays used are static. Thus the arrays are dimensioned to accommodate the largest molecular system and basis set even for a comparatively small system. Attempts to implement dynamic arrays in FORTRAN have produced unwieldy results.

In the past efficient programming techniques and optimum use of computer resources, such as core store, drum and disc storage and tape, have often taken second place to the implementation of a working programme. Most modern programmes are derived from two approaches, which differ in their basic design philosophies: POLYATOM, which was written by Barnett, Harrison[44] et. al. working at M.I.T., is suitable for a university environment, requiring comparatively small core storage and running

time which is segmented. The programme is divided into four sections - integral preparation, integral evaluation, diagonalisation of Fock matrix and calculation of molecular properties, each of which requires about 200k bytes of core and is written without overlay structure. One of the essential differences between POLYATOM and IBMOL is the organisation of the integrals. In POLYATOM, a general integral list is generated which must carry information concerning integral and type lables, whereas in IBMOL all integrals are computed and organised sequentially. IBMOL, which was written by Clementi and co-workers[45] at the IBM Special Research Laboratory, San Jose, is dependent upon the availability of large scale, dedicated computers. Thus the programme utilises core storage to the full, requiring about 500k bytes of core with overlay structure. The programmes updated successors IBMOL IV and IBMOL V are the versions which have been used for the ab initio calculations reported in this thesis. They are written ih the comparatively fast FORTRAN H, though some sub-routines are in the even faster, but lower level, ASSEMBLER, and divide into four parts

(i) Integral Evaluation

(ii) Transformation

(iii) S.C.F. procedure

(iv) Mulliken population analysis[46] and dipole moments. This part is generally run as a separate programme, requiring the final eigen-vectors from the previous section (which are output in card form) as input.

The organisation of the IBMOL IV is shown in flow chart form in fig. 1.3.

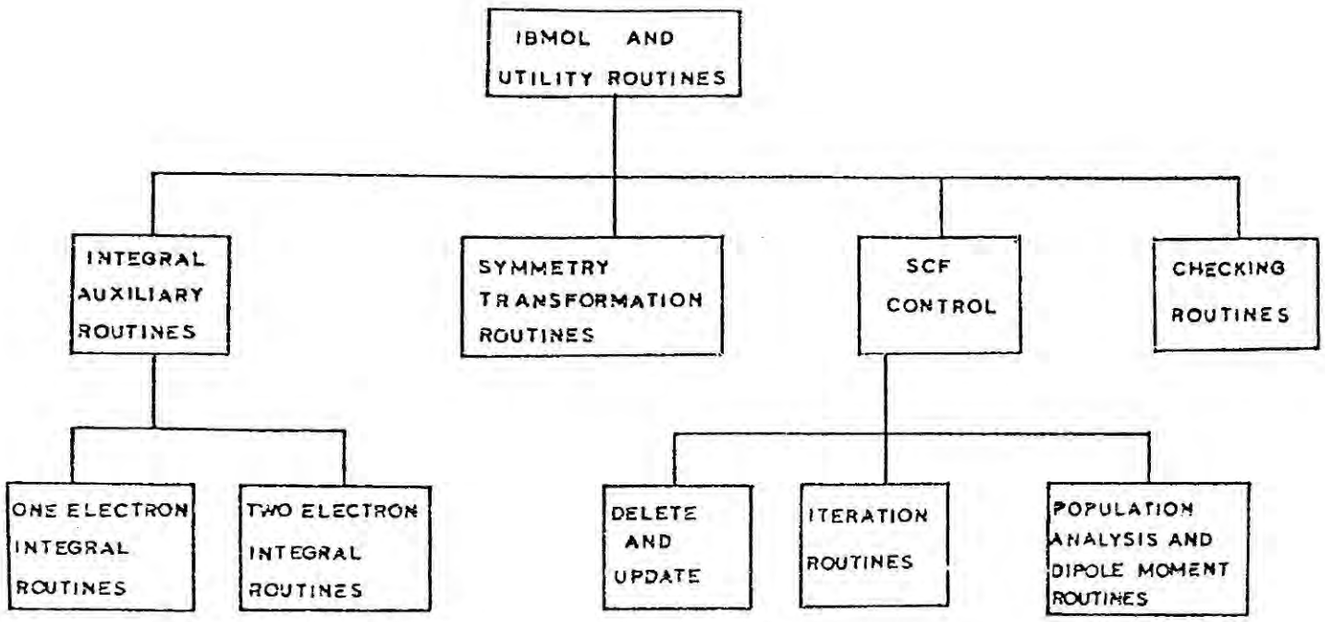


Fig 1.3 Flow chart for IBMOL IV

Coordinates are input in atomic units followed by basis set data. Gaussian functions are used and these may be uncontracted or contracted as required and hence the input is a set of exponents and a set of contraction coefficients. The number of integrals to be evaluated is approximately proportional to the fourth power of the number of basis functions used and hence storage and handling of the integrals of a system with even only a moderate basis set is problematic. In IBMOL V the integrals are organised into a supermatrix form as follows. The one electron integrals are ordered into a N by N matrix (where N is the number of basis functions), and if symmetry adapted functions are used, then many can be set to zero immediately, and the matrix takes a block diagonal form. The two electron

integrals may be divided into the two groups of coulomb and exchange integrals (known as J and K respectively).

If there are two representations Γ_λ and Γ_μ then the coulomb integral may be defined as

$$J_{\lambda_{ij},\mu kl} = \int (\text{charge distribution for electron (1)}) \times (\text{charge distribution electron (2)}) \times \frac{1}{r_{12}} \partial\tau_1 \partial\tau_2 \qquad (1.133)$$

$$= \int \left(\psi_{\lambda i}^{(1)} \psi_{\lambda j}^{(1)}\right) \times \left(\psi_{\mu k}^{(2)} \psi_{\mu l}^{(2)}\right) \times \frac{1}{r_{12}} \partial\tau_1 \partial\tau_2 \qquad (1.134)$$

and for each J integral, a K integral is defined by transposition of indecies

$$K_{\lambda_{ij},\mu kl} = \tfrac{1}{2} \left(\int \psi_{\lambda i}^{(1)} \psi_{\mu k}^{(1)} \psi_{\lambda j}^{(2)} \psi_{\mu l}^{(2)} \frac{1}{r_{12}} \partial\tau_1 \partial\tau_2 \right.$$

$$\left. + \int \psi_{\lambda i}^{(1)} \psi_{\mu l}^{(1)} \psi_{\mu k}^{(2)} \psi_{\lambda j}^{(2)} \frac{1}{r_{12}} \partial\tau_1 \partial\tau_2 \right) \qquad (1.135)$$

Thus $(\sum_\lambda n_\lambda (n_\lambda + 1)/2)$ charge distributions are calculated and organised into a vector of this length, called a supervector. This is done for both electron (1) and electron (2) and then the J (or the K) integrals are organised into a matrix of dimension $(\sum_\lambda n_\lambda (n_\lambda + 1)/2) \times (\sum_\lambda n_\lambda (n_\lambda + 1)/2)$ called the J (or K) supermatrix. A copy of the integrals is kept on magnetic tape as it may be decided, at a later date, to perform another calculation on the same system with some of the centres having changed coordinates, this being facilitated by the MOVE routine. With IBMOL IV it was also possible to add new atoms but this is not possible using IBMOL V owing to the supermatrix organisation.

Using IBMOL V, three types of basis function contraction are possible. These are

(i) Functions of the same type, upon the same centre. This is the contraction, usual with gaussians, discussed previously in section 1.4.

(ii) Functions of different type, but on the same centre. This is used for construction of hydridised orbitals.

(iii) Functions of mixed type on mixed centres. This is used for the generation of symmetry adapted functions.

For the SCF part a set of starting 'trial' vectors is required. After three iterations of the Fock matrix diagonalisation process, an efficient extrapolation routine is called up and hence the starting vectors are not required to be of great accuracy. Generally these are taken from a previous semi-empirical calculation. The point at which convergence (CNDO/II) of the energy components is judged to have been reached is set by the value specified on the input. When the Fock matrix is thus judged to be diagonalised the programme terminates and the output file is ready to be printed. The relative timing within the SCF section may be divided

~ 30% Fock matrix diagonalisation

~ 70% Assembly of Fock matrix

and hence it is clear that the integral files must be stored on a fast random access facility.

The output format has changed somewhat with the introduction of IBMOL V but both are basically similar. The output takes the form

(i) Printout of Input data

(ii) Printout of one electron integrals

(iii) Information about SCF part - no. of iterations etc.

(iv) Final values of total and electronic energies

(v) Final set of eigenvalues and eigenvectors

It is also usual to use the facility in both programmes to have a card output of the final set of vectors, which may then be used as part of the input into a programme to perform a Mulliken[46] population analysis.

IBMOL IV was implemented on the Northumbrian Universities Multiple Access Computer IBM 360/67 which had initially 512K bytes of core, backed up by a drum store, eight disc drives and two magnetic tape drives. Input was via a card reader which could also punch cards for output if

required (a useful facility for punching a set of final vectors for use as input to the Mulliken population analysis programme), and output was via a line printer. During the time when IBMOL IV was being used for the calculations described in this thesis the computer had the above mentioned core size of 512k bytes which was insufficient core space to implement the programme, and hence it was necessary to generate a virtual machine by means of continuous transfer of information between core and drum store. This resulted in the availability of close on a million bytes of virtual core, but at a reduced efficiency of ~80%. IBMOL IV was an interactive programme in which safety copies of integrals could be organised from the console.

IBMOL V was implemented on the IBM 360/195 at the Atlas Computing Laboratory.

II ELECTRON SPECTROSCOPY FOR CHEMICAL APPLICATIONS (ESCA)

1.9 ESCA - An Introduction

The technique of Electron Spectroscopy for Chemical Applications (ESCA)[47] is a comparatively new and very powerful molecular probe, which has a wide range of application in many fields of chemistry. The approach divides into two main areas - low energy electron spectroscopy, (Ultraviolet Photoelectron Spectroscopy[48] or UPS), in which only valence orbitals are investigated, and high energy electron spectroscopy, (E.S.C.A.)[47], in which the inner shell or core electrons are the main concern. Since the work reported in this thesis has been performed with the use of the latter technique only, the former will not be considered further. E.S.C.A. involves the measurement of the kinetic energies of electrons expelled from the core levels of atoms or molecules by photoionisation due to an incident beam of monochromatic X-radiation of accurately known energy.

The electron binding energy is obtained as the difference between the quantum energy of the exciting radiation $h\nu$, and the measured kinetic energy, E_{kin}

$$E_B = h\nu - E_{kin} \qquad (1.136)$$

In core electron spectroscopy the atoms in a molecule retain their atomic individuality and show characteristic binding energy ranges by which they are readily identified. The overall intensities of peaks arising from photoionisation of core electrons may be used for quantitative analysis.

Traditionally core electrons are regarded as being unaffected by the chemical environment in which the atom is situated, however these levels are subject to chemical shifts,[49], analysis of which can yield valuable information. As a spectroscopic tool ESCA has several advantages.

The sample may be a solid, liquid or gas and the technique is

essentially non-destructive.

Sample requirement is modest, in favourable cases 1mgm of a solid, 0.1μl of liquid or 0.5cc of a gas (STP).

The technique has high sensitivity, is independent of the spin properties of any nucleus and is applicable in principle to any element of the periodic table with the exceptions of hydrogen and helium where the core levels are also the valence levels.

The information obtained is directly related to the electronic structure of a molecule and the theoretical interpretation is relatively straight forward.

Information can be obtained on both the core and valence energy levels of molecules.

1.10 Basic Processes involved in ESCA

In ESCA samples are irradiated with X-rays of known energy, typically $MgK_{\alpha_{1,2}}$ (1253.7eV) or $AlK_{\alpha_{1,2}}$ (1486.6eV) when electrons having a binding energy less than the energy of the exciting radiation may be ejected, allowing the study of both core and valence electrons. Thus in the emission of a 1s electron from a gaseous sample, the kinetic energy of the photoelectron will be given by

$$KE = h\nu - BE - E_r \qquad (1.137)$$

where E_r is the recoil energy of the sample. This last term is usually negligible except where high energy X-rays are used with light atoms. For the species studied in this work, energy differences due to recoil energy may be ignored.

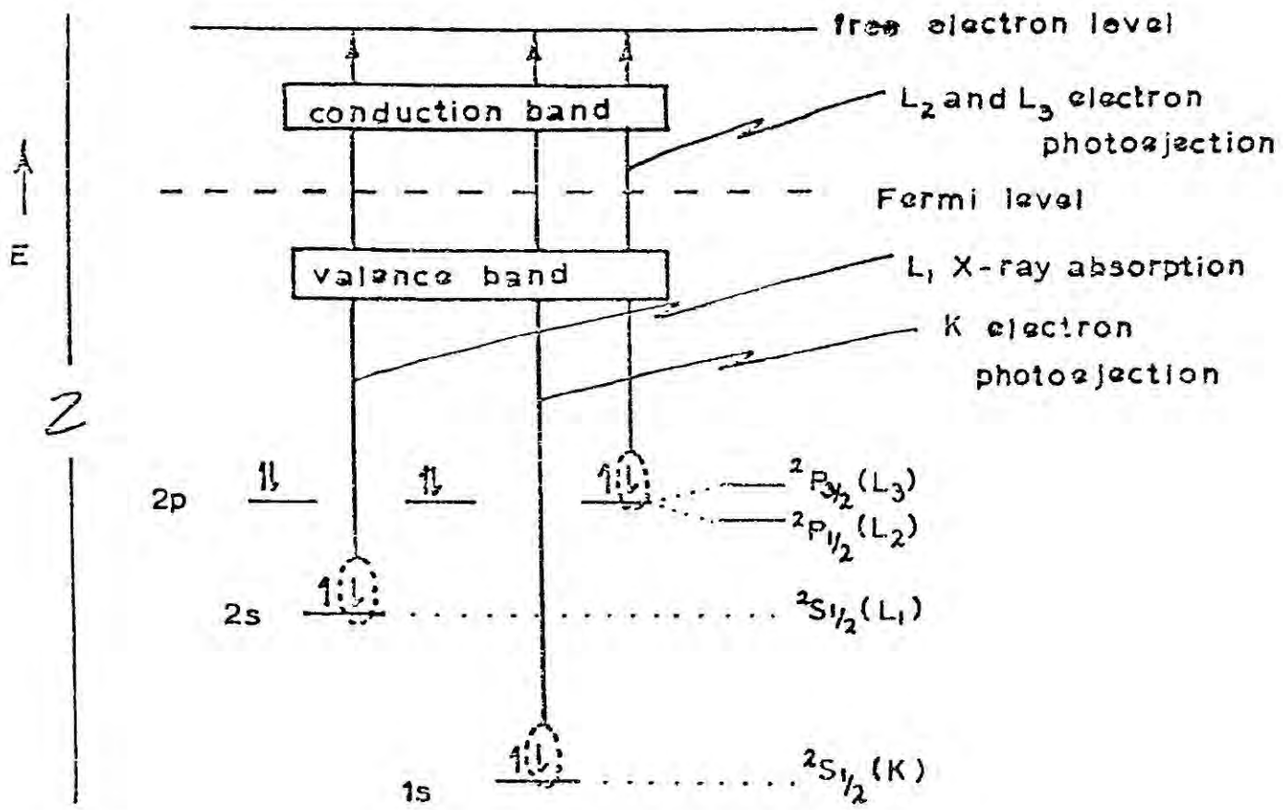


Fig. 1.4 The photoionisation process

Electronic relaxation will follow the process of photoionisation, fig. 1.4 and this may occur by means of one of two processes, shown diagramatically in Fig. 1.5

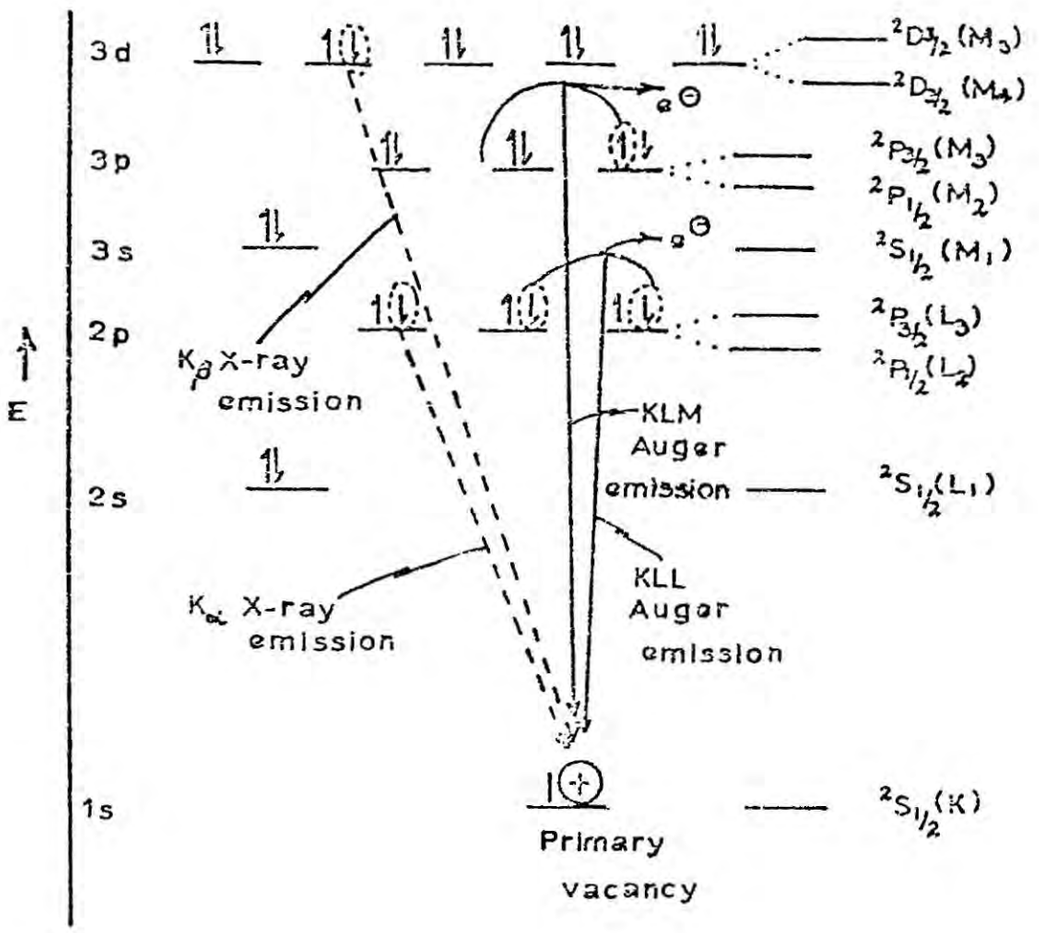


Fig. 1.5 Diagram of relaxation processes following photoionisation.

These processes are

(i) X-ray emission[50,51]

$$(A^+)^* \longrightarrow A^+ + h\nu_2$$

(ii) Auger process[52 - 55]

$$(A^+)^* \longrightarrow A^{++} + e_2^-$$

The relative proportion of these which occur is dependent upon the atomic number of the element, and both are studied as spectroscopic tools. In process (i), the energies of the emitted X-rays give information on the differences in energy levels in the sample (subject to the selection

rules $\Delta l = \pm 1$, $\Delta j = \pm 1, 0$). A similar technique is X-ray absorption spectroscopy[47] where the X-ray beam is passed through the sample and the intensity of the transmitted radiation measured as a function of wavelength. Absorption edges are produced corresponding to the energy required for the excitation of an electron from one of the inner shells to the lowest unoccupied level. In process (ii), Auger electron emission, an electron from an outer shell once again fills the lower vacancy but the energy change in the process results not in photo-emission but in the expulsion of another electron to leave a doubly ionised state. This process is more probable for elements of lower atomic number than is X-ray emission but may be readily distinguished from the primary photoelectrons since their kinetic energy will be independent of the energy of the exciting radiation.

The initial ionisation of a core electron will result in the sudden perturbation of the valence cloud, which may lead to the simultaneous excitation ("shake-up") or emission ("shake-off") of an outer electron,[56,57,58] from which arise satellite peaks corresponding to electrons of lower kinetic energy than the main photoionisation peak

$$KE = h\nu - BE - \bar{E} \qquad (1.138)$$

where $\bar{E}$ is the energy of the shake-up or shake-off process.

The natural definition of the binding energy of an electron in a free atom or molecule is the energy required to remove the electron from a given level to infinity (i.e. the vacuum level). For a solid sample, however, the situation is not so straightforward, since the outer electronic levels are broadened into bands and a potential barrier exists at the surface. Fig. 1.6 shows that the measurement of a binding energy of an electron in a solid sample will be given by

$$BE = h\nu - \phi_{sp} - T_{sp} - E_r \qquad (1.139)$$

where ϕ_{sp} is the spectrometer work function and T_{sp} is the kinetic energy

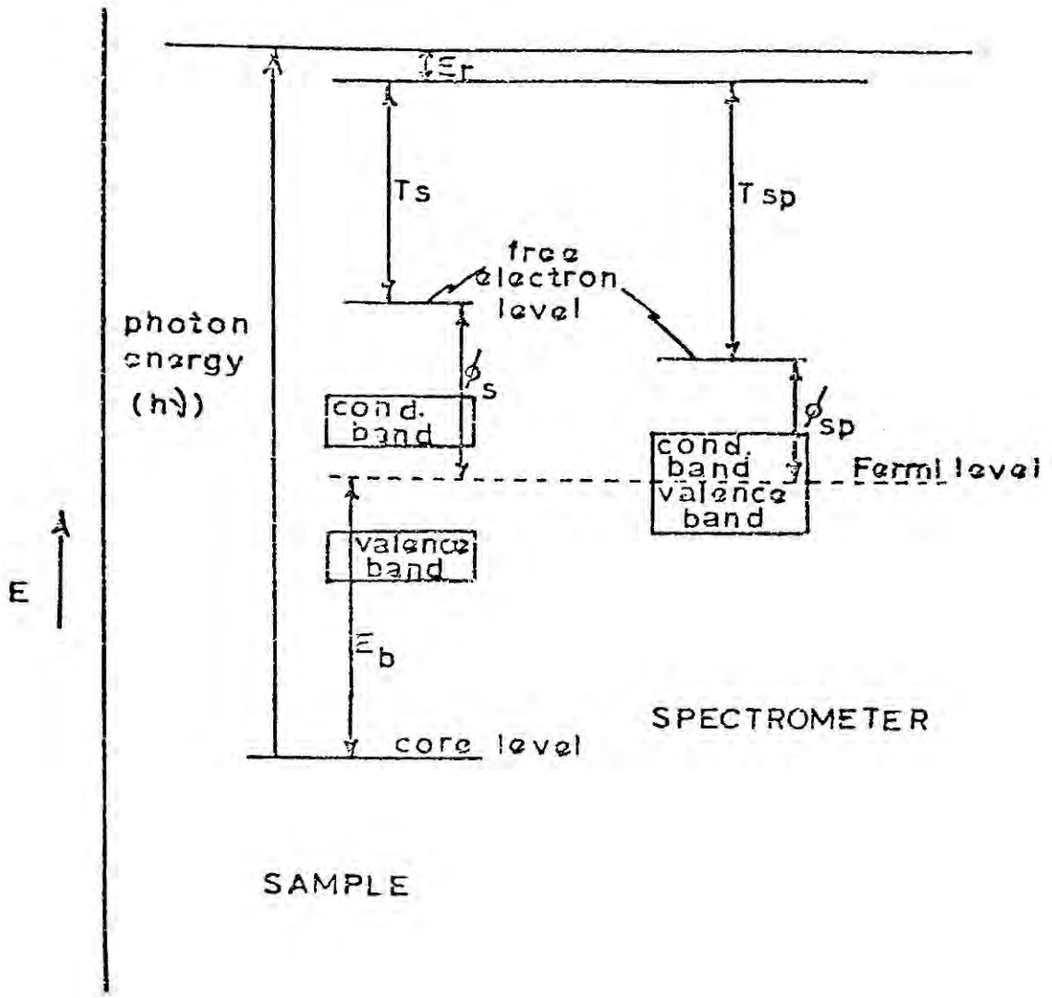


Fig. 1.6 Measurement of core binding energy in a solid sample

of the photoelectron in the viscinity of the spectrometer. The most convenient reference level for photoelectrons arising from solid samples is thus the Fermi level, E_f, defined for metals by

$$\int_0^{E_f} N(E)dE = N \text{ where } N(E) = Z(E)F(E) \qquad (1.140)$$

where Z(E) is the density of states for electrons (i.e. the number of energy levels between E and E + ΔE), F(E) is the Fermi probability distributions, the probability that a Fermi particle in a system in thermal equilibrium at temperature T will be in a state with energy E.

$$F(E) = (e^{(E-E_f)/kT} + 1)^{-1} \quad (kT \ll E_f) \qquad (1.141)$$

N is the total number of electrons in the system and the electrons fill the available states up to the Fermi level. Thus in fig. 1.6 photoionisation is occuring from a core level in a solid sample which is in electrical contact with the spectrometer. Thus the Fermi levels of sample and spectrometer are the same and any differences between the work functions of the sample and spectrometer gives a difference in macropotential[59,60] and an electric field arises in the space between the sample and the spectrometer. Hence the kinetic energy of the photoelectron when it enters the analyser is slightly different from the energy it possessed on emerging from the sample. Thus the former kinetic energy is measured and taking zero binding energy to be at the Fermi level and ignoring recoil energy gives

$$BE = h\nu - \phi_{sp} - T_{sp} \qquad (1.142)$$

Thus using the Fermi level as the reference means that the binding energy will be independent of the sample work function but dependent upon the spectrometer work function, a constant correction. In order for the Fermi levels to adjust to thermodynamic equilibrium a sufficient number of free charge carriers must be present. In the case of insulating samples a sufficient number of free charge carriers is formed during X-radiation[47] (this includes all the species studied in this thesis), however a surface build up of charge may occur, thus shifting all the energies of the emitted electrons significantly (typically one or two eV for the systems studied in this work). For core ionisations from atom A in a molecule AX the binding energies in solid (e) and vapour (b) may be related by

$$AX(s) \longrightarrow AX(g) \qquad (a)$$

$$AX(g) \longrightarrow {}^{*}AX^{+}(g) + e^{-}(v) \qquad (b)$$

$${}^{*}AX^{+}(g) \longrightarrow {}^{*}AX^{+} \text{ (in solid AX)} \qquad (c)$$

$$e^{-}(v) \longrightarrow e^{-} \text{ (Fermi)} \qquad (d)$$

$$AX(s) \longrightarrow {}^{*}AX^{+} \text{ (solid AX)} + e^{-} \text{ (Fermi)} \qquad (e)$$

where * represents a core level vacancy. The binding energy difference between solid and vapour therefore will depend on the sample work function (d) and also the energy required to move a molecule of AX from the solid (a) and the energy of placing the core ionised species $*AX^+$ back into the solid. However, for molecules in the absence of strong interactions in the solid phase the shifts in binding energy are similar to those in gases, but the actual binding energies are higher in gases due to the difference in reference level. This is shown diagramatically in fig. 1.7.

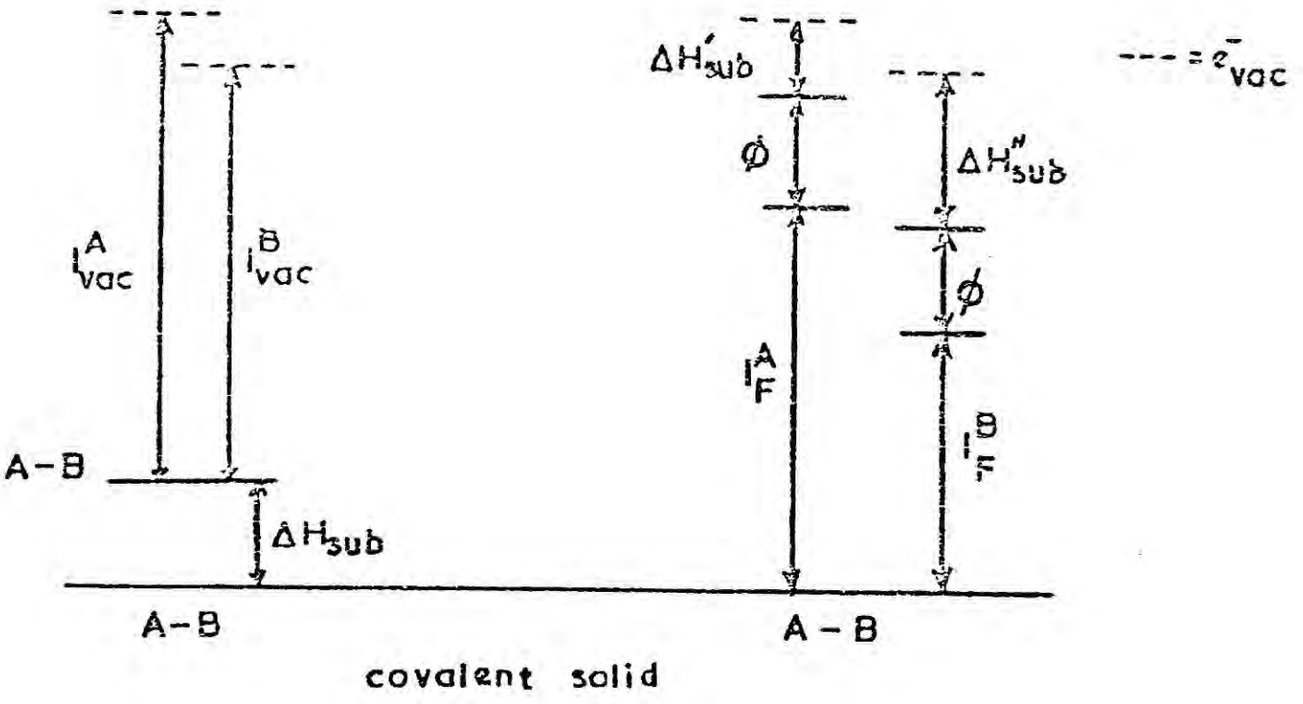


Fig. 1.7 Relationship between binding energies measured in the solid and gas phases.

The shift in binding energy will be given by the difference in ionisation potentials measured with reference to the vacuum level

$$\Delta = (I_A - I_B)_{vac}$$
$$= (I_A - I_B)_F + (\Delta H'_{sub} - \Delta H''_{sub}) \qquad (1.143)$$

For different samples the shift will be given by

$$\Delta = (I_A - I_X)_{vac}$$
$$= (I_A - I_X)_F + (\phi_A - \phi_X) + (\delta_A - \delta_X)$$
$$+ (\Delta H_X - \Delta H_A) + (\Delta H'_A - \Delta H'_X) \qquad (1.144)$$

However, for closely related materials in contact with the spectrometer

$$(\phi_A - \phi_X) \simeq 0 \qquad (1.145)$$

$$(\delta_A - \delta_X) \simeq 0 \qquad (1.146)$$

$$(\Delta H_X - \Delta H_A) + (\Delta H'_A - \Delta H'_X) \simeq 0 \qquad (1.147)$$

$$\text{i.e. } (I_A - I_X)_{vac} \equiv (I_A - I_X)_F \qquad (1.148)$$

1.11 Theoretical Interpretation of Core Binding Energies

Theoretical treatments are feasible only for isolated molecules and in the last section the relationship between solid phase and gas phase measurements was discussed. The absolute binding energies are not of relative importance but it is the relative binding energies for which theoretical interpretation is required.

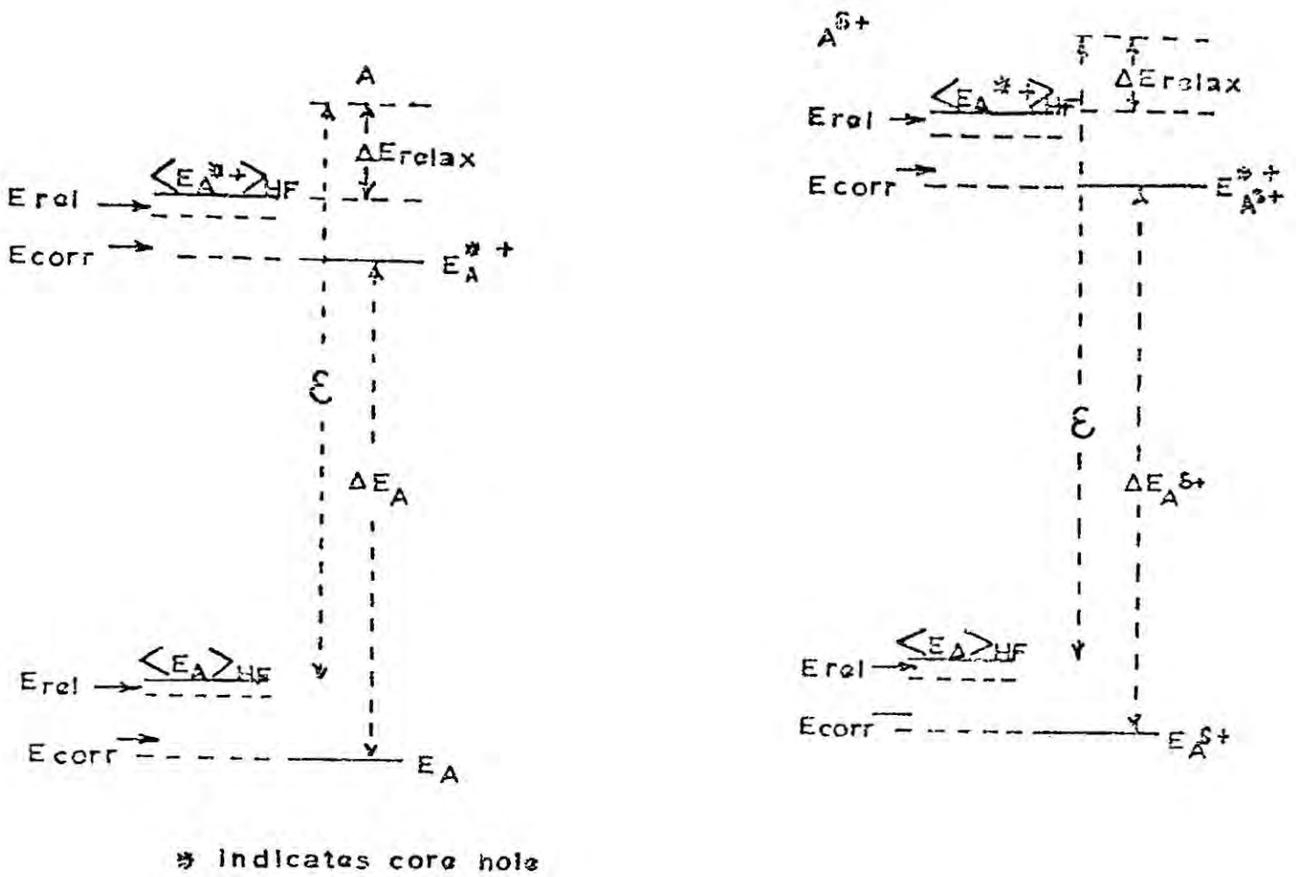


Fig. 1.8 Relationship between experimental, Koopmans and hole state binding energies.

Fig. 1.8 shows the relationship between experimental, Koopmans' theorem and hole state binding energies. Thus the energies calculated for the ground state will require relativistic and correlation energy corrections to yield the absolute energies, and similarly this will apply also to the ionised state, with the added complication that the energy of the ionised state will be calculated too high since an application of Koopmans' theorem does not allow for the relaxation of an ionised state due to electronic reorganisation. Richards[61] has made an examination

of the approximations inherent in the interpretation of photoelectron spectra by means of molecular orbital calculation and the application of Koopmans' Theorem.

From fig. 1.8 it may be seen that

$$I^{A}_{pot}{}^{1} = (E_A - E_A^{*+}) = \Delta E_A \tag{1.149}$$

$$= (\langle E_A \rangle_{HF} - \langle E_A{}^{*+} \rangle_{HF}) + (E^{A}_{corr} - E^{A*+}_{corr}) + (E^{A}_{rel} - E^{A*+}_{rel}) \tag{1.150}$$

$$\simeq (\varepsilon^{A}_{HF} - \Delta E_{relax}) + (E^{A}_{corr} - E^{A*+}_{corr}) + (E^{A}_{rel} - E^{A*+}_{rel}) \tag{1.151}$$

$$I^{A\delta+}_{pot}1 = (E_A\delta+ - E_A^{*+}\delta+) = \Delta E_A\delta+ \tag{1.152}$$

Shift in Binding Energy

$$\Delta = I^{A\delta+}_{pot} - I^{A}_{pot} = \Delta E_A\delta+ - \Delta E_A \tag{1.153}$$

$$= \Delta(\Delta E^{A\delta+}_{HF} - \Delta E^{A}_{HF}) + \Delta(\Delta E^{A\delta+}_{corr} - \Delta E^{A}_{corr}) + \Delta(\Delta E^{A\delta+}_{rel} - \Delta E^{A}_{rel}) \tag{1.154}$$

$$= \Delta(\varepsilon^{A\delta+} - \varepsilon^{A})_{HF} + \Delta(\Delta E^{A\delta+}_{relax} - \Delta E^{A}_{relax})$$

$$+ \Delta(\Delta E^{A\delta+}_{corr} - \Delta E^{A}_{corr}) + \Delta(\Delta E^{A\delta+}_{rel} - \Delta E^{A}_{rel}) \tag{1.155}$$

But for core orbitals one can approximate

$$\Delta(\Delta(\Delta E_{corr})) \simeq 0 \tag{1.156}$$

$$\Delta(\Delta(\Delta E_{rel})) \simeq 0 \tag{1.157}$$

$$\Delta(\Delta(\Delta E_{relax})) \simeq 0 \tag{1.158}$$

and hence the shift in the binding energy of the core orbitals may be taken as

$$\Delta = \Delta(\varepsilon^{A\delta+} - \varepsilon^{A})_{HF} \tag{1.159}$$

This result is of crucial importance for the interpretation of core binding energy measurements by means of quantum mechanical calculation, and table 1.3 shows some experimental and calculated binding energies to illustrate this. Whilst the absolute magnitudes of the energies are

Table 1.3 C1s Shifts relative to Methane (eV)[66]

	ΔE_{exp}[56,67]	$\Delta E_{Koopmans}$
C_2H_6	-0.2	0.2
C_2H_4	-0.1	0.9
C_2H_2	0.4	1.4
C_2H_4O	2.0	2.4
CH_3OH	1.9	2.0
HCO_2H	5.0	6.0
CO_2	6.8	8.3
CO	5.4	5.5
Cyclopropane	-0.2	0.4

Table 1.4 Pair Correlation energies for first row atoms.

	au	eV
1s - 1s	-0.0409	-1.113
1s - 2s)) 1s - 2p)	-0.0027	-0.073
bond - bond	-0.0409	-1.11
bond - bond[1]	-0.0118	-0.321
lonepair	-0.0257	-0.699
lonepair - lonepair[1]	-0.0358	-0.974
lonepair - bond	-0.0301	-0.819

not in good agreement, the shifts relative to methane are described quite well by calculation.

In table 1.4 is shown some values for pair correlation energies in first row atoms, which can be seen to be far from negligible. An interesting series results if the different types of correlation interaction are placed in descending order. The series shows the same order as that described by Gillespie and Nyholm[62] to discuss the geometry of molecules. Examination of the table shows that the pair correlation energies for core levels are intra-orbital whereas for the valence shell correlation energies between different orbitals can be substantial. Thus it may be concluded that overall correlation energy corrections for core orbitals are of less importance than for valence orbitals, where the change of correlation energy on ionisation depends markedly on the localisation characteristic of the orbital.

A detailed study of the ionised states of the CH_4 molecule with a basis set approaching the Hartree Fock limit has been made by Clementi and Popkie[36]. These have demonstrated that for the 1s hole state the correlation energy is exactly the same as for the neutral molecule. In this example, then, correlation energy makes no contribution to the binding energy of the 1s electron which is given within experimental error as the energy difference at the Hartree Fock limit. However, the difficulty and expense of such calculations provides impetus for the development of computationally simpler models.

Table 1.5 shows the radial maxima of the orbitals of the neon atom as a function of ionised state. The first removal leaves the radial maximum for the other electron almost unchanged whereas the valence electrons relax to accommodate the effect of the rising effective nuclear charge. Hence, relaxation energy neglected in Koopmans'Theorem is associated essentially with valence electrons. To describe adequately shifts by non-empirical treatments it is necessary, therefore, to have

Table 1.5 Radial maxima of the orbitals of the Neon atom as a function of ionised state[68].

	Ne atom	Ne^+ 2p hole	Ne^+ 2s hole	Ne^+ 1s hole
$\langle r \rangle 1s$	0.1576	0.1576	0.1578	0.1545
$\langle r \rangle 2s$	0.8921	0.8603	0.8536	0.8171
$\langle r \rangle 2p$	0.9652	0.8759	0.8841	0.7993

an adequate description of the valence shell. This requires a physically balanced basis set. This is not necessarily a well conditioned function of the absolute size of the basis set, for example, fig. 1.9 shows schematically how shifts in core binding energies derived from Koopmans' theorem might depend upon the basis set used. Thus it is

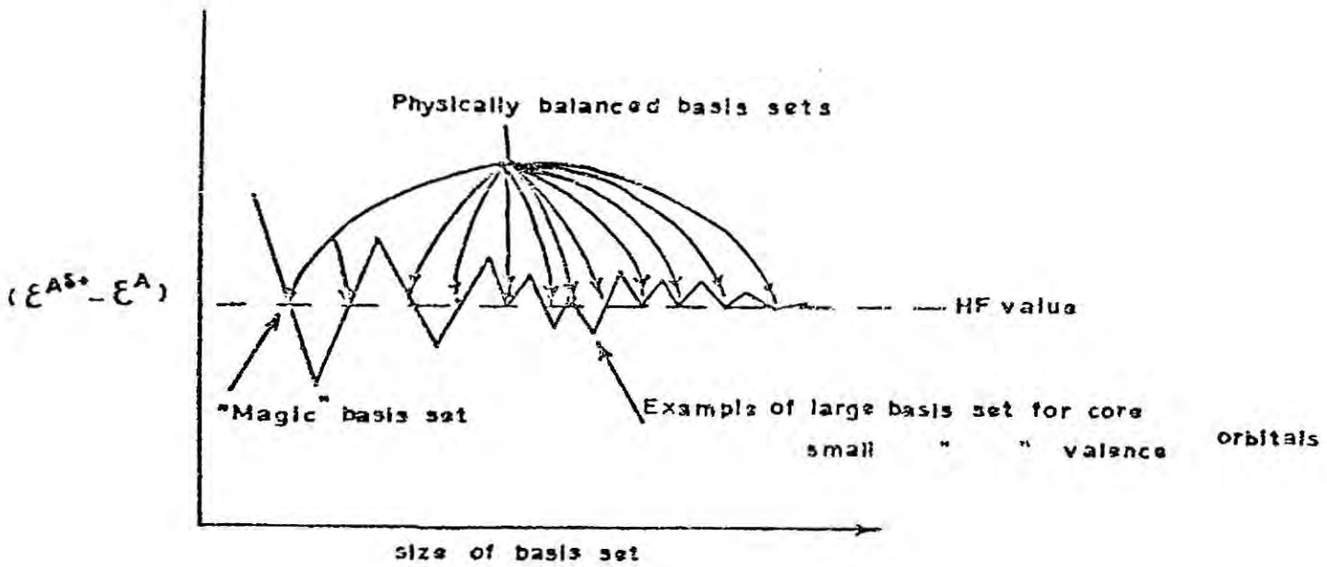


Fig 1.9 Dependence of calculated chemical shift upon the nature of the basis set used.

possible to have a relatively small basis set with an inadequate description of cores (and therefore a poor total energy) but a reasonable description of the valence orbitals which gives a good approximation to the shifts at the Hartree Fock limit. By contrast, it is also possible to have a

large basis set with very good description of the cores (and therefore a good total energy) with an adequate description of the valence orbitals.

The various theoretical models available for the interpretation of shifts in molecular core binding energies will now be discussed in further detail.

(i) Ab Initio calculation upon molecule and ion, i.e. calculation of Hole States. The binding energy of a core electron can be expressed as

$$E_{core} = E_i - E_f \tag{1.160}$$

where E_i and E_f are the total energies of the initial and final ionised state respec. Both may be expressed as

$$E_i = E_i^{HF} + E_i^{corr} + E_i^{rel} \tag{1.161}$$

$$E_f = E_f^{HF} + E_f^{corr} + E_f^{rel} \tag{1.162}$$

where E_i^{HF} - energy of initial state at HF limit

E_i^{corr} - contribution due to electron correlation

E_i^{rel} - contribution due to relativistic effects

$$\text{hence Ecore} = (E_i - E_f)^{HF} + (E_i - E_f)^{corr} + (E_i - E_f)^{rel} \tag{1.163}$$

For removal of a 1s electron from a carbon atom (^3P), approximate magnitudes are

$$(E_i - E_t)^{HF} \sim 285 \text{ eV} \tag{1.164}$$

$$(E_i - E_t)^{corr} \sim 0.8 \text{ eV} \tag{1.165}$$

$$(E_i - E_t)^{rel} \sim 0.4 \text{ eV} \tag{1.166}$$

and clearly the correlation and relativistic corrections are small (for small z). The shift is expressed by

$$\Delta\text{shift} = ({}^1E - {}^2E)$$

$$= ({}^1E_i - {}^1E_f - {}^2E_i + {}^2E_f)^{HF} + ({}^1E_i - {}^1E_f - {}^2E_i + {}^2E_f)^{corr}$$

$$+ ({}^1E_i - {}^1E_f - {}^2E_i + {}^2E_f)^{rel} \tag{1.167}$$

The second and third terms depend upon differences between small quantities and anyway the correlation and relativistic terms are not greatly affected by changes in electronic environment, hence the shifts are dominated by the first term. Calculations within these approximations can usually predict shifts to within a few tenths of an eV.

(ii) Ab Initio calculations on a molecule and application of Koopmans theorem[13]. The necessity for the development of computationally less expensive models has already been stressed, and, to some extent, application of Koopmans Theorem meets this requirement. For a closed shell molecule described by

$$\Psi_M = |\psi_a \ldots\ldots \psi_k| \tag{1.168}$$

the total energy will be given by

$$E_M = \sum_{r=a}^{k} \varepsilon_r - \sum_{\text{pairs } rs} (2J_{rs} - K_{rs}) + V_{nn} \tag{1.169}$$

and the orbital energies may be expressed as

$$\varepsilon r = \left\langle \psi_r \left| -\tfrac{1}{2}\nabla_1^2 - \sum_n \frac{Z_n}{r_{n1}} \right| \psi_r \right\rangle + \sum_{s=a}^{k} (2J_{rs} - K_{rs}) \tag{1.170}$$

$$= H_{rr}^{core} + \sum_{s=a}^{k} (2J_{rs} - K_{rs}) \tag{1.171}$$

hence

$$\sum_{r=a}^{k} \varepsilon_r = \sum_{r=a}^{k} H_{rr}^{core} + \sum_{r=a}^{k}\sum_{s=a}^{k} (2J_{rs} - K_{rs}) \tag{1.172}$$

$$\text{now, } \sum_{r=a}^{k}\sum_{s=a}^{k} (2J_{rs} - K_{rs}) = 2 \sum_{\text{pairs } rs} (2J_{rs} - K_{rs}) \tag{1.173}$$

$$\text{thus, } \sum_{r=a}^{k} \varepsilon_r = \sum_{r=a}^{k} H_{rr}^{core} + 2 \sum_{\text{pairs } rs} (2J_{rs} - K_{rs}) \tag{1.174}$$

$$E_M = \sum_{r=a}^{k} H_{rr}^{core} + \sum_{\text{pairs } rs} (2J_{rs} - K_{rs}) + V_{nn} \tag{1.175}$$

If an electron is removed from one spin orbital ψ_a, then the energy

of the ionised state (assuming the wavefunctions of the other electrons remain unchanged) will be given by

$$\Psi_{M^+} = |\psi_b \ldots\ldots \psi_k| \qquad (1.176)$$

$$E_{M^+} = \sum_{r=1}^{k} H_{rr}^{core} + \sum_{\substack{\text{pairs } rs \\ r\neq a \; s\neq a}} (J_{rs} - K_{rs}) + V_{nn} \qquad (1.177)$$

$$\text{hence } E_M - E_{M^+} = H_{aa}^{core} + \sum_{s=a}^{k} (2J_{as} - K_{as}) \qquad (1.178)$$

$$= \varepsilon_a$$

i.e. the ionisation potential may be equated to the orbital energy. In practise, of course, in going to the ionised state there is a change in the wave function, and this relaxation, or reorganisation, produces a lowering of energies. Thus use of Koopmans' theorem predicts values for ionisation potentials that are too large. However, as discussed earlier, for core electrons the degree of localisation is largely independent of the atomic chemical environment, making the reorganisation energy approximately constant for a given core level. Thus

$$\Delta_{shift} = \Delta\varepsilon + \Delta E_{reorg} + \Delta E_{corr} + \Delta E_{rel} \qquad (1.179)$$

and by an approximate cancellation of errors

$$\Delta_{shift} \simeq \Delta\varepsilon \qquad (1.180)$$

Thus for limited series of closely related molecules, Koopmans' Theorem gives an adequate interpretation of chemical shifts of core levels but the method will break down when relaxation energies do not remain fairly constant between members of the series in question.

(iii) Charge Potential model. The charge potential model relates core electron binding energies on the atom from which core ionisation takes place and the potential from the charges in the remainder of the molecule. Since the core electrons are localised upon their respective atoms then the change in electronic environment in going from one molecule to another is due to changes in the distribution of the valence electrons. This distribution can be regarded as providing a potential against which work

must be done to pull a core electron from an atom.

Initially this approach was taken to the limit, by saying that the lower the electronic density on the atom being considered, the more difficult would be the removal of a core electron. Whilst this is a reasonable approach in a specific series, generally to consider that the potential arises only from the local valence electron distribution is too gross an approximation. A severe limitation arises from the description of charge distribution in molecules where an arbitrary division overlap densities is made and absolute magnitudes depend on the nature of the basis set. A better approach is to treat the valence electron distribution as continuous and hence define a potential at the atom which may be computed directly from the wave function. Schwartz [63] has shown this to be

$$\Delta E_{\text{core be shift}} = \Delta\Phi_{\text{valence}}$$

$$\Phi_{\text{valence}} = 2 \sum_{i \neq \text{core}} \left\langle \phi_i \left| \frac{1}{r_n} \right| \phi_i \right\rangle + \sum_{m \neq n} \frac{Z_m^*}{R_{mn}} \qquad (1.182$$

Z_m^* is effective reduced nuclear charge

This is probably the most rigorous approach but that of Manne [56] is of greater utility for chemists. In this latter approach the valence electron distribution is treated as a set of point charges centred on each nucleus. Considering the expression for orbital energy

$$\varepsilon_r = \left\langle \psi_r \left| -\tfrac{1}{2}\nabla_i^2 - \sum_n \frac{Z_n}{r_{nl}} \right| \psi_r \right\rangle + \sum_{s=a}^{k} (2J_{rs} - K_{rs}) \qquad (1.183)$$

$$= \left\langle \psi_r \left| -\tfrac{1}{2}\nabla_i^2 - \frac{Z_m}{r_{ml}} \right| \psi_r \right\rangle + J_{rr} + \left\langle \psi_r \left| \sum_{n \neq m} \frac{Z_n}{r_{nl}} \right| \psi_r \right\rangle$$

$$+ \sum_{s \neq r}^{k} (2J_{rs} - K_{rs}) \qquad (1.184)$$

If ε_r corresponds to a core level on atom m then the first term in (1.184) is a constant for the atom, independent of valence distribution. The second term also is independent of valence distribution. The second term also is independent of valence distribution. Furthermore, in the expansion

$$\sum_{s=r}^{k} (2J_{rs} - K_{rs})$$

the one centre coulomb and exchange integrals involving other core orbitals on atom m will be invarient to valence distribution, and exchange integrals involving other core orbitals on atom m will be invarient to valence distribution, and exchange integrals between the core orbital on m and core and valence orbitals on other atoms are sufficiently small to be ignored. Thus (1.184) may be written

$$\varepsilon_r = E_o + \left\langle \psi_r \left| -\sum_{n=m} \frac{Z_n}{r_{n1}} \right| \psi_r \right\rangle + \sum_{s=r}^{k} 2J_{rs} \qquad (1.185)$$

Where E_o is a constant, and the summation of coulomb integrals extends over all valence orbitals on other atoms. The second term represents the attraction between a core electron in ψ_r which is localised on m, and the nuclei of the other atoms of the molecule. Thus the term can be rewritten

$$\sum_{n \neq m} - \frac{Z_n}{r_{nm}}$$

To evaluate the third term, first consider the terms involving the core orbital ψ_j on atom n

$$J_{rj} = \left\langle \psi_r(1)\psi_j(2) \left| \frac{1}{r_{12}} \right| \psi_r(1)\psi_j(2) \right\rangle \qquad (1.186)$$

This represents the repulsion between ψ_r^2 on m and ψ_j^2 on m and can be rewritten $\frac{1}{r_{nm}}$.

When ψ_i is a valence orbital J_{ri} is written

$$J_{ri} = \left\langle \psi_r(1)\psi_i(2) \left| \frac{1}{r_{12}} \right| \psi_r(1)\psi_i(2) \right\rangle \qquad (1.187)$$

and expanding ψ_i in terms of the basis set

$$J_{ri} = \sum_p c^2_{ip} \, \psi_r(1)\phi_p(2) \left| \frac{1}{r_{12}} \right| \psi_r(1)\phi_p(2) \qquad (1.188)$$

this gives rise to two types of integral, depending upon whether or not ϕ_p is centred upon atom m or not. Thus the integrals involving charge clouds centred on different atoms can be separated out

$$J_{ri} = \sum_{\substack{\text{sum over } \phi_p \\ \text{on atoms } m \neq n}} c^2_{ip} \left\langle \psi_r(1)\phi_p(2) \left| \frac{1}{r_{12}} \right| \psi_r(1)\phi_p(2) \right\rangle$$

$$+ \sum_{\substack{\text{sum over } \phi_p \\ \text{on atom m}}} c^2_{ip} \frac{1}{r_{nm}} \qquad (1.189)$$

These can be put back in (1.185) to give

$$\varepsilon_r = E_o + \sum_{\substack{\text{all atoms} \\ n(\neq m)}} -\frac{Z_n}{r_{nm}} + \sum_{\substack{\text{all core} \\ \text{orbitals on} \\ \text{atoms } n(\neq m)}} \frac{2}{r_{nm}} + \sum \sum_{\substack{\text{all valence MO's} \\ \text{\& all constituent AO's} \\ \text{in the LCAO's} \\ \text{on atoms } n(\neq m)}} \frac{2c_{ip}^2}{r_{mn}}$$

$$+ \sum_{\substack{\text{all valence AO's} \\ \text{on m}}} c_{ip}^2 \left\langle \psi_r(1)\phi_p(2) \left| \frac{1}{r_{12}} \right| \psi_r(1)\phi_p(2) \right\rangle \qquad (1.190)$$

collecting terms common to each atom

$$\varepsilon_r = E_o + \sum \left(\frac{-Z_n}{r_{nm}} + \frac{2}{r_{nm}} \right) + \sum \frac{2c_{ip}^2}{r_{nm}}$$

$$+ \sum c_{ip}^2 \left\langle \psi_r(1)\phi_p(2) \left| \frac{1}{r_{12}} \right| \psi_r(1)\phi_p(2) \right\rangle \qquad (1.191)$$

However $\sum 2 + \sum 2c_{pi}^2$ is the total electronic population on a given atom n and hence $-Z_n + \sum 2c_{ip}^2$ represents the charge q_n on n.

Thus

$$\varepsilon_r = E_o + \sum_{n\neq m} \frac{q_n}{r_{nm}} + \sum 2c_{ip}^2 \quad \psi_r(1)\phi_p(2)\left|\frac{1}{r_{12}}\right|\psi_r(1)\phi_p(2) \qquad (1.192)$$

In the third term $\sum 2c_{ip}^2$ represents the total valence electron population on atom m and since this is proportional to the charge (1.192) can be written

$$\varepsilon_r = E_o + kq_m + \sum_{n\neq m} \frac{q_n}{r_{nm}} \qquad (1.193)$$

In theory k should represent the average repulsion between a core and valence electron on atom m though in view of the approximations involved in the derivation of (1.193) it is probably best to regard it as a parameter dependent on the definition of atomic charge, and basis set used. The form of the equation is readily interpreted in chemical terms as it relates the binding energy of a core level to the charge distribution in the molecule. In the absence of a potential from the remainder of a particular core level would be proportional to the charge on that atom. In practise, it is only for a closely related series of molecules, when the third term in (1.193) remains effectively constant that binding energies correlate directly with charges on atoms.

1.12 Spectral Line Widths and Resolution

The resolution of the spectrophotometer analyser (see instrumentation, appendix 2), $\Delta E/E$, where E is the energy of the electrons, depends upon the mean radius of the analyser hemispheres R and the combined widths of the entrance and exit slits W.

$$\frac{\Delta E}{E} = \frac{R}{W} \qquad (1.194)$$

Resolution can be improved by reducing slit widths (thus reducing signal intensity), increasing the hemisphere radius or by retarding the electrons prior to their entry into the analyser. The overall resolution depends also on contributions other than from the analyser

$$(\Delta E_m)^2 + (\Delta E_x)^2 + (\Delta E_l)^2 + (\Delta E_s)^2 \qquad (1.195)$$

where ΔE_x is the width of the Xray line inducing emmission

ΔE_l is the width of the natural energy distribution of the electron energy level

ΔE_s is the width of the broadening due to spectrometer aberation and depends on the emission energy E and the slit width.

The natural line width at half maximum height of the core level under investigation ΔE_s and that of the incidental radiation ΔE_x (unless monochromatization is employed) depend on the uncertainty principle [64]

$$\Delta E . \Delta t \simeq h/2\pi \qquad (1.196)$$

where Δt is the lifetime of the state. The photoelectric process is believed to occur in a time interval of the order of 10^{-18} sec, [64] whereas nuclear relaxation times are of the order of 10^{-13} sec,[65] and the hole created in the core will have a lifetime of approximately 10^{-16} sec.[64] Thus the process may be regarded as sudden compared with nuclear (but not electronic) motion.

CHAPTER II

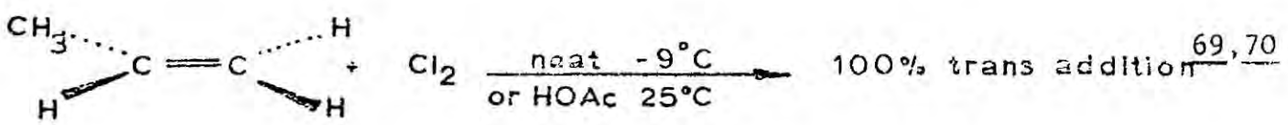
LCAO-MO-SCF CALCULATIONS UPON PROTOTYPE POTENTIAL ENERGY SURFACES FOR ELECTROPHILIC ADDITION TO ALKENES

2.1 Introduction

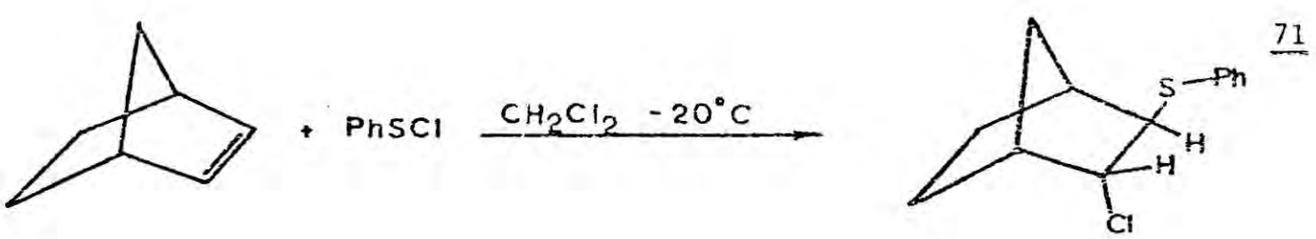
Carbon carbon double bonds undergo a wide variety of addition reactions, under both free radical and ionic conditions, the importance of which are manifest in many branches of chemistry, not least being synthetic organic chemistry. As outlined below, the stereochemistry of the electrophilic additions can allow important insight into the mechanism of these reactions and valuable data have thus been collected. These have led to an overall picture in which the first process is attack by an electrophile. The stereochemistry of overall reaction now hinges on the mechanistic steps which follow since the structure of the resulting carbonium ion intermediate will crucially effect the way in which the nucleophile can now attack with minimum activation energy. Hence one can see that to describe the process adequately, information on two chemical species are required, namely intermediates and transition states. The experimentalist can probe the former with a variety of tools, nmr in particular has greatly facilitated the description of carbonium ion intermediates, however ambiguities in interpretation sometimes arise. The transition state, on the other hand is clearly beyond the reach of any experimental approach for, by definition, it is the point on the reaction profile of least stability and hence of shortest lifetime. For this reason, information regarding the transition state can only be inferred from experimental studies, which give data referring to differences between the initial and transition state e.g. studies of substituent effects. Hence it becomes apparent that the description of electrophilic addition to alkenes can much benefit from the application of a study by theoretical methods, when in

principle both intermediates and transition states may be investigated and described and possible reaction profiles may be compared. For these reasons in this work a study was made of various prototype systems of electrophilic olefin addition by means of both non empirical and semi emprirical LCAO MO SCF quantum mechanical calculation, and this chapter details theoretical studies of prototype systems for electrophilic addition to alkenes to shed light on aspects which are not amenable to direct experimental observation.

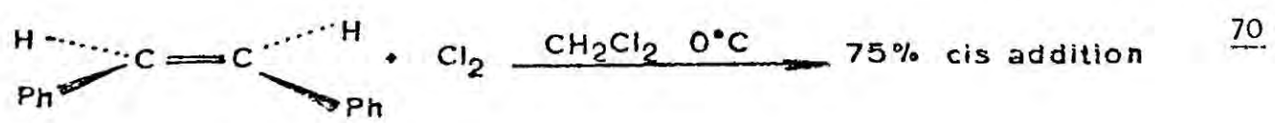
As already mentioned, experimental evidence has shown that these reactions are generally stereochemically precise and examples of both stereospecific trans and stereoselective as addition can be found e.g.



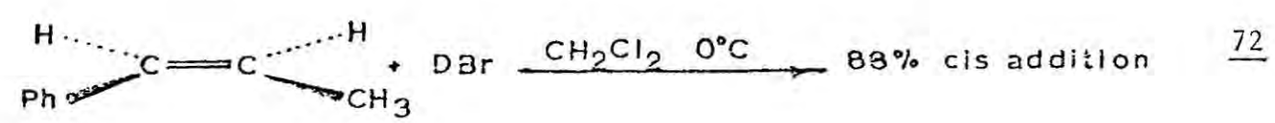
[69, 70]



[71]



[70]



[72]

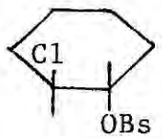
Clearly there are several factors which can possibly affect the percentage yields of various products obtained, notably the structure of the olefin, the nature of the adding species and the polarity of the solvent in which the reaction is carried out. From an examination of table 2.1.

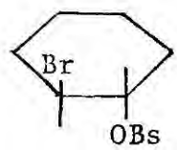
Table 2.1 Stereochemistry of Alkene addition

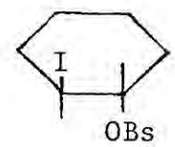
Adding Species	Conditions	Alkene	% trans addition	ref
DBr	CH_2Cl_2, 0°	1 Ph propene	12	72
F_2	CCl_3F, -126°C	1 Ph propene	22[a], 27[b]	73
Cl_2	CH_2Cl_2, 0°C	1 Ph propene	25[a], 33[b]	70
Br_2	CCl_4, 2-5°C	1 Ph propene	83[a], 88[b]	74
INCO	Et_2O	β deuterostyrene	100	75
ArSCl	CCl_4, 25°C	1 Ph Propene	100	74

a from cis olefin b from trans olefin

It is seen that in general when the adding species is formally Br^+, I^+ and RS^+, addition proceeds in a trans manner. This is usually taken to indicate the involvement of a bridged species, i.e. formation of an onium ion. Further evidence is found for the relative stabilities of halonium ions from the study of halogen as a neighbouring group in solvolysis reactions.[76] Neighbouring sulphur, iodine and bromine are seen to assist ionisation but there is no evidence for participation by neighbouring chlorine or fluorine. For example, in the solvolysis of a series of trans-2-halocyclohexyl brosylates[77]





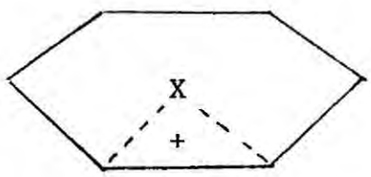


rate enhancement trans/cis	4	800	2,700,000

a distinct enhancement in the rates of the trans compounds relative to the respective cis compounds is seen, and the relative enhancement shows the trend

$$Cl < Br < I$$

Thus despite the greater steric compression which would be expected in the iodine case it must be concluded that the solvolysis of the trans compounds passes through the formation of a bridges intermediate species



where the relative stabilities of these species are in the order

$$Cl < Br < I$$

When X is F or Cl clear cut rate enhancements are not apparent. Furthermore, returning to addition to alkenes, when the electrophile is formally H^+, F^+ or Cl^+, examples of both stereospecific trans and stereoselective cis additions are found. This is usually interpreted as involving the formation of either bridged ion or classical ions, depending upon the olefin structure, though the relative stability of bridged ions with respect to classical ions

$$H < F < Cl$$

may be inferred. Considerable data have been collected by Olah and co-workers, and several bridged ions have been prepared and observed by low temperature nmr.[78] Again, however, this work is less conclusive when the adding species are not Cl, Br or I.

In figure 2.1 are sumarised the basic points which it was desired to investigate by means of theoretical calculation. From this several problems may be itemised:

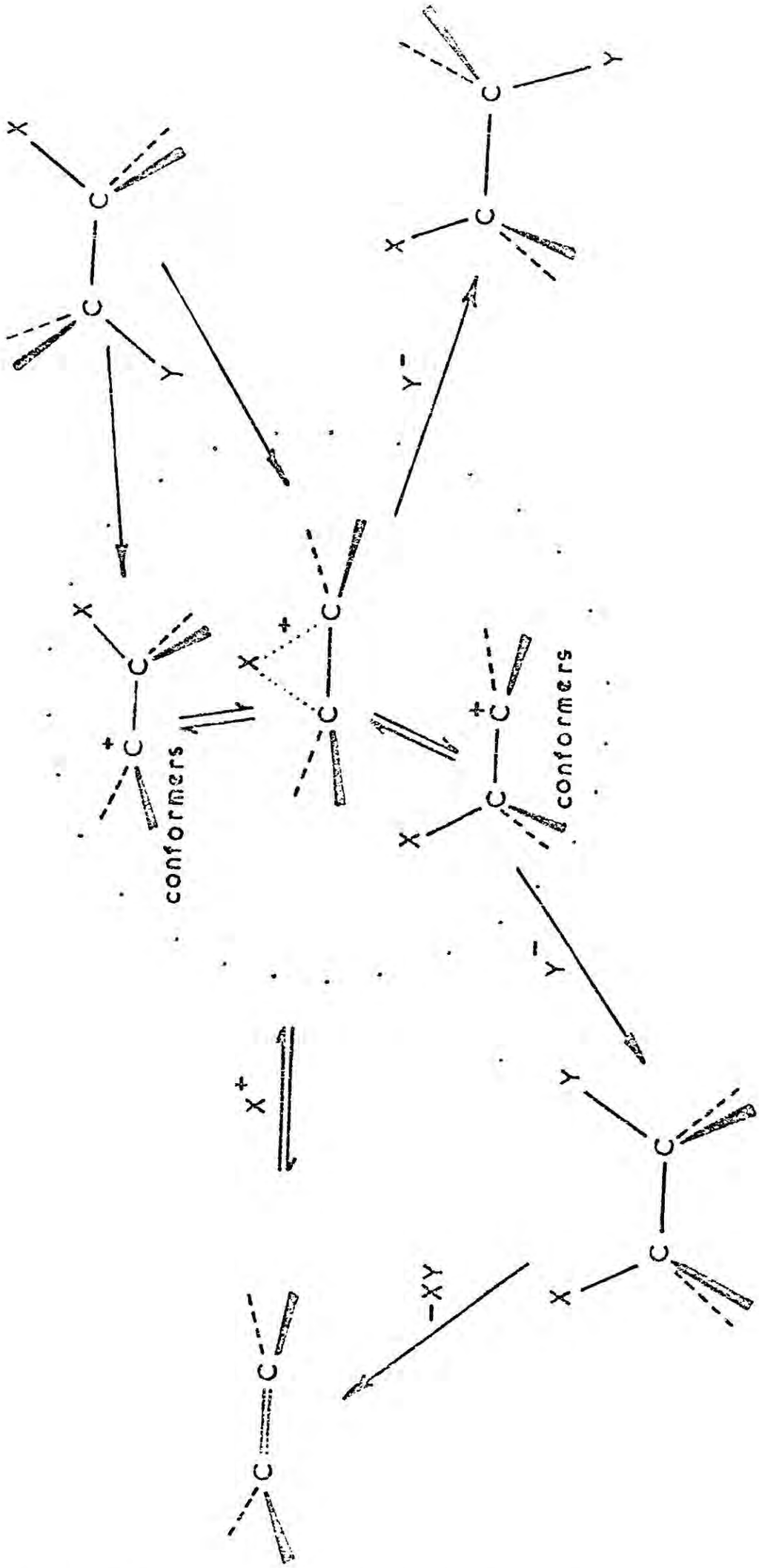


Fig 2.1 Outline of the processes examined in this study

(i) Ease of formation of bridged or classical ions as a function of olefin and electrophile and the direction of addition.

(ii) Equilibria between bridged and classical ions, and the conformational aspects of the latter.

(iii) The favoured direction of approach of a nucleophile to a classical ion, and the related process of elimination.

Some of these problems have been investigated here for the cases of H, F and Cl by means of ab initio and semi-empirical calculation. The ab initio approach has two main limitations which are

(i) The limitations of the Hartree Fock one electron model, which have been discussed in chapter I

(ii) The limitations of computer capability and time available.

Since the molecular systems studied were in closely related series with the same number of electron pairs, then in discussing interconversions within these series correlation energy changes should be relatively small. The second limitation is of rather greater importance, however, as this will result in a compromise in the choice of basis set and in the extent of geometry optimisation. For these reasons the use of a fairly limited basis set was necessary and this was tested for physical 'reasonableness' by computing bond lengths and rotation barriers in related hydrocarbons, thus providing a means for comparison with experimental data. To reduce further the computing work load, chemical "intuition" was used in order to reduce the multi-dimensional analysis inherent in full geometry optimisation.

Theoretical calculations on one molecule should yield results which predict the behaviour and nature of isolated, i.e. gas phase, molecules. With these data it should then be possible to go on further to understand more about solvent effects when the system is transferred

to a solvent environment. An example may be drawn from measurements of the acidities of the alcohols indicated below, which, in solution, show the trend

$$Me > Et > {}^{i}Pr > {}^{t}Bu$$

The original conclusion drawn from these data was that the methyl group was a better electron releaser than hydrogen. However theoretical calculations indicate that in fact overall the methyl group is electron accepting and that the energy differences would predict a reversal of the above order. These theoretical observations were in agreement with gas phase measurements [80] of alcohol acidity which also showed this apparently reversed order, and clearly the observed behaviour in solution is a result of enhanced solvation of the anions in the order

$$MeO^{-} > EtO^{-} > {}^{i}PrO^{-} > {}^{t}BuO^{-}$$

This example clearly demonstrates the ability of theoretical calculation to shed insight of the true nature of solvation effects in some systems.

2.2 Rotation Barriers in Ethane and Ethyl Cation

There are two major features of interest in the computation of rotation barriers; firstly the magnitude of the barrier and secondly the relative magnitudes of the component contributions to the total. Absolute magnitudes of rotational barriers in many substituted and unsubstituted hydrocarbons are readily available experimentally and are well documented and hence comparison with these provides a test of the adequacy of a given basis set for the reproduction of such barriers. Although in principle one can measure barriers to rotation in simple carbonium ions the experimental difficulties are such that none have

been reported and hence calculation of these can yield date of considerable importance for the organic chemist. A further great advantage of non-empirical calculation on conformational processes is that detailed analysis of component attractive and repulsive contributions to barriers to rotation or inversion is possible.

Ethyl cation has been the subject of numerous theoretical investigations, however at the stage at which this work was initiated no ab-initio calculations had been performed. Results which have been obtained in this and other work are collected in table 2.2. That internal rotation may be understood adequately within the Hartree-Fock model has been shown by a calculation [87] in which electron correlation was allowed for, when the rotational barrier decreased by only 0.13 Kcal $mole^{-1}$. The majority of the calculations of table 2.2 have indicated rotational barriers which are reasonable close to the experimental value of 2.93 Kcal $mole^{-1}$, the main exception being the calculation employing a FSGO basis set.

In this work the basis set chosen [88] (see Appendix) was C(7,3/3,1), H(3,1) making a total of 50 primitive and 18 contracted functions for ethane. The CH bond length, assumed invarient in hydrocarbons, was fixed at 1.093Å and similarly the HCH bond angle was taken as rigid at 109.5°. With these restrictions the energy of this species was minimised with respect to the length of the CC bond. The rotational energy barrier then calculated was 2.72 Kcal $mole^{-1}$, which is in good agreement with the experimental value, and hence the basis set was assumed adequate for the examination of conformational processes. In agreement with all previous workers the staggered form is predicted to be the most stable and it is of great interest to decompose the rotational barrier into net attractive and repulsive contributions.

Table 2.2 Internal Rotation in Ethane and Ethyl Cation

	Calc^{d} Barrier	Eror for stag.	comments	basis set	ref
Ethane	3.32	-78.305478	rigid rot^{n} assumed	STO 3G†	81
	3.33	-78.862315	pot^{l} cst^{s} evaluated	STO 4G	
	2.90	-78.30618	stag D_{3d}; r_{cc} 1.538A r_{ch} 1.086Å <HCH 108.2° [exp. r_{cc} 1.531Å r_{ch} 1.096A <HCH 107.8°]	STO 3G	82
	2.80	-79.11582	stag D_{3d}; r_{cc} 1.529Å r_{ch} 1.083Å <HCH 107.7°	STO 431G	
	2.58	-78.978110	Rigid rot^{n}	C10s6pH4sGTO CGTO‡(2s2p2s)	83
	5.17	-67.347295	Rigid rot^{n}Exp$^{\ell}$geom also calc^{d} r_{cc} 1.397Å	Minimal FSGO‡‡	84
	3.14	-79.203142	Rigid rot^{n}Exp$^{\ell}$geom	C9sFpH4sGTO	85
	3.65	-79.23770	Stag D_{3d} r_{cc} 1.551Å <HCH 107.3° Eclip D_{3o} r_{cc} 1.570Å <HCH 107.0°	C11s7pH6sGTO S3p3s CGTO augmented by 3d x z, $3d_{yz}$ 2p polar^{n} fn_{s}	86
Ethyl Cation	0.0		rigid rot^{n}	STO 3G	82
	0.22		flexible rot^{n}	STO 3G	

† STOnG least squares expanison of STO basis set in terms of nGTO per STO

‡ CGTO refers to contraction of GTO

‡‡ Floating Spherical Gaussian Orbitals

The total energy of a molecular system may be taken as a sum of four terms, nuclear repulsion potential Vnn, nuclear-electron attractions Vne, electron repulsions Vee and electronic kinetic energy T:

$$Etot = Vnn + Vne + Vee + T \qquad (2.1)$$

The variation of each term may be studied during conformational change. However, it must be remembered that at present it is only possible to describe qualitatively the origin of the conformational energy barriers in terms of energy components. This is due to the use of a decomposition of the total energy into components whose comparative changes (phase, amplitude) on changing between conformations are not much affected by improving the wave function, by scaling the energy terms to force agreement with the virial thesren [89] and by small geometrical perturbations [90]. Lehn [91], [92] has described several two component decompositions of the total energy into overall attractive and repulsive terms

attractive	repulsive
Vne	Vnn + Vee + T
Vnn + Vee + T	Vee
Vne + Vee + T	Vnn
Vne + Vee + Vnn	T

Thus the energy barrier may be described in terms of the variations ΔVatt and Δ Vrep. In this work the second of these decompositions has been used, and the results for ethane are shown in table 2.3

Table 2.3 Rotational Components of Ethane

	Etot (au)	(Vne + T) (au)	Vee (au)	Vnn (au)
stag	-78.906247	-186.13054	65.377007	41.847286
eclip	-78.901915	-186.15409	65.397540	41.854639

$|\Delta Vatt| = 0.0162au$ $|\Delta Vrep| = 0.0205au$ $\Delta Etot = 2.72 Kcalmole^{-1}$

It is clear from the table that the barrier is dominated by the net repulsive components and hence the obvious pictorial representation of the barrier as repulsive interaction in the eclipsed conformation is essentially correct. More recent work by Allen[83] is also in agreement with these conclusions.

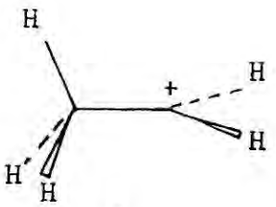
As indicated previously, in the case of reactive intermediates, such as carbonium ions, the experimental conformational data is less readily obtainable and hence the theoretical results are particularly valuable. The ethyl cation has been studied fairly extensively at varying stages of sophistication but at an ab initio level it has been much less stuedied than ethane. Results of previous calculations are shown in table 2.2, and the conclusion is that the rotational barrier is very small. In this work the same basis set was used as for ethane, and the CH bond lengths were set at 1.093 and 1.084Å and <HCH angles at 109.5^{o} and 120^{o} for sp^3 and sp^2 environments respectively. The energy was again minimised with respect to the CC bond length and the rotational barrier was found to be effectively zero for rigid rotation. The 'barrier' has been analysed in terms of attractive and repulsive components and it can be seen that (table 2.4) all the energy components are virtually rotationally independent.

Table 2.4 Rotational Components of Ethyl Cation

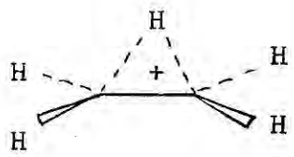
	Etot (au)	(Vne + T) (au)	Vee (au)	Vnn (au)
stag	-77.911158	-170.37442	55.45621	37.00706
eclip	-77.911153	-170.37441	55.45619	37.00706

2.3 Ethyl cation and Bridge protonated ethylene

The simplest model for electrophilic addition to olefins is that formally generated by the addition of a proton to the ethylene molecule. Two possible extremes may be considered for this process; the ethyl cation (I) and bridge protonated ethylene (II)



I



II

This system has been investigated by theoretical methods of varying stages of approximation, beginning in 1964 with an EHT treatment by Hoffmann.[93] This method, however, is theoretically unsound for charged species and subsequent investigations [94] have used the semi empirical methods which are formal approximations to the ab initio treatment, such as CNDO, INDO, MINDO and NDDO. However, Dewar [94] has shown that in all these methods, the stability of bridged ions is considerably over-estimated (due no the neglect of large numbers of electron repulsion and exchange integrals, and hence an ab initio treatment is necessary in order to obtain reliable results. Dewar et al [94] applied this method (see table 2.5) and obtained the result that ethyl cation was 9Kcal $mole^{-1}$ more stable than bridge protonated ethylene. This work relied, however, on approximate geometries derived from NDDO calculations.

Table 2.5 Results of some earlier investigation [94] of the ethyl cation - protonated ethylene system.

	Method	r_{cc}(Å)	dist of H from CC bond centre	Erel(Kcal $mole^{-1}$)
Ethyl cation	NDDO	1.43	-	(0.0)
	ab initio	1.48	-	(0.0)
Prot^{d} Ethylene	NDDO	1.41	1.0	-33.2
	ab initio	1.46*	1.11*	9.0

* assumed geometry.

In this work the contracted gaussian basis set was the aforementioned C(7,3/3,1) H(3/1). The CC bond length in the ethyl cation was optimised as described earlier, whilst for bridge protonated ethylene, the CC bond length, distance of the bridging hydrogen to the centre of the C-C bond, and the out of plane bending of the ethylenic hydrogen atoms were investigated as functions of the total energy. The ethylenic CH bond lengths were fixed as 1.084Å and the <HCH angles assumed rigid at 120°. Figure 2.2 shows the total energies of the ethyl cation and bridge protonated ethylene as functions of their respective CC bond lengths and figure 2.3 shows the total energy of bridge protonated ethylene as a function of the distance of the distance of the bridging hydrogen to the CC bond centre. From these may be deduced the optimised bond lengths, summarised in table 2.6.

The CC bond length for ethyl cation using this basis set is seen to be comparable with that for ethane, whilst that for bridge protonated ethylene is rather longer than that of ethylene itself. Out of plane bending of the ethylenic hydrogen atoms of bridge protonated ethylene was found to case a slight raising of total energy,

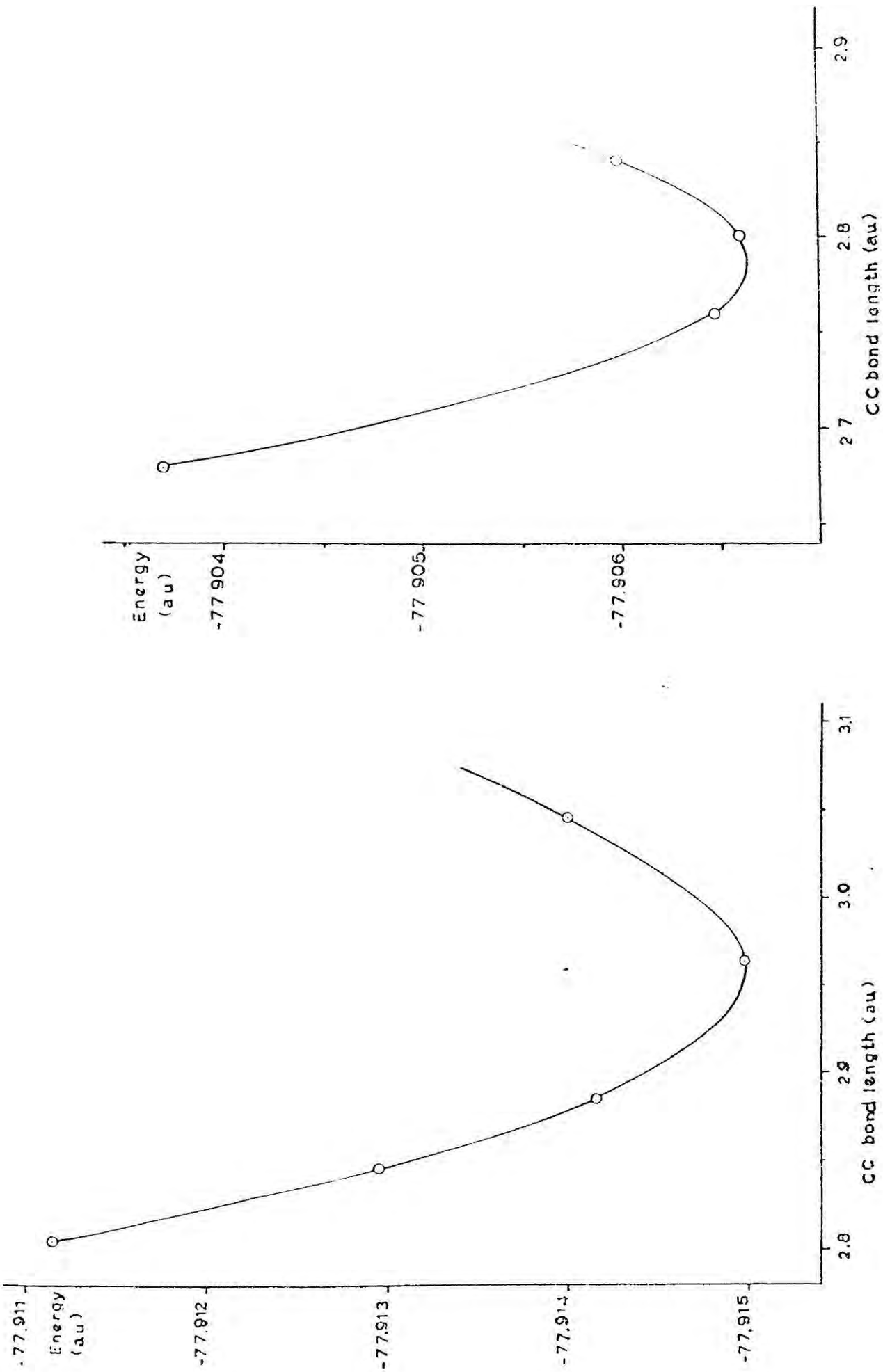


Fig 2.2 Total energy of Ethyl Cation (left) and Bridge-protonated Ethylene (right) as functions of CC bond length

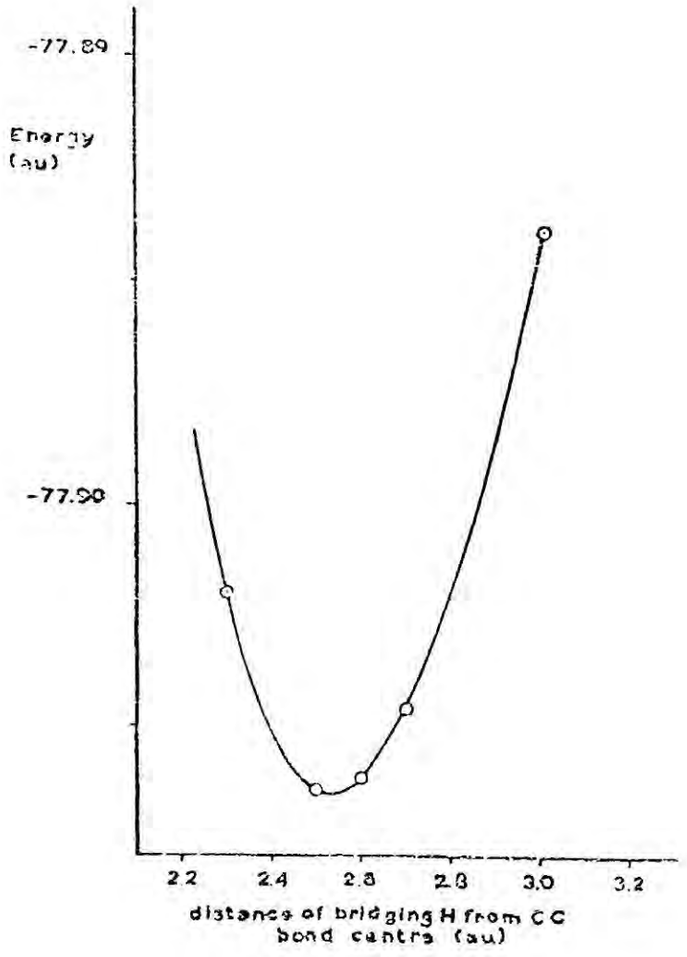
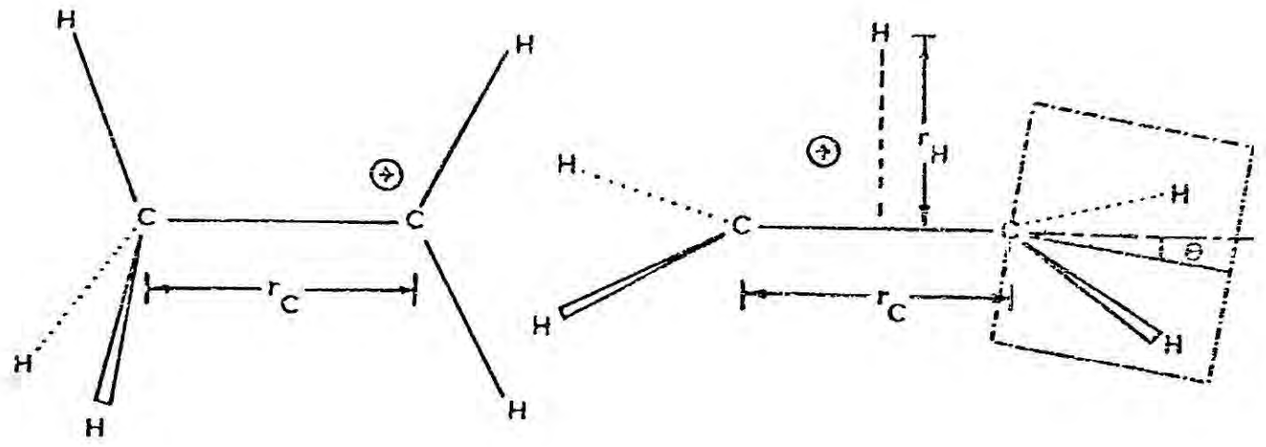


Fig 2.3 Total energy of Bridge-protonated Ethylene as a function of the distance of the bridging H from the CC bond centre



	r_C 1.37Å	r_C 1.47Å r_H 1.34Å θ 0°
ΔE (Kcal mole^{-1})	(0.0)	5.2 3.4 *

Table 2.6 Partially optimised geometries of Ethyl Cation and Bridge-protonated Ethylene

<Out of plane	ΔE (Kcal mole^{-1})
3.0°	0.0
5.0°	0.2 Kcal mole^{-1}
7.0°	0.5 Kcal mole^{-1}

though in subsequent investigations [95,96] some workers have obtained slight displacements of these hydrogens (3°) to produce slightly greater stabilisation (0.3 Kcal mole^{-1}). In some recent work, Pople et al have employed split valence basis set calculations for the examination of various simple carbonium ions including classical forms of the ethyl cation and bridge protonated ethylene. A sequential process of increasing basis set size was used with the result that after extensive geometry optimisation and the use of an STO 6-31G basis set which included extra functions of p and d symmetry, then the symmetrically bridged non-classical ion becomes marginally more stable than the classical forms. However, with the exception of these results all the non empirical calculations reproduce the result that the classical ethyl cation is lower in energy than the bridge- protonated ethylene molecule, but nevertheless the energy difference between the two species is seen to be fairly small and is expected to become even smaller if calculations employing polarisation functions and a very large basis set are performed.

Using the geometries derived above, CNDO and INDO calculations have been performed on ethyl cation and bridge-protonated ethylene and the results obtained predict that the bridge-protonated form should be more stable than the classical ethyl cation (by 28.5 and 29.8 Kcal mole^{-1} respectively). This is in accord with the observation earlier of the overestimation of the stabilities of bridged ions by these methods.

2.4 Interconversion of ethyl cation and bridge protonated ethylene

The small energy difference between the ethyl cation and bridge protonated ethylene obtained in this work raises the possibility of protonated ethylene as an intermediate in electrophilic addition to olefins.(cf 98) However, before coming to any firm conclusions it is important to gain some understanding as to how easy is the transformation between these species.

Thus calculations were carried out along an idealised reaction co-ordinate in which the bridging hydrogen in protonated ethylene slides along the bond such that the hydrogen moves in a straight line from its initial position to its final position in ethyl cation. This corresponds to a continuous change in C_2H bond lengths. The CC and C(2)H bond lengths, and <HC(2)H bond angles were assumed to change continuously in going from protonated ethylene to ethyl cation (see fig. 2.4). Calculations were performed at four intermediate points using the C(7,3/3,1) H(3/1) basis set employed previously. The results are shown graphically in fig. 2.5.

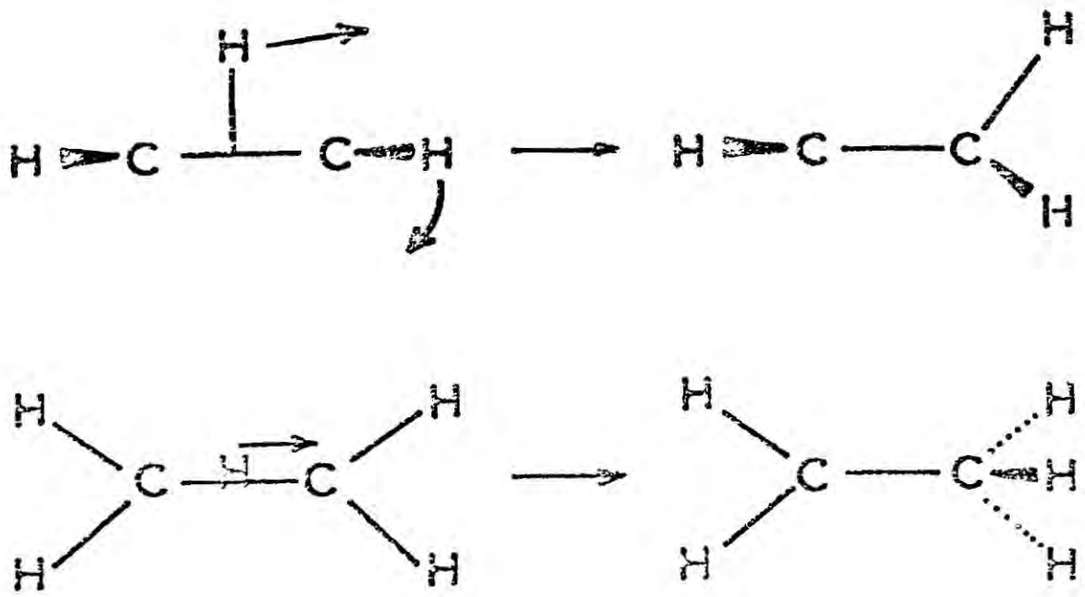


Fig 2.4 Diagram of atomic movements involved in the transformation of Bridge-protonated Ethylene to Ethyl Cation.

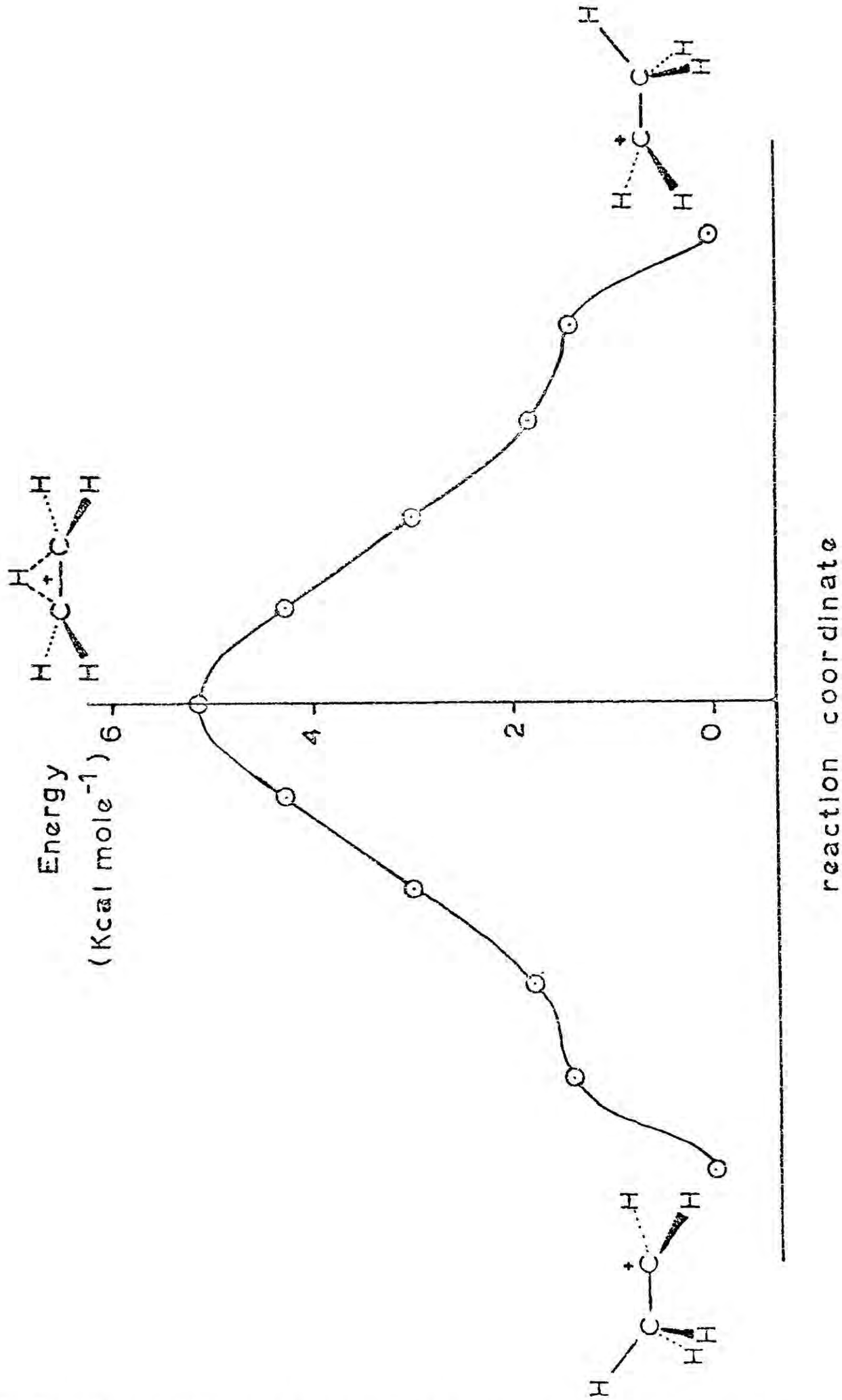


Fig 2.5 Reaction coordinate for a 1, 2 - hydrogen shift in Ethyl cation via Bridge-protonated Ethylene

The interesting result emerges that it is possible to transform bridge protonated ethylene to ethyl cation without an activation barrier. Thus it may be postulated that in the gas phase, protonated ethylene, rather than being an intermediate, is the transition state for scrambling the hydrogen atoms of ethyl cation. As will be shown later, this is also the likely result in solution also, and the activation energy obtained is compatible with the results of Brouwer et al [99] who examined the system

by nmr at various temperatures. The four methyl groups appear equivalent (δ 2.75ppm, j = 5 cps) and unbroadned down to $-85^{o}C$. This is equivalent to K $>10^{4}sec^{-1}$. Further evidence comes from the mass spectrometric investigation of CH_3CD_3 [100], where $\overset{+}{C}H_2CD_3$ and $CH_3CD_2^+$ produced complete scrambling of H and D, as shown by the appearance of metastables for CH_2D and CD_2H in statistical ratios. Thus it may be argued that the scrambling is fast compared to the time of flight of the parent ion and hence scrambling is complete at dissociation.

Recent extended basis set calculations by Pople et al [97] have suggested that the classical ions may not be close to local minina in the potential energy surface and that the bridged form may be the equilibrium structure at the Hartree-Fock limit.

Using CNDO and INDO treatments a rather different picture emerges. The profile is essentially inverted since the bridge protonated

ethylene is predicted to be more stable than the ethyl cation but also mimina of ~ 6kcal $mole^{-1}$ are seen on either side of the bridge-protonated form. This would suggest a predominant species with an essentially bridging proton but displaced slightly from the centre of the CC bond.

2.5 Solvent effects in ethyl cation - $prot^d$ ethylene system

The rate constant for a reaction is expressible as

$$K = f \exp(-\Delta Eg - \Delta Esolv(T) / RT) \qquad (2.2)$$

where f is the ratio of partition functions, ΔEg is the gas phase activation energy and ΔE solv(T) is the solvation energy change. This last term depends mainly on the size of the ion and is not readily calculable. It may be divided into three components

(a) Cavitation energy - associated with the 'hole' created in the solvent by the ion

(b) Orientation term - due to the changes in average orientation of solvent molecules by the ion

(c) Interaction term - due to the average intermolecular forces which appear between the solvated ions and the partially orientated solvent molecules.

Due to the comparable size of ethyl cation and protonated ethylene, only the interaction term is really of importance in the determination of solvation energy differences. This term can itself be decomposed into three parts.

i) part due to dispersion forces

ii) part due to isotropic charge-dipole or dipole-dipole interactions

iii) part due to anisotropic interactions such as hydrogen bonding.

Terms i) and iii) can be considered to be very similar in the two ions and hence the major term contributing to the difference in ΔEsdu(T) is likely to be the isotropic interaction term. This may be estimated approximately from the formula [101], [102], [103]

$$\Delta Esolv = -\Delta\sum \frac{q_i q_j}{2r_{ij}} \left(1-\frac{1}{D}\right) \qquad (2.3)$$

where the q_i's are the charges on each atom and r_{ij} is the interatomic distance between atoms i and j except when i and j are identical. In this case r_{ij} represents the effective radius of atom i. D is an effective dielectric constant for the solvent.

From the gross atomic charges and <r> values (from atomic Hatree Fock calculations) for the effective radii, the calculated solvation energies for the system were

ethyl cation	-95.0 kcal mole^{-1}
protonated ethylene	-92.6 kcal mole^{-1}

(For a solvent of high D, i.e. $1 >> \frac{1}{D}$)

Hence the activation barrier for scrambling the hydrogens of ethyl cation is predicted to be slightly larger (by ~ 2.4kcal mole^{-1}) in solution than in the gas phase. The predicted greater solvation energy of the ethyl cation also reverses the ordering of the stabilities of classical and bridged forms as obtained by Pople.[97] Thus it may be concluded that in solution the ethyl cation - bridge protonated ethylene potential energy surface will be qualitatively as described here.

The absolute magnitudes of these solvation energies would appear reasonable if, to take a highly simplified model, the charge in bridge protonated ethylene is evenly distributed and it is assumed that the molecule overall presents a spherical surface to the solvent then taking

an average value for the mean radius of 1.34A (i.e. the distance of the hydrogen from the CC bond centre), the solvation energy would approximate to that of K^+ (Ionic radius 1.33Å). The enthalpy of hydration for this ion has been estimated [104] to be -84.0 kcal $mole^{-1}$. In view of the approximations involved, this may be regarded as close agreement.

The approximations inherent in the assumptions of the Hartree Fock formalism may also be considered. In general it has been found that in discussing processes in which the number of electron pairs do not change, correlation energy corrections are relatively small. It may be anticipated that since the classical ion and the bridged species possess the same number of electron pairs that correlation energy differences should also be relatively small. However, the calculated energy difference between the two species is also relatively small and since there is a change in the degree of connectivity in proceeding from one to the other small changes in correlation energy might become of some importance. An attempt has been made to investigate this possibility. Snyder and Basch [34] have shown how crude estimates of correlation energy differences may be obtained from pair correlation energies (established from atomic data) and appropriate population analyses

$$Ecorr = \sum_i \varepsilon_{ii} \tfrac{1}{2}\rho_i + \sum_{i<j} \varepsilon_{ij}\rho_i\rho_j \qquad (2.4)$$

where ρ_i is the atomic orbital electron density for orbital i. The atomic orbital pair population of orbital i is then $\frac{1}{2}\rho_i$. The atomic orbital pair correlation energy E_{ij} is the correlation energy of a pair of electrons in the same (i=j) or different ($i \neq j$) atomic orbitals. From atomic data they have prepared a set of pair correlation energies for first row atoms and hydrogen, the relevent data is given in table 2.7a

Table 2.7a Pair correlation energies in au

Pair		Atom		
i	j	C	F	H
1s	1s	-0.0409	-0.0398	-0.0409
1s	2s	-0.0015	-0.0014	
1s	2p	-0.0015	-0.0016	
2s	2s	-0.0284	-0.0119	
2s	2p	-0.0139	-0.0084	
2p	2p	-0.0258	-0.0258	
2p	2p'	-0.0123	-0.0123	

From the form of equation 2.4 it can be seen that Ecorr may be split into two terms E_{corr}^{intra} and E_{corr}^{intra} depending on whether the pairs of orbitals are located on the same or different atoms. From atomic data only E_{corr}^{intra} may be estimated. However from extended basis set calculations and estimates of ΔE_{corr}^{intra} for a series of hydrogenation reactions, Snyder and Basch showed that ΔE_{corr}^{inter} was in general of the same order of magnitude and opposite in sign, and hence may not be neglected. A roughly exponential dependence of ΔE_{corr}^{inter} on inter-atomic distance between the first row atoms was, however,evident from their data. In the case of the ions studied in this work only order of magnitude estimates of ΔE_{corr}^{intra} have been made, and these are shown in table 2.7b

Table 2.7b Estimates of ΔE_{corr}^{intra} for bridged and classical ions (in a.u.)

Atom		Bridge-protonated ethylene	ethyl cation
C	1s - 1s	-0.0818	-0.0818
	1s - 2s	-0.0080	-0.0081
	1s - 2p	-0.0183	-0.0181
	2s - 2s	-0.0378	-0.0384
	2s - 2p	-0.1128	-0.1132
	2p - 2p	-0.0785	-0.0780
	2p - 2p'	-0.0760	-0.0759
H	1s - 1s	-0.0663	-0.0665
	TOTAL	-0.4795	-0.4800

The difference in ΔE_{corr}^{intra} between the two species is estimated to be -0.0005 au i.e. -0.31 kcals with respect to the classical ion and for comparison the total correlation energy for ethylene has been estimated [105] to be -0.5191± 0.0114 a.u. Thus correlation energy differences are likely to make a relatively minor contribution.

2.6 Approach of a Nucleophile to Ethyl Cation

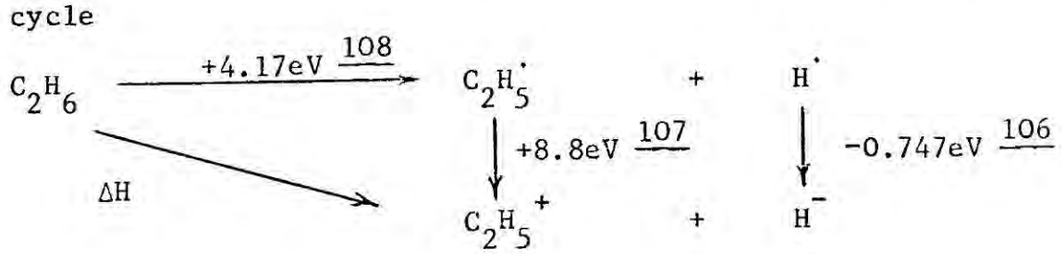
Once the ethyl cation has been generated by electrophilic addition of a proton to an ethylene molecule, the next stage in the reaction will be the attack at the electron defficient carbon atom by a nucleophile. It is of considerable interest to consider the stereochemistry of this nucleophilic attack. Two distinct approaches of a nucleophile of the staggered conformation of ethyl cation, which represent extrema of directions of approach, may be visualised - where the nucleophile approaches the electron defficient carbon in a trans or cis coplanar arrangement, shown in fig 2.6.

The overall enthalpy change for the process may be calculated from the total energies of ethane, ethyl cation and H^-. The energy of H^- may be taken from that of the hydrogen atom (by definition -0.5au) plus its electron affinity [106] (-0.747eV). Thus the enthalpy change for the process

$$\underset{-77.911153au}{C_2H_5^+} + \underset{-0.527au}{H^-} \longrightarrow \underset{-78.906247au}{C_2H_6}$$

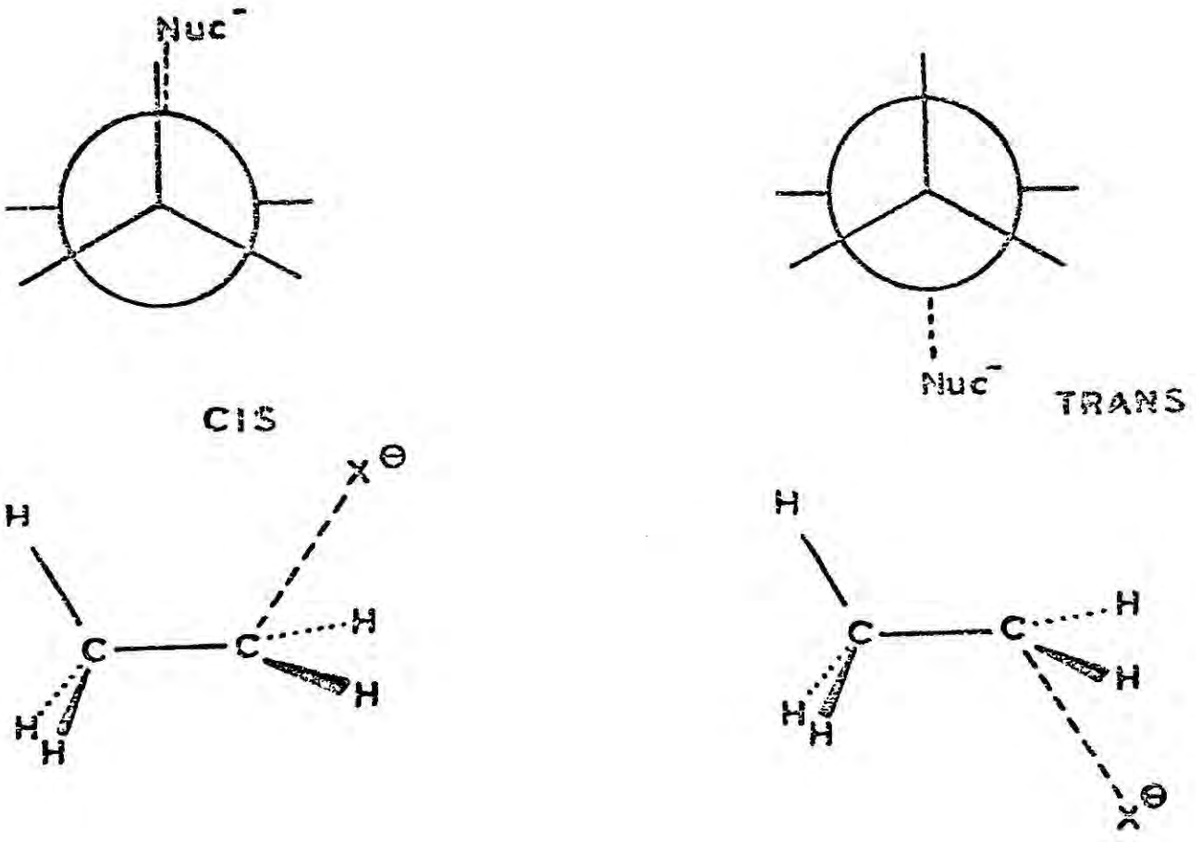
is calculated to be -288.4 kcal $mole^{-1}$

An independent thermochemical estimate can be made by consideration of the Born cycle



The enthalpy change thus calculated is -281.0 kcal mole^{-1}, which is in good agreement with the above figure.

Figure 2.6 Diagram of approach of prototype nucleophile to Ethyl Cation



Since the quenching of the carbonium ion by the nucleophile is so exothermic, the transition state for the reaction should approximate to the isolated ions (Hammond-Polanyi postulate),[109] and hence the interest is in the initial stages of the approach of the nucleophile. This being the case it is of interest to interpret the long range interaction, corresponding to an extensively stretched CH bond i.e. to heterolytic fission. Such a situation is not well described within the Hartree-Fock formalism, however the problem may be circumvented within the Hartree-Fock approach by artificially localising the negative charge upon the nucleophile. The hydride ion has been used as a prototype nucleophile and to ensure the above charge localisation the basis functions for this hydrogen atom were considered as if they belonged to a seperate symmetry representation. Considering the hydride nucleophile and the two directions of approach it may be seen that the reaction products would be ethane in either staggered or eclipsed conformation. Using the C(7,3/3,1), H(3/1) basis set, calculations were performed along two reaction co-ordinates, in which the hydride ion approached the electron deficient carbon along projections of the eventual C-H bonds which would exist in the product ethane.

The results are shown in table 2.8, and clearly indicate that although the final product of the trans approach (i.e. staggered ethane) is more stable, an initial cis approach is favoured. Thus, by the principle of microscopic reversibility it can be postulated that an initial cis elimination will be preferred. It has been shown [110] that in the gas phase elimination of hydrogen halide from ethanes the process occurs in a stereospecifically cis manner. Thus it may be inferred that the transition state for pyrolysis of a substituted halide will look very much like a fully developed carbonium ion and halide

ion in an ion pair. This provides the closest experimental test of the theory.

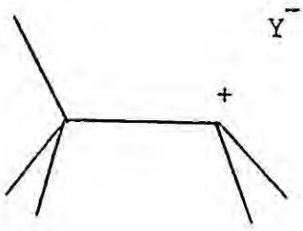


Table 2.8 Relative energy of the $C_2H_5^+$ H^- system as a function of C^+ .....H^- separation.

distance of H^- from electron defficient C	ΔE (kcal $mole^{-1}$) (cis - trans)
∞	0.0
4.0	-1.65
3.5	-2.11
3.0	-2.66
2.0655 (ethane)	2.72

2.7 Proton addition to Fluoroethylene

The next system investigated in this work was that generated by the electrophilic addition of a proton to fluoroethylene. This can produce two classical ions, 1-and 2-fluoroethyl cations, and the bridge protonated fluoroethylene and the work followed a similar pattern to that described in the previous section.

2.8 Conformational processes in fluoroethane and 1- and 2- fluoroethyl cations

In this work a basis set comparable to the ethyl cation-protonated ethylene work was used i.e. (7,3/3,1) for carbon and fluorine, and 3/1 for hydrogen. The geometry used was that optimised previously for ethane, with a hydrogen atom replaced by fluorine at a CF bond length [111] of 1.33Å. Bond angles were again assumed invariant at 109.5°. The rotational barrier obtained was 2.60 kcal $mole^{-1}$, the staggered conformer being the more stable. This shows that the barrier is, indeed, very close to that for ethane (2.72 kcal $mole^{-1}$), despite the much larger total energy of fluoroethane (staggered, -177.53239 au).

Table 2.9 Rotational components for fluoroethane

	Etot(au)	$E_{(1elec)}$ $(Vne+T)_{(au)}$	$E_{(2elec)}$ $(Vee)_{(au)}$	Enuc(Vnn) (au)
stag	-177.53239	-400.15366	143.00673	79.614547
eclip	-177.52825	-400.23156	143.05534	79.647979

$|\Delta Vatt| = 0.0445$ au

$|\Delta Vrep| = 0.0486$ au

$\Delta Etot = 2.60$ kcal $mole^{-1}$

From an examination of table 2.9 it is evident that the barrier is repulsively dominated, and though the barrier is very close to that for ethane the absolute magnitudes of the components are much larger in fluoroethane. Thus it appears as if an extra potential energy term is being added (approximately equal in both) to Vatt and Vrep. This can be attributed to the difference in effective potential between fluorine and hydrogen atoms. Since the radius of the fluorine atom is comparable to that of hydrogen, the effect being observed is that of lowering the

potential well around one of the rotating atoms due to the high charge density of fluorine. These results and conclusions are in good agreement with the work of Allen[83] who has since also analysed the rotational barrier of fluoroethane.

Earlier it was noted that the barrier to rigid rotation in ethyl cation is indistinguishable from zero, and this is in accordance with the organic chemists picture of essentially free rotation in simple alkyl carbonium ions. Thus it was of interest to examine the conformational processes for the fluoro substituted ethyl cations.

The geometries for both ions were obtained from that previously optimised for ethyl cation, with a hydrogen atom at C_1 or C_2 being replaced by a fluorine atom at a CF bond length of 1.33Å and the bond angles remaining fixed at 109.5° and 120° for sp^3 and sp^2 environments respectively.. The basis set used was the same C,F(7.3/3.1) H(3/1) set used previously, and calculations were performed upon distinct staggered and eclipsed conformations. A total of three calculations on 1-fluoroethyl cation and four on 2-fluoroethyl cation allows the construction of energy diagrams corresponding to intervals of 30° in the angle of rotation about the C-C bonds.

The results for the 1-fluoroethyl cation are shown in fig 2.7 and two points of interest arise. The first is the very low barrier to rotation (0.62 kcal $mole^{-1}$), and the second is the fact that the most stable conformer is that with hydrogen eclipsing fluorine. It is interesting to note that this feature was also reproduced by semi empirical calculations (CNDO and INDO SCF MO) performed in this work, though the predicted barrier was rather smaller in both calculations (0.17 and 0.18 kcal $mole^{-1}$ respectively). The barrier may be compared

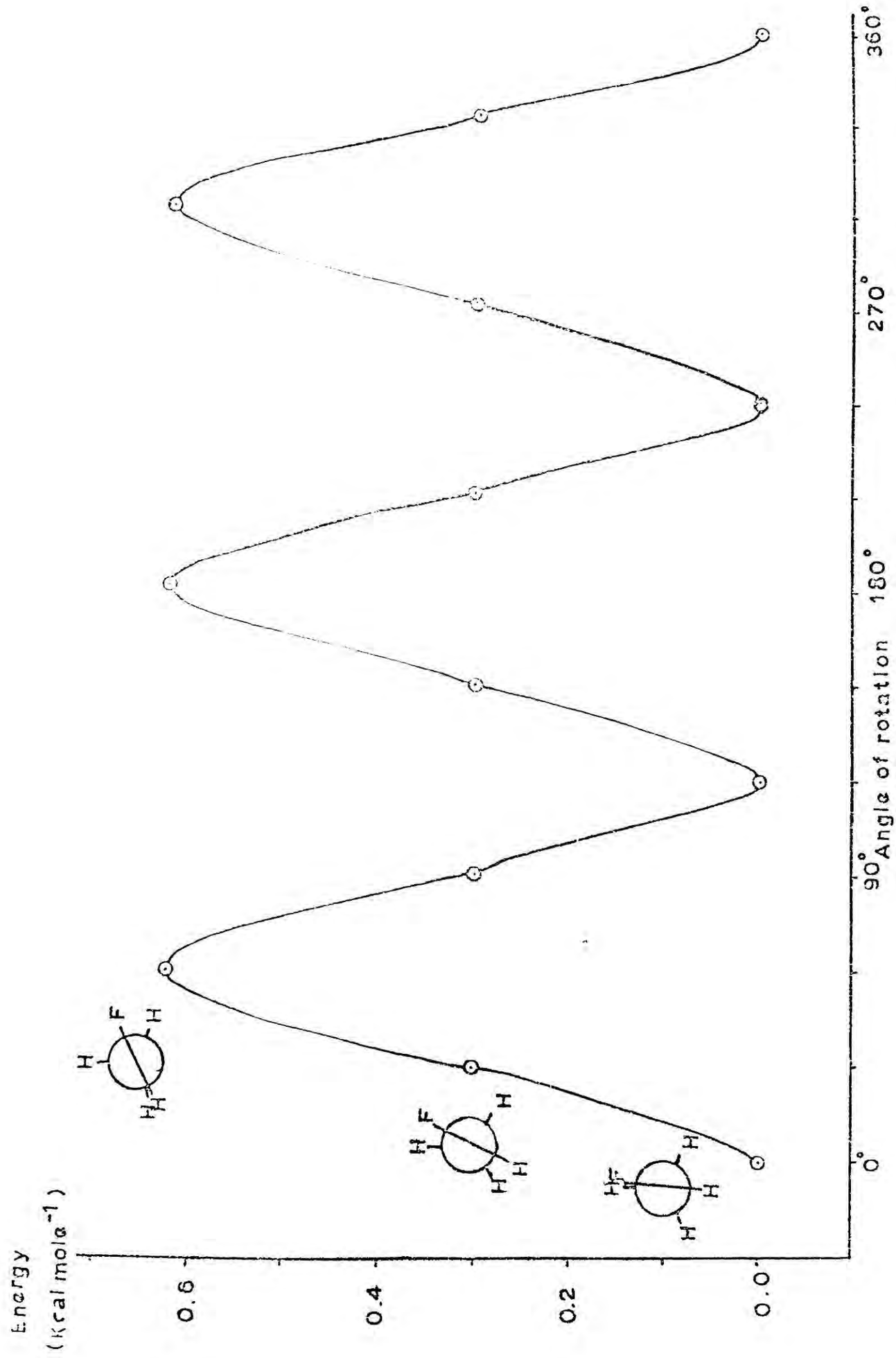


Fig 2.7 Conformational behaviour of the 1- fluoroethyl cation

with the isoelectronic acetaldehyde molecule,[112] which has a rotational barrier of 1.15 kcal $mole^{-1}$, and the most stable conformer with hydrogen eclipsing oxygen. Clearly it is again of interest to analyse the energy barrier of 1-fluoroethyl cation in terms of its attractive and repulsive components. The results of the analysis are shown in table 2.10.

Table 2.10 Rotational components for 1-fluoroethyl cation

rotamer	Etot(au)	E_{1elec}(au)	E_{2elec}(au)	Enuc(au)
0^o	-176.58818	-373.98804	126.28470	71.11516
60^o	-176.58719	-373.96845	126.27900	71.10226

$|\Delta Vatt| = 0.00669$au

$|\Delta Vrep| = 0.0057$ au

Δ Etot $= 0.62$ kcal $mole^{-1}$

In contrast to both ethane and fluoroethane the barrier is seen to be attractively dominated, and this has also been shown to be the case for acetaldehyde.

The results for the 2-fluoroethyl cation are shown in fig 2.8. The most stable conformer has hydrogen eclipsing fluorine, and the least stable conformer is staggered. From simple considerations the reverse might have been expected. In marked contrast to ethyl and 1-fluoroethyl cations, there is a substantial rotational barrier (10.53 kcal $mole^{-1}$) and this is dominated by attractive components - see table 2.11.

rotamer	Etot(au)	E_{1elec}(au)	E_{2elec}(au)	Enuc(au)
0^o	-176.52545	-377.73242	128.80992	72.39706
90^o	-176.50867	-377.40754	128.57972	72.31914

$|\Delta Vatt| = 0.2470$au

$|\Delta Vrep| = 0.2302$au

$\Delta Etot = 10.53$ kcal $mole^{-1}$

Table 2.11 Rotational components for the 2-fluoroethyl cation

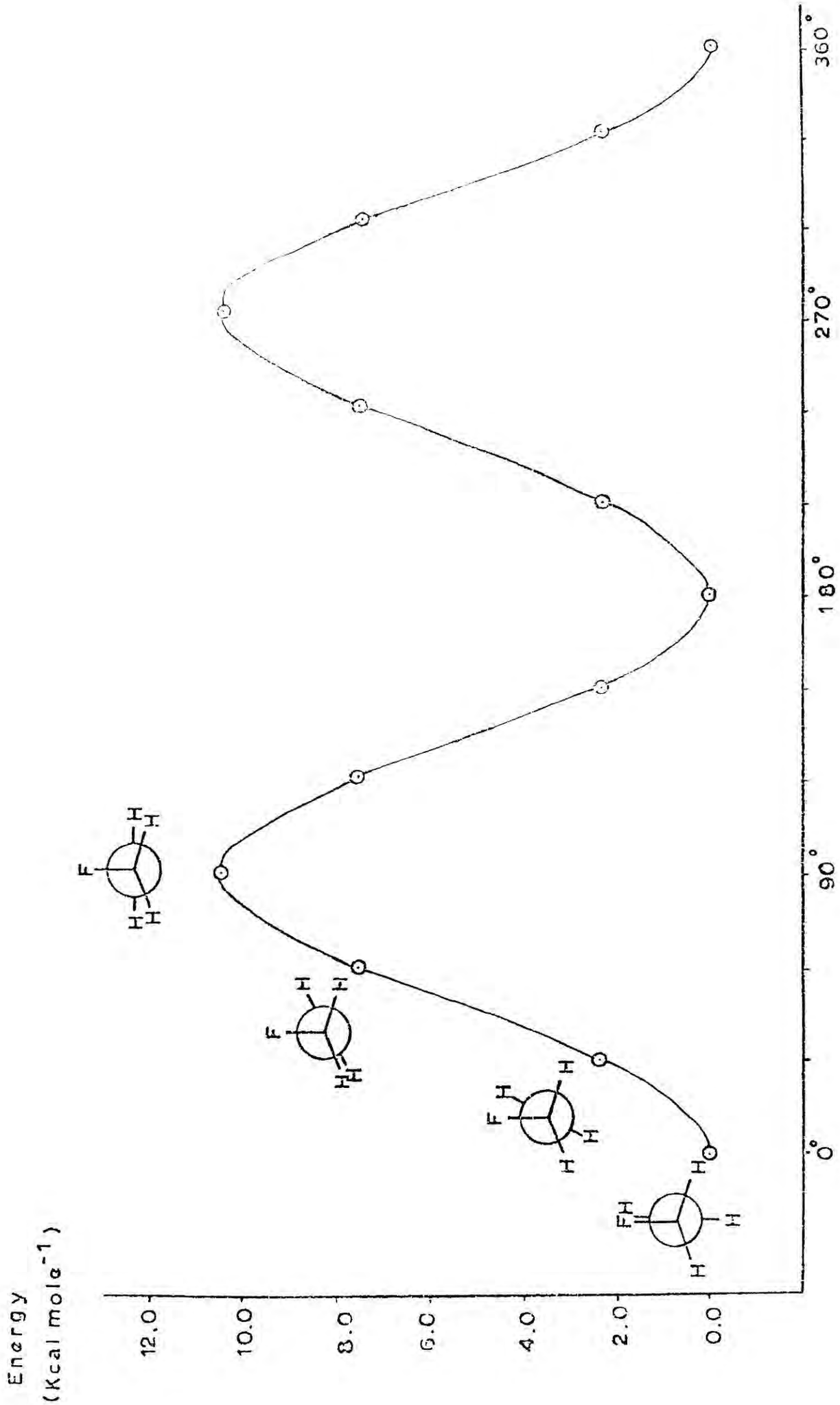


Fig 2.8 Conformational behaviour of the 2- fluoroethyl cation

In this latter respect the system shows similarity to the behaviour of the 1-fluoroethyl cation. INDO and CNDO calculations once again predict the same ordering of stabilities of the conformers as the non-empirical calculations and a very large rotational barrier is predicted (26.1 and 23.2 kcal $mole^{-1}$ respectively). A more recent calculation by Radom and co-workers[114] with an STO-3G basis set, has produced a rotational barrier of 9.31 kcal $mole^{-1}$, which is in good agreement with that produced by the abinitio calculations here, with slightly different geometry and basis set.

Pictorially the stability of the eclipsed conformations of both the 1- and 2- fluoroethyl cations might perhaps be attributed to favourable interactions between the eclipsed H and F atoms. However a Mulliken population analysis shows little evidence for this.

Ion	H-F overlap population
1- fluoroethyl	-0.00127
2- fluoroethyl	-0.00062

The bond overlap populations between the eclipsing atoms is small and negative for both ions. Clearly the high rotational barrier of the 1- fluorethyl cation will have important ramifications for any reaction of this ion.

2.9 Thermochemical Stability of 1- and 2- Fluoroethyl Cations

The 1 fluoroethyl cation is calculated to be 39.36 kcal $mole^{-1}$ more stable than the 2-fluoroethyl cation, which can be attributed to the substantial +M effect [115] of fluorine attached to an electron defficient centre. This can be seen on examination of population analyses - table 2.12

Table 2.12 Population analysis results for fluoroethane and the 1- and 2- fluoroethyl cations.

	gross atomic charge on F
Fluoroethane	-0.28
1 fluoroethyl cation	-0.10
2 fluoroethyl cation	-0.22

A further measure of the electronic effect of fluorine may also be obtained by comparing the relative energies for the processes

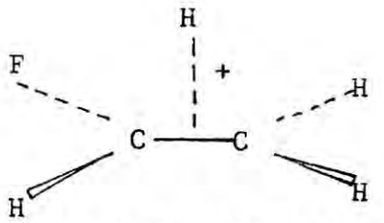
$RH \rightarrow R^+ + H^-$

		ΔErel(kcal mole^{-1})
$CH_3 - CH_3$	$\longrightarrow$ $CH_3 - CH_2^+ + H^-$	0.0
$CH_3 - CH_2F$	$\nearrow$ $CH_3 - \overset{+}{C}HF + H^-$	-29.5
	$\searrow$ $\overset{+}{C}H_2 - CH_2F + H^-$	+9.9

The stabilising and destabilising influence of substituting fluorine for hydrogen at the 1 and 2 positions respectively in ethyl cation is clearly shown. From mass spectrometric appearence potentials it has been shown that for the methyl cation, replacement of H by F stabilises the resulting ion by 27 ± 3 kcal mole^{-1}, which is in good agreement with that calculated here for the 1-fluoroethyl cation. Carbonium ion stabilisation by fluorine has also been studied recently be means of ion cyclotron resonance spectroscopy. [117] Pople [114] has studied the effects of substituents in a β position to the electron defficient centre of the ethyl cation in terms of the population of the formally vacant 2p orbital at the carbonium electron defficient centre. He has thus shown that a high/low 2p (C^+) population will lead to a preferentially staggered/eclipsed conformation and that the 2p(C^+) population in the staggered form may be treated as a measure of the relative hyperconfugative abilities of the C - X bond.

2.10 Interconversion of classical ions via bridge protonated fluoroethylene

Electrophilic addition of a proton to fluoroethylene can also produce the bridge protonated species.



It is of interest to examine the reaction co-ordinate relating 1-and 2-fluoroethyl cations through bridge protonated fluoroethylene. The geometry of the bridged ion was taken to be that previously optimised for the corresponding unsubstituted bridged ion with a proton replaced by fluorine at a C - F bond length of 1.33Å. Starting from the energetically preferred eclipsed conformation of 2-fluoroethyl cation, the transformation to the bridged ion involves moving H(3) [H(4)] in a straight line from its initial position to its final position in bridge protonated fluoroethylene, whilst rotating H(4) [H(3)] into the the F - C(2) - C(1) plane. This, in fact corresponds to a continuous change in the C(2) - H(3) [H(4)] bond length. The C - C bond length and C(2) - H(4) [H(3)] bond angles were assumed to change continuously throughout the transformation. This transformation may be represented diagramatically as in fig 2.9.

The results are shown in fig 2.10.

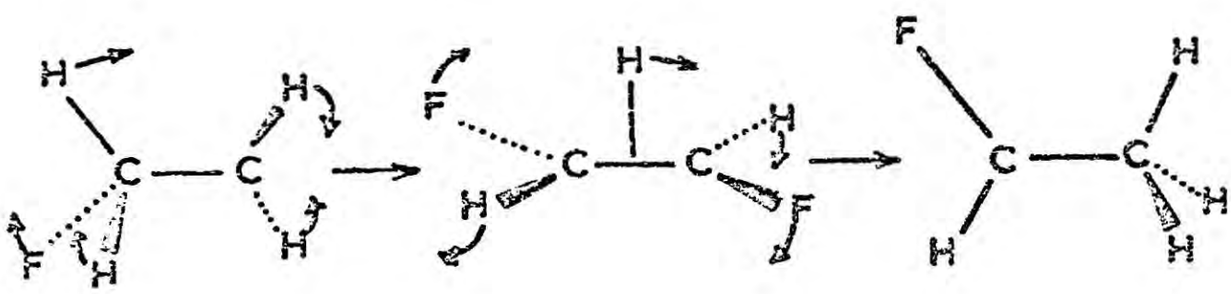


Fig 2.9 Diagram of the atomic movements involved in the transformation of 2- fluoroethyl cation to 1- fluoroethyl cation via the bridge-protonated form.

The interesting feature evident from fig 2.10 is that it is possible to transform 2-fluoroethyl cation into 1-fluoroethyl cation without activation barrier, through bridge protonated fluoroethylene. The bridged ion is calculated to be 12.10 kcal $mole^{-1}$ more stable than 2-fluoroethyl cation but since no activation energy is required in transformation to 1-fluoroethyl cation a discrete existance of this ion is not predicted. Thus, the overall electrophilic addition of HX to the olefin should yield CH_3CHFX, as is observed experimentally. The activation energy for the transformation of the 1- into the 2- fluoroethyl cation will be large, and the transition state must approximate very closely to the 2-fluoroethyl cation.

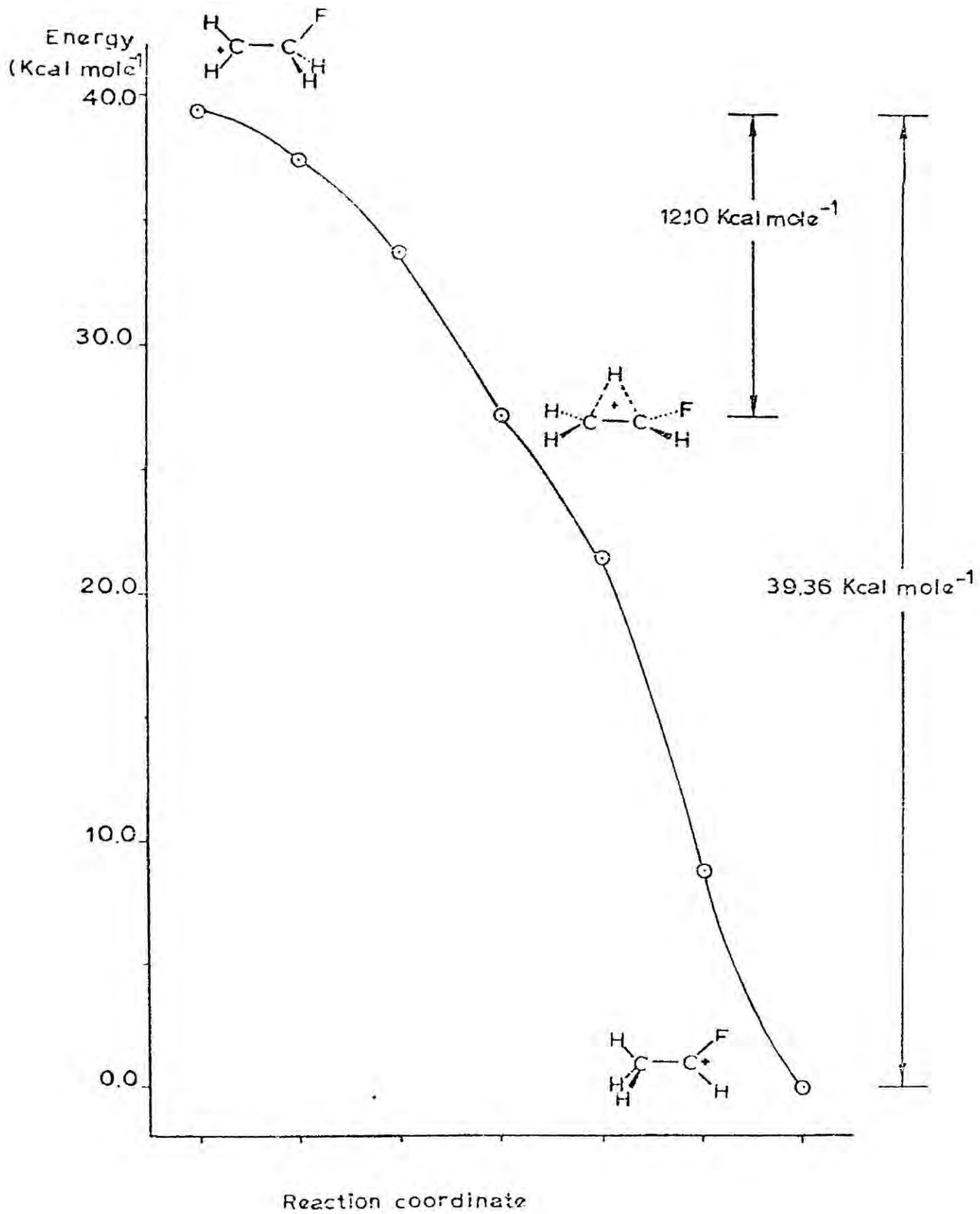


Fig 2.10 Reaction coordinate for the transformation of 2- fluoroethyl cation to 1- fluoroethyl cation via bridge-protonated fluoroethylene

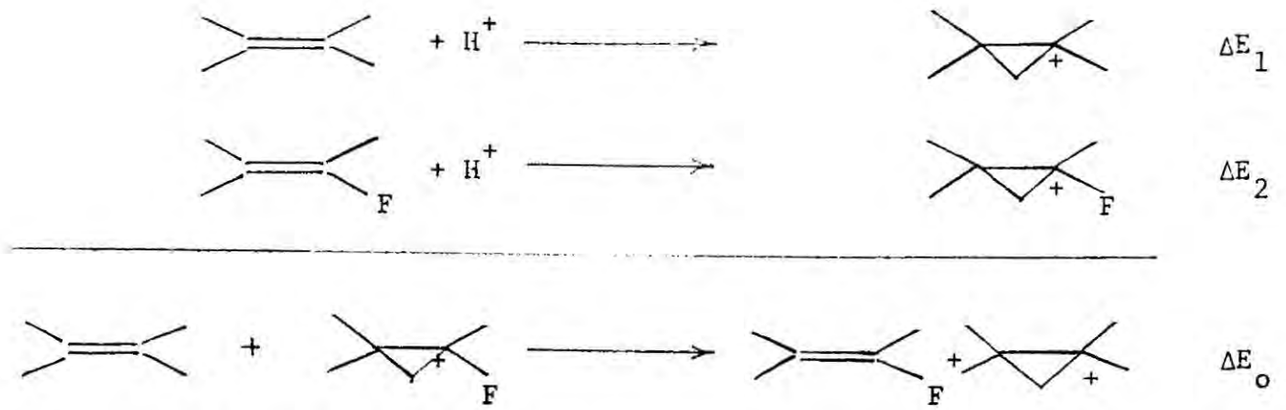
Calculations have been made of approximate solvation energies for these species with the results shown in table 2.13. As expected the delocalisation of the positive charge by fluorine in the 1- fluoroethyl cation considerably lowers the solvation energy with respect to the 2- fluoroethyl cation.

Table 2.13 Calculated solvation energies for the $C_2H_4F^+$ system.

	Esolv (kcal $mole^{-1}$)
1-fluoroethyl	-112.4
2-fluoroethyl	-126.6
protᵈfluoroethylene	-127.0

On this basis it may be seen that the energy separation between the 1-and 2-fluoroethyl cations, and that between the bridge protonated fluoroethylene and the 1-fluoroethyl cation should both be reduced, whilst that between the 2-fluoroethyl cation and the bridge protonated fluoroethylene molecule should be left virtually unchanged. Hence the net result in solution should be essentially the same but with a slightly reduced activation barrier. Thus, starting from $CH_2\overset{+}{C}D_2$ only $CHD_2\overset{+}{C}HF$ should be produced with no scrambling at C(2). This is in contrast to the ethyl cation discussed previously. Furthermore, the bridged ion is no longer the transition state as for the transformation suggested in the case of the ethyl cation rearrangement.

Some measure of the electronic effect of replacing hydrogen by fluorine in the bridged ion can be obtained by comparing the relative energies for formation of the bridged ions from the corresponding ethylenes, i.e. by a consideration of the energy change for the isodesmic processes.



Taking protonated ethylene as standard the relative energy for the production of protonated fluoroethylene is calculated to be +7.0 kcal $mole^{-1}$. Thus production of the bridged ions from the corresponding ethylenes is inhibited by replacing hydrogen by fluorine. Conversely, starting from the corresponding ethanes it is relatively (~7.5 kcals $mole^{-1}$) easier to produce the bridge protonated fluoroethylene.

In figure 2.11 are shown the bond overlap populations for the bridged ions. The interesting point is that whilst the bridging hydrogen is symmetrically placed along the CC axis, the CH bond overlap populations are unequal. This inequality is biased in favour of the formation of 1-fluoroethyl cation, which gives a simple pictorial interpretation of the driving force for the transformation. Closely related to bond overlap populations obtained in non-empirical treatments are partitioned bond overlap populations defined by equation (2.5) for CNDO SCF MO treatments.[118]

$$T_{\mu\nu} = 2 \sum P_{\mu\nu} S_{\mu\nu} \qquad (2.5)$$

It may be seen that essentially the same result may thus be derived from a CNDO treatment since the partitioned bond overlap populations are

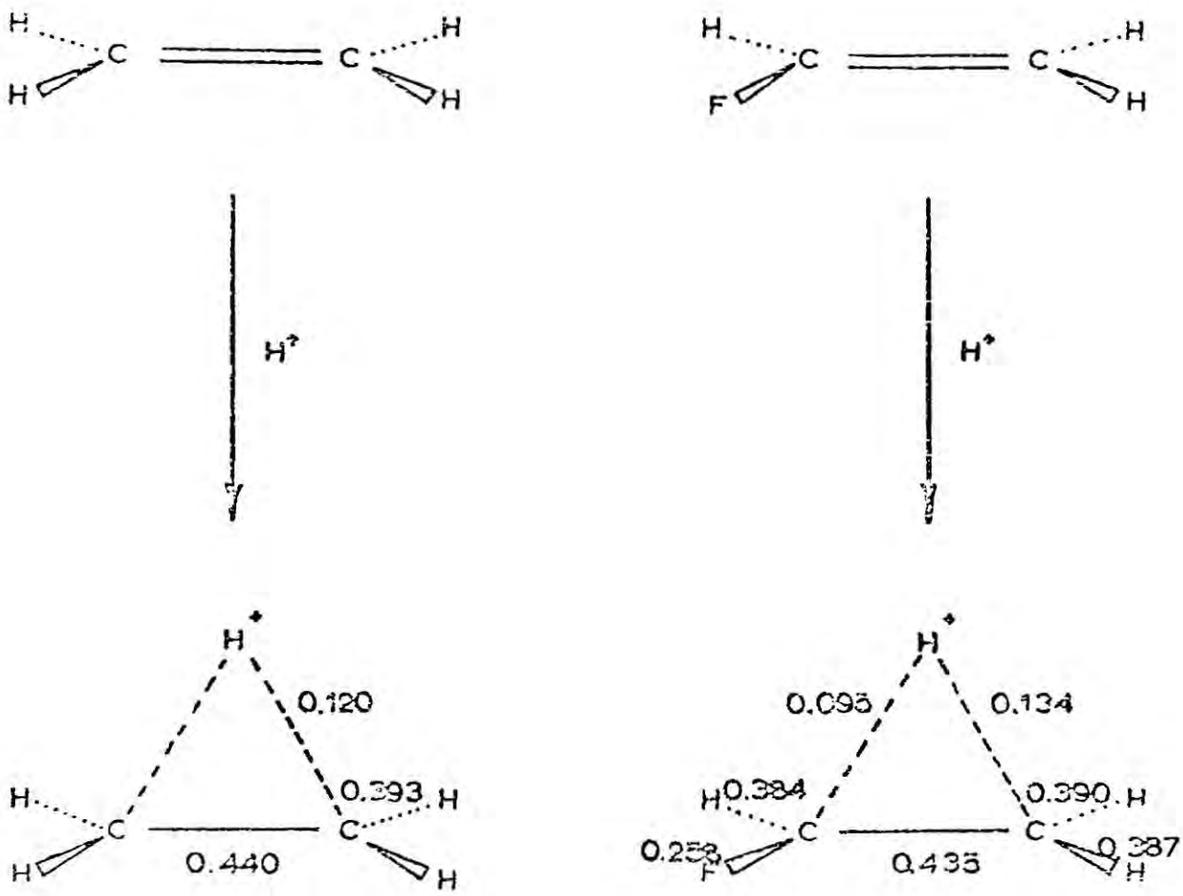


Fig 2.11 Bond overlap populations for bridge-protonated ethylene and bridge-protonated fluoroethylene

2.11 The Fluoronium Ion

Electrophilic addition of the hypothetical F^+ species, or a suitable precursor, could result in an initial production of the classical 2-fluoroethyl cation, or a bridged fluoronium ion. A cursory examination of available experimental data provides, at most, very slight evidence for the latter possibility, but theoretically the fluoronium ion represents a most interesting structure, worthy of investigation. The geometry of the carbon hydrogen framework was

assumed to remain essentially the same as that in bridge protonated ethylene, and, using the C,F(7,3/3,1)H(3/1) basis set, the energy of the species was studied as a function of the distance of the bridging fluorine from the centre of the CC bond, see fig 2.12.

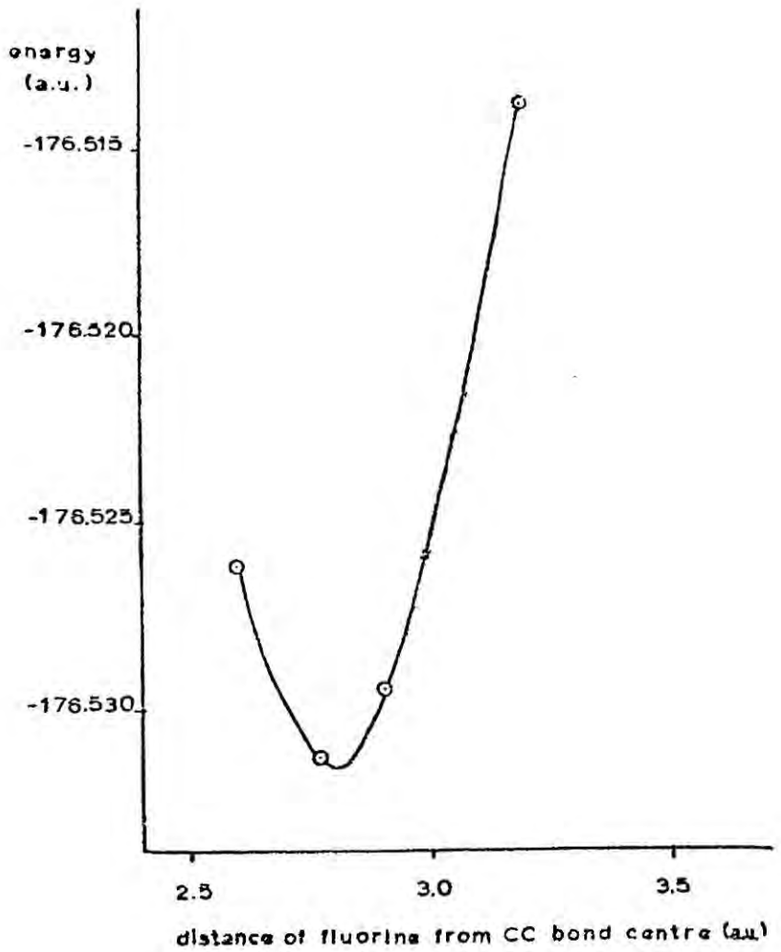


Fig 2.12 Total energy of fluoronium ion as a function of the distance of the F atom from the centre of the CC bond

In this way a partially optimised geometry was obtained in which the fluorine atom was at 1.466Å from the CC bond centre. The total energy of the species suggests that it would be ~ 3.6kcal $mole^{-1}$ more stable than the 2-fluoroethyl cation. Furthermore it is most likely that more complete geometry optimisations for both ions together with the

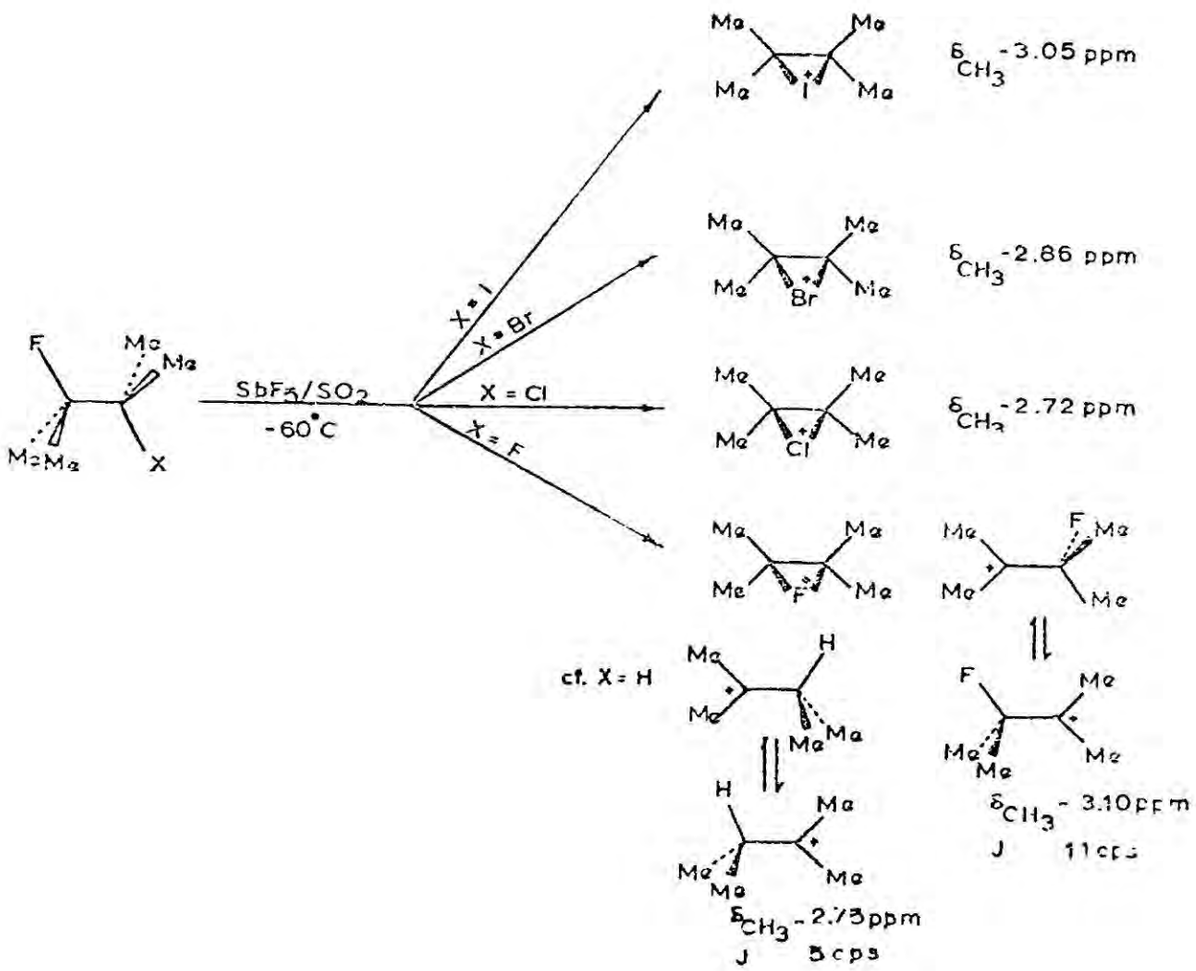


Fig 2.13 Summary of the work of Olah 78, 119 concerning the ionisation of 2-halo-3-fluoro-2,3-dimethyl butanes in SO_2/SbF_5

inclusion of polarisation functions in the basis set would increase this gap, thus making 3.6 kcal mole^{-1} a lower limit for the relative stability of the bridged ion. These points will be discussed in greater detail later. The calculated relative energies of the various isomeric $(C_2H_4F)^+$ species are sumarised in table 2.14.

Species	Relative Energy (kcal mole^{-1})
1 fluoroethyl *	0.0
2 fluoroethyl *	39.36
Fluoronium Ion	35.78
Prot^d fluoroethylene	27.26

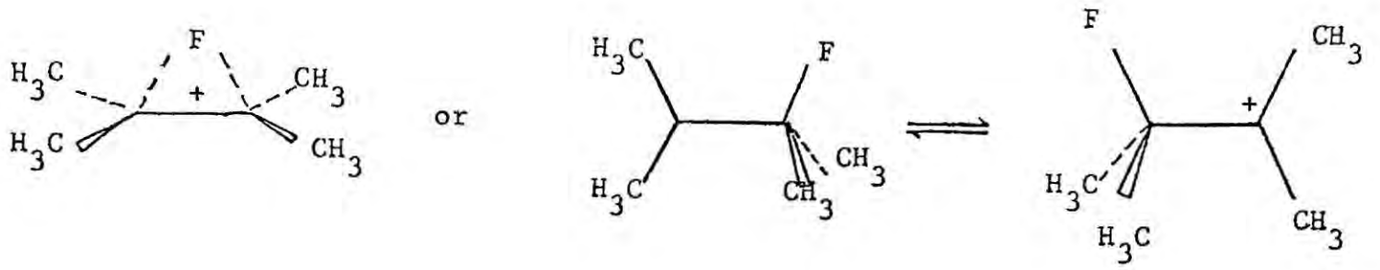
Table 2.14 Relative energies of isomeric $C_2H_4F^+$ species

In view of the indicated stability of the fluoronium ion it is of interest to investigate in some detail the experimental data relecant to this point. Figure 2.13 summarises some of Olah's [78], [119] work on the ionisation of 2-halo 3-fluoro 2,3-dimethyl butanes in SO_2SbF_5 The existance of bridged iodonium, bromonium and chloronium ions are strongly supported by the nmr data. However in the case of fluorine, Olah rationalises his data not in terms of a fluoronium ion, but of a rapidly equilibrating pair of classical ions with an intramolecular 1, 2 fluorine shift. Were this interpretation of the nmr data the only one possible it would be very difficult to envisage this transfer occuring without involving an intermediate species very akin to the fluoronium ion. Thus it might be argued that the classical and bridged ions are close in energy - as results here would indicate - and that this is analogous to the case of bridge protonated ethylene and ethyl cation. Olah's analysis of the chemical shift and coupling constant data are given in fig 2.14 but these are by no means incompatible with the

* most stable conformers

postulate of a bridged fluoronium ion and this alternative interpretation of the data is included in fig 2.14

Figure 2.14



No change to $-90^{o}C$

δ_{CH_3} = −3.10ppm J = 11cps

CCl_3 δ_F = +150ppm

$(Me)_2$ C(F)–C(F) $(Me)_2$

CCl_3F δ_F = +181ppm

$$\text{predicted } J = \left[J_{\substack{C-C \\ F\ H}} + J_{\substack{C-C-C \\ F\quad H}} \right]$$

$$\simeq \tfrac{1}{2}\left[20 + 2\right] \simeq 11\text{cps}$$

$$\beta_{CH_3} = \tfrac{1}{2}\left[\delta_{C-CH_3} + \delta_{\substack{C-CH_3 \\ F}} \right]$$

$$\simeq \tfrac{1}{2}\left[-4.4 + -1.7\right]$$

$$= 3.1\text{ppm}$$

A prime difficulty implicit in Olah's intepretation of the data is the existance of other possible competative rearrangements. For example, 1, 2 Methyl shifts, which are known to be extremely facile, could lead to formation of the ion $C(Me)_3C^{+}M_eF$, which is clearly inconsistent with the nmr data. However, independent evidence would suggest that if classical ions are involved then this should be the thermodynamically most stable species. For example fig 2.15 shows some experimental data concerning the ionisation of 1- fluoro- 2- methyl- 2- proponol in magicacid, together with a postulated mechanism to explain the formation of the product ion in which a fluorine atom and a methyl group stabilise the electron deficient centre. This ion II is thermodynamically more stable

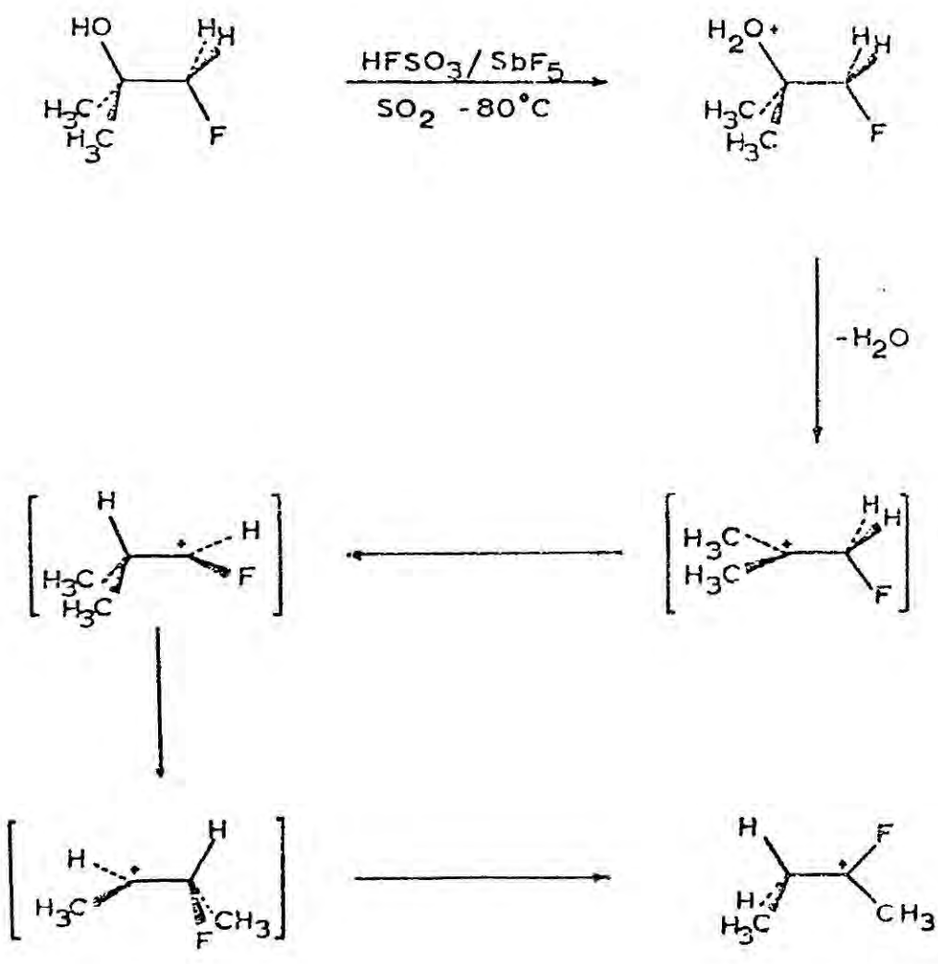


Fig 2.15 Proposed mechanism for the rearrangement of the carbonium ion initially produced upon ionisation of 1-fluoro-2-methyl-2-propanol in magic acid

than the initially formed ion (I) with two methyl groups attatched to the electron deficient centre and the fluorine atom attached to the other carbon exerting a destabilising influence as indicated by the relative stabilities of the 1- and 2- fluoroethyl cations. Thus even if it assumed that the stabilising effect of a methyl group at an electron deficient centre is the same as for a fluorine atom, the energetic preference for II over I, which a priori might be considered as the two most stable carbonium ions in this reaction sequence, would still be $\sim$10 kcal mole^{-1}.

Clearly, resolution of the problem would be aided by calculations along an idealised reaction coordinate relating the classical ions through the bridged fluoronium ion. In an idealised reaction coordinate this corresponds to sliding the fluoronium atom in 2-fluoroethyl cation along a straight line along the CC bond to its final position in the fluoronium ion, which involves a continuous variation of CF bond length. Simultaneously the hydrogen atoms attatched to C(2) are moved towards the plane perpendicular to the F - C(2) - C(1) plane, and the hydrogens attatched to C(1) are rotated into this plane. The basis set used was the previously employed C,F(7,3/3,1) H(3/1). The results are shown diagramatically in fig 2.16.

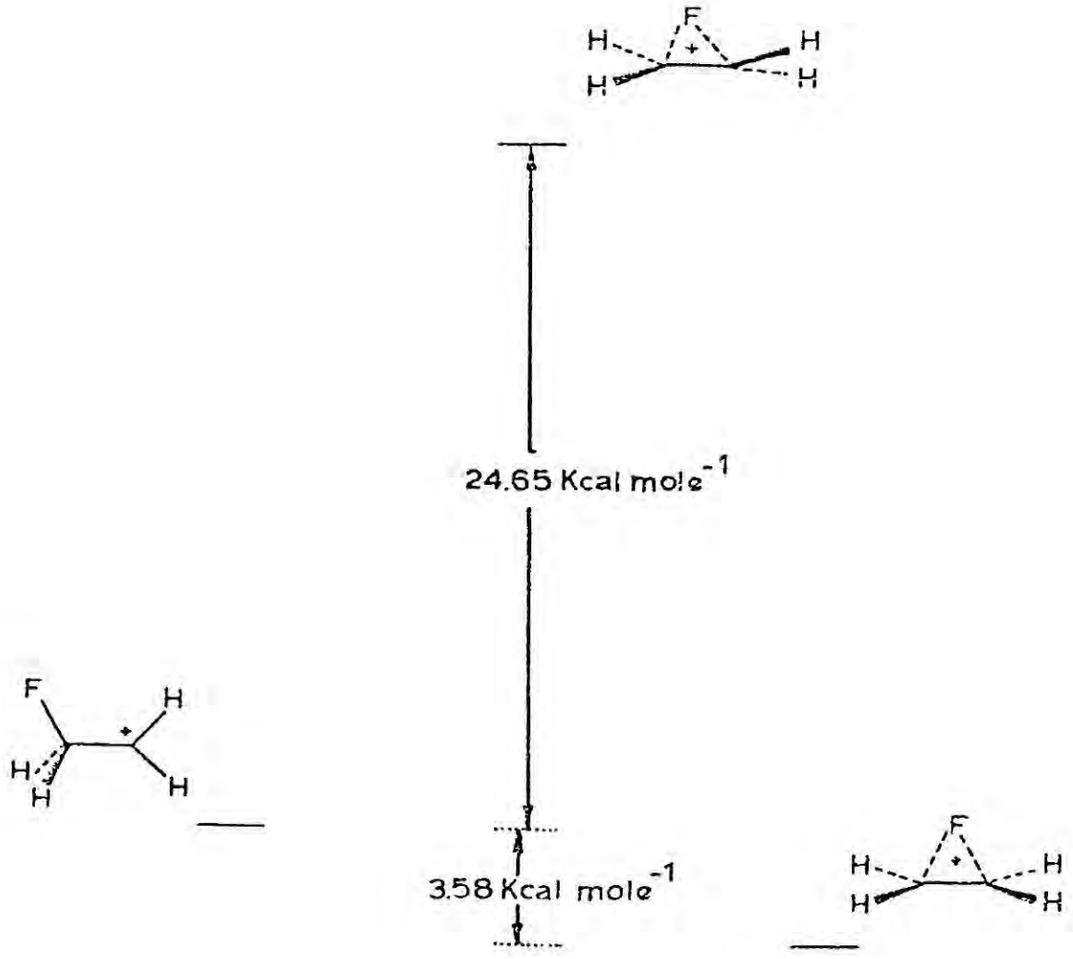


Fig 2.16 Diagram showing the considerable rise in total energy in the transformation between 2-fluoroethyl cation and the fluoronium ion.

Though these results are probable not accurate quantitavely (the energies are quoted to two decimal places, representing the limit of computational accuracy rather than indicating any proposed limits of accuracy imposed by choice of geometries and basis set), the qualitative picture they present should be reliable. Clearly the important result is that a substantial activation barrier is involved in transferring the fluorine atom in an intramolecular manner.

As postulated in section 2.5 the small energy difference between the classical and the bridged species may mean that correlation energy differences between the 2-fluoroethyl cation and the fluoronium ion will assume importance. Hence values of E_{corr}^{intra} have been calculated

from the atomic pair correlation energies of Snyder and Basch (table 2.7a) and these are shown in table 2.15

Table 2.15 Estimates of ΔE_{corr}^{intra} for bridged and classical ions (in a.u.)

	Atom	Fluoronium ion	2-Fluoroethyl cation
F	1s - 1s	-0.0398	-0.0398
	1s - 2s	-0.0055	-0.0055
	1s - 2p	-0.0164	-0.0165
	2s - 2s	-0.0117	-0.0115
	2s - 2p	-0.0851	-0.0839
	2p - 2p	-0.0663	-0.0666
	2p - 2p'	-0.1083	-0.1092
C	1s - 1s	-0.0818	-0.0818
	1s - 2s	-0.0079	-0.0078
	1s - 2p	-0.0172	-0.0171
	2s - 2s	-0.0375	-0.0368
	2s - 2p	-0.1055	-0.1029
	2p - 2p	-0.0740	-0.0735
	2p - 2p'	-0.0688	-0.0688
H	1s - 1s	-0.0512	-0.0533
	TOTAL	-0.7770	-0.7750

ΔE = 0.002au = 1.26 kcal $mole^{-1}$ with respect to the classical ion.

It is evident from these data that correlation energy corrections may well be of similar magnitude to the SCF calculated energy difference between the fluoronium ion and the classical 2-fluoroethyl cation.

These calculations, taken in conjunction with all the available experimental data, suggest that the only consistent interpretation is that the fluoronium ion represents a local minimum on the $C_2H_4F^+$ potential energy surface. The weight of evidence presented here then suggest that the ionisation of 2, 3-difluoro -2,3-dimethyl butanes in SbF_5/SO_2 as studied by Olah produces a bridged fluoronium ion rather than a rapidly

equilibrating pair of classical 2-fluoroethyl cations.

2.12 Potential Energy Surface for $(C_2H_4Cl)^+$ system

So far some aspects of the potential energy surface for the $(C_2H_4X)^+$ system have been examined for X = H and F. Some very distinct differences have emerged between these two systems and it is possible to draw several conclusions concerning substituent effects, particularly with regard to the stabilisation of electron defficient centres in carbonium ions. Hence it is of considerable interest to pursue further the investigation of such effects and thus to examine the replacement by chlorine of a hydrogen atom in the prototype system. In this way it should be possible to define the substituent effects with greater precision and recognise some emerging trends.

2.13 Conformational Processes in Chloroethane and 1- and 2-Chloroethyl Cations.

Chloroethane

The first molecular system and process to be investigated in this section was the conformational processes in chloroethane, calculations which also served to test the adequacy of the basis set used for chlorine. The basis set consisted [121] of 75 primitive gaussian functions, optimised 10s 6p for chlorine, 7s 3p for carbon and 3s for hydrogen, which were reduced to 27 contracted functions corresponding to contracted sets of 4s 2p on chlorine, 3s 1p on each carbon and 1s on each hydrogen. Preliminary calculations carried out on the 2-chloroethyl cation with an STO - 3G basis set and including d orbitals on chlorine indicated that d orbital participation is negligible as far as the classical ions are concerned, and hence this almost certainly applies also to chloroethane.

The geometry used was the same carbon hydrogen framework optimised for ethane, taking a CCl distance of 1.77Å,[111] and bond angles were fixed at 109.5°. The calculations were performed on the staggered and eclipsed conformations of chloroethane, and the barrier to rotation obtained is shown in table 2.16, where comparison has also been made with the results of the calculations upon fluoroethane and ethane which used comparable basis sets.

Table 2.16 Barriers to Internal Rotation in Ethane and haloethanes.(kcal $mole^{-1}$)

	Calculated	Experimental
Ethane	2.72	2.93 [122]
Fluoroethane	2.60	3.33 [123]
Chloroethane	3.33	3.69 [124]

The barriers are in reasonable agreement with the experimental values, and indicate that the basis set for chlorine is adequate for discussion of conformational processes. The component energy terms are shown in table 2.17.

Table 2.17 Rotational components for Chloroethane

	Etot(au)	(Vne + T)(au)	Vee(au)	Vnn(au)
stag	-537.475518	-944.019214	303.498782	103.04491
eclip	-537.470215	-944.153704	303.576414	103.10707

$|\Delta Vatt| = 0.072au$

$|\Delta Vrep| = 0.078au$

$\Delta Etot = 3.33kcal\ mole^{-1}$

The attractive and repulsive terms have also been plotted for the three compounds, fig 2.17. Clearly although the rotational barriers are similar in the series, the absolute magnitudes of the component attractive

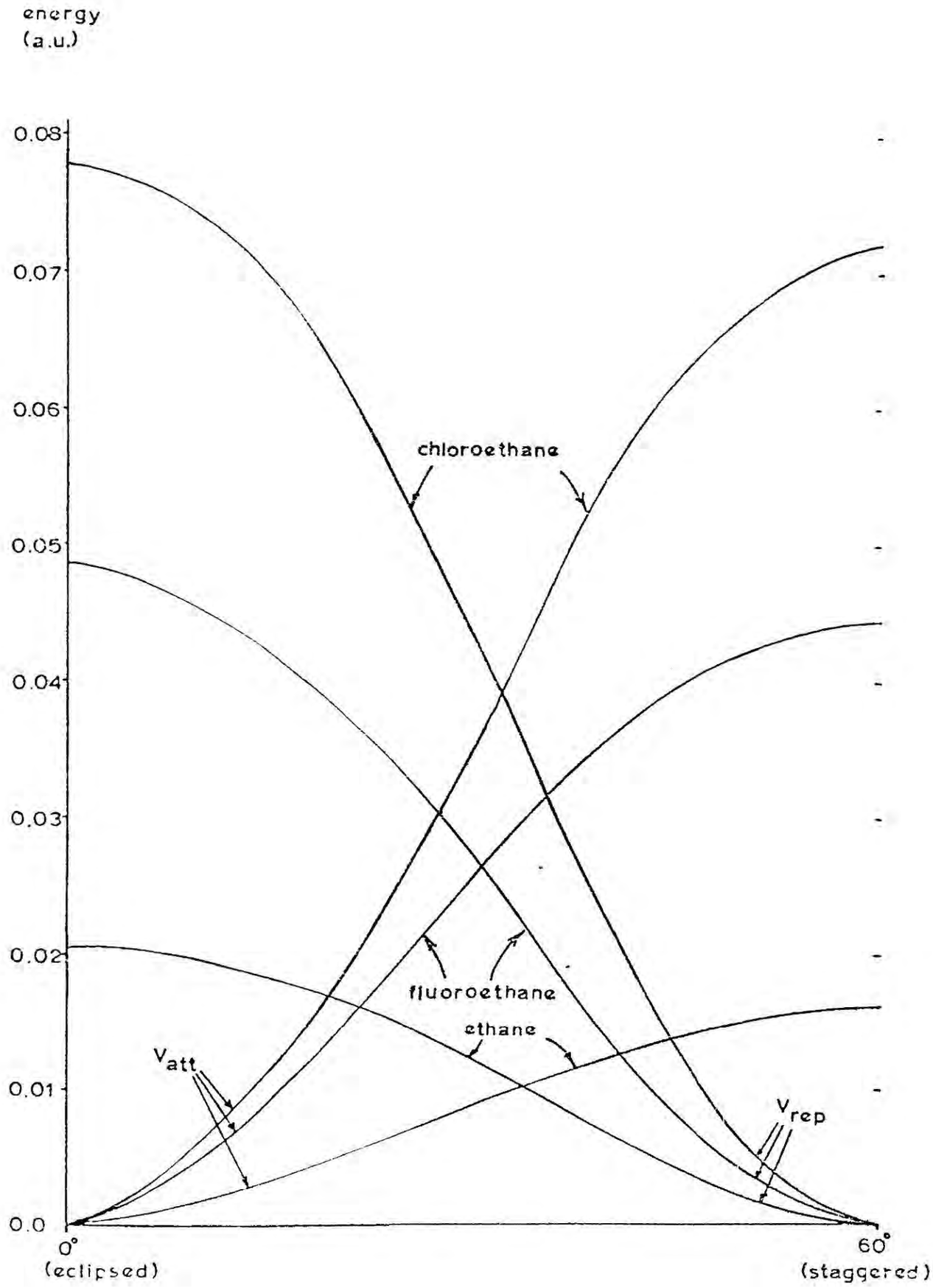


Fig 2.17 Variation with conformation of attractive and repulsive components of energy barriers of ethane, fluoroethane and chloroethane.

and repulsive energy terms increase progressively in going from ethane to fluoroethane to chloroethane, which may be compared with the situation in the fluoro and chloro substituted cations discussed shortly.

The geometries for the chloroethyl cations were derived from that optimised for ethyl cation, with a hydrogen at C_1 or C_2 being replaced by a chlorine atom at a standard CCl distance of 1.77Å, and bond angles once again being assumed rigid at 109.5° and 120° for sp^3 and sp^2 environments respectively. The basis set was the same as the above for chloroethane i.e. chlorine (10s 6p/ 4s 2p), carbon (7s 3p/3s 1p) and hydrogen (3s / 1s) i.e. a total of 72 primitive gaussian functions contracted to 26 functions.

1-Chloroethyl cation

The results for the 1-chloroethyl cation, together with those for the 1-fluoroethyl cation for comparison, are shown in fig 2.18. There are two main points of interest. Firstly, the small rotational barrier (0.53 kcal $mole^{-1}$) which is very comparable to the 1-chloroethyl cation and secondly that the most stable conformer has chlorine eclipsing hydrogen and the least stable conformer is staggered, which is the exact parallel of the case for the 1-chloroethyl cation. The rotational barrier has been analysed in terms of component attractive and repulsive terms as before and the results are shown in table 2.18.

Table 2.18 Rotational components for the 1-chloroethyl cation

Conformation	Etot(au)	(Vne + T) (au)	Vee(au)	Vnn (au)
H eclip Cl	-536.519439	-911.061590	282.408806	92.133345
H eclip H	-536.518589	-911.023762	282.400473	92.104701

$|\Delta Vatt| = 0.00918au$

$|\Delta Vrep| = 0.00833au$

$\Delta Etot = 0.53kcal\ mole^{-1}$

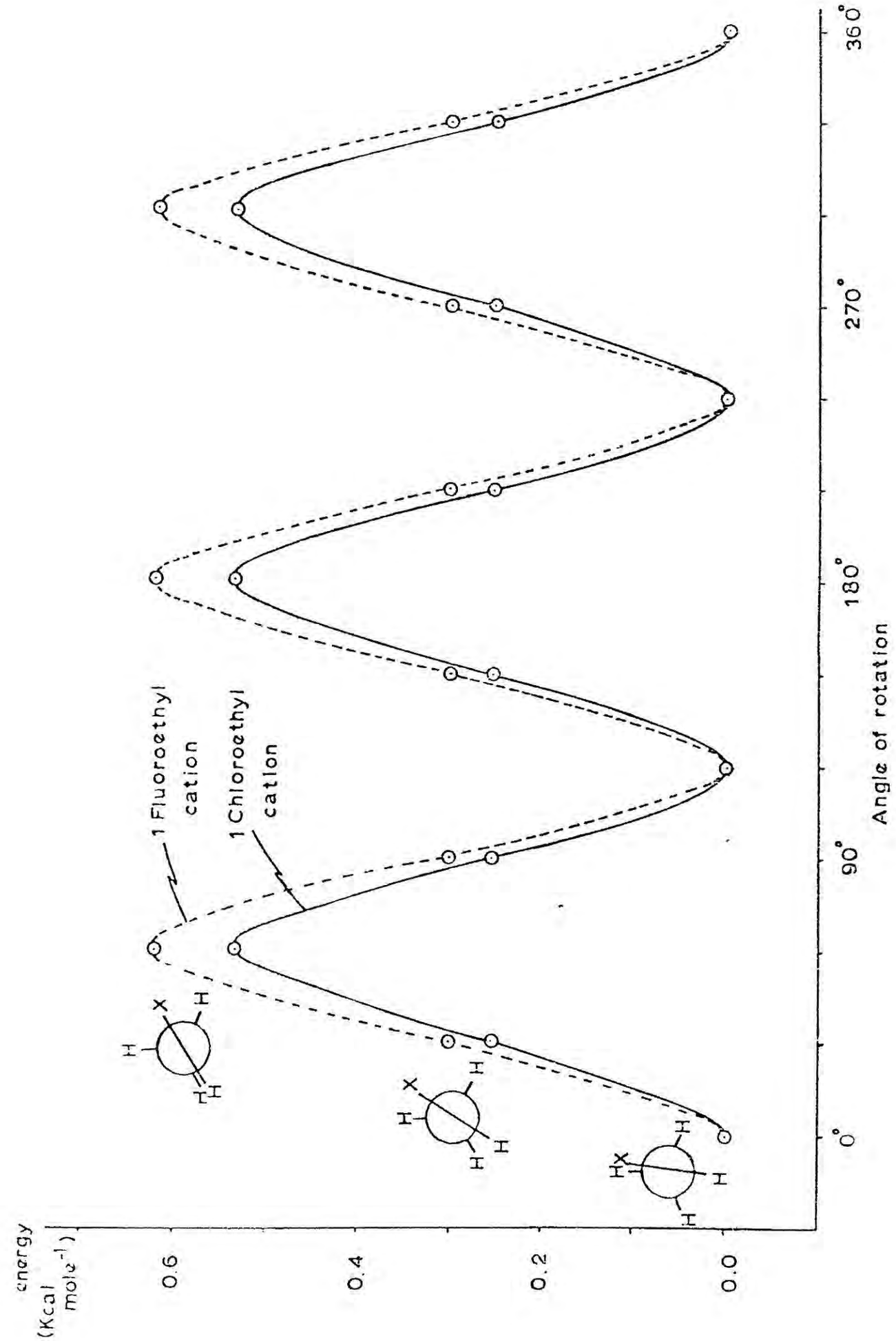


Fig 2.18 Conformational behaviour of the 1- chloroethyl cation

Clearly the barrier is attractively dominated, as was the case for the 1-and 2-fluoroethyl cations. Examination of table 2.4 and 2.10 show that as in the case of the substituted ethanes the magnitude of the energy components increases in going from the ethyl cation to the 1-fluoroethyl cation to the 1-chloroethyl cation. CNDO calculations have been performed upon the 1-chloroethyl cation and the same ordering of stabilities of conformers is predicted as by non empirical calculation. The rotational barrier produced, 0.31kcal $mole^{-1}$, is also in good agreement with the ab initio result.

2-Chloroethyl Cation

The results for the 2-chloroethyl cation are shown diagramatically in fig 2.19. The rotational barrier is seen to be slightly larger (0.91 kcal $mole^{-1}$) than in the 1-chloroethyl cation, and now the most stable conformer is staggered. These results contrast quite strikingly with those obtained for the 2-fluoroethyl cation, where the calculated barrier is much larger and the most stable conformer has fluorine eclipsing hydrogen. These results suggest that the dominating influence upon the rotational barrier is different from that in the

Table 2.19 Rotational components for the 2-chloroethyl cation

Conformation	Etot(au)	(Vne + T) (au)	Vee(au)	Vnn(au)
HH // HH	-536.473015	- 916.756102	285.906789	94.776299
Cl eclip H	-536.471559	-917.552254	286.535764	94.544931

$|\Delta Vatt| = 0.62752$

$|\Delta Vrep| = 0.628975au$

$\Delta Etot = 0.91kcal\ mole^{-1}$

1-chloroethyl cation and the 1- and 2- fluoroethyl cations, and the energy components are shown in table 2.19. By comparison with both the 1 substituted cations, table 2.10 and 2.18, the absolute magnitudes

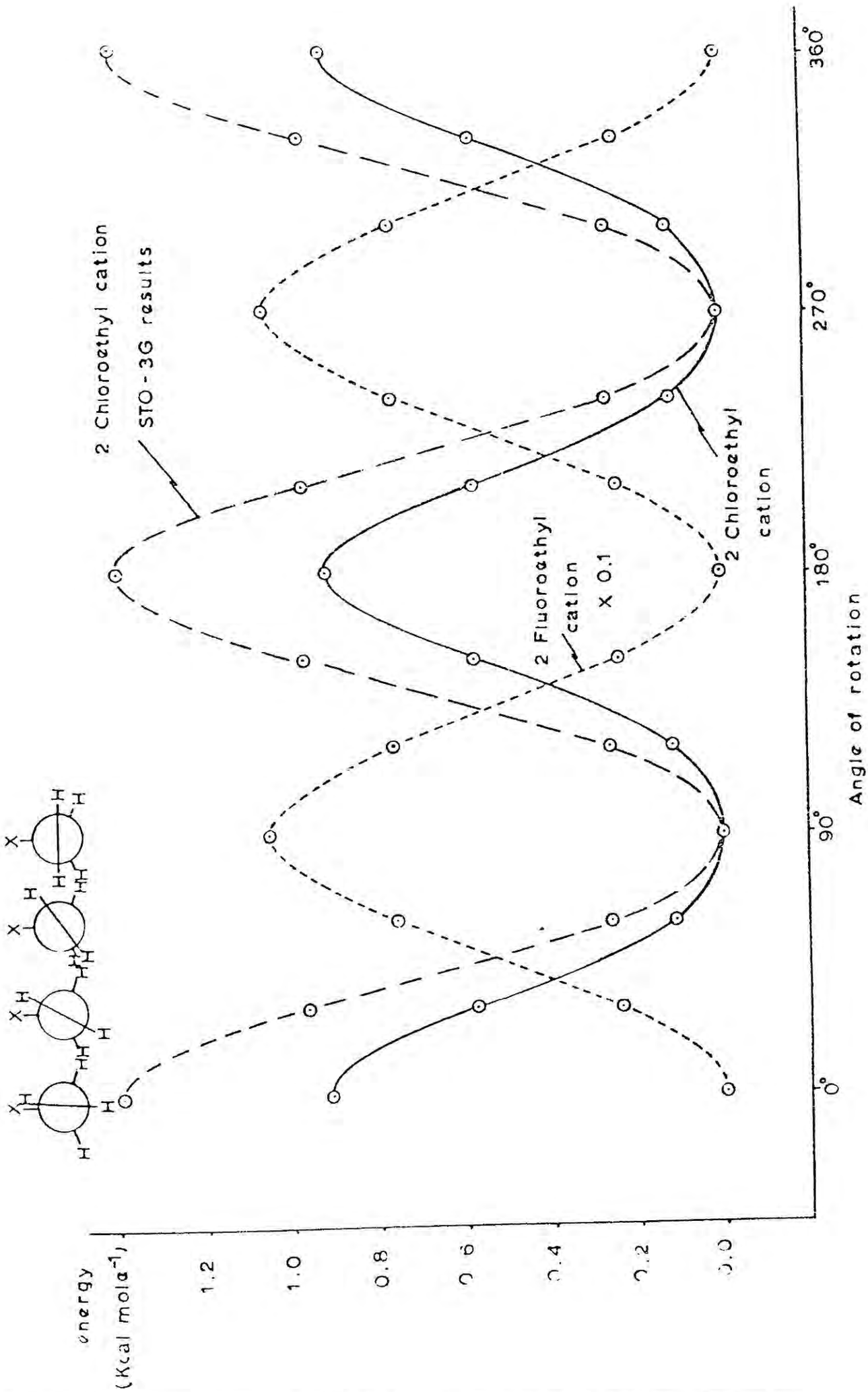


Fig 2.19 Conformational behaviour of the 2- chloroethyl cation, including results obtained with the use of an STO-3G basis set

of both the attractive and repulsive components are much larger for both the 2 substituted ethyl cations, that for chloro again being larger than for the fluoro species. The change in attractive components is going from the staggered (HH parallel to HH) to eclipsed (H eclipsing F or Cl) conformations is opposite in sign for the fluoro as compared to the chloro species, being larger in absolute magnitude for the latter. For the 2-chloroethyl cation, however, the change in repulsive terms is dominant so that overall the barrier to rotation is now much smaller than for the 2-fluoroethyl cation. Rotational barriers in several 2 substituted ethyl cations have now been investigated by Pople and co-workers [125] with the computationally inexpensive STO 3G basis set, and the results are shown in table 2.20.

Table 2.20 Calculated Barriers to Rotation in some 2 substituted ethyl cations [125]

XCH_2CH_2	Rotational Barrier (kcal $mole^{-1}$)	Most stable conformer
H	0.00	-
F	10.53	Eclip (Feclip H)
Cl	1.40	staggered
CH_3	2.52	staggered
CH_2CH_3	3.73	staggered
CH_2F	2.11	staggered
CH_2OH	0.91	staggered
CH_2CN	0.87	staggered

With the exception of the 2-fluoroethyl cation the rotational barrier in simple 2 substituted ethyl cations tend to be fairly small, which accords with the classical view of organic chemists of free rotation in simple carbonium ions. To investigate the comparison between calculations employing different basis sets, the 2- chloroethyl cation was also investigated in an STO 3G basis set including 3d orbitals on chlorine (exponent 1.8). The results are included in fig 2.19 and the component analysis is shown in table 2.21, where good agreement is shown with table 2.19.

Table 2.21 Rotational components (STO 3G) for 2-chloroethyl cation

Conformation	Etot (au)	(Vne + T)(au)	Vee(au)	Vnn(au)
HH // HH	-531.504024	-912.074439	286.194116	94.376266
Cl eclip H	-531.501796	-912.860089	286.813362	94.544931

$|\Delta Vatt| = 0.6170au$

$|\Delta Vrep| = 0.6192au$

$\Delta Etot = 1.398kcal\ mole^{-1}$

Part of the difference in calculated barrier to rotation may arise from the use of unoptimised exponents in the STO 3G basis set. (Exponents were evaluated by application of Burns Rules [126] see Appendix V). For none of the occupied orbitals is there significant d orbital participation on chlorine and for the classical chloroethyl cations the role of 3d orbitals is that of providing polarisation functions for the s, p basis set.

CNDO calculations have been performed on the 2-chloroethyl cation, reproducing the ab initio results of the relative stabilities of the various conformers. The predicted rotational barrier, 5.16 kcal $mole^{-1}$, is rather large in comparison with the ab initio result, but the relative barriers

2- chloroethyl > 1-chloroethyl

is predicted in agreement with the result of the non empirical treatment.

2.14 Relative Energies of 1- and 2- substituted ethyl cations

For both the chloro and fluoro substituted ethyl cations, the 1- substituted carbonium ion is more stable due to the stabilisation of the electron defficient centre by delocalisation from the lone pair on the halogen. The energy differences between the 1- and 2- substituted

species are 39.36 kcal mole^{-1} and 29.23kcal mole^{-1} for the fluoro and chloro species respectively. The energy difference for the fluoro species is somewhat larger than that for the chloro species and this can be ascribed largely to the larger sigma inductive effect of fluorine which tends to destabilise the 2- fluoroethyl cation relative to the 1- fluoroethyl cation more than the relative effect of chlorine in the respective chloro cations. This point will be illustrated further in section 2.16.

2.15 Interconversion of the 1- and 2- Chloroethyl Cations

In the electrophilic addition of a proton to chloroethylene, the production of the bridge protonated species is a theoretical possibility, which will also be a potential intermediate in the inverconversion of the classical 1-and 2-chloroethyl cations. The geometry of the bridged ion was taken to be that previously optimised for the corresponding unsubstituted bridged ion, with a hydrogen atom replaced by chlorine at a CCl bond distance of 1.77Å. Starting from the energetically preffered (HH parallel to HH) conformation of 2-chloroethyl cation the transformation to the bridge protonated ion involves (see fig 2.20) moving H(1) in a straight line from its initial to its final position in bridge protonated chloroethylene, whilst Cl and H(2) are moved into plane and H(3) and H(4) are rotated into plane, both movements involving a continuous change in dihedral angle. The second stage of the transformation involves moving H(1) in a straight line from its initial position in the 1-chloroethyl cation, whilst Cl and H(2) are rotated into the vertical plane and H(3) and H(4) are moved to their final positions, again by means of continuous variation of dihedral angles. In both stages of the transformation, the CC bond length and

all CH bond lengths are varied continuously and variations of bond angles are also continuous. The results are shown graphically in fig 2.21.

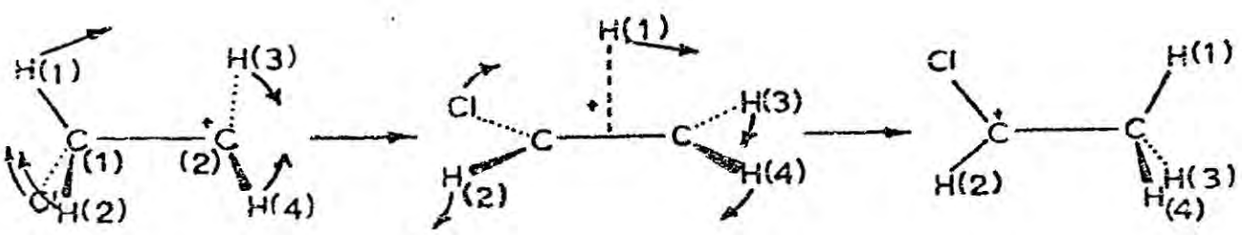


Fig 2.20 Diagram of atomic movements involved in the transformation of 2- chloroethyl cation via bridge-protonated chloroethylene

Several features of interest emerge from this reaction profile and from comparison of this with those derived previously for the hydrogen and fluorine analogues. For the interconversion of the unsubstituted ethyl cations an activation barrier of 5.2 kcal $mole^{-1}$ was found, where the bridge protonated ethylene was the transition state. For the fluoroethyl cations inverconversion no activation barrier was found for the transformation of the 2- to 1- fluoroethyl cations, and it was apparent

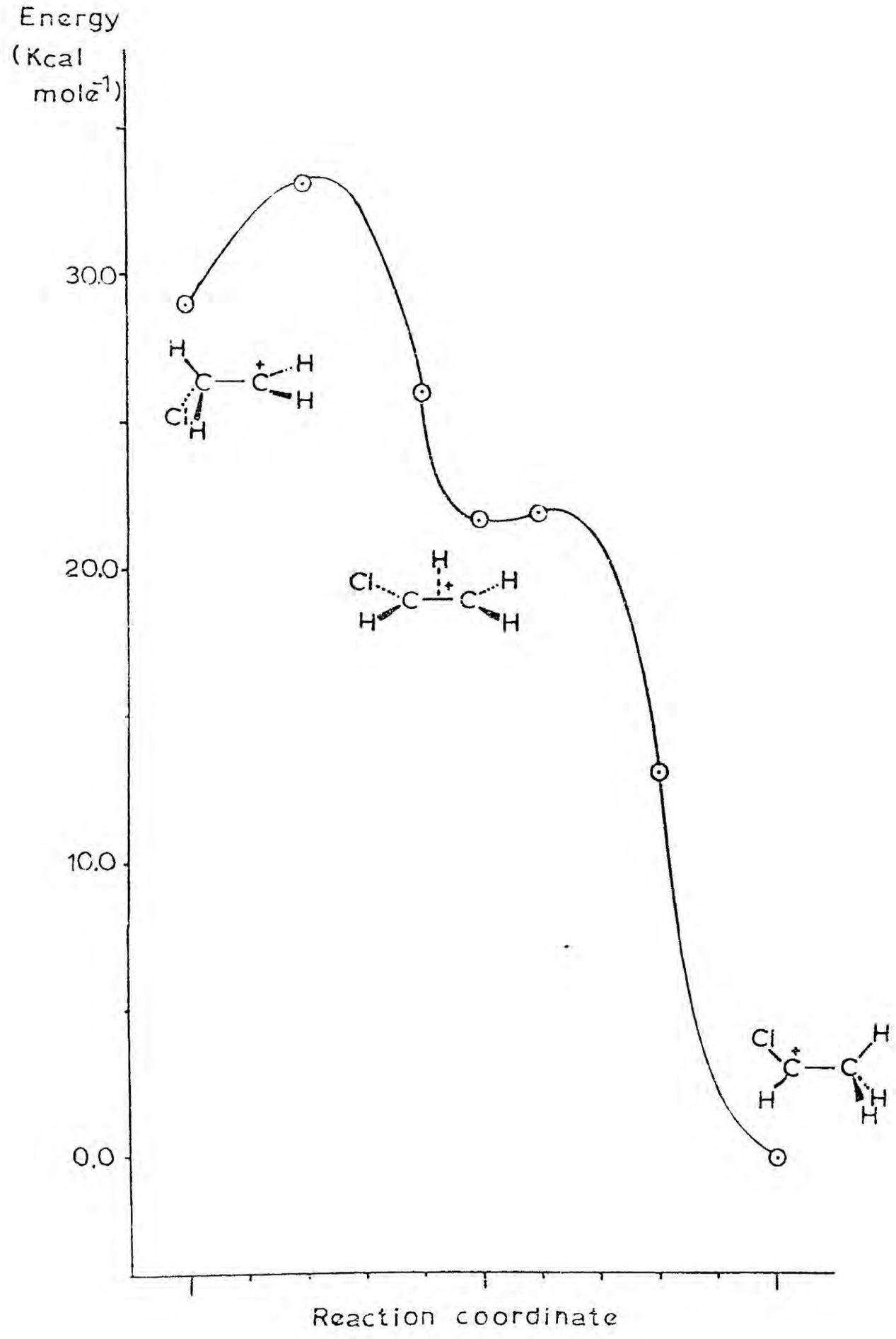
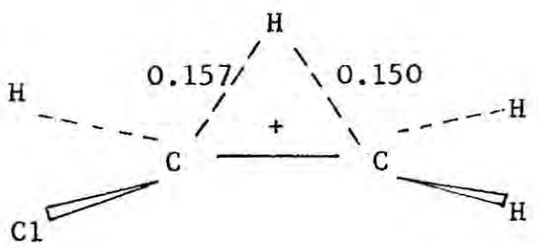


Fig 2.21 Reaction coordinate for the transformation of 2- chloroethyl cation to 1- chloroethyl cation via bridge-protonated chloroethylene

that the bridge protonated fluoroethylene should not have a discrete existence. The shape of the potential energy cross section for the interconversion of the chloroethyl cations is similar to that for the interconversion of the fluoroethyl cations, byt there are some notable differences. Firstly there is an activation barrier of about 4.3 kcal $mole^{-1}$ for the overall interconversion which is present in the first stage i.e. in transforming 2-chloroethyl cation to bridge protonated chloroethylene. Secondly there is a slight activation barrier (about 0.3 kcal $mole^{-1}$) in converting the bridge protonated chloroethylene to the 1-chloroethyl cation, though clearly this will not be of any importance and will not ensure a discrete existance for the bridge protonated species. As in the case of fluoroethylene, it is predicted that electrophilic addition of HX should yield CH_3CHClX, which is in accord with experimental observation. Again, the activation energy required for the reverse transformation, i.e. from 1-to 2-chloroethyl cations, will be very large (about 33.4 kcal $mole^{-1}$) but now the transition state (as in the case of the forward reaction) will be intermediate in character between the 2-chloroethyl cation and bridge protonated chloroethylene. In the light of more recent work [127] on the effect of polarisation functions it seems likely that the relative stabilities of bridged ions will tend to be underestimated if they are not included. In the calculations reported here, the net effect of inclusion of polarisation functions in the examination of the $C_2H_4Cl^+$ system is likely to be the observation of a much better defined local minimum for bridge-protonated chloroethylene, and could also produce a local minimum in the case of the fluorosystem. It seems likely that these results will be qualitatively the same in solution, although the overall energy difference between the 1- and 2- chloroethyl cations

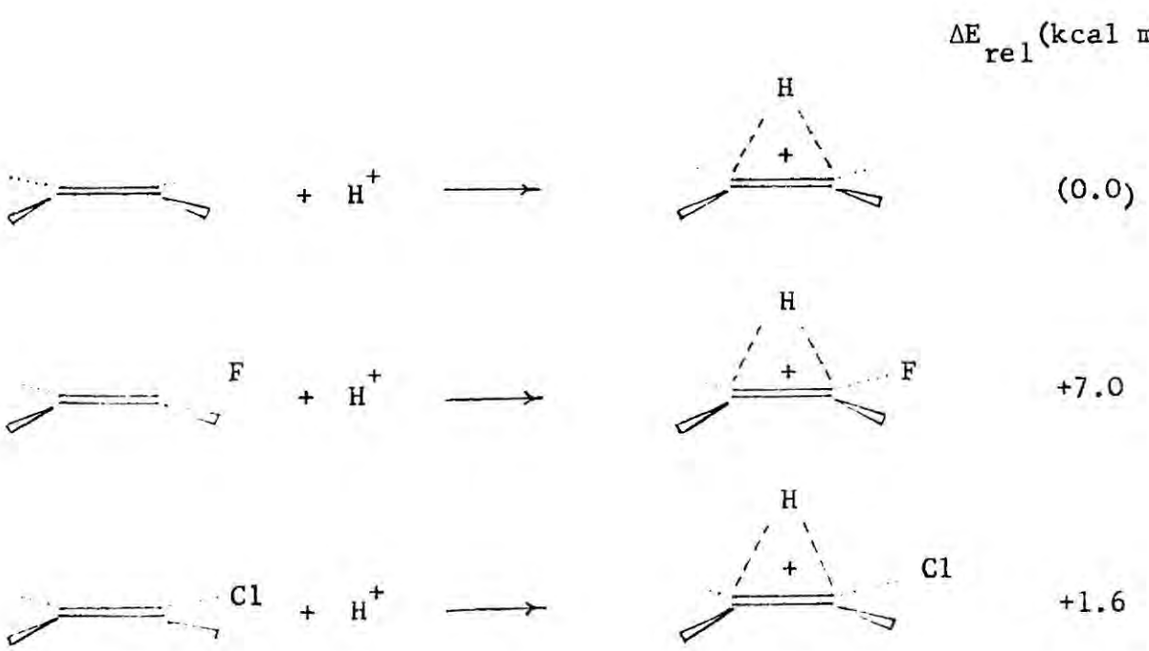
will probably be reduced slightly. It is clear that, as in the case of the 2-fluoroethyl cation, that in the case of $CH_2Cl\overset{+}{C}D_2$ the only product of transformation will be CHD_2C^+HCl i.e. no scrambling at C(2).

As in the case of the bridged protonated fluoroethylene, some insight into the reaction driving force may be obtained by examination of CNDO/II partitioned bond overlaps. The populations can be seen to be biased in favour of the formation of the 1-chloroethyl cation



2.16 Thermochemical Stabilities

A calculation using a comparable basis set has been performed upon chloroethylene and hence the examination of the destabilising effect of halogen substitution in the bridge protonated species may be extended.



Whilst the effect of chlorine substitution is to destabilise the bridge protonated olefin it can be seen that it is less effective in this than is fluorine. This is qualitatively in accord with expectation.

The size of the energy difference between 1- and 2- chloroethyl cations (29.13 kcal $mole^{-1}$) indicates the large stabilising effect of the chlorine attached to the electron deficient centre, but again this stabilisation is less efficient than in the fluoroethyl cation, where the energy difference is larger (39.36 kcal $mole^{-1}$). This is also demonstrated by a further examination of the relative energies for the process $RH \rightarrow R^+ + H^-$

		ΔE_{rel} (kcal $mole^{-1}$)
CH_3CH_3	$\longrightarrow$ $CH_3\overset{+}{C}H_2$	(0.0)
CH_3CH_2F	$CH_3\overset{+}{C}HF$	-29.5
	$CH_2F\overset{+}{C}H_2$	+9.9
CH_3CH_2Cl	$CH_3\overset{+}{C}HCl$	-22.0
	$CH_2Cl\overset{+}{C}H_2$	+7.2

The overall effect of chlorine substitution parallels that for fluorine substitution. From mass spectometric appearence potentials, Martin,Taft and Lampe have estimated [116] that for the methyl cation replacement of H by F or Cl stabilises the resulting ion by 27 $\pm$ 3 and 30 $\pm$ 4 kcal $mole^{-1}$ respectively. These values are in reasonable agreement with those reported here.

2.17 The Chloronium Ion

Addition of the electrophile Cl^+ or a suitable precursor to ethylene could, in an analogous fashion to F^+ addition, produce either the classical 2-chloroethyl cation or the bridged chloronium ion. The experimental evidence of Olah, summarised earlier in fig 2.13, demonstrates

the existence of the tetramethyl-ethylene-bridged chloronium ion fairly unequivocaly

As in the fluoronium ion work, computer limitations have prevented the full optimisation of the chloronium ion geometry. Thus as a reasonable starting point the geometry of the ethylene fragment was taken once again as that optimised previously for bridge protonated ethylene and then the energy of the species was investigated as a function of the distance of the chlorine atom from the centre of the CC bond, using the same basis set as for the chloroethyl cations, see fig 2.22.

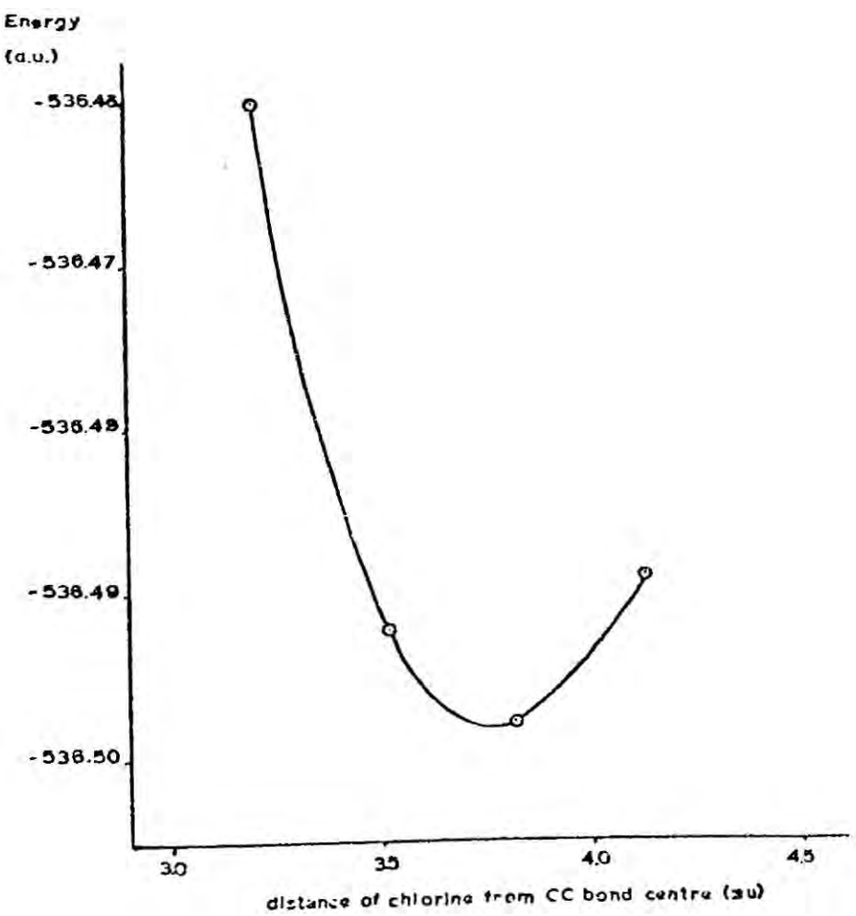


Fig 2.22 Total energy of the chloronium ion as a function of the distance of the Cl atom from the CC bond centre

The thus partially optimised geometries of both the chloronium and fluoronium ions are shown in fig 2.23, together with the microwave geometries [128] of oxirane and thiirane for comparison, although the protonated forms of these, which are isoelectronic with the fluoronium and chloronium ions respectively, would provide more instructive comparison, were their geometries available. In both cases the computed carbon-halogen bond lengths are considerably longer (~0.6Å) than the respective carbon oxygen and carbon sulphur bond lengths in oxirane and thiirane. The carbon carbon bond length (1.468Å) employed in these calculations falls within the same range as those in the oxygen and sulphur heterocycles.

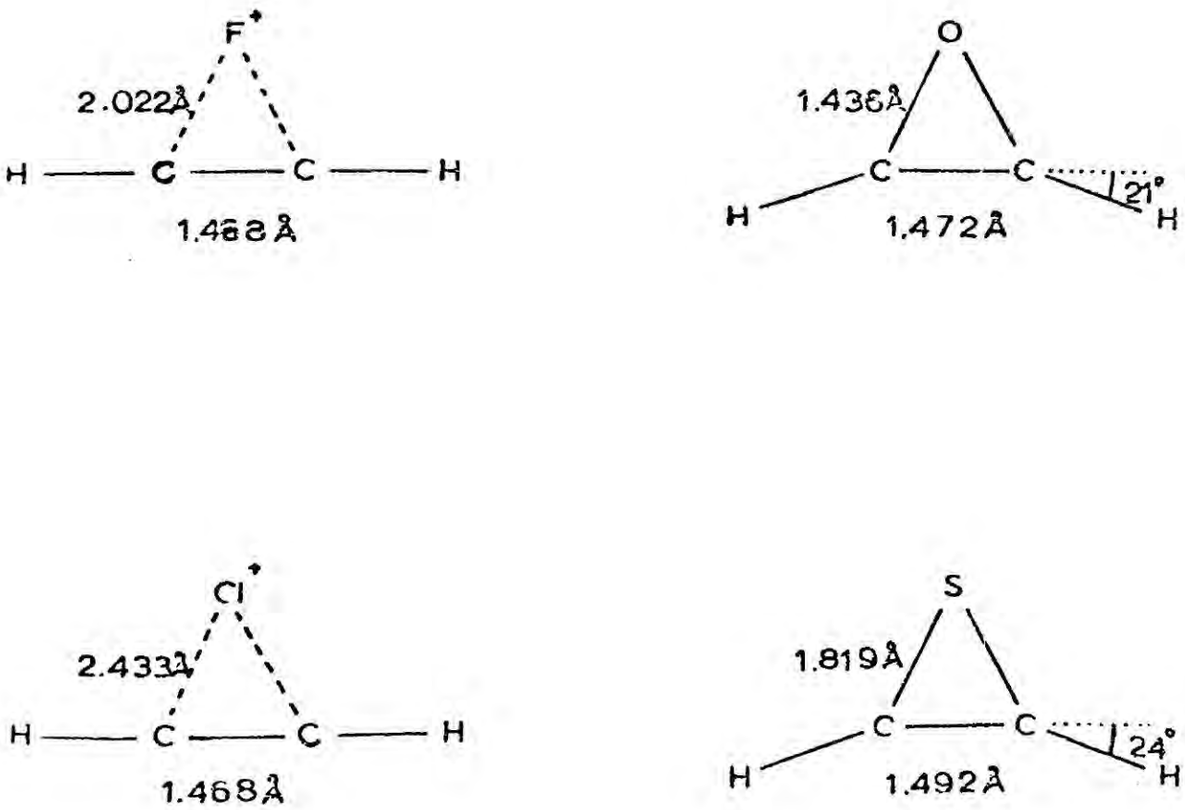


Fig 2.23 Partially optimised geometries of fluoronium and chloronium ions and the experimental (microwave) [128] geometries of oxirane and thiirane

For the latter the hydrogens are bent back by $\sim 21^{o}$ and 24^{o} respectively and the neglect of this factor in the partial geometry optimisation for the bridged halonium ions is likely to be the single most important deficiency. However on the basis of the much larger bond lengths between carbon and heteroatom for the bridged ions the out of plane bending of the hydrogens should be quite small and in the case of bridge protonated ethylene itself the calculations reported earlier (section 2.3) indicate that small deformations from planarity are energetically inexpensive. It seems likely, therefore, that the out of plane bending in the bridged ions will be small and that the calculations reported here provide a reasonable approximation to the geometry of these systems.

For bridged halonium ions the halogen is formally in a valance state higher than normal and thus d orbitals might be expected to become relatively more important compared with the classical ions. A detailed calculation on S protonated thiirene [129] (which is isoelectronic with the analogous chlorine bridged acetylene species), with an extended basis set show that d orbitals on second row atoms do not play a major rôle in understanding structure and bonding in such species. Both the neglect of the 3d orbitals in the bridged chloronium ion and the incomplete geometry optimisations will tend to underestimate the energies of the bridged ions with respect to the classical substituted ethyl cations and hence the calculations should indicate a lower limit for the relative stability of the bridged with respect to the classical ions.

Fig 2.24 shows the relative energies of the bridged and classical species for the $[C_2H_4X]^+$ (X = H,F,Cl) system obtained in these calculations and it may be seen that the relative stability of the

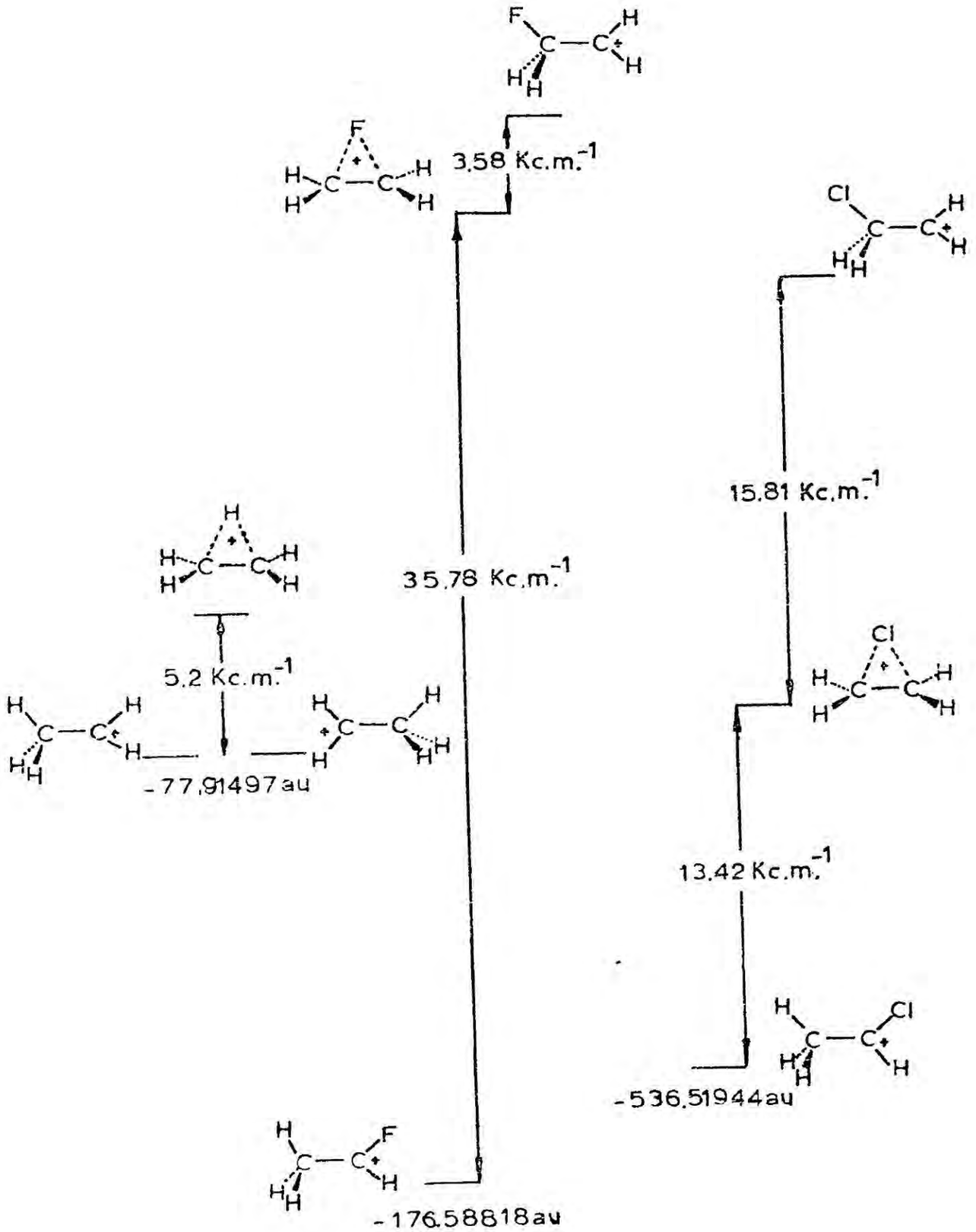


Fig 2.24 Relative energies of classical to bridged ions for the unsubstituted, fluoro and chloro systems

chloronium ion with respect to the 2-chloroethyl cation is now rather greater than was the case for the fluoronium ion, and inclusion of d orbitals would be expected to increase this effect.

Thus the results obtained in this work indicate that for the isolated molecules both the fluoronium and chloronium ions are more stable than their respective 2 substituted ethyl cations and that the order of stability of the bridged ions with respect to the classical ions in H<F<Cl, which is in keeping with available experimental data. The chloronium ion, like the fluoronium ion, is predicted to be less stable than the corresponding 1 substituted ethyl cation.

CHAPTER 3

EXPERIMENTAL AND THEORETICAL ASPECTS OF MOLECULAR CORE BINDING ENERGIES OF CARBONYL COMPOUNDS

3.1 Introduction

As indicated in Chapter 1, Xray photo electron spectroscopy (ESCA) can provide valuable information on the structure and bonding in molecules. Furthermore, the core binding energies thus derived may be used as a sensitive test of the wave functions calculated theoretically by non-empirical molecular orbital methods.

In this work an experimental and theoretical study was made upon the closely related series of molecules CH_3COX, where X = H, CH_3, OH, OCH_3, NH_2, $NHCH_3$, $COCH_3$, COOH, CN and $OCOCH_3$. This series is of particular interest to the organic chemist, carbonyl compounds occupying central importance in a large part of synthetic organic chemistry.[130] The electronic structures and the chemical properties - most importantly basicity and reactivity towards nucleophiles are drastically influenced by the nature of the attached group X. Amides are the most basic of these compounds and the generally recognised order of basicity in solution is

amides > ketones > aldehydes > carboxylic acids > acid chlorides.

This trend should be reflected in the core binding energies of both the carbon and oxygen atoms. A study of a series of closely related molecules such as these acetyl compounds has the added advantage that it is simple to measure the shift in core levels within a given molecule, although by studying the samples as thin films on a gold backing and monitoring the fermi level of the latter the absolute binding energies for the core levels of these compounds have been obtained. The relatively small size of these molecules also offers the possibility of non-empirical calculations with a physically balanced basis set.

Two approaches have been made to relate the experimental binding energies to theoretical calculations the theoretical background for which has been covered in Chapter 1. Firstly, orbital energies obtained from the ab initio calculations have been compared directly with the measured binding energies, assuming Koopmans' theorem. Secondly, the charge distributions obtained from the ab initio treatment have been correlated with the measured binding energies using the charge potential model of Siegbahn et al. A correlation has also been performed with the charge distributions obtained from CNDO/2 SCF MO calculations.

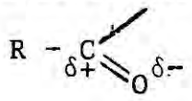
3.2 Experimental

Spectra were recorded on an AEI ES 100 spectrometer, using Mg $K\alpha_{1,2}$ radiation. These measurements were made by Clark and Barber [131] using equipment at A.E.I., Manchester. Liquid samples (~0.2µl.) were introduced via a heated reservoir shaft and insertion lock assembly. By adjusting the temperature of the reservoir a convenient leak rate across a metrosil plug located in the shaft was maintained, and samples were studied as condensed films on a cryogenic sample holder on a gold backing. Solid samples were mounted in a capillary tube and, by employing an appropriate probe assembly and insertion locks, were sublimed directly onto the cooled probe. The conducting backing helped reduce sample charging effects, but some small effects remain and the binding energy differences between different samples are estimated to be correct to within ±0.3 eV. The gold $4f_{7/2}$ level at 84 eV was used

as a reference level. Overlapping peaks were deconvoluted using a Du Pont 310 curve resolver.

3.3 Qualitative Discussion

Core binding energy shifts in molecules are largely determined by the local charge distribution and thus the shifts tend to parellel intuitive ideas concerning this distribution e.g. the large shift of the C_{1s} level of the carbonyl carbon atom



However, since valence electron distributions in molecules are continuous functions, the assignment of 'charges' to atoms within a molecule is somewhat arbitrary and depends on how the overlap density is partitioned between atoms. Nevertheless, within the limitations to the calculation of such charge density, the concept can be of great value in organic chemistry. An organic chemist obtains an intuitive idea of charge distribution in a system from a study of its chemistry. Thus he might make deductions about the charge density on a given atom from the ease of nucleophilic or electrophilic attack at that centre. However, what he is calling charge density is really the energetics of the nucleophile/electrophile approach, i.e. a measure of the potential experienced by the approaching species, which will depend not only on the atom concerned but also on the charges on the other atoms in the molecule. From the discussion of the charge potential model (Chapter 1) it is evident that there are close similarities between the factors determining 'intuitive charge distributions' and those determining

shifts in core binding energies. This shows the value of ESCA in organic chemistry since the measured shifts are closely related to the organic chemists intuitive charge distributions rather than actual charge distributions.

The measured C_{1s} and O_{1s} binding energies are presented in Table 3.1. The assignment of individual core levels is discussed fully in the next section. Clearly the most striking feature is the large difference between $\underline{C}H_3$ and $\underline{C}O$ binding energies for all the compounds. In every case the carbonyl binding energy is greater than that of the methyl carbon, the differences ranging from 4.6 eV for pyruvonitrile to 2.9 eV in acetaldehyde. The carbonyl C_{1s} binding energies are strongly dependent upon the nature of the attached group X, varying from 290.2 eV in pyruvonitrile to 288.0 eV in acetaldehyde. The methyl C_{1s} binding energies are less influenced by the nature of X, varying from 286.6 eV in pyruvonitrile to 285.0 eV in acetamide. In general the $\underline{C}O$ binding energies appear to parellel the electron donating properties of the group X - the lowest binding energies are in acetaldehyde, acetone and the two amides, and the highest binding energies are in the molecules with an attached oxygen, acetic acid, methyl acetate and acetic anhydride. A high binding energy is also observed in pyruvonitrile. The relatively large $\underline{C}O$ binding energy in methyl acetate is a feature of interest, the inductive effect of the methoxy oxygen overcoming the lone pair donation. The methyl C_{1s} binding energies approximately parellel the carbonyl C_{1s} binding energies, demonstrating a second-order shift effect.

Table 3.1 Measured C_{1s} and O_{1s} core binding energies for CH_3COX (eV)

X	C_{1s}	O_{1s}
H	288.0 (CO) 285.1 ($\underline{C}H_3$)	533.4
CH_3	288.3 (CO) 285.3 ($\underline{C}H_3$)	533.4
$COCH_3$	288.6 (CO) 285.5 ($\underline{C}H_3$)	535.3
CN	290.2 (CO) 287.6 ($\underline{C}N$) 286.6 ($\underline{C}H_3$)	534.9
$OCOCH_3$	290.0 (CO) 285.7 ($\underline{C}H_3$)	535.5 (COC) 534.2 ($C\underline{O}$)
OH	289.9 (CO) 285.6 ($\underline{C}H_3$)	534.4 (OH) 533.1 ($C\underline{O}$)
OCH_3	289.8 (CO) 287.0 ($\underline{O}CH_3$) 286.0 ($\underline{C}H_3$)	535.8 ($\underline{O}CH_3$) 534.6 ($C\underline{O}$)
COOH	290.0 (COOH) 289.0 (CO) 285.7 ($\underline{C}H_3$)	535.6 ($C\underline{O}$) 535.2 (COOH) 534.2 ($CO\underline{O}H$)
NH_2	288.5 (CO) 285.0 ($\underline{C}H_3$)	532.0
$NHCH_3$	288.6 (CO) 286.2 ($\underline{N}CH_3$) 285.4 ($\underline{C}H_3$)	532.9

The carbonyl oxygen 1s binding energies are spread over a 3.6 eV range, and again follow the carbonyl C_{1s} binding energies. The amides show very low O_{1s} levels which are due to lone pair donation from the nitrogen. The relative order of the O_{1s} binding energies in the two amides might not be that expected from a consideration of the methyl group inductive effect, although the ab initio calculations predict the observed order. In pyruvic acid, biacetyl and pyruvonitrile, where the carbonyl groups are adjacent to a second carbonyl group, it is interesting to note that the carbonyl oxygen binding energy is relatively high.

Siegbahn and co workers[56] have obtained C_{1s} binding energies measured in the gas phase for some of this series of molecules and the two sets of results, together with those for ethanol for comparison, are compared in Table 3.2. The relationship between solid state and gas phase measurements was discussed in Chapter 1 and it was shown that they could be related by means of an appropriate Born cycle (Fig.1.7). Thus, in the absence of any strong intermolecular interactions the difference should only be one of reference level. The table shows that the reference level difference is about 5.6 ± 0.4 eV and the correlation between the two sets of results is shown in Fig. 3.1. The graph has a slope of 0.9, indicating that shifts in the solid phase are slightly greater than those in the gas phase. The correlation is very good ($r^2 = 0.97$), and much of the scatter can be attributed to sample charging effects (± 0.3 eV) in the solid phase. The good correlation indicates that for a limited series of molecules, measurements of core binding energies made in the solid state may be discussed in terms of calculations on isolated molecules.

Table 3.2 Comparison of measured C_{1s} binding energies with those obtained from gas phase measurements [56]

Molecule		Solid	Gas
CH_3CH_2OH	CH_3	284.8	290.9
	CH_2	286.2	292.3
CH_3COCH_3	CH_3	285.3	291.2
	CO	288.3	293.8
CH_3CHO	CH_3	285.1	291.3
	CO	288.0	293.9
CH_3COOCH_2	CH_3	286.0	291.2
	CO	289.8	294.8
	OCH_3	287.0	292.4
CH_3COOH	CH_3	285.6	291.4
	CO	289.9	295.4

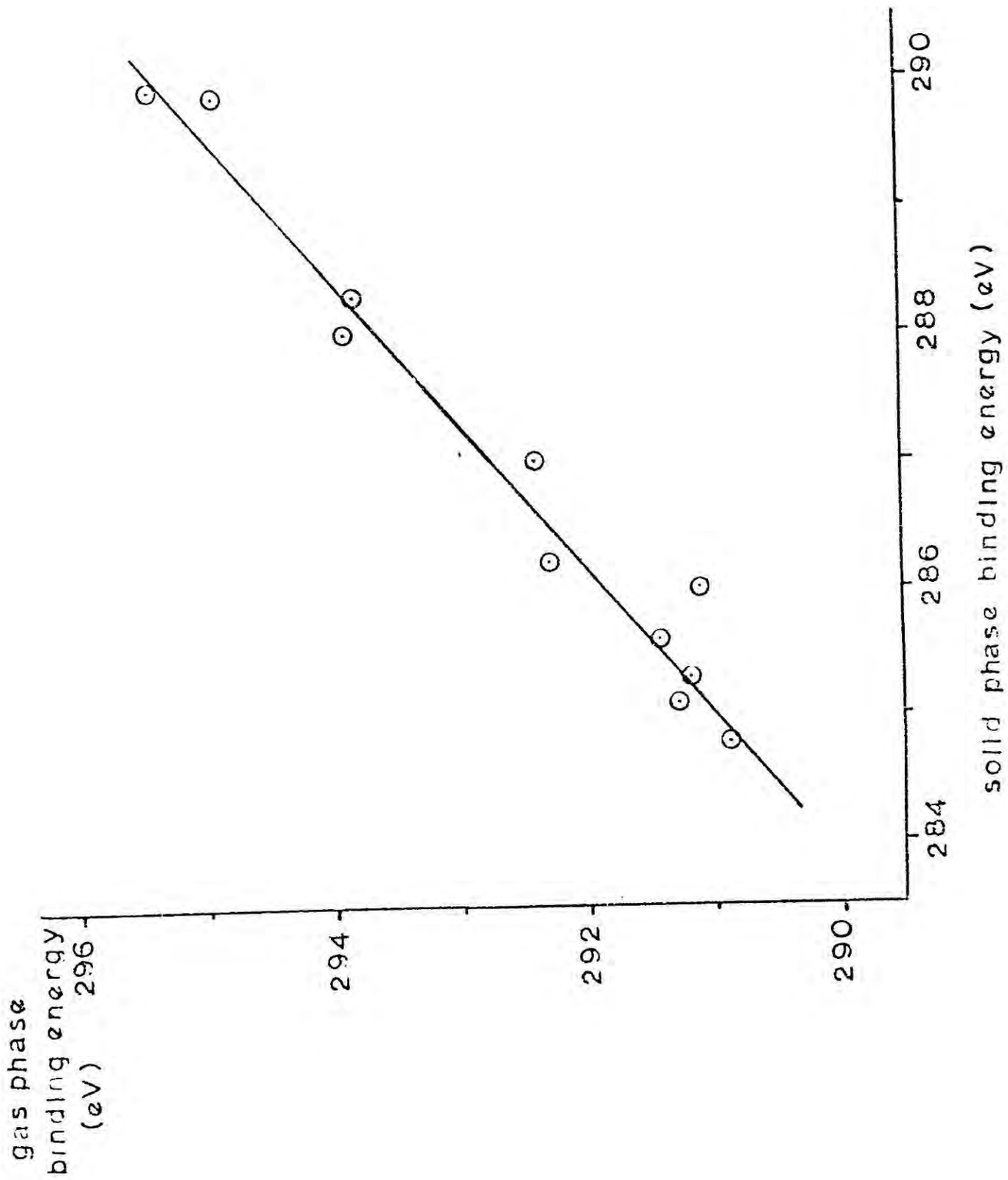


Fig. 3.1. Plot of solid phase and gas phase measurements of binding energies

3.4 Quantitative Discussion

Within the Hartree-Fock formalism, core binding energies may be examined as follows:

i) Hole state calculations

i.e. $E_{B.E.} = (E_{molecule} - E_{ion}) + \Delta E_{corr} + \Delta E_{rel}$

thus allowing for the relaxation of valence electrons on core ionisation.

ii) Koopmans' Theorem.

iii) Potential models.

iv) Equivalent cores method.

With molecules of the size of the acetyl compounds studied in this work, it has not been practicable to investigate at the non empirical level other than by calculations on the neutral molecule. Hence this discussion is in terms of

i) Koopmans' theorem, and

ii) The charge potential model.

The charge potential model is related to Koopmans' theorem and therefore, in principle, both neglect electronic reorganisation accompanying photoionisation. Since the electronic structures of these molecules are fairly similar it might be expected that relaxation at the $\underline{C}H_3$ and $\underline{C}O$ carbons to be somewhat similar through the series, although they could be quite different for the two types of environment.

In this work ab initio calculations have been performed on the series of acetyl compounds and orbital energies and atomic charges (as defined by a Mulliken population analysis) calculated.

CNDO/II SCF MO all valence electron calculations have also been performed and comparison has been made between the correlation with this charge distribution and that using the ab initio charge distribution.

(i) Application of Koopmans' theorem.

The largest molecule studied, acetic anhydride, represents a 13 nuclei, 54 electron problem, and it is evident that a fairly limited basis set must be used. The basis set used in this work was (5,2/3,1) for carbon, oxygen and nitrogen and (2/1) for hydrogen.

The calculated binding energies are shown in Table 3.3. In correlating these values with the measured binding energies (see Table 3.1) it must be remembered that the calculations refer to isolated i.e. gas phase molecules, whereas the experimental results are derived from measurements made in the condensed phase. As discussed previously, in the absence of strong intermolecular forces this does not result in serious errors other than a change in reference level. However, in the case of acetic acid polymerisation by hydrogen bonding occurs in the shift observed between the two oxygen atoms. It is also to be noted that the oxygen $_1$s assignment made by Siegbahn and co-workers[56] is incorrect, predicting a higher binding energy for the carbonyl oxygen. A plot of the calculated binding energies against the gas phase binding energies measured by Siegbahn[56] is shown in Fig. 3.2, and a plot of the calculated binding energies against the solid phase binding energies is shown in Fig. 3.3. Several interesting features may be distinguished from these graphs. The two slopes are in good agreement (1.13 and 1.17 respectively) and are quite close to the

Table 3.3 Calculated (ab initio) core binding energies for for CH_3 COX (eV).
(cf experimental values, Table 3.1)

X	C_{1s}	O_{1s}
H	313.07 (CO) 310.86 ($\underline{C}H_3$)	564.47
CH_3	313.11 (CO) 310.85 ($\underline{C}H_3$)	563.83
$COCH_3$	313.55 (CO) 310.69 ($\underline{C}H_3$)	565.20
CN	315.09 (CO) 312.11 ($\underline{C}N$) 311.72 ($\underline{C}H_3$)	566.10
$OCOCH_3$	314.74 (CO) 311.23 ($\underline{C}H_3$)	565.38 (COC) 564.10 ($\underline{C}O$)
OH	314.36 (CO) 310.99 ($\underline{C}H_3$)	564.86 (OH) 561.89 ($\underline{C}O$)
OCH_3	314.46 (CO) 312.49 ($\underline{O}CH_3$) 311.08 ($\underline{C}H_3$)	564.96 (OCH_3) 562.86 ($\underline{C}O$)
COOH	314.62 (COOH) 314.05 ($\underline{C}O$) 310.85 ($\underline{C}H_3$)	565.53 (CO) 565.06 (COOH) 562.87 ($CO\underline{O}H$)
NH_2	313.50 (CO) 310.86 ($\underline{C}H_3$)	561.41
$NHCH_3$	314.10 (CO) 312.32 ($\underline{N}CH_3$) 310.55 ($\underline{C}H_3$)	562.53

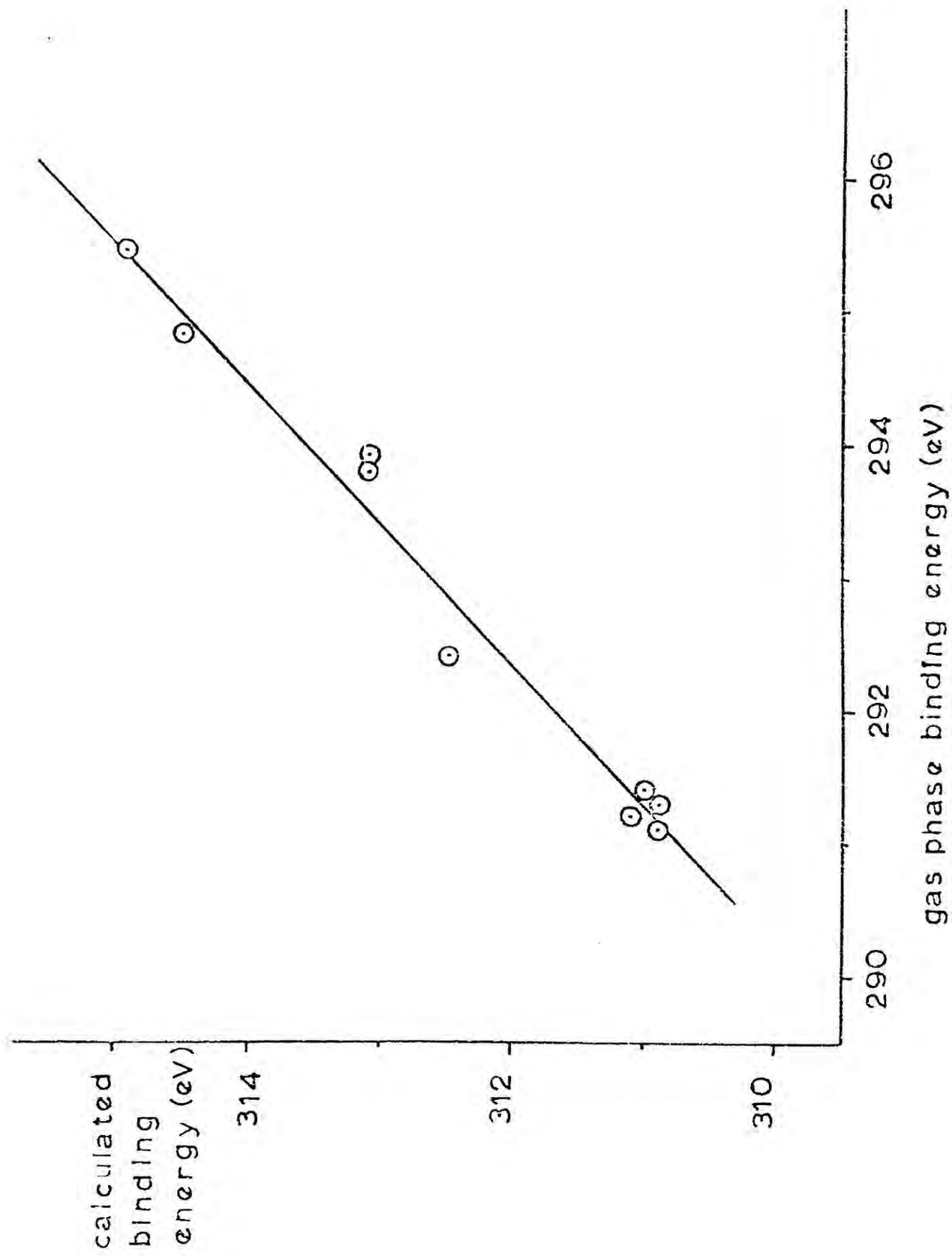


Fig. 3.2. Plot of calculated and measured (gas phase) binding energies

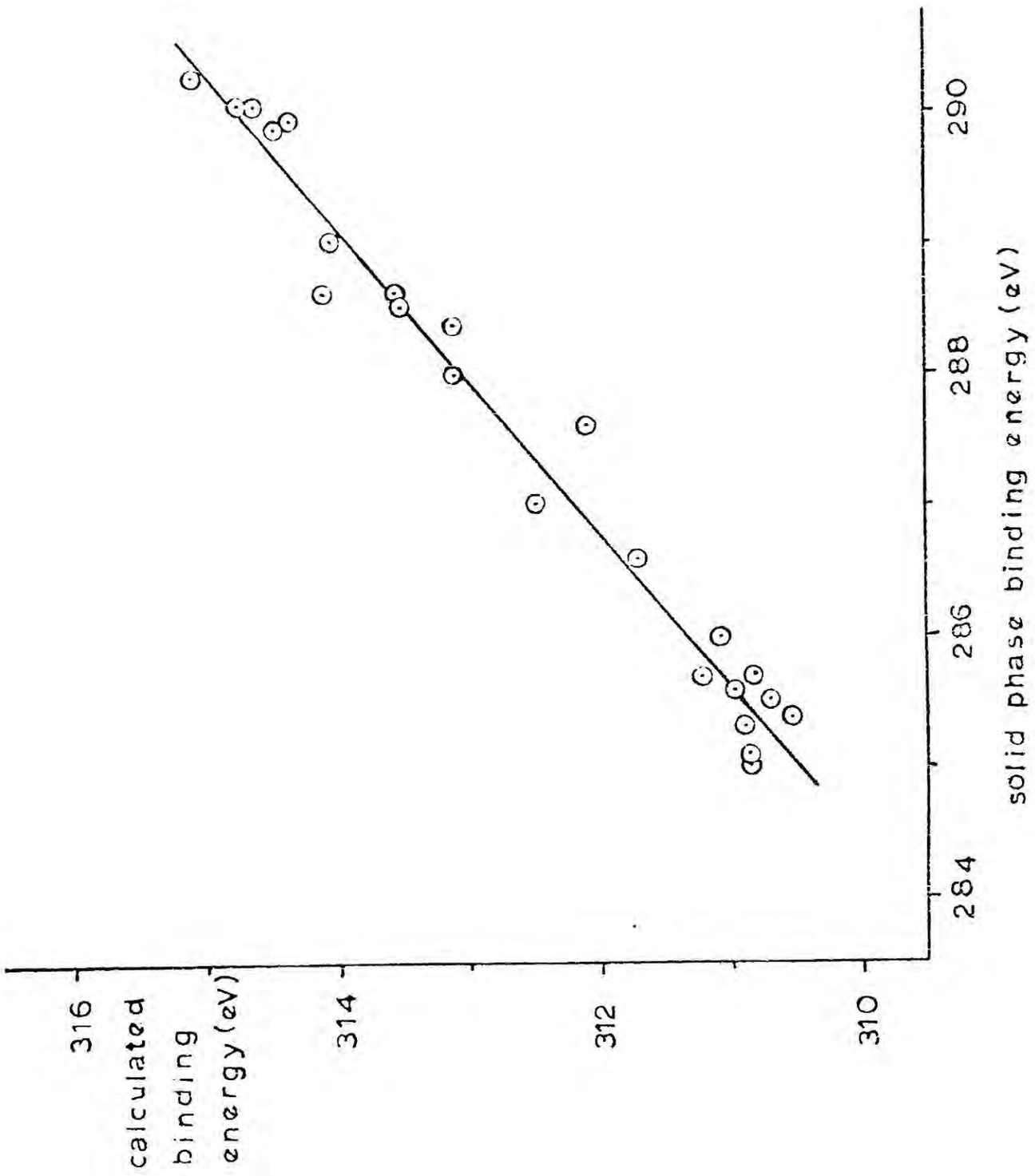


Fig. 3.3. Plot of calculated and measured (solid phase) binding energies

theoretical value of 1.0, and the correlation coefficients, r^2, of 0.97 and 0.96 respectively are very good. This clearly indicates the value of using orbital energies for the assignment of experimental core binding energies. The axes of both graphs are displaced relative to each other. The displacement of ~20 eV in Fig. 3.2 is composed of the reorganisation energy, correlation energies and relatavistic energies. The displacement of ~25.5 eV in Fig. 3.3 is composed of this same 20 eV plus an additional ~ 5.5 eV arising from the difference in the reference levels in the gaseous and solid phases.

Both the close binding energies of the methyl and the carbonyl carbons in acetaldehyde and acetone and the low carbon binding energies of the amides are confirmed by the calculations.

A comparison of the observed methylcarbonyl carbon shifts with the calculated shifts is shown in Table 3.4. With the exception of N-methyl acetamide, the observed shifts are larger than those calculated, the most extreme case being that of pyruvonitrile, with a 1.2 eV discrepancy between the two. This suggests that relaxation energies are different for the two types of carbon, and that relaxation energies may differ between compounds. Interestingly, the calculations predict that introduction of a methyl group into acetamide should increase the shift by 1.0 eV. Experimentally, however, a shift of 0.3 eV occurs in the opposite sense, which would be predicted from consideration of the inductive effect of a methyl group. The calculated shift occurs mainly

Table 3.4 Comparison of measured core binding energy shifts between carbonyl and methyl carbon atoms with those calculated on a basis of Koopmans' Theorem.

X	Observed shift	Calculated shift
H	2.9	2.2
CH_3	3.0	2.2
$COCH_3$	3.1	2.9
CN	4.6	3.4
$OCOCH_3$	4.3	3.5
OH	4.3	3.4
OCH_3	3.8	3.4
COOH	3.3	3.2
NH_2	3.5	2.6
$NHCH_3$	3.2	3.6

through a 0.6 eV increase of the carbonyl carbon binding energy, though no shift is found experimentally.

(ii) Correlations with charge distributions.

To summarise some of the theory covered in Chapter 1, in the electrostatic potential model of Siegbahn and co-workers, shifts in core binding energies may be related to the charge distribution by

$$E^i = E^i_o + kq_i + \sum_{j \neq i} \frac{q_j}{r_{ij}} \qquad (3.1)$$

where E^i_o is a reference level and the second term represents the potential from the charge on the atom in question. The charge dependence parameter, k, depends upon the definition of atomic charge and the basis set employed and is approximately

$$\langle \phi_c(1)\phi_v(2) \frac{1}{r_{12}} \phi_c(1)\phi_v(2) \rangle$$

The last term is an intramolecular Madelung potential which accounts for the potential from charges in the rest of the molecule.

a) Ab initio charge distributions. Early attempts were made to relate core binding energies to charge densities, without a consideration of the potential arising from the remaining atoms in the molecule. However, neglect of the Madelung potential is, in general, not valid. Fig. 3.4 shows a plot of binding energies against 'charge density' showing the expected overall trend (higher charge - higher binding energy) but considerable

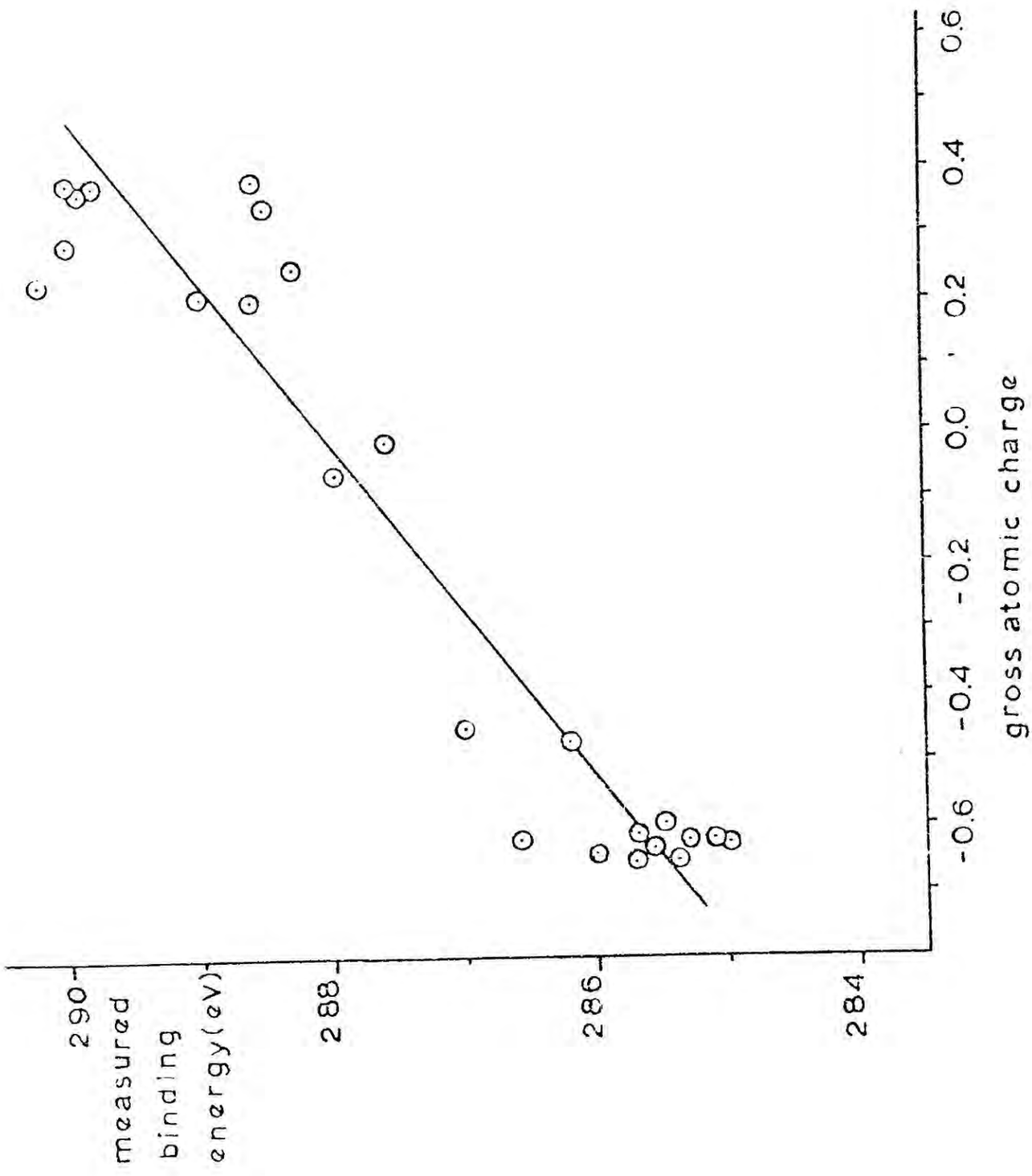


Fig. 3.4. C_{1s} binding energies against charge densities at the respective atoms.

scatter. Inclusion of the Madelung potential, as will be demonstrated later in this section, gives rise to a rather better correlation. Table 3.5 shows the ab initio charges for the C_{1s} levels together with the relative binding energies $E - E_o$ calculated using equation (3.1), and also the relative shifts $\Delta(E - E_o)$ within a given molecule. With the exception of acetaldehyde, the carbonyl carbon charges all lie within the range 0.19 - 0.38. The very low calculated charge in acetaldehyde is compensated by a negative Madelung potential of corresponding low magnitude giving a relative binding energy similar to that in acetone. The methyl carbon charges lie in the range -0.45 - -0.66. These values are much lower than would be expected on a simple physical picture, and are an artefact of this type of calculation. The relative $\underline{C}H_3$ - $\underline{C}O$ shifts may be compared with the experimental shifts, and those calculated using orbital energies given in Table 3.4. The correlation is quite good, that with the shifts calculated from orbital energies being slightly better. Thus it is clear that calculated ab initio charge distributions are reliable for the assignment of core levels.

The value of k used in calculating the relative binding energies may be found from a plot of the atomic charge against the measured binding energy, corrected for the Madelung potential, Fig. 3.5. The slope is 19.4 ± 0.4 eV/unit charge in reasonable agreement with the theoretical value of 22.0 eV/unit charge (calculated with the use of Slater orbitals), and the correlation ($r^2 = 0.99$) is very good.

Table 3.5 Ab initio charges, Madelung potentials and calculated core binding energy shifts for CH_3COX

X	Atom	q_i	$\sum_{i\neq j} \frac{q_j}{r_{ij}}$ (eV)	E - Eo (eV)	Δ(E - Eo) (eV)
H	$\underline{C}O$	0.032	-1.73	-1.11	2.88
	$\underline{C}H_3$	-0.623	8.19	-3.99	
CH_3	$\underline{C}O$	0.241	-6.06	-1.38	2.34
	$\underline{C}H_3$	-0.624	8.39	-3.72	
$COCH_3$	$\underline{C}O$	0.194	-3.51	0.25	3.01
	$\underline{C}H_3$	-0.595	8.28	-3.26	
CN	$\underline{C}O$	0.218	-3.88	0.35	3.01
	$\underline{C}H_3$	-0.624	9.45	-2.66	-2.01
	$\underline{C}N$	-0.020	-0.26	-0.65	
$OCOCH_3$	$\underline{C}O$	0.368	-6.65	0.49	3.53
	$\underline{C}H_3$	-0.657	9.71	-3.04	
OH	$\underline{C}O$	0.350	-6.82	-0.03	3.28
	$\underline{C}H_3$	-0.635	9.01	-3.31	
OCH_3	$\underline{C}O$	0.365	-7.20	-0.12	3.25
	$\underline{C}H_3$	-0.645	9.14	-3.37	-0.67
	$O\underline{C}H_3$	-0.453	6.09	-2.70	
COOH	$\underline{C}O$	0.197	-3.28	0.54	3.80
	$\underline{C}H_3$	-0.616	8.69	-3.26	4.49
	$\underline{C}OOH$	0.227	-4.14	1.23	
NH_2	$\underline{C}O$	0.339	-7.67	-1.09	2.44
	$\underline{C}H_3$	-0.630	8.69	-3.53	
$NHCH_3$	$\underline{C}O$	0.374	-7.88	-0.62	3.28
	$\underline{C}H_3$	-0.652	8.75	-3.90	-0.83
	$N\underline{C}H_3$	-0.478	6.20	-3.07	

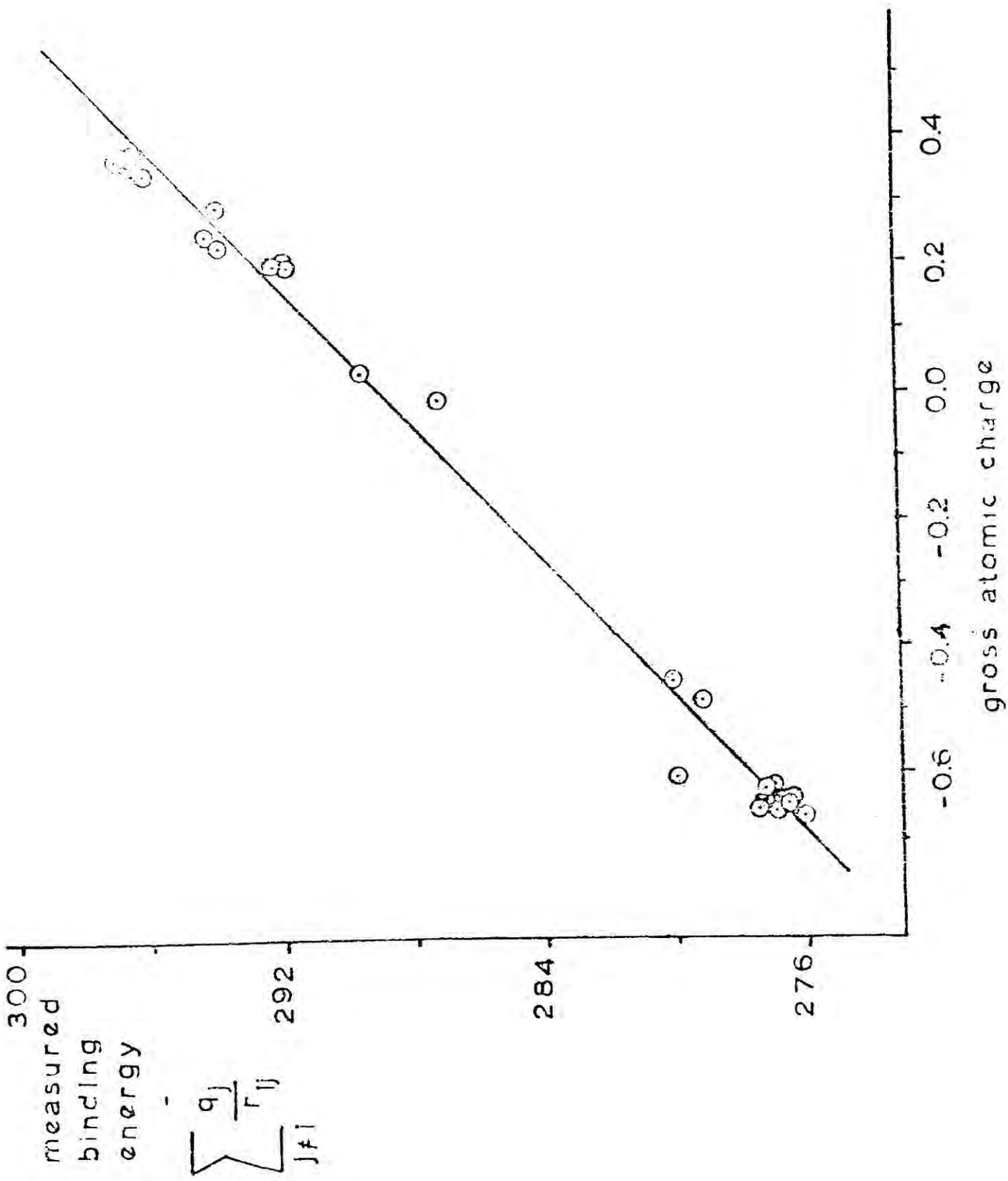


Fig. 3.5. C_{1s} binding energies corrected for Madelung potentials against ab initio charge density at the respective atoms.

b) CNDO/2 charge distributions. For ab initio calculation, a molecule of the size of acetic anhydride represents a 'big molecule', expensive of computer time and toward the upper limit as regards routine ab initio calculation. Thus it is of some considerable interest to examine the feasibility of the use of more approximate SCF - MO theories as an aid to the assignment of molecular core binding energies. Hence a comparison has been made of the relative shifts predicted by CNDO/2SCF MO calculations with the corresponding ab initio results.

The calculated charges, relative binding energies and relative shifts are shown in Table 3.6. With the exception of acetaldehyde, the carbonyl carbon charges are very close to the ab initio charges; however the methyl carbon charges are considerably smaller than the corresponding ab initio charges, and are rather closer to what would be expected intuitively. It is interesting to note that a 'normal' carbonyl carbon charge is predicted by the CNDO/2 method.

The $E - E_o$ values are consistently about 4eV higher than the ab initio values, but the relative shifts are quite close. In fact, if these shifts are compared with those in Table 3.4 it can be seen that CNDO shifts are closer to those calculated using the ab initio orbital energies than are the shifts calculated using the ab initio charge distribution. For example, the $\underline{C}H_3 - \underline{C}O$ shift of 3.01 eV in pyruvic acid is much closer to both the observed and orbital energy shifts than is the abnormally high value of the ab initio shift of 3.8 eV. The high observed shifts in acetic

Table 3.6 CNDO/II charges, Madelung potentials and calculated core binding energy shifts for CH_3COX.

X	Atom	q_i	$\sum_{i\neq j} \frac{q_j}{r_{ij}}$ (eV)	$E - E_o$ (eV)	$\Delta(E - E_o)$ (eV)
H	$\underline{C}O$	0.223	-2.93	2.47	2.28
	$\underline{C}H_3$	-0.065	1.76	0.19	
CH_3	$\underline{C}O$	0.240	-3.27	2.54	2.50
	$\underline{C}H_3$	-0.069	1.71	0.04	
$COCH_3$	$\underline{C}O$	0.196	-2.07	2.67	2.89
	$\underline{C}H_3$	-0.094	2.06	-0.22	
CN	$\underline{C}O$	0.246	-2.87	3.08	3.00
	$\underline{C}H_3$	-0.067	2.20	0.58	2.73
	$\underline{C}N$	-0.003	0.42	0.35	
$OCOCH_3$	$\underline{C}O$	0.382	-5.19	4.05	3.72
	$\underline{C}H_3$	-0.090	2.51	0.33	
OH	$\underline{C}O$	0.392	-5.41	4.08	3.87
	$\underline{C}H_3$	-0.090	2.39	0.21	
OCH_3	$\underline{C}O$	0.387	-5.36	4.01	3.86
	$\underline{C}H_3$	-0.092	2.38	0.15	-1.63
	$\underline{O}CH_3$	0.144	-1.71	1.78	
COOH	$\underline{C}O$	0.211	-1.59	3.52	3.01
	$\underline{C}H_3$	-0.066	2.11	0.51	4.06
	$\underline{C}OOH$	0.359	-4.12	4.57	
NH_2	$\underline{C}O$	0.361	-5.39	3.35	3.30
	$\underline{C}H_3$	-0.083	2.09	0.05	
$NHCH_3$	$\underline{C}O$	0.376	-5.32	3.78	3.84
	$\underline{C}H_3$	-0.096	2.26	-0.06	-1.74
	$\underline{N}CH_3$	0.117	-1.15	1.68	

anhydride, acetic acid and methyl acetate are predicted by both methods, but all three types of calculation incorrectly predict a higher $\underline{C}H_3$ carbon 1s binding energy in N-methyl acetamide than in acetamide. It must be noted that the relative ordering of methyl and nitrile C_{1s} levels is incorrectly predicted by the CNDO/2 method, and the $\underline{C}H_3$ - $\underline{C}N$ shift is 4.74 eV greater than that calculated from ab initio charges.

Further comparison between the results of the ab initio and the CNDO/2 calculations may be made by plotting the gross atomic charges (Fig. 3.6) and the Madelung potentials (Fig. 3.7) derived from the two methods. The graphs show that the trends obtainable are the same from both calculations, but considerable scatter is evident.

Fig. 3.8 shows a plot of the CNDO/2 charge against the observed binding energy corrected for the Madelung potential. A good correlation is obtained (r^2 = 0.98) and the slope of 24.2 ± 0.7 eV/unit charge, although significantly different from the ab initio slope, is in good agreement with the theoretical value.

Thus it can be clearly seen that CNDO/2 theory is adequate in giving good correlations with experimental results, and also in assigning experimental molecular core binding levels, although the results for pyruvonitrile show that some care is needed in assignments in certain systems.

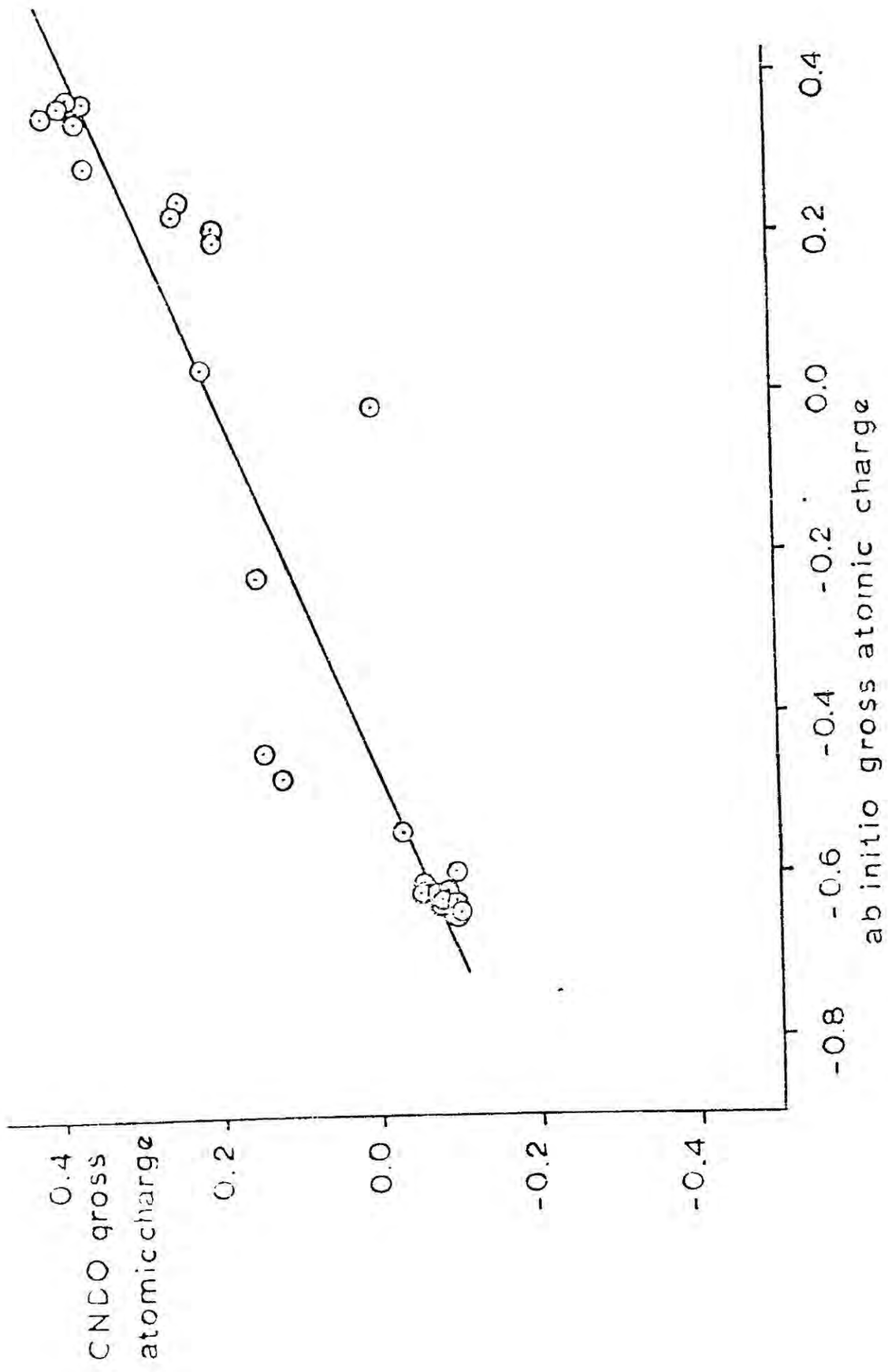


Fig. 3.6. Plot of ab initio and CNDO/II gross atomic charges for carbon atoms.

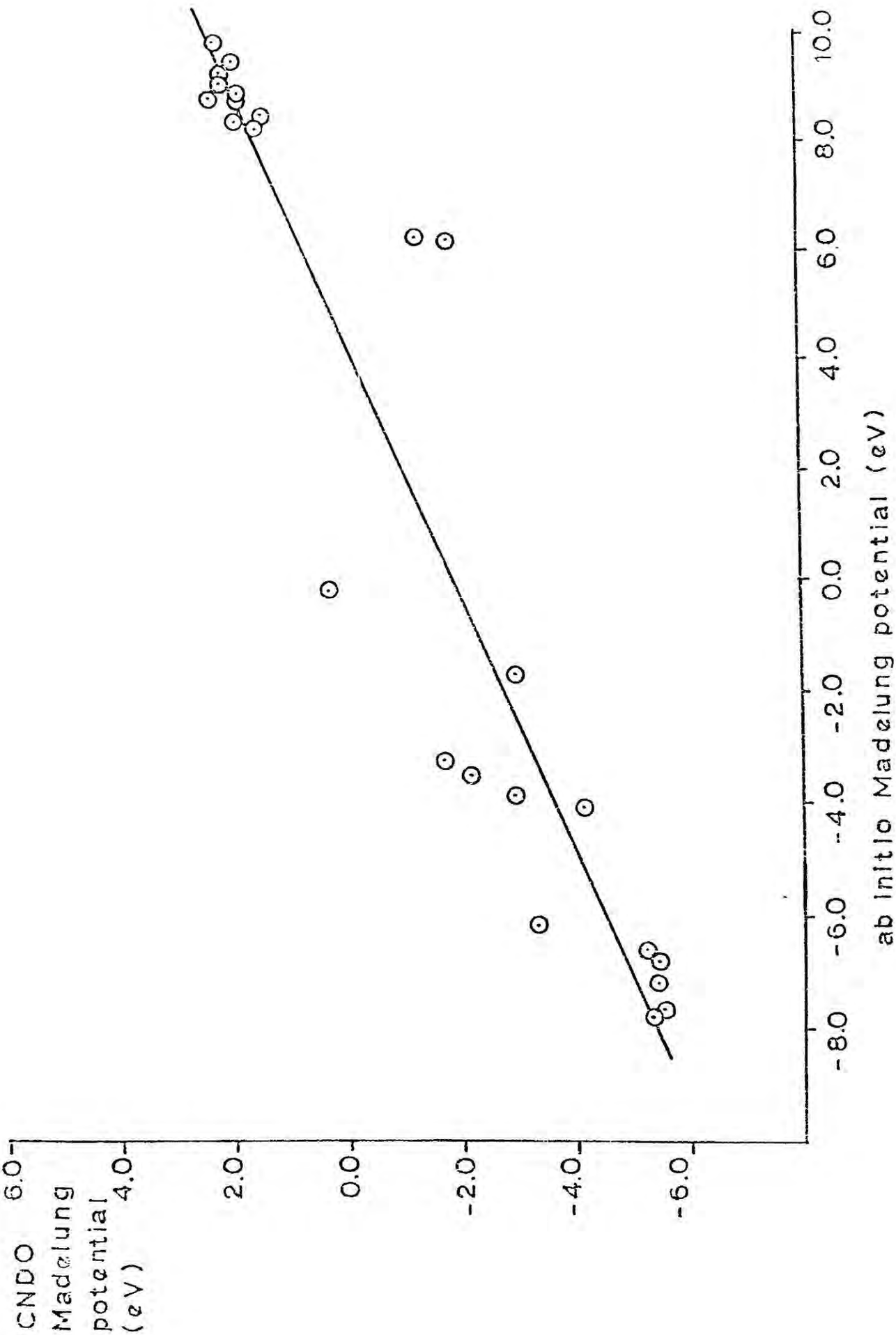


Fig. 3.7. Plot of ab initio and CNDO/II Madelung potentials for carbon atoms.

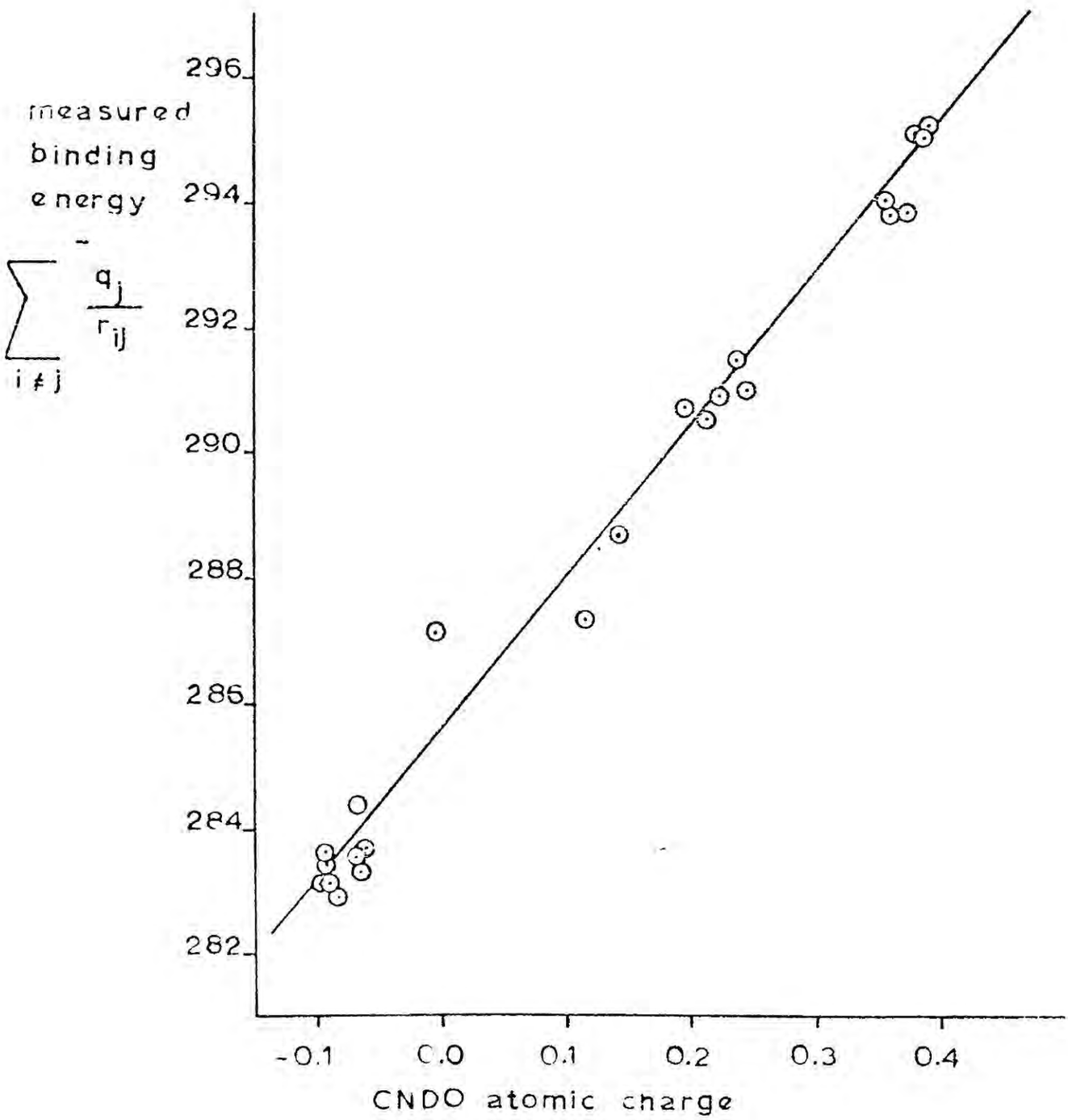


Fig. 3.8. C_{1s} binding energy corrected for Madelung potentials against CNDO/II charge density at the respective atoms.

CHAPTER IV

EXPERIMENTAL AND THEORETICAL ASPECTS OF MOLECULAR CORE BINDING ENERGIES IN A SERIES OF FIVE MEMBERED RING HETEROCYCLES

4.1 Introduction

The series of heterocyclic compounds chosen for this investigation all contain a five membered ring with one or more heteroatoms. This results in a series presenting an interesting variation of properties with respect to the relative positions and nature of the heteroatom(s). Some members of this series are widespread in nature and hence of interest from this point of view, but all are of considerable intrinsic interest to the organic chemist. They are formally 6π electron systems with varying degrees of aromaticity and their reactivity and patterns of substitution are of considerable importance in the study of theoretical organic chemistry and to organic chemists generally. From an ESCA viewpoint, these compounds present a series for which the collection of molecular core binding energy data may provide a good test of available methods of assignment; the comparatively small size of these molecules enabling the computation of non-empirical wave functions and energies. Thus comparison is possible between assignments based upon non-empirical and semi-empirical charge densities, invaluable information for the investigation of larger species.

The basicities of some of the series are shown in Table 4.1, together with the pK_b of pyridine for comparison. The pK_b values show particularly great variation. The low value of pyrrole would be expected but the difference between the values of pyrazole and imidazole is interesting. The boiling points of these compounds are also given in Table 4.1 as here again large differences are seen.

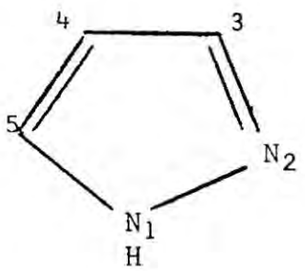
	Basic pK	Boiling point (°C)
Pyridine	5.6	115
Pyrrole	-3.8	130-131
Pyrazole	2.5	187
Imidazole	7.2	256
Thiazole	2.5	117

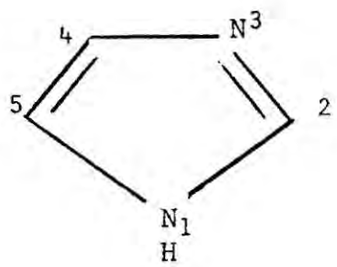
Table 4.1 pk and boiling point data for some heterocycles[132]

This indicates profound differences in hydrogen bonding structure, which in turn make acquisition of data on electronic density desirable. It is known that the difference between pyrazole and imidazole may be explained by the formation of hydrogen bonded dimers and chains respectively.

It is also of interest to relate data pertaining to reactivities of aromatic systems with ESCA data. For example, the five membered ring heterocyclic systems are extremely reactive towards electrophiles, in which case the transition state for electrophilic substitution will be relatively early (compared with, say, benzene) and will to some extent, therefore, reflect the potential experienced by the approaching electrophile. Hence a priori there may well be a tendency for the site of substitution to reflect the potential at a given site, and therefore the reactivity pattern has features in common with factors determining shifts in

core binding energies.

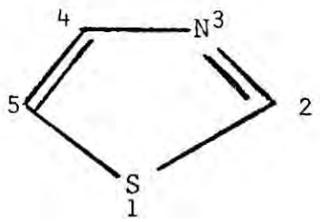




In concentrated sulphuric acid, pyrazole is nitrated at the 4 position,[133] as in the case of bromination, sulphonation and base catalysed iodination.[134] Imidazole is more susceptible to electrophilic reagents and nitration and iodination lead to 4(5) substitution.[133, 135] In alkaline solution, diazonium coupling occurs at the 2 position. The low boiling point of Furan (31.4° C) compared with the other compounds of similar molecular weight reflects the absence of hydrogen bonding. The ring system is especially susceptible to electrophilic substitution and attack at the 2(5) position is predominent. Under some circumstances it can behave as a nucleophile.[136] Isoxazole is intermediate in boiling point between furan and pyrrole. Electrophilic attack occurs at the 4 position. Phenyl isoxazoles are preferentially nitrated in the phenyl ring showing a net deactivating effect of the 'pyridine' nitrogen over the activating 'furan' oxygen.

Pyrrole is very reactive towards electrophiles and substitutes at the 2(5) position.

Thiazole undergoes electrophilic substitution at position



5, and, to a lesser extent, at position 4.

4.2 Experimental

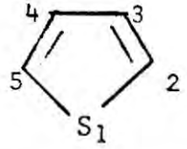
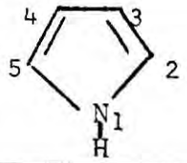
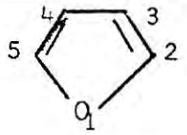
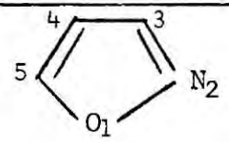
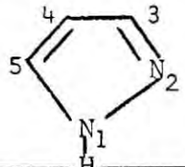
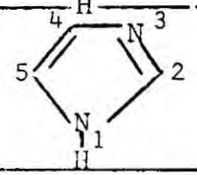
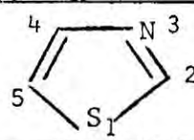
Spectra were measured on the AEI ES 100 electron spectrometer, using Mg $K\alpha_{1,2}$ radiation. Samples were studied as thin films on gold and, under the experimental conditions employed, the gold $4f_{7/2}$ core level at 84 eV binding energy had a half width of 0.95 eV. Spectra were deconvoluted where necessary using the Du Pont 310 curve resolver, the line width of 1.1 eV (for C_{1s} levels) and peak shape (almost gaussian) for each channel being taken from a previous study of benzene.

4.3 Qualitative Discussion

The measured core binding energies are tabulated in Table 4.2, and some interesting trends may be seen. These may be conveniently examined by means of the following scheme.

(i) Replacement of ring CH by N. The shifts in core binding energies brought about by this substitution may be correlated with the nature of the initial heteroatom.

Table 4.2 Measured molecular core binding energies (eV)

		C_{1s}								
		C2(C5)	C3(C4)							
Thiophene		285.0	285.0		$S2_s$	228.6				
					$S2p_{1/2}$	165.3	$S2p_{3/2}$	164.3		
Pyrrole		285.7	284.8		N_{1s}	401.2				
Furan		286.7	285.6		O_{1s}	535.8				
		C3	C4	C5						
Isoxazole		286.9	285.9	287.4	N_{1s}	402.1	O_{1s}	536.6		
Pyrazole		286.3	285.5	287.1	N_{1s} (1)	402.4	N_{1s} (2)	401.1		
		C2	C4	C5						
Imidazole		287.2	285.7	286.5	N_{1s} (1)	402.3	N_{1s} (3)	400.7		
Thiazole		286.8	287.2	286.1	N_{1s} (3)	400.5	S_{2s}	229.7		
							$S_{2p1/2}$	166.3	$S_{2p3/2}$	165.3

a) Thiophene to thiazole.

Heteroatom	S binding energy increases 1.1 eV
Adjacent CH	Average binding energy increases 2.0 eV

b) Pyrrole to pyrazole.

Heteroatom	N binding energy increases 1.2 eV
Adjacent CH	Average binding energy increases 1.5 eV

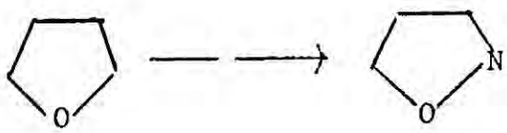
c) Pyrrole to imidazole

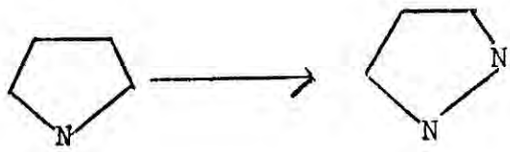
Heteroatom	N binding energy increases 1.1 eV
Adjacent CH	Average binding energy increases 1.6 eV

d) Furan to isoxazole

Heteroatom	O binding energy increases 0.8 eV
Adjacent CH	Average binding energy increases 1.0 eV

In all cases, replacement of a ring CH by N has led to increased core binding energies of the remaining ring atoms of the molecules. This provides an indication of the overall electronic drift from the ring atoms to nitrogen. Examination of the shifts in core binding energy of the heteroatom reveals an overall increase. The net effect of replacing ring CH by N as far as the attached heteroatom is concerned is larger in the case of a nitrogen heterocycle than for an oxygen heterocycle.

i.e.  ΔO_{1s} BE 0.8 eV

 ΔN_{1s} BE 1.2eV

These data may be rationalised on the basis of the electronegativity of the heteroatom. Thus the effect of replacing ring CH by the more electronegative N substituent is comparatively less the greater the electronegativity of the heteroatom. The shifts in core binding energies of adjacent ring CH shows a marked trend with the nature of the heteroatom.

$$O < NH < S$$

This is readily interpretable in terms of the differences in electron density on the ring CH, which is influenced by the nature of the heteroatom i.e.

CH electron density as a function of heteroatom $\quad S > NH > O$

Thus the greater the electronic density present on the carbon atom, the greater the change which will be brought about by the substitution of a ring CH by the comparatively electronegative N, and hence the greater will be the chemical shift in the carbon core level.

Examination of the binding energies of the nitrogen atoms of pyrazole and imidazole shows a further trend, i.e.

by an average of 1.4 eV. This reflects the effect of the large potential exerted by the positively charged hydrogen atom at a short distance from the nitrogen atom (~1.0 Å), as is indicated by the high value of the Madelung potential at the nitrogen, shown in the next section. Clearly this has considerable ramifications in terms of the effect of hydrogen bonding on the nitrogen core levels.

(ii) Replacement of ring NH by O or S

a) Imidazole to thiazole.

Heteroatom	N binding energy decreased 0.2 eV
Adjacent CH	Average binding energy decreased 0.4 eV

These results are in accord with simple electronegativity considerations.

b) Pyrazole to isoxazole.

Heteroatom	N binding energy increased 1.0 eV
Adjacent CH	Binding energy increased 0.3 eV

As would be expected, the nitrogen shift here is rather greater than in (a) as a result of the greater electronegativity of O than N and also the fact that the N is now directly adjacent to the replacement.

c) Pyrrole to thiophene.

Adjacent CH	Binding energy decreased 0.7 eV
Distal CH	Binding energy increased 0.2 eV

Here again is seen both electronegativity and proximity effects.

(iii) Replacement of ring O by S

Furan to Thiophene.

Adjacent CH	Binding energy decreased 1.7 eV
Distal CH	Binding energy decreased 0.6 eV

These effects are once again in accord with electronegativity and proximity considerations and the comparison with (iic) shows greater and less ambiguous core level shifts. This can be ascribed to the greater electronegativity charge i.e.

$$\text{shift } (S \rightarrow O) > \text{shift } (S \rightarrow N)$$

Comparison of reactivity data (see section 4.1) with core

binding energy data suggests a correlation of site of electrophilic attack with a lower binding energy, and hence a greater electron density, for pyrazole, imidazole and isoxazole. This is in accordance with a simple model of attack at the site of greatest negative charge thus suggesting an early transition state. However, the reactivities towards electrophiles of the single heteroatom compounds show the trend

Furan > Pyrrole > Thiophene

and a pattern of 2(5) substitution, indicating a difference from the above mentioned compounds. The binding energies for the adjacent (α) carbons are higher than those of the β carbons suggesting that the overall electron density at the α carbon must be considerably lower than at the β position. Comparison with the experimental results can be rationalised on the basis of either

(i) That the transition state is well advanced and hence does not approximate well to the isolated molecule plus electrophile, or

(ii) That if the transition state is relatively early, the interaction of the electrophile with the more polarisable π system is of greater importance.

4.4 Quantitative discussion

Non-empirical LCAO-MO-SCF calculations in comparable contracted gaussian basis sets have been reported on pyrrole,[137] thiophene,[138,139,140] isoxazole, pyrazole and imidazole.[141] These calculations refer to

isolated molecules, i.e. gas phase, whereas the measurements reported here were obtained from samples examined as thin films upon gold i.e. solid phase. For pyrazole and imidazole extensive hydrogen bonding undoubtedly exists in the crystal lattice, and the N_{1s} core levels in particular would be expected to be modified such that the shifts within a given molecule might become smaller. For example, ab initio calculation (i.e. application of Koopmans' theorem) predicts a shift between the N_{1s} levels of pyrazole of 2.7 eV, whereas the observed shift is only 1.3 eV. In N-methyl pyrazole, however, where hydrogen bonding is no longer possible, the shift in the N_{1s} levels is 2.3 eV (that this result is largely due to removal of hydrogen bonding is indicated by measurements made on N-methyl pyrrole, which suggest that the electronic effect of replacing hydrogen by methyl is quite small as far as the nitrogen is concerned.

Plots of the ab initio orbital energies, for the carbon and nitrogen core levels, against the experimental binding energies for pyrrole, pyrazole, imidazole and isoxazole are shown in Fig. 4.1. The correlation within a given molecule for C_{1s} levels is very good, though overall there is some scatter which may be due, in part, to differential charging effects between samples (though this should be minimised by studying the samples as thin films on a conducting backing). For the N_{1s} levels the correlation is also quite reasonable, however the calculated shifts are somewhat larger than those measured which, in the cases of pyrazole and imidazole, may be assumed to result from the hydrogen bonding present in the lattice.

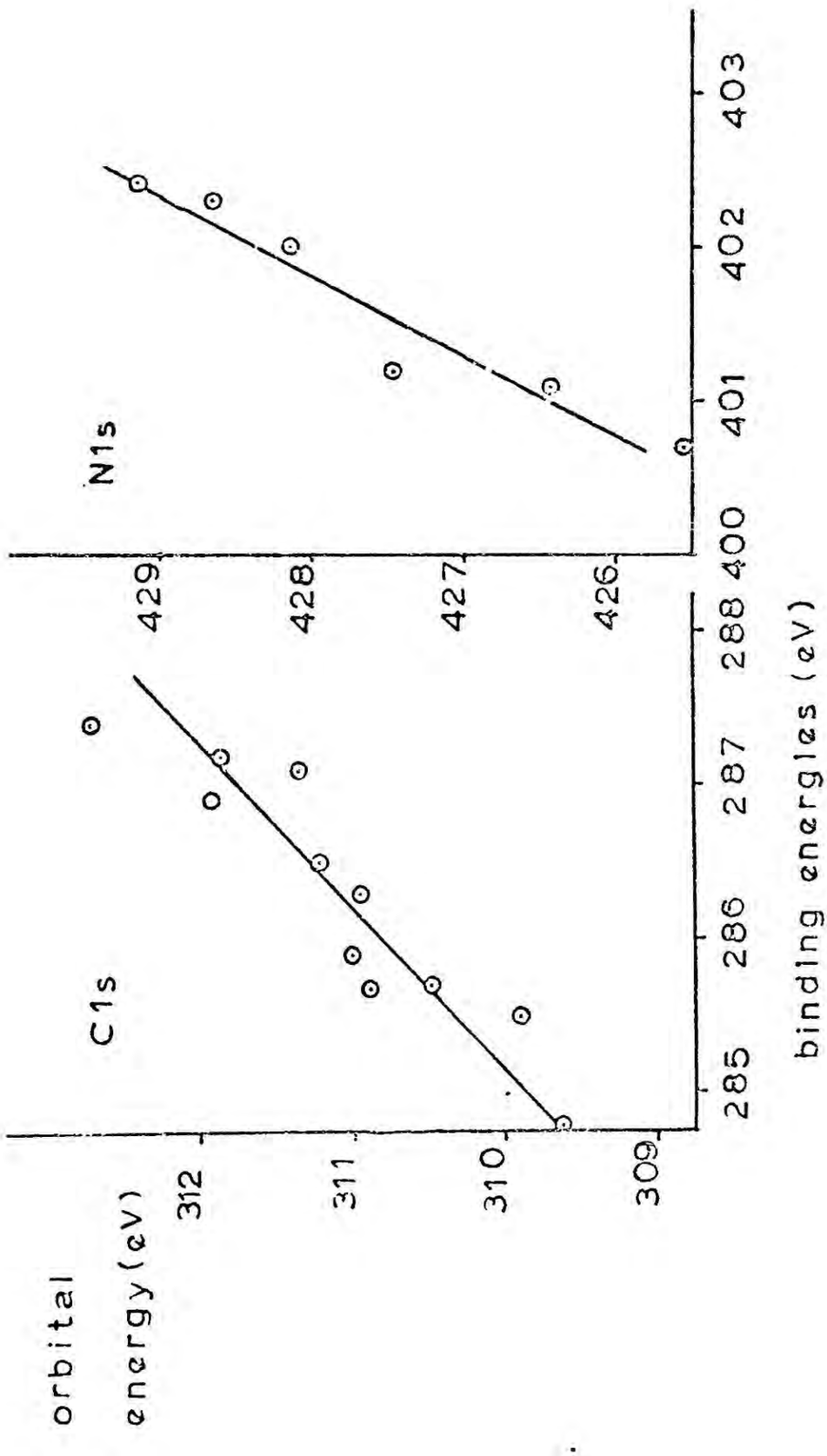


Fig. 4.1. Ab initio orbital energies against measured core binding energies of carbon and nitrogen atoms.

Two non-empirical calculations have been published on thiophene, which differ considerably in the predicted C_{1s} shift between C2(C5) and C3(C4). The calculations of Clark and Armstrong[138] suggest that these carbons should have very similar binding energies, whilst those of Siegbahn et al[139] predict a shift of 0.7 eV. The measured carbon core binding energy, however, shows a broadening of only 0.1 eV relative to that of benzene. Gelius et al[142] have measured the C_{1s} level in gaseous thiophene, thus with enhanced resolution, and have estimated the chemical shift between C2(C5) and C3(C4) to be 0.34 ± 0.12 eV. The best ab initio calculation on thiophene of Gelius et al[140] estimates a shift of 0.58 eV. The calculation of Clark and Armstrong is in best agreement with the experimental results, suggesting the possibility of the use of core binding energy measurements for the comparison of limited basis set calculations. The calculations of Clark and Armstrong employed two basis sets, one which included d orbital participation on the sulphur and one which did not. Their results indicated that d orbital participation in thiophene is not of importance and, since the correlation with the experimental data indicate that their basis sets were physically well suited to the calculation, this conclusion is likely to be correct.

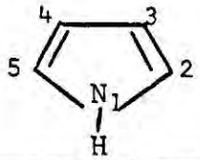
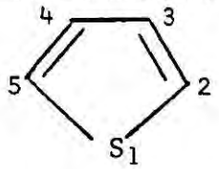
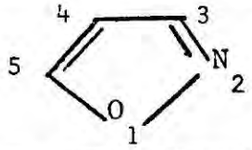
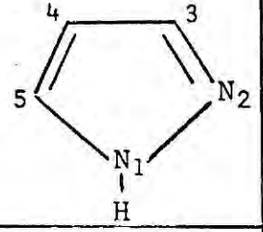
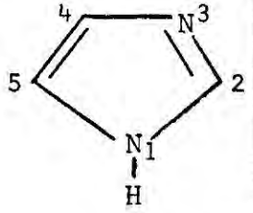
The non-empirical calculations have also been used to calculate Madelung potentials for these molecules, and hence express the core binding energies in terms of the ab initio charge distribution

$$E_i = E_i^o + kq_i + \sum_{i \neq j} \frac{q_j}{r_{ij}} \qquad (4.1)$$

The ab initio charges and Madelung potentials for the carbon and hetero atoms of these molecules are shown in Table 4.3, and Fig. 4.2 shows plots of experimental binding energies, corrected for Madelung potentials, against charge for the carbon and nitrogen core levels of pyrrole, pyrazole, imidazole and isoxazole. The charges are taken from Mulliken population analyses performed on these molecules. The correlations are reasonable and for the C_{1s} levels the value of k = 17.1 eV/unit charge may be compared with k = 18.8 eV/unit charge obtained from extended basis set calculations.[56] The N_{1s} levels show rather less scatter and give a value for k = 17.6 eV/unit charge.

It has also been of interest to perform CNDO/II calculations on this series of compounds, and use the atomic charge densities thus calculated to attempt an assignment of the experimental core binding energies. The CNDO/II charges and Madelung potentials are shown in Table 4.4. A value for k of 25 eV/unit charge has been found appropriate to CNDO/II charge distributions, in the series of acetyl compounds presented in the previous chapter and in other work.[143,144] Using this value together with the CNDO/II charge distributions assignments have been made for the series. The calculated relative binding energies, together with the experimental values, are displayed in Table 4.5. The assignment is in agreement with that by means of non-empirical calculation in the majority of cases, and in some cases the agreement of calculation and experimental results is remarkable. The agreement for pyrazole and imidazole is the poorest, and here again this is probably in part due to effects

Table 4.3 Non-empirical charges and Madelung Potentials

	Carbon atoms			Hetero atoms		
	atom	qi	$\sum_{j\neq i} \frac{q_j}{r_{ij}}$ (eV)	atom	qi	$\sum_{j\neq i} \frac{q_j}{r_{ij}}$ (eV)
Pyrrole 	2	-0.105	-1.281	1	-0.408	3.603
	3	-0.255	0.544			
Thiophene 	2	-0.396	3.221	1	-0.384	-4.952
	3	-0.249	-0.081			
	2*	-0.302	2.229	1*	0.227	-3.530
	3*	-0.242	0.355			
Isoxazole 	3	-0.157	1.003	1	-0.208	0.124
	4	-0.270	2.461	2	-0.032	-2.337
	5	-0.036	-0.512			
Pyrazole 	3	-0.157	-0.856	1	-0.348	3.513
	4	-0.266	0.467	2	-0.131	-2.177
	5	-0.114	-0.459			
Imidazole 	2	-0.001	-2.236	1	-0.445	5.072
	4	-0.169	-0.259	3	-0.260	-0.046
	5	-0.127	-0.482			

* Using an extended basis set which included 30 functions on sulphur

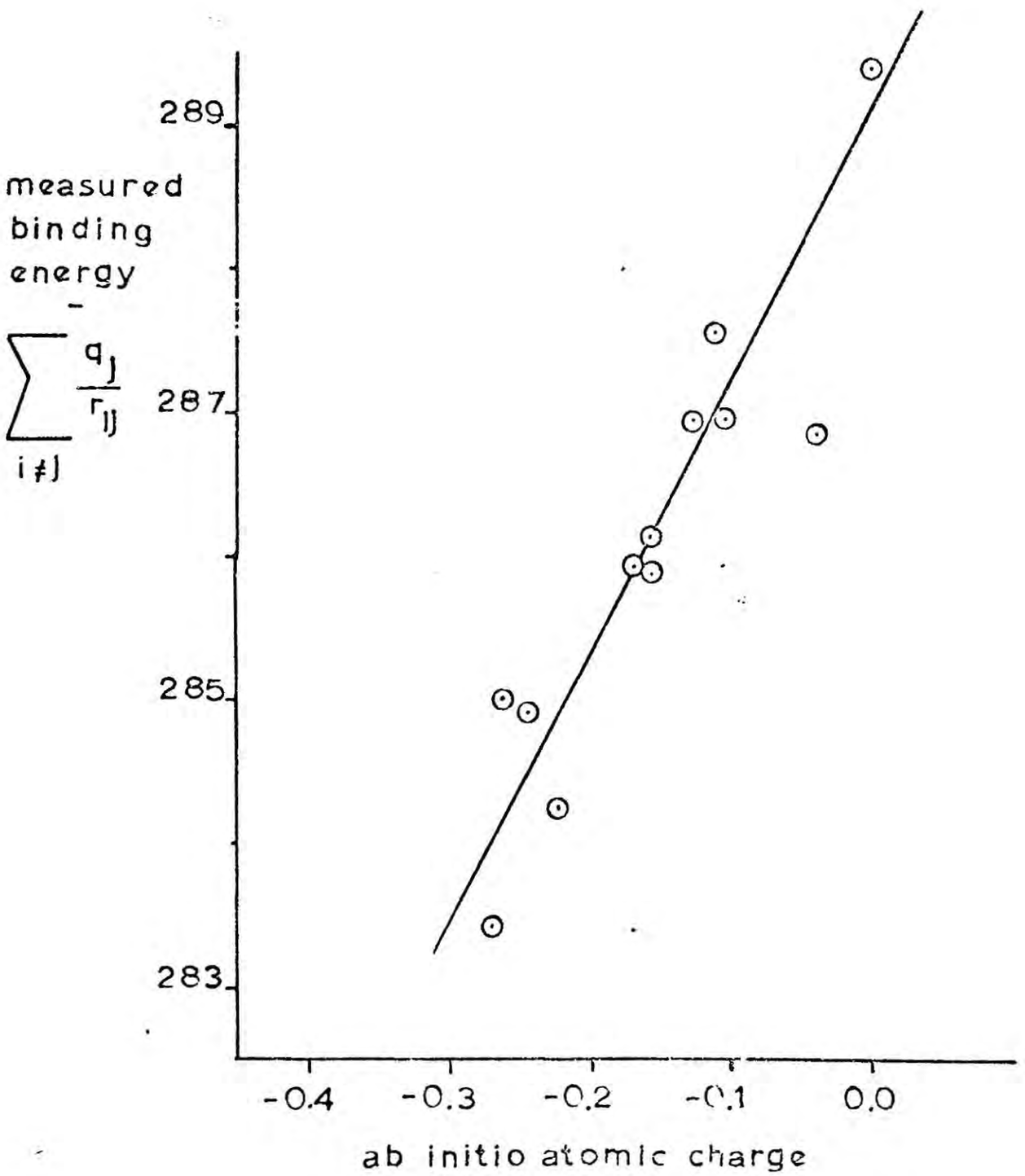


Fig. 4.2(a). Measured C_{1s} binding energies corrected for Madelung potentials against ab initio atomic charge.

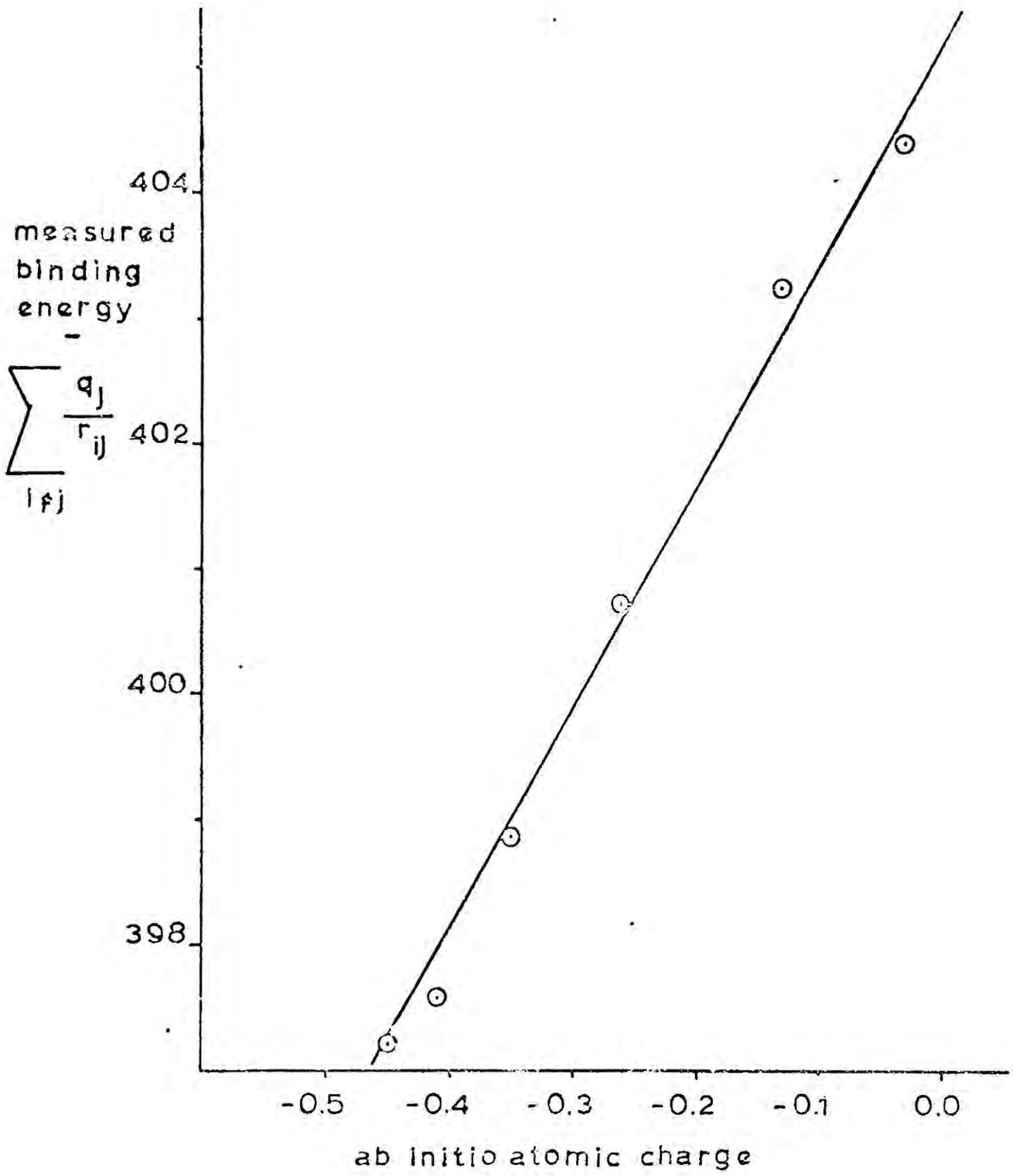


Fig. 4.2.(b) Measured N_{1s} binding energies corrected for Madelung potentials against ab initio atomic charge.

Table 4.4 CNDO/II charges and Madelung Potentials

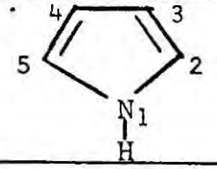
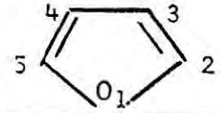
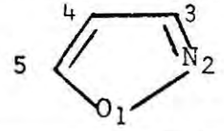
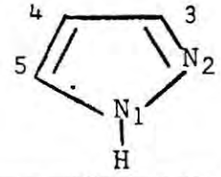
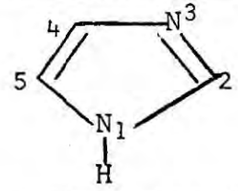
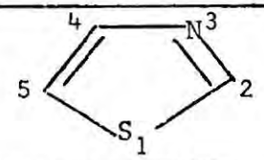
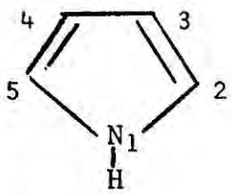
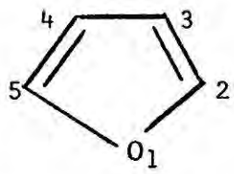
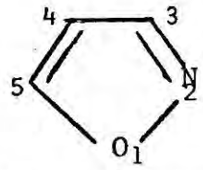
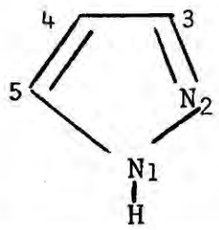
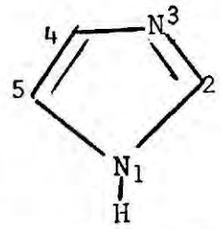
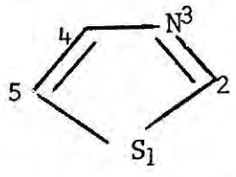
	Carbon atoms			Hetero atoms		
	atom	q_i	$\sum_{j\neq i}\frac{q_j}{r_{ij}}$(eV)	atom	q_i	$\sum_{j\neq i}\frac{q_j}{r_{ij}}$(eV)
Pyrrole 	2 3	0.058 -0.055	-1.007 0.212	1	-0.102	1.877
Furan 	2 3	0.101 -0.029	-1.476 0.573	1	-0.175	1.901
Isoxazole 	3 4 5	0.061 -0.039 0.103	-0.703 0.896 -1.192	1 2	-0.102 -0.067	0.719 0.212
Pyrazole 	3 4 5	0.081 -0.035 0.001	-0.809 -0.292 0.658	1 2	0.054 -0.217	0.175 2.54
Imidazole 	2 4 5	0.127 0.043 -0.010	-1.902 -1.239 -0.016	1 3	-0.093 -0.164	2.096 1.533
Thiazole 	2 4 5	0.069 0.103 -0.050	-1.012 -1.236 0.741	1 3	-0.063 -0.109	0.408 1.421

Table 4.5 Calculated and Experimental Relative Core Binding Energies of Carbon atoms (eV)

	Atom	Calculated Relative Binding Energy	Experimental
Pyrrole 	2	(0.0)	(0.0)
	3	-1.6	-0.9
Furan 	2	(0.0)	(0.0)
	3	-1.2	-1.1
Isoxazole 	3	(0.0)	(0.0)
	4	-1.0	-1.0
	5	+0.6	+0.5
Pyrazole 	3	+0.5	-0.8
	4	-1.9	-1.6
	5	(0.0)	(0.0)
Imidazole 	2	+1.4	+1.5
	4	(0.0)	(0.0)
	5	-0.1	+0.8
Thiazole 	2	(0.0)	(0.0)
	4	+0.6	+0.4
	5	-1.2	-0.7

resulting from the hydrogen bonding present in the solid structures. The agreement in the cases of furan and isoxazole is very close.

In Fig. 4.3 the Madelung potential corrected binding energies are plotted against CNDO/II charge for the C_{1s} levels. A good straight line is obtained, the slope of which indicates a value for k of 25.4 eV/unit charge.

A further point of interest is to examine the separate sigma and pi charges on the carbon and hetero atoms in these molecules, see Table 4.6. These were taken from the diagonal elements of the SCF bond order matrix of the CNDO/II output. It is interesting to note that quite large charges can be built up, although, in general, σ and π components of charge distribution tend to be mutually compensatory and thus the total charge at each centre remains comparatively small. The most pronounced effect occurs on the heteroatom centres, where the large negative charge which builds up in the sigma system is usually almost, but not totally, compensated by a slightly smaller positive charge on the pi system. In the case of the carbon atoms the magnitudes of these charges are smaller and generally a positive charge on the sigma system is matched by a slightly smaller negative charge on the pi system. This demonstrates the total invalidity of using an SCF method other than one including all the valence electrons for the calculation of charge densities and assignment of molecular core binding data. It also shows how methods which are basically pi electron SCF methods and do not consider sigma-pi interaction, such as the Del Re method, are also unfit for this use.

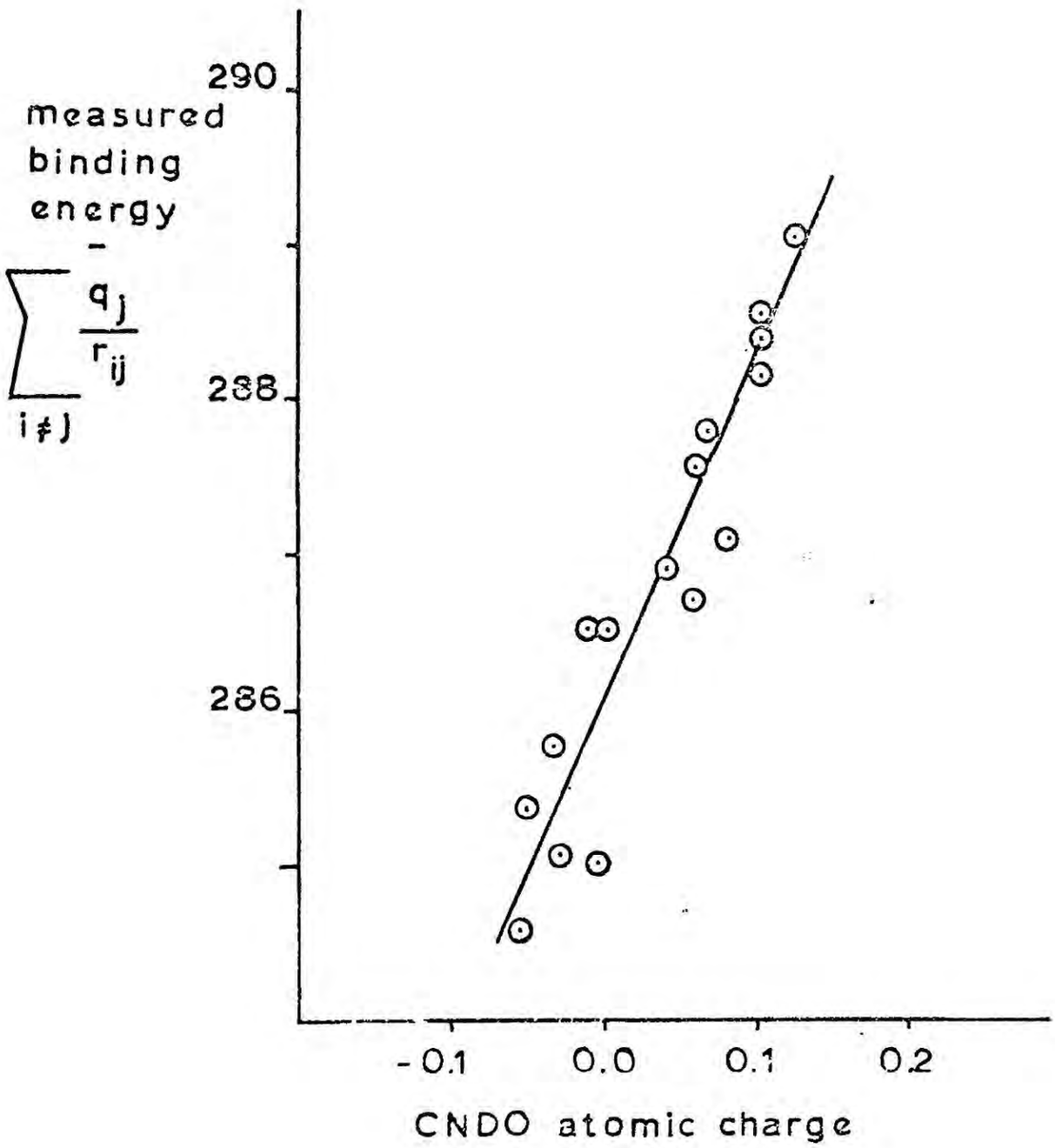
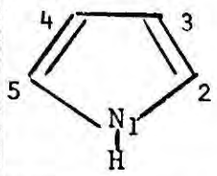
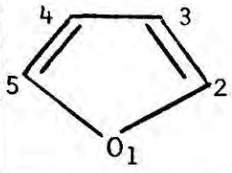
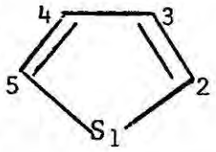
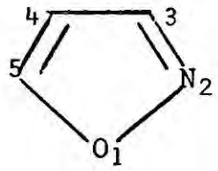
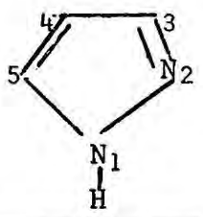
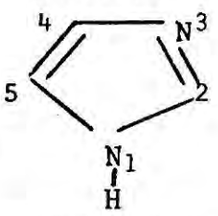
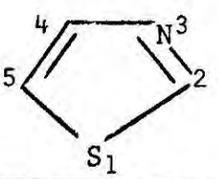


Fig. 4.3. Measured C_{1s} binding energies corrected for Madelung potentials against CNDO/II atomic charge.

Table 4.6 σ and π CNDO/II charges

	σ charges		π charges	
	Atom	Charge	Atom	Charge
Pyrrole 	1	-0.442	1	0.320
	2	0.133	2	-0.075
	3	0.030	3	-0.085
Furan 	1	-0.422	1	0.248
	2	0.167	2	-0.066
	3	0.029	3	-0.058
Thiophene 	1	-0.076	1	-0.016
	2	-0.007	2	-0.030
	3	0.016	3	0.022
Isoxazole 	1	-0.373	1	0.271
	2	0.132	2	-0.199
	3	0.041	3	0.021
	4	0.025	4	-0.064
	5	0.132	5	-0.029
Pyrazole 	1	-0.528	1	0.581
	2	0.225	2	-0.442
	3	0.049	3	0.032
	4	0.122	4	-0.087
	5	0.086	5	-0.085
Imidazole 	1	-0.502	1	0.409
	2	0.176	2	-0.049
	3	-0.008	3	-0.156
	4	0.107	4	-0.064
	5	0.130	5	-0.140
Thiazole 	1	-0.106	1	0.043
	2	0.047	2	0.022
	3	-0.051	3	-0.058
	4	0.069	4	0.034
	5	-0.008	5	-0.041

CHAPTER V

EXPERIMENTAL AND THEORETICAL STUDIES OF THE MOLECULAR CORE BINDING ENERGIES OF SOME PYRIMIDINE NUCLEIC ACID BASES AND RELATED COMPOUNDS OF BIOLOGICAL INTEREST

5.1 Introduction

The pyrimidine bases are of immense importance from a biological standpoint, as constituents of the nucleic acids, DNA and RNA, which are vital to the function and replication of living cells.[145] Some of the aza and fluoro substituted forms also have a possible use in the chemotherapy of cancer tissue,[146] where the rapidly growing cells incorporate them into nucleic acids resulting in a lethal synthesis. From a chemical point of view this series is interesting as it contains several smaller series within it, trends within which may be correlated with core binding energy data. The spectra have all been measured in the solid state and hence much valuable information on the form of these compounds in the solid state has been forthcoming. In particular greater insight into intermolecular hydration and tautomeric form has been obtained, information upon which could be important in biological studies.

5.2 Experimental

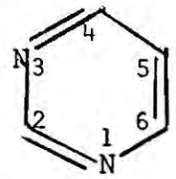
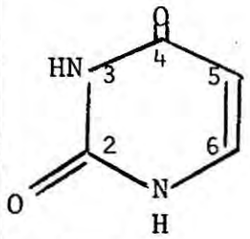
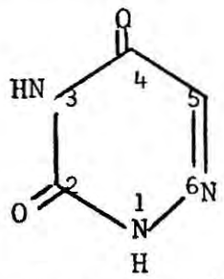
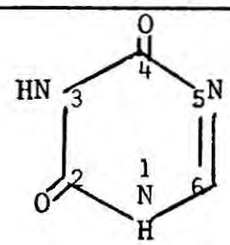
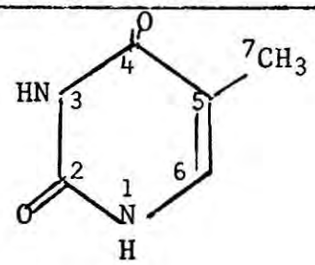
Spectra were recorded upon the AEI ES 100 spectrometer, using Mg $K\alpha_{1,2}$ radiation. Involatile samples were studied both as powders on Scotch tape and as pressed discs. Better resolution and counting statistics were obtained with the pressed disc technique. 1,3-dimethyl uracil was sufficiently volatile to be sublimed directly into the sample chamber and was studied as a thin film condensed on gold. Calibration of the energy scale for the samples studied as powders and pressed discs was accomplished by studying separately samples for which a residual C_{1s} signal from Scotch tape or from a

trace of hydrocarbon on the surface of the disc were observable. The C_{1s} binding energies for these two references were established separately with respect to the gold $4f_{7/2}$ level at 84 eV. Absolute binding energies are estimated to be accurate to ± 0.3 eV. Spectra, where necessary, were deconvoluted using the Du Pont 310 curve resolver, using line widths and line shapes derived from the previous studies of heterocyclic molecules. Assignments of core levels were made with the aid of theoretical calculations (see later).

5.3 Qualitative Discussion

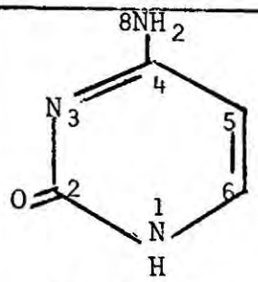
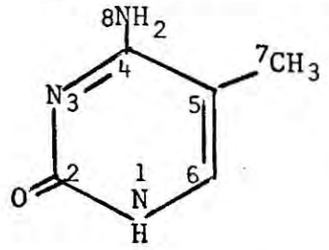
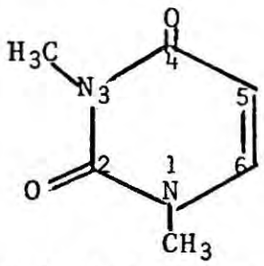
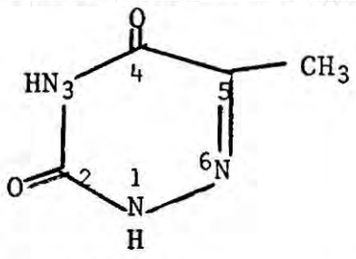
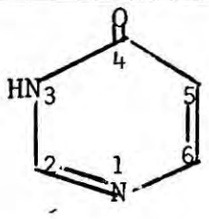
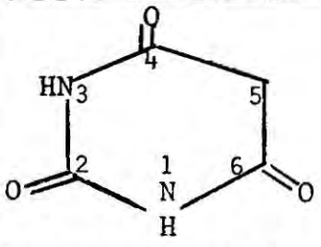
The measured binding energies are shown in Table 5.1 and are also shown diagramatically in Fig. 5.1. The various structural series which may be investigated are shown in Fig. 5.2 where a connecting line indicates a particular modification relating the two species. In general shifts in core binding energies both within a given molecule and with respect to another molecule may be rationalised on the basis of simple electronegativity considerations. Thus in pyrimidine and 1.3 dimethyl uracil, for which there are no complications due to other possible tautomeric forms, the most tightly bound carbon 1s levels are those attached to the most electronegative atoms. In pyrimidine, C3 attached to two nitrogen atoms is more tightly bound than C4(C6) which is only attached to one nitrogen, which is, in turn, more tightly bound than C5. In 1.3 dimethyl uracil, the carbonyl carbons attached to both oxygen and nitrogen are the most tightly bound, followed by the methyl carbons attached to nitrogen.

Table 5.1 Measured molecular core binding energies (eV).

	Pos^n	C_{1s}	ΔC_{1s}*	N_{1s}	ΔN_{1s}*
 Pyrimidine	2 4(6) 5 1(3)	287.4 286.7 285.6	(0.0) (0.0) (0.0)	 400.5	 (0.0)
 Uracil	2 4 6 5 1 3	290.0 289.0 286.9 285.4	2.6 2.3 0.2 -0.2	 401.5 401.1	 1.0 0.6
 6-aza-uracil	2 4 5 1 6 3	289.7 288.8 286.2	2.3 2.1 0.6	 401.9 401.5 401.1	 1.4 - 0.6
 5-aza-uracil	2 4 6 1 3 5	289.8 289.5 287.9	2.4 2.8 1.2	 401.8 401.1 400.0	 1.3 0.6 -
 Thymine	2 4 6 5 7 1 3	289.7 288.5 286.6 285.7 285.3	2.3 1.8 -0.1 0.1 -	 401.6 400.9	 1.1 0.4

* These values show the shift in binding energy with respect to the corresponding atom in pyrimidine, where relevant.

Table 5.1 Continued

	Pos^n	C_{1s}	ΔC_{1s}*	N_{1s}	ΔN_{1s}*
 Cytosine	2	289.3	1.9		
	4	288.2	1.5		
	6	287.0	0.3		
	5	285.8	0.2		
	1			401.3	0.8
	8			400.6	-
	3			399.8	-0.7
 5-methyl cytosine	2	289.2	1.8		
	4	288.1	1.4		
	6	286.9	0.2		
	5	286.1	0.5		
	7	285.6	-		
	1			401.7	1.2
	8			400.6	-
	3			399.6	-0.9
 1.3-dimethyl uracil	2	288.9	1.5		
	4	287.9	1.2		
	N-methyl	286.9	-		
	6	286.8	0.1		
	5	285.3	-0.3		
	1			401.3	0.8
	3			400.8	0.3
 6-aza thymine	2	289.8	2.4		
	4	289.0	2.3		
	5	286.6	1.0		
	7	285.4	-		
	1			402.3	1.8
	6			401.8	-
	3			401.5	1.0
 4-oxy pyrimidine	4	288.8	2.1		
	2	287.9	0.5		
	6	286.7	0.0		
	5	286.0	0.4		
	3			401.5	1.0
	1			400.2	-0.3
 Barbituric acid	2	290.1	2.7		
	4	289.2	2.5		
	5	286.2	0.6		
	1(3)			401.6	1.1

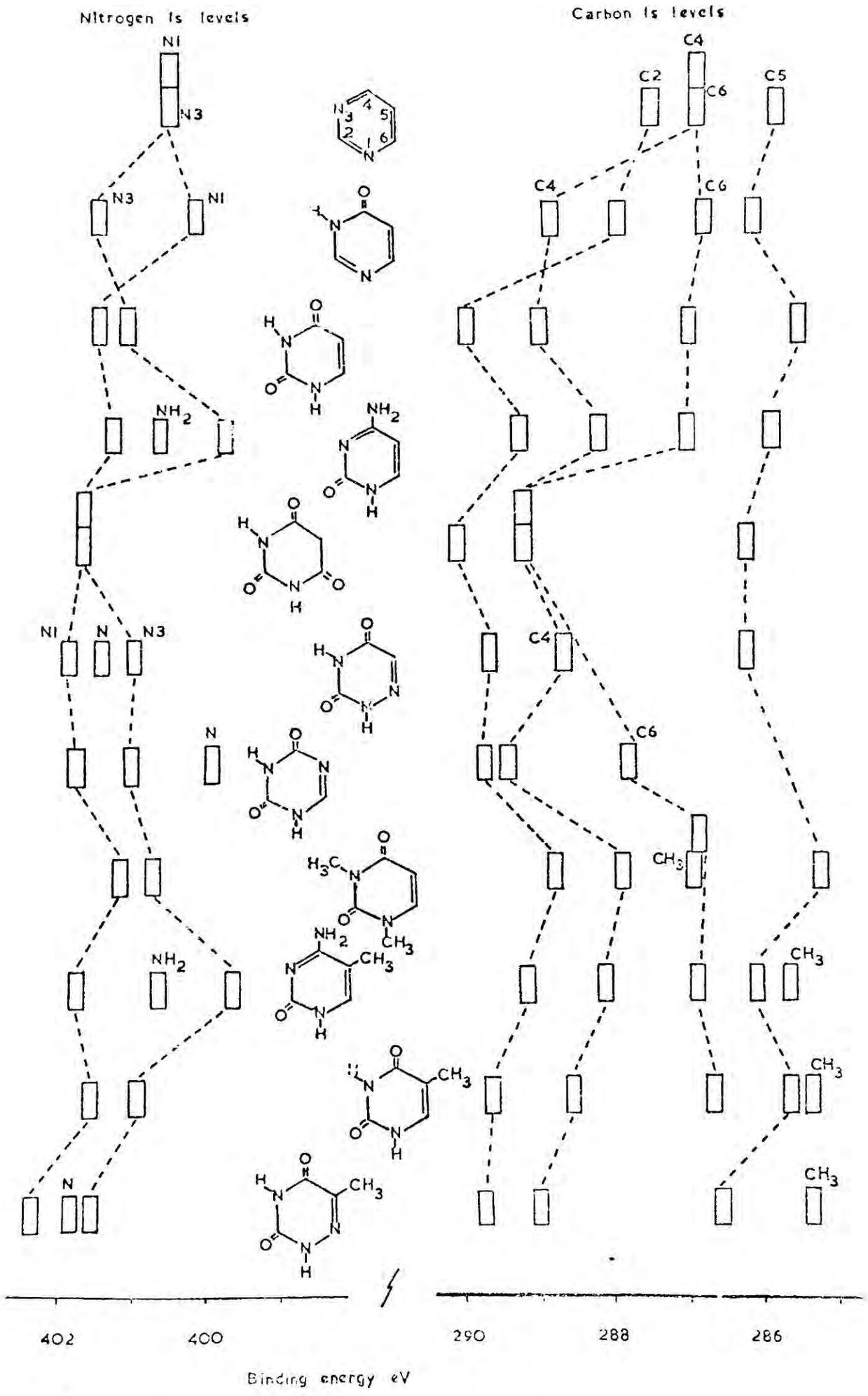


Fig. 5.1. Measured C_{1s} binding energies.

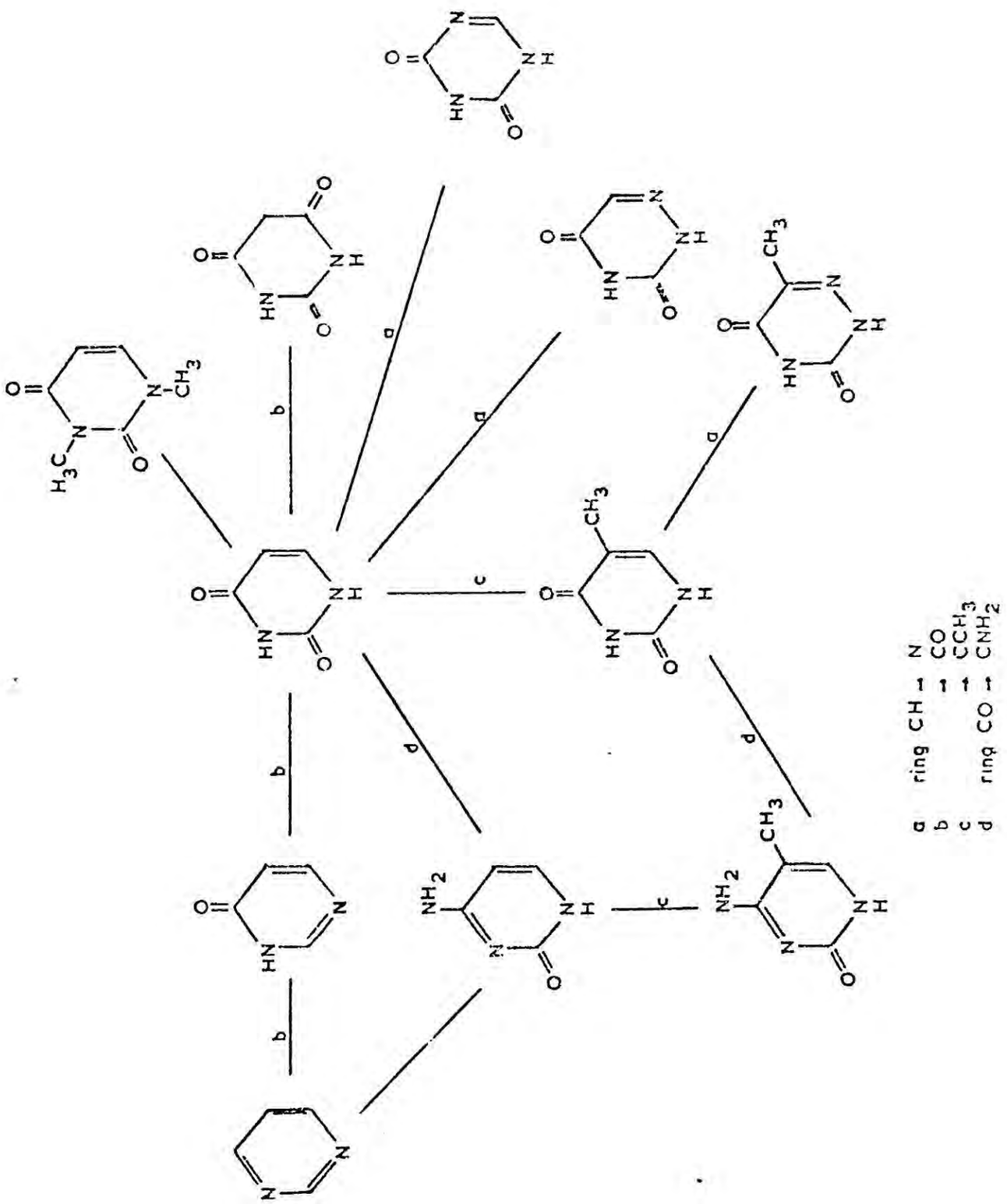


Fig. 5.2. Structural sub-series present within the main series of pyrimidine compounds.

For the other molecules there is an added complication due to the possibility of tautomerism. X-ray diffraction studies have indicated that in the solid state uracil,[147] thymine[148] and barbituric acid[149] exist in the keto form, as indicated in Fig. 5.2. The previous studies on heterocyclic molecules indicated that the core level for a'pyrrole type' nitrogen with hydrogen or an alkyl group attached is more tightly bound than for a pyridine type nitrogen. This is also evident in comparing 1.3 dimethyl uracil (lactam or pyrrole type nitrogen) with pyrimidine (pyridine type nitrogen). The higher binding energies for the core levels of the nitrogens in uracil as compared with pyrimidine then suggest that it possesses the keto structure in the solid state. This is unambiguously settled by a consideration of the C_{1s} levels, where comparison with pyrimidine shows that C2 and C4 shift to higher binding energy by 2.6 eV and 2.3 eV respectively. This is inconsistent with a formulation as a dihydroxy pyrimidine since the shifts expected for replacing ring hydrogens by hydroxyl groups are ~1.5 eV (e.g. diethyl ether). The absolute binding energies for both C_{1s} and O_{1s} levels are, however, in close agreement with those found in the previous study of acetamide. These arguments also apply to thymine and barbituric acid and will be quantified in a later section.

Thus in each case the observation of a single nitrogen with high binding energy (corresponding to lactam type nitrogen) and characteristic high C_{1s} and low O_{1s} binding energies associated with the C = O structural feature confirms the keto structure in the solid state. For the aza substituted uracils similar arguments apply and in each case a keto structure in the solid state is predicted.

Examination of Fig. 5.2 reveals that several structural series are evident, as mentioned earlier, and these are conveniently discussed separately.

(i) Effect of replacement of ring CH by N.

Considering first the series uracil, 5 aza uracil and 6 aza uracil, the shifts in core binding energies may, in general, be understood in terms of simple electronegativity considerations. Thus in going from uracil to 5 aza uracil the core binding energy of C6 increases by 1.0eV, whilst replacement of CH by nitrogen meta to the two lactam type nitrogens has little effect on their N_{1s} core binding energies which are closely similar to those in uracil. The core level for C4 shifts by a smaller amount (0.5 eV) to higher binding energy. For 6 aza uracil the largest shifts are again for the atoms directly bonded to the CH group replaced by nitrogen. Thus N_1 and C_5 both increase in binding energy by 0.4 eV and 0.8 eV respectively. The distinction between the structurally isomeric 5 and 6 aza uracils can be made quite easily from their spectra, on the basis of either the C_{1s} and N_{1s} core levels. Similar trends are apparent in going from thymine to 6 aza thymine, and the atoms directly attached to the 6 position undergo substantial increases in binding energy.

(ii) Effect of changing ring CH environment to a carbonyl group.

This is shown quite strikingly in Table 5.1 and Fig. 5.1. The average shift in C_{1s} binding energy in changing pyrimidine type ring CH to a carbonyl group is 2.4 eV, which may be compared with a shift of 1.5 eV when a ring hydrogen is replaced by an amino group.

(iii) Effect of methyl substitution.

Methyl substitution has a relatively small effect when the group is attached to carbon; the net effect being to raise the binding energy of the attached carbon by 0.4 eV. N-methylation of uracil has a much larger overall effect, lowering the binding energies of the C2 and C4 C_{1s} levels by 1.1 eV, and of the N_{1s} levels by 0.3 eV. Previous work on N-methyl pyrrole and pyrrole itself, both studied as thin films on gold, suggest that the electronic perturbation of the methyl group is relatively small. The much larger effect apparent in going from uracil to 1.3 dimethyl uracil is most readily understandable in terms of the extensively hydrogen bonded structure of uracil, which was studied as a pressed disc. It is also significant that the O_{1s} binding energies in 1.3 dimethyl uracil are somewhat higher than for the other compounds.

Using normally accepted line widths it has not been possible to obtain deconvolutions of the O_{1s} spectra, which must be a result of extensive intermolecular hydrogen bonding and hydration. Hence the centroids of the relatively broad peaks have been measured and are shown in Table 5.2. Due to the lack of deconvolution analysis of these data is difficult but a few features are worthy of comment. For the majority of these compounds the centroids lie close together in the region 533.1 - 533.5 eV but there are a few notable exceptions. Cytosine, with its single ketonic oxygen atom shows a lower binding energy than the average, 0.4 eV below that in uracil, whilst barbituric acid, with the three oxygen atoms has an average

Table 5.2 Mean O_{1s} binding energies

Compound	Binding Energy
Uracil	533.1
6-aza uracil	533.1
5-aza uracil	533.3
Thymine	533.1
Cytosine	532.7
5-methyl cytosine	533.1
Dimethyl uracil	533.5
6-aza thymine	533.3
Barbituric acid	533.8
4-oxypyrimidine	533.4

oxygen binding energy of 0.7 eV greater than that in uracil. These results are in accord with predictions based upon simple electronegativity considerations. Substitution of a ring CH by N is seen to have little effect upon the O_{1s} binding energy, although the substitution in the 5 position of uracil, where it might be expected to exert its greatest effect, is seen to raise the mean binding energy by 0.2 eV.

5.4 Quantitative Discussion

When this work was performed it was not feasible to perform ab initio calculations upon a series of molecules of this size, although a few calculations have been performed upon some members of

the series in other laboratories and subsequently in this laboratory. In many cases core binding energy assignments are evident by comparison with other series of molecules and by simple electronegativity considerations, and these have been confirmed by non-empirical calculations on certain key molecules. Non-empirical calculations are available for thymine and cytosine, and in the case of 5-aza uracil an STO 4-31 G calculation has been carried out, and these provide valuable comparison with assignments based on the charge potential model using CNDO/II charges.

The calculated charges and Madelung potentials are tabulated in Table 5.3, and a plot of binding energy, corrected for the Madelung potential, against charge for the C_{1s} and N_{1s} levels is shown in Fig. 5.3.

For the C_{1s} levels, correlation is extremely good ($r^2 = 0.998$) and the calculated slope of 22.8 eV/unit charge is in good agreement with the theoretical value of 22.0. This can be compared with values of k found in previous and other work: for the acetyl compounds and also a series of binuclear aromatic compounds and their perfluoro derivatives[143] a value of 25.0 eV/unit charge has been calculated, and for a series of six membered nitrogen heterocycles and their perchloro and perfluoro derivatives[150] a value of 22.4 has been observed.

The correlation for the N_{1s} levels is rather less meaningful ($r^2 = 0.716$) and the calculated slope of 16.1 ± 2.3 eV/unit charge is lower than would be expected. It may be inferred from these

Table 5.3 CNDO/II charges, Madelung Potentials and Assignments

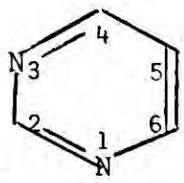
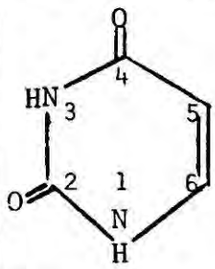
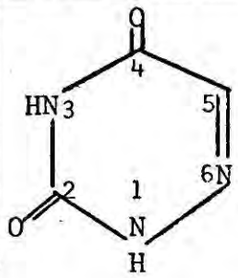
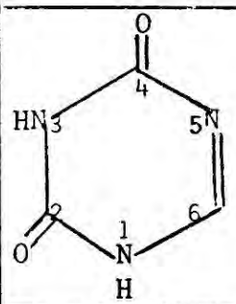
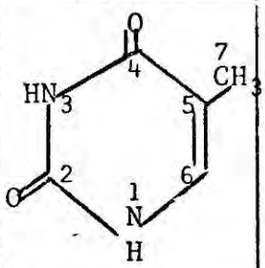
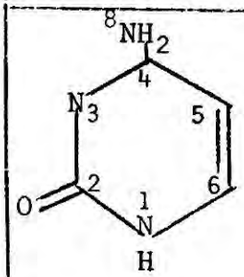
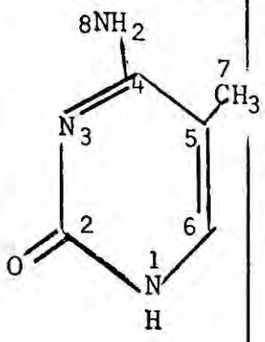
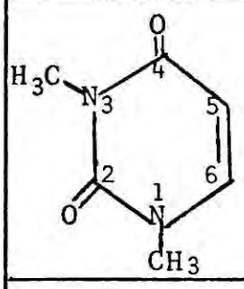
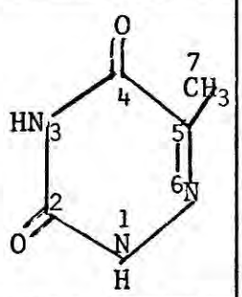
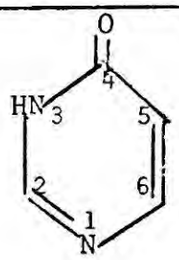
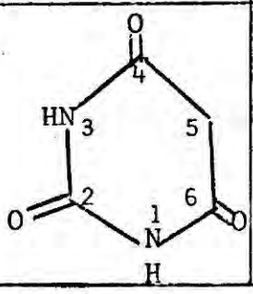
	Pos^n	q_i	$\sum_{i\neq j} \frac{q_i}{r_{ij}}$ (eV)	kq_i* (eV)	$(E-E^o)$* (eV)	kq_i† (eV)	$(E-E^o)$† (eV)
	2	0.203	-3.00	5.08	2.08		
	4	0.142	-1.81	3.55	1.74		
	5	-0.062	1.65	-1.55	0.10		
	1	-0.180	2.63	-4.50	-1.87	-3.60	-0.97
	2	0.456	-5.35	11.40	6.06		
	4	0.389	-5.70	9.73	4.03		
	6	0.174	-0.80	4.35	3.55		
	5	-0.175	3.76	-4.38	-0.62		
	1	-0.136	5.05	-3.40	1.65	-2.72	2.33
	3	-0.183	5.23	-4.58	0.65	-3.66	1.57
	2	0.440	-5.54	11.00	5.47		
	4	0.349	-4.61	8.73	4.11		
	5	-0.057	2.28	-1.43	0.86		
	1	-0.127	4.51	-3.18	1.34	-2.54	1.97
	3	-0.234	5.86	-5.85	0.01	-4.68	1.18
	6	0.007	0.69	0.18	0.86	0.14	0.83
	2	0.455	-5.50	11.38	5.88		
	4	0.439	-6.40	10.98	4.57		
	6	0.272	-2.14	6.80	4.67		
	1	-0.177	6.29	-4.43	1.86	-3.54	2.75
	3	-0.234	6.57	-5.85	0.72	-4.68	1.89
	5	-0.282	5.11	-7.05	-1.94	-5.64	-0.53
	2	0.446	-5.81	11.15	5.34		
	4	0.358	-5.02	8.95	3.93		
	6	0.134	-0.45	3.35	2.90		
	5	-0.101	2.80	-2.53	0.27		
	7	0.014	0.14	0.35	0.49		
	1	-0.173	5.37	-4.33	1.04	-3.46	1.91
	3	-0.231	5.30	-5.78	-0.48	-4.62	0.68
	2	0.425	-7.36	10.63	3.26		
	4	0.316	-4.41	7.90	3.49		
	6	0.183	-1.46	4.58	3.11		
	5	-0.165	3.40	-4.13	-0.73		
	1	-0.155	4.39	-3.88	0.51	-3.10	1.29
	3	-0.374	4.76	-9.35	-4.59	-7.48	-2.72
	8	-0.201	4.44	-5.03	-0.59	-4.02	0.42

Table 5.3 continued

	Posn	q_i	$\sum_{i \neq j} \frac{q_i}{r_{ij}}$ (eV)	kq_i* (eV)	(E-E°)* (eV)	kq_i† (eV)	(E-E°)† (eV)
	2	0.424	-7.36	10.63	3.26		
	4	0.305	-4.12	7.63	3.50		
	6	0.159	-1.11	3.98	2.86		
	5	-0.114	2.82	-2.85	-0.03		
	7	0.006	0.33	0.15	0.48		
	1	-0.156	4.30	-3.90	0.40	-3.12	1.18
	3	-0.371	4.73	-9.28	-4.54	-7.42	-2.69
	8	-0.202	4.48	-5.05	-0.57	-4.04	0.44
	2	0.439	-5.71	10.98	5.26		
	4	0.363	-5.40	9.08	3.68		
	6	0.161	-0.92	4.03	3.11		
	5	-0.168	3.50	-4.20	-0.70		
	1	-0.122	0.94	-3.05	-2.11	-2.44	-1.50
	3	-0.188	-0.44	-4.70	-5.14	-3.76	-4.20
	2	0.435	-5.70	10.95	5.25		
	4	0.339	-4.36	8.48	4.12		
	5	0.006	1.34	0.15	1.49		
	7	-0.007	0.92	-0.18	0.74		
	1	-0.131	4.24	-3.28	0.97	-2.62	1.62
	3	-0.234	5.76	-5.85	-0.09	-4.68	1.08
	6	-0.034	1.08	-0.85	0.23	-0.68	0.40
	2	0.23	-2.18	5.75	3.58		
	4	0.36	-5.42	9.00	3.58		
	6	0.15	-1.96	3.75	1.79		
	5	-0.14	2.71	-3.50	-0.79		
	1	-0.22	3.47	-5.50	-2.03	-4.40	-0.92
	3	-0.16	4.51	-4.00	0.51	-3.20	1.31
	2	0.45	-5.78	11.25	5.47		
	4	0.36	-4.18	9.00	4.82		
	6	0.37	-4.39	9.25	4.86		
	5	-0.14	4.56	-3.50	1.06		
	1	-0.26	6.59	-6.50	0.09	-5.20	1.39
	3	-0.25	6.40	-6.25	0.15	-5.00	1.40

* Based upon $k_c = k_n = 25.0$

† Based upon $k_n = 20.0$

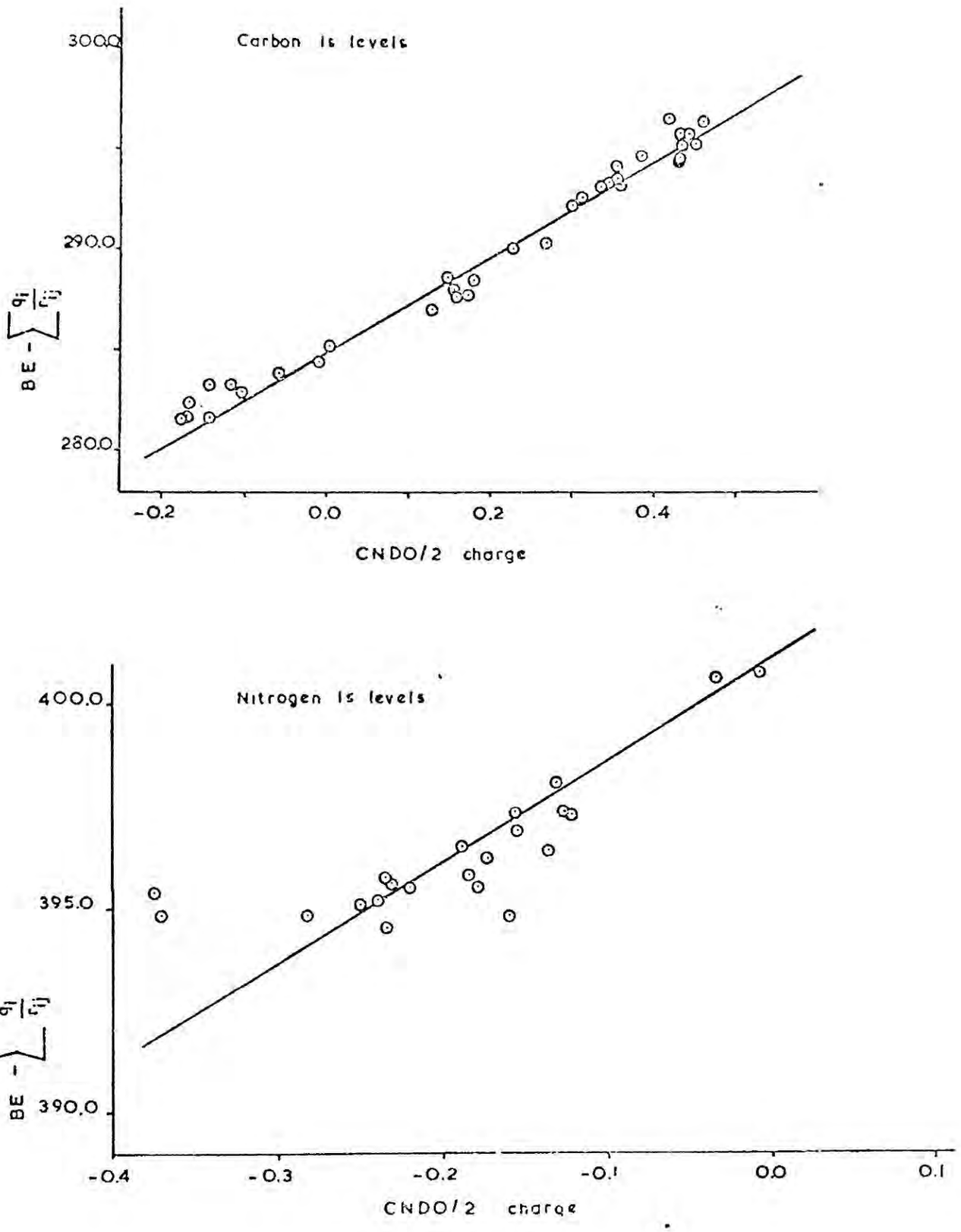


Fig. 5.3. Measured carbon and nitrogen core binding energies corrected for Madelung potential against CNDO/II charge.

smaller shifts that these molecules show considerable hydrogen bonding in the solid state, as is evident from other experimental data. It has been noted previously[151] that the K value for nitrogen atoms is heavily dependent upon the environment of those atoms. Thus if the correlation is made for a series of N1 binding energies with a uracil lactam type environment then the correlation becomes better and the K value more reasonable i.e.

$$k = 21.94 \pm 4.1 \text{ eV/unit charge} \qquad (r^2 = 0.88)$$

In view of the apparent variability of the k value for nitrogen within the series as a whole, the assignments for the nitrogen atoms have been recalculated using a value of k_N = 20.0 eV/unit charge, and the appropriate values of kq_i and $(E - E^o)$ are shown in Table 5.3. However, in no case does this charge result in differences in assignment.

A further point of interest may be found on examination of Table 5.3. For purposes of assignment, a value of k = 25 eV/unit charge was used for both carbon and nitrogen, and the kq values are shown in the table. In most instances the Madelung potential term is merely a correction to the charge dependence, however, in some cases it may become comparable in magnitude, or even dominant. This is often the case at C5 through the series, largely as a result of the small charge on the atom, being remote from the more electronegative nitrogen and oxygen atoms.

Pullman[152] has performed ab initio calculations upon cytosine and thymine and it is interesting to analyse these results and correlate them with the experimental results quoted here. Fig. 5.4 shows a plot of ab initio orbital energy against measured binding energies for the C_{1s} and N_{1s} levels. The correlation for the carbon levels is reasonable, though there is a lot of scatter. This would be expected with the approximations inherent in the use of Koopmans' theorem. The correlation for the nitrogen atoms are very bad, however, and it would appear that the points from the two molecules have separate correlations. This could arise from differences in charging effects between the two compounds, but is more likely to be a result of the limited basis set used in the calculations and differences in hydrogen bonding. The validity of assignments based upon the application of Koopmans' theorem clearly rests upon an assumption of a constant value for the reorganisation energies of the atoms concerned. In the case of pyridine recent work[153] indicates that relaxation energies for different sites within the molecule are likely to be closely similar. Detailed studies of hole states of pyridine and comparison with Koopmans' theorem showed that for the C_{1s} levels the reorganisation energies at different sites are all within 0.1 eV. Consequently, it is likely that Koopmans' theorem should provide a good description of the shifts in core binding energies for pyrimidine also. A recent ab initio calculation upon pyrimidine is available[154] and shifts in C_{1s} levels calculated with an assumption of Koopmans' theorem are in reasonable agreement

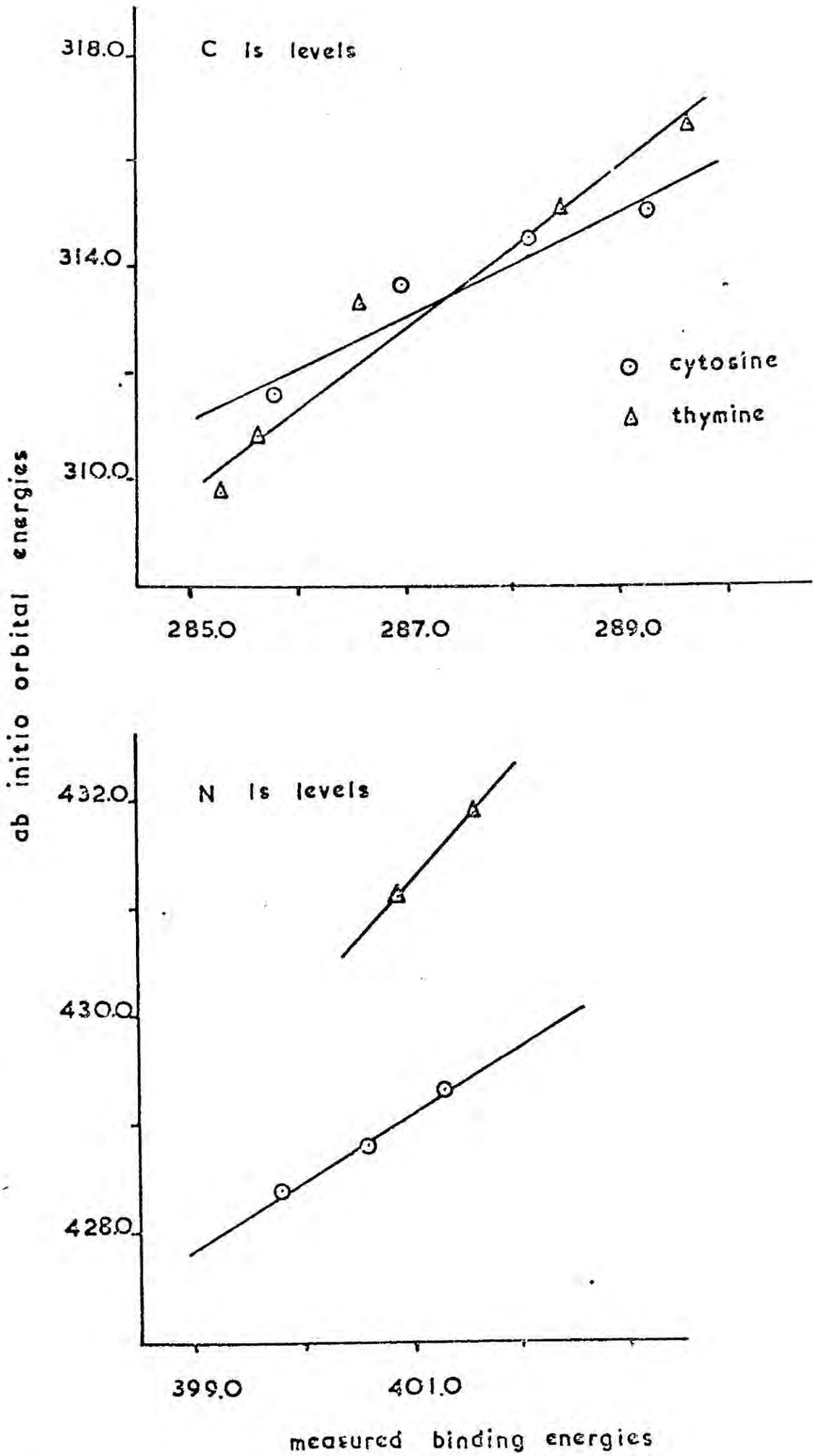
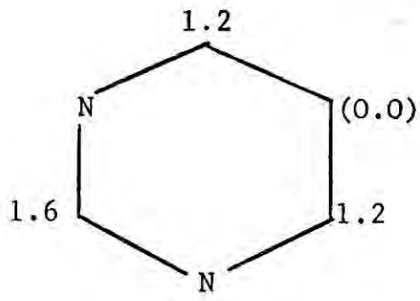
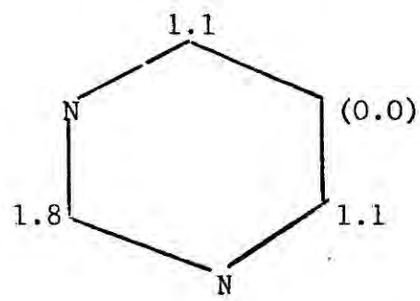


Fig. 5.4. Orbital energy of C_{1s} and N_{1s} against measured core binding energies for cytosine and thymine.

with measured core binding energy shifts. These non-empirical



Ab initio



Measured

data, together with the thymine and cytosine calculations and also an STO 4 - 31 G calculation for 5-aza uracil, obtained in this laboratory, have been used to assign C_{1s} and N_{1s} core levels in these compounds by means of Koopmans' theorem. These are shown in Table 5.4 along with the CNDO/II charge potential assignments, and it may be seen that agreement is virtually complete, thereby allowing further confidence in the charge potential model as a means of assignment.

The ab initio cytosine and thymine results have also been used for the calculation of Madelung potentials and core shifts, tabulated in Table 5.5, and plots of measured binding energies, corrected for the Madelung potential, against the atomic charges are shown in Fig. 5.5. The correlations are reasonably good, giving the values k_C = 20.3 eV/unit charge and k_N = 23.3 eV/unit charge.

5.5 Tautomeric Form

Information concerning which tautomeric form a particular molecule is in in the solid state is of considerable interest in

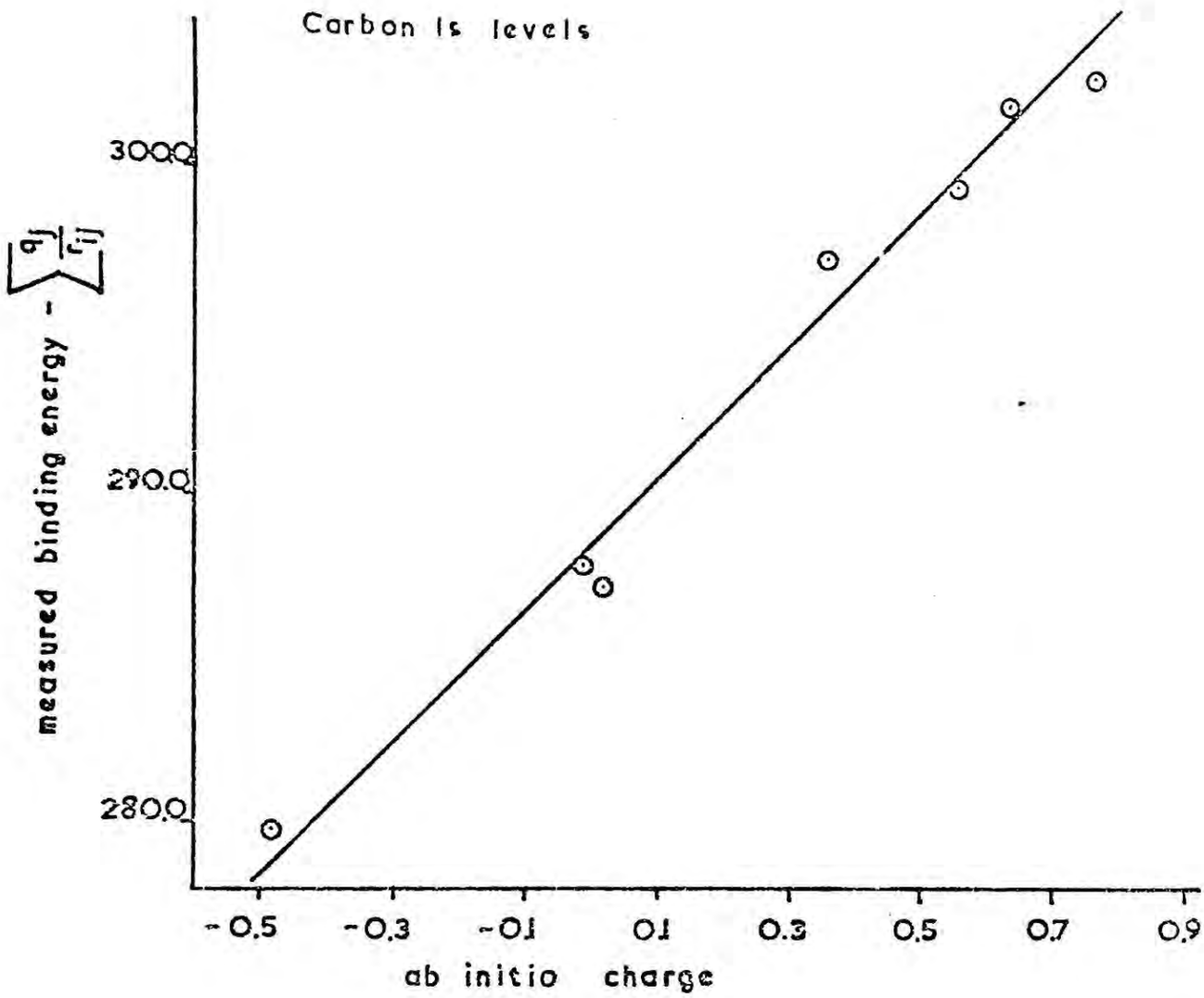
Table 5.4 Core binding energy assignments for Cytosine, Thymine and 5-aza Uracil based upon the charge potential model using CNDO/II populations and comparison with those derived from non-empirical calculations with an assumption of Koopmans' Theorem.

Compound	Assignments (descending order of chemical shift)	
	Charge Potential (CNDO/II)	Koopmans' (ab initio)
Cytosine[152]	2	2
	6	6
	4	4
	1	1
	3	3
	5	5
Thymine[152]	2	2
	4	4
	6	6
	7	5
	5	7
	1	1
	3	3
5-aza Uracil	2	2
	6	6
	4	4
	1	1
	3	3
	5	5
Pyrimidine[154]	2	2
	4(6)	4(6)
	5	5
	1(3)	1(3)

Table 5.5 Ab initio charges, Madelung potentials and calculated core binding energy shifts for thymine and cytosine

	Atom	q_i	$\sum_{i \neq j} \frac{q_j}{r_{ij}}$	$E - E^o$ *
Thymine				
	C 2	0.77	-13.11	4.60
	4	0.56	-10.30	2.58
	6	0.02	0.28	0.74
	5	-0.14	- 0.62	-3.84
	N 1	-0.71	11.95	-2.25
	3	-0.72	11.26	-3.14
Cytosine	C 2	0.64	-12.50	2.22
	4	0.36	- 9.00	-0.72
	5	-0.01	6.00	5.77
	6	-0.48	- 0.94	-11.98
	N 1	-0.60	9.37	-2.63
	7	-0.72	4.22	-10.18
	3	-0.46	10.86	1.66

* $k_C = 23.0$
$k_N = 20.0$



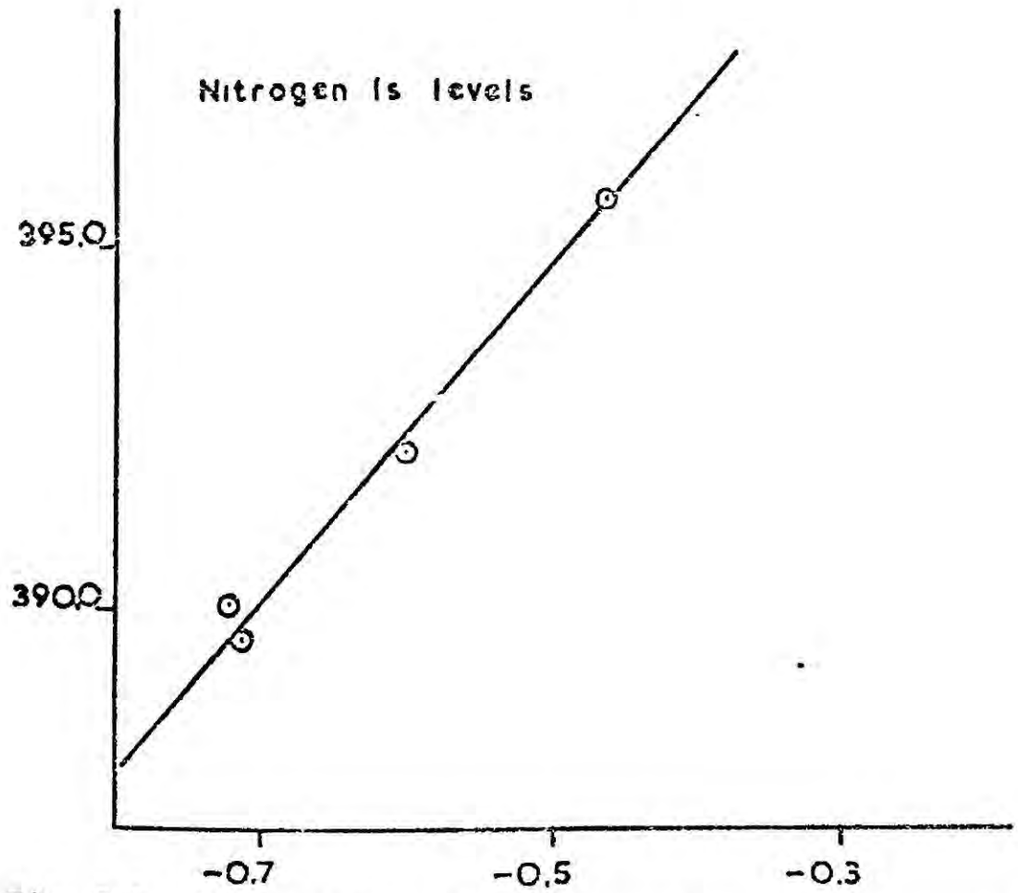
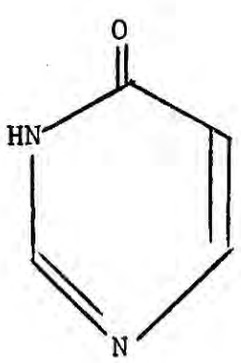


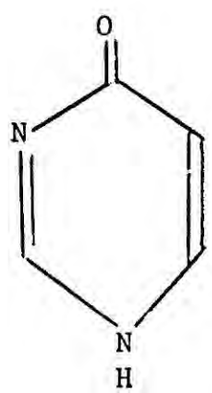
Fig. 5.5. Measured core binding energies corrected for Madelung potential against ab initio charge.

biology. The assumption of the usually accepted tautomeric forms for the purine and pyrimidine bases of the nucleic acids is essential for the understanding of the fidelity of replication, transcription and translation of the genetic code. Furthermore the possible existence of other tautomeric forms allows some insight into possible means of chromosomal mutation. It is also interesting to speculate that 'unusual' tautomeric forms of these bases may be involved in the control of protein synthesis in the mammalian cell, together with the associated problem of differentiation. There is a certain amount of experimental evidence for tautomerism in the pyrimidines. NMR evidence tends to suggest that uracil and thymine exist virtually totally in the di-keto form but data suggest that cytosine[155,156,157] exhibits several tautomeric forms, depending upon the solvent in which it is dissolved or from which it is crystallised. NMR[158] has also been used to study the equilibrium between 4-oxypyrimidine and its enolic tautomer.

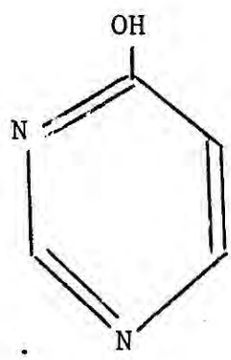
In the case of 4-oxypyrimidine, in solution an equilibrium is established between the tautomeric forms I and II with I being the major component. The substantial differences in binding energies



I



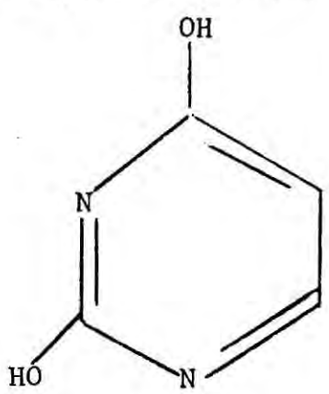
II

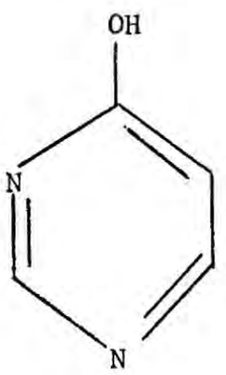


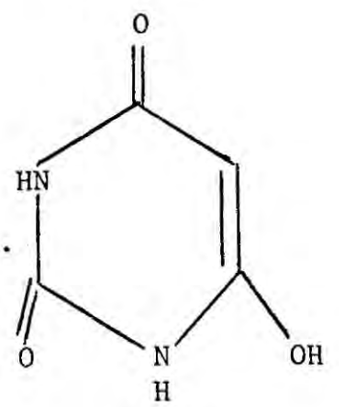
III

for the two nitrogen core levels characteristic of pyridine and lactam type environments supports either structure I or II in the solid phase, with perhaps structure I being slightly favoured on the basis of a comparison of the O_{1s} binding energy with other compounds. The 4 hydroxypyrimidine structure (III) is ruled out by both the N_{1s} and C_{1s} levels. A case can be made on the basis of simple electronegativity considerations on the N_{1s} for structure I. For this structure one might expect the binding energy for the pyridine type nitrogen to be similar to that in pyrimidine itself, whilst the lactam nitrogen should be similar to N1 in uracil. For structure II since the pyridine type nitrogen is now attached to the electron deficient carbonyl carbon, an increase in binding energy with respect to pyrimidine might be expected and by a similar argument the (NH) nitrogen somewhat lower in binding energy.

The problem as to the tautomeric form present in the solid state can be settled more satisfactorily by direct calculation of shifts in core binding energies based upon the charge potential model. Thus CNDO/II SCF calculations were performed upon the hydroxy tautomers of uracil, 4 oxypyrimidine and barbituric acid. If the predicted shift between the nitrogen binding energies in







both keto and hydroxy forms are compared with the measured shifts then it is seen (Table 5.6) quite strikingly that the keto forms are predominant. Furthermore by using the CNDO/II SCF data and a standard linewidth on the curve resolver, the spectrum of barbituric acid was synthesised in both keto and enol forms, Fig. 5.6.

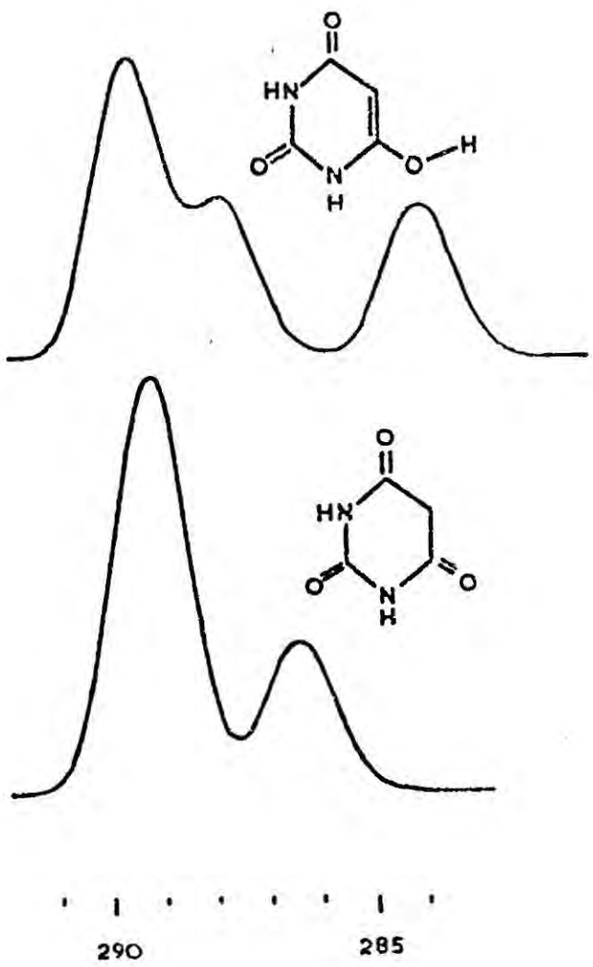


Fig. 5.6 Synthesised spectra of enol and keto forms of barbituric acid. These may be compared with the measured spectrum in Appendix I.

It is clear that the predicted spectrum of the keto tautomer is quite similar to the experimentally obtained spectrum, whilst that of the enol tautomer is different.

Table 5.6 Shift between ring nitrogen atoms

	CNDO/II SCF keto forms (eV)	CNDO/II SCF enol forms (eV)	Measured shift
Uracil	0.72	0.0	0.4
4 Oxypyrimidine	2.18	0.03	1.3
Barbituric Acid	0.00	1.46	0.0

5.6 Hydration in Solid State

The state of hydration of these molecules in the solid state is another topic of biological interest, and again it would be interesting to see if information concerning this is obtainable from ESCA. Using the curve resolver, the areas under the C_{1s} and O_{1s} peaks for all the compounds were measured and ratios calculated. These were compared with theoretical values calculated on the assumption that no water is present in the lattice. The two sets were plotted graphically, Fig. 5.7, and it can be seen that for most

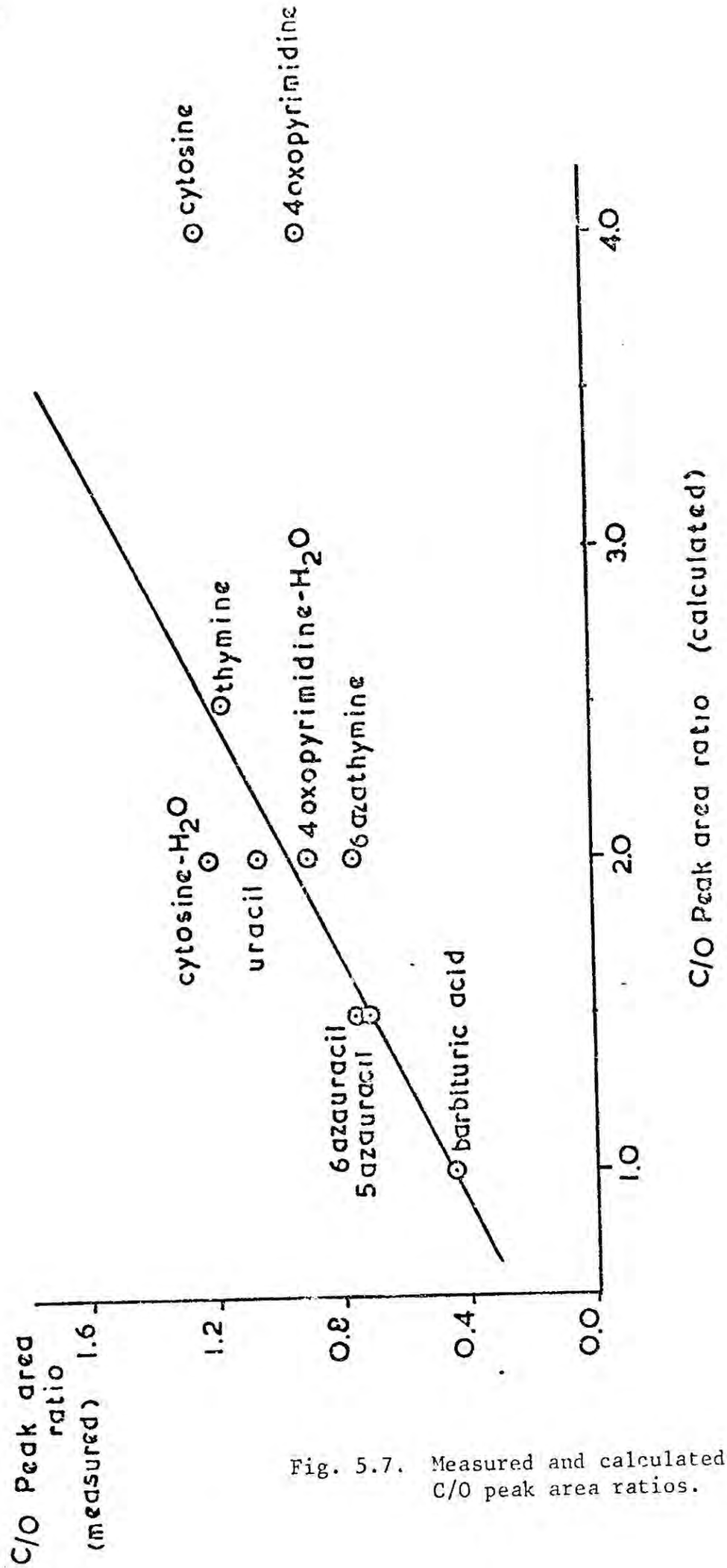


Fig. 5.7. Measured and calculated C/O peak area ratios.

cases the points lie around a straight line of slope 2.0. However, some points were conspicuously off this line, notably cytosine and 4 oxypyrimidine. Hence a new 'calculated' ratio was derived assuming the presence of another oxygen atom per molecule, whereupon the new points lie quite close to the straight line. Thus it is concluded that each molecule of these compounds is associated with one water molecule in the crystal lattice.

The factor of 2.0 between experimental and theoretical ratios arises from the escape depth dependence of the overall intensities of inelastic peaks. Thus this factor will include

(i) differences in sensitivity arising from instrumentation,

(ii) inherent differences in cross sections for photo-ionisation, and

(iii) differences in escape depths of electrons, leading to the production of elastic peaks.

APPENDICES

APPENDIX I E.S.C.A. SPECTRA

(i) Five membered ring heterocycles

Pyrrole

Furan

Thiophene

Pyrazole

Imidazole

Isoxazole

Thiazole

(ii) Pyrimidine Bases

Uracil

4-Oxy Pyrimidine

5-Aza Uracil

6-Aza Uracil

Thymine

6-Aza Thymine

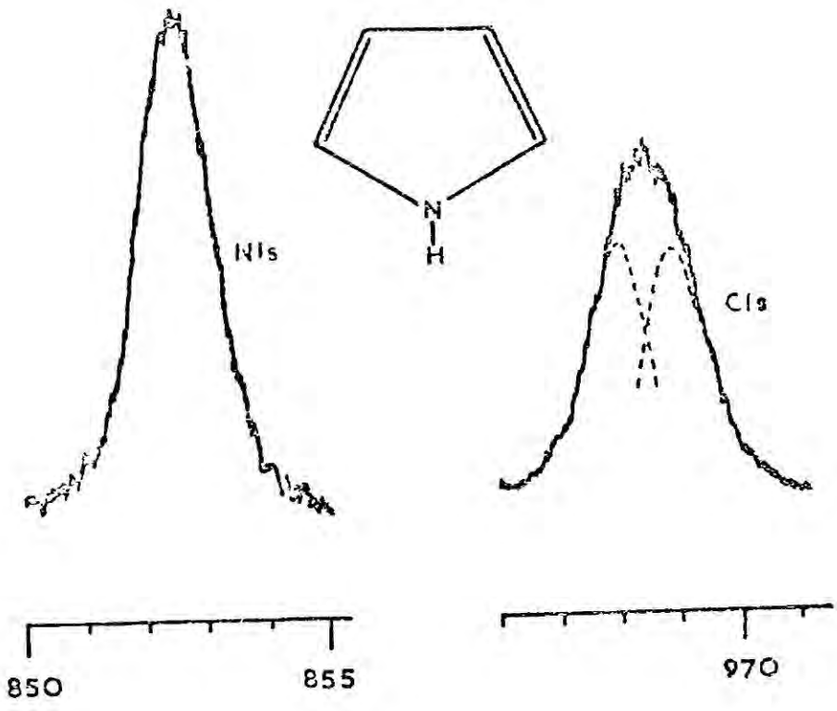
1,-3-Dimethyl Uracil

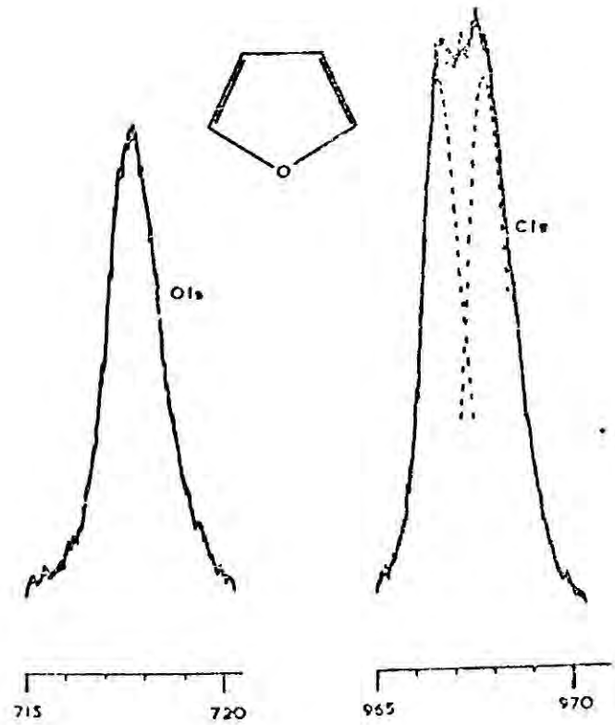
Cytosine

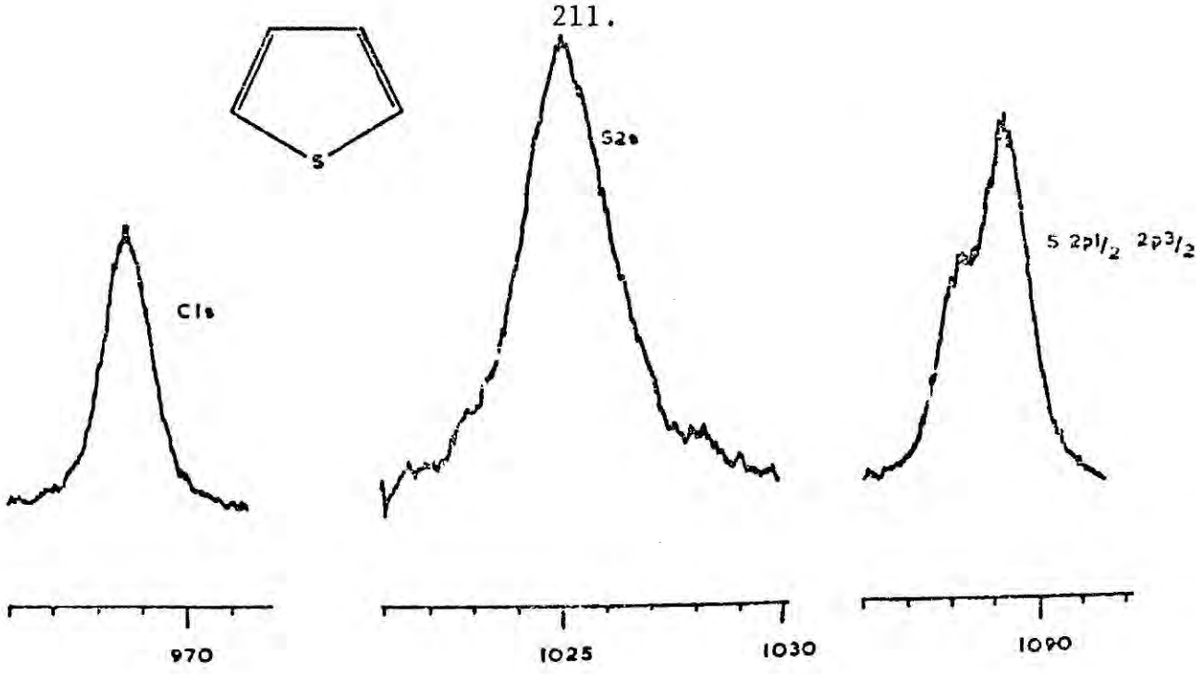
5-Methyl Cytosine

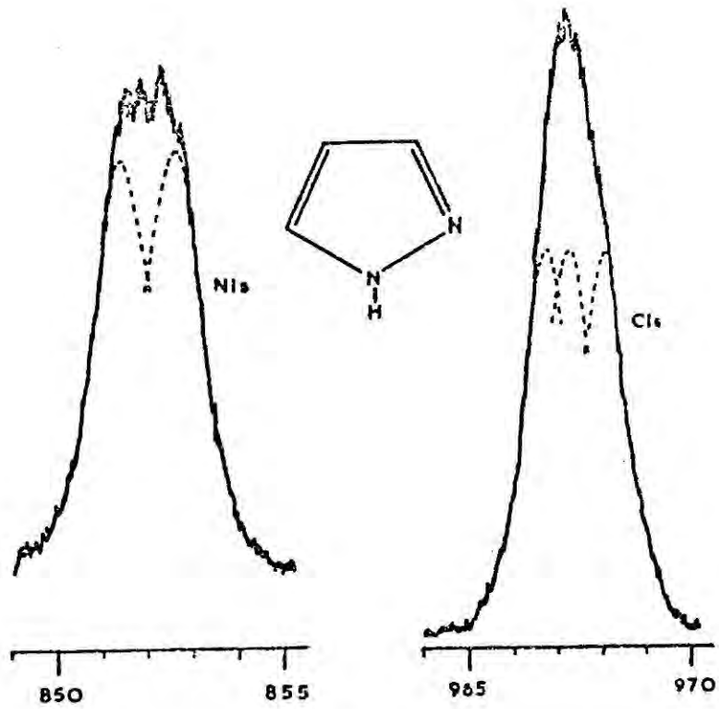
Barbituric Acid

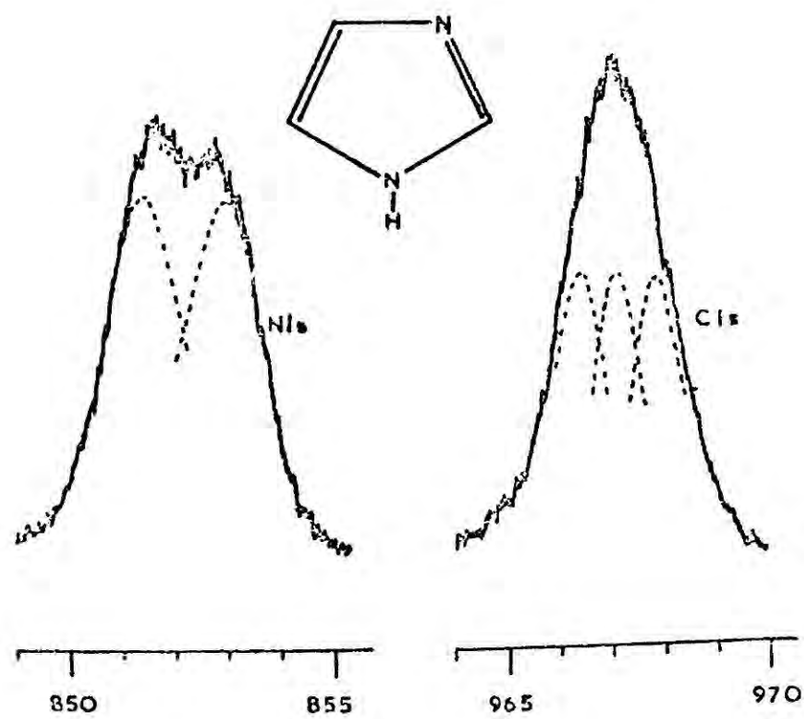
FIVE MEMBERED RING HETEROCYCLES

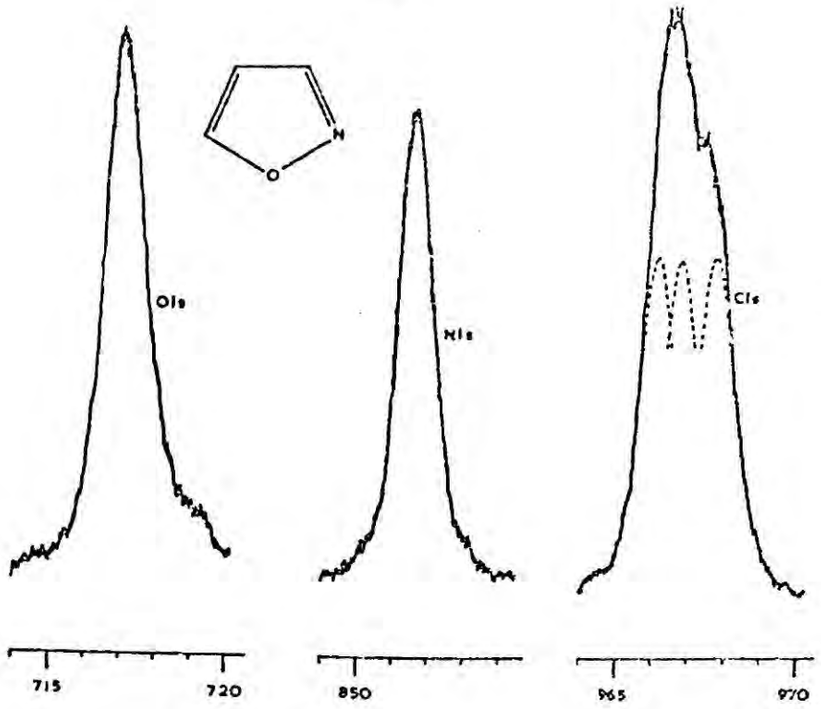


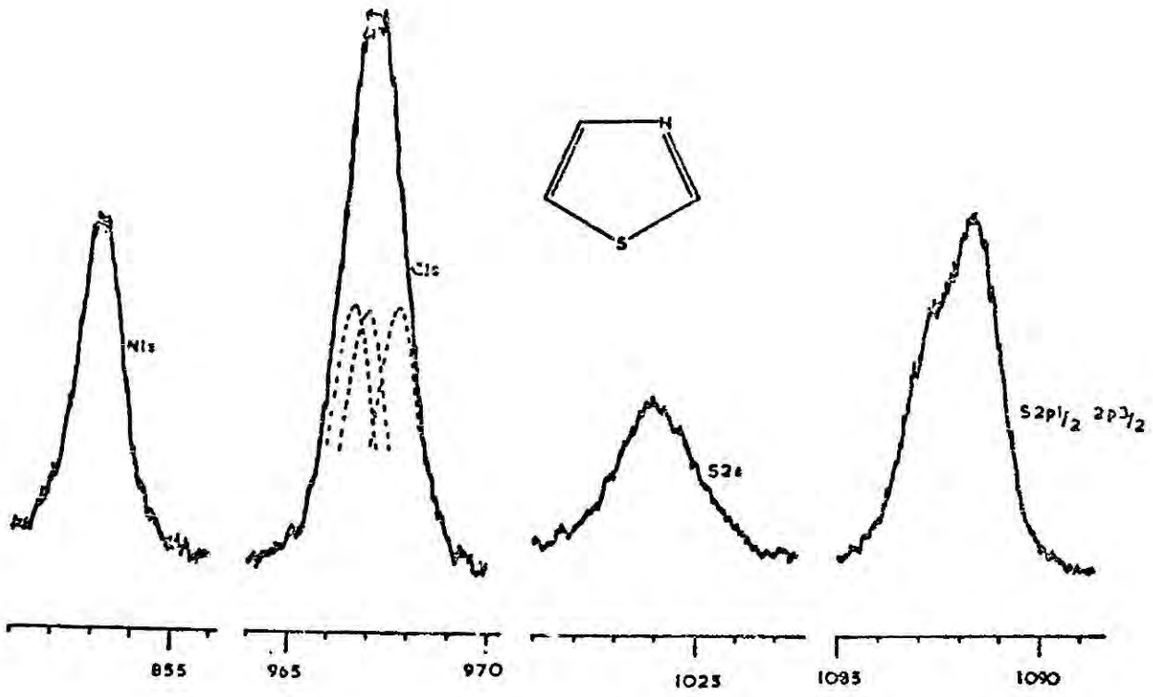



S
C1s
S2s
S 2p1/2 2p3/2
970
1025
1030
1090

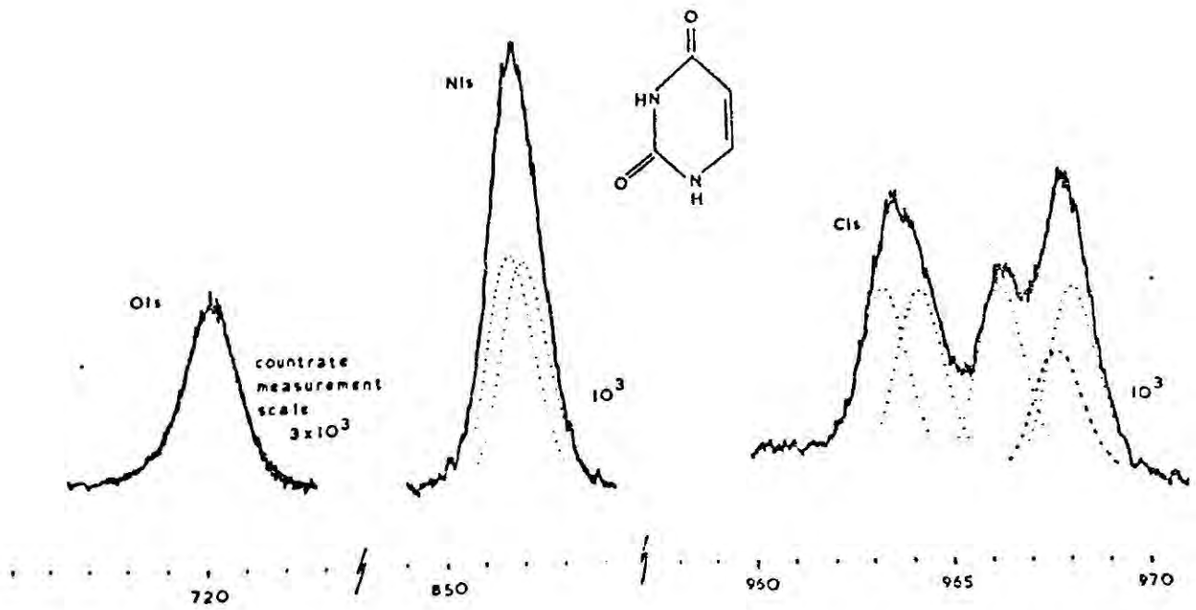

N1s
N
N
H
C1s
850
855
965
970

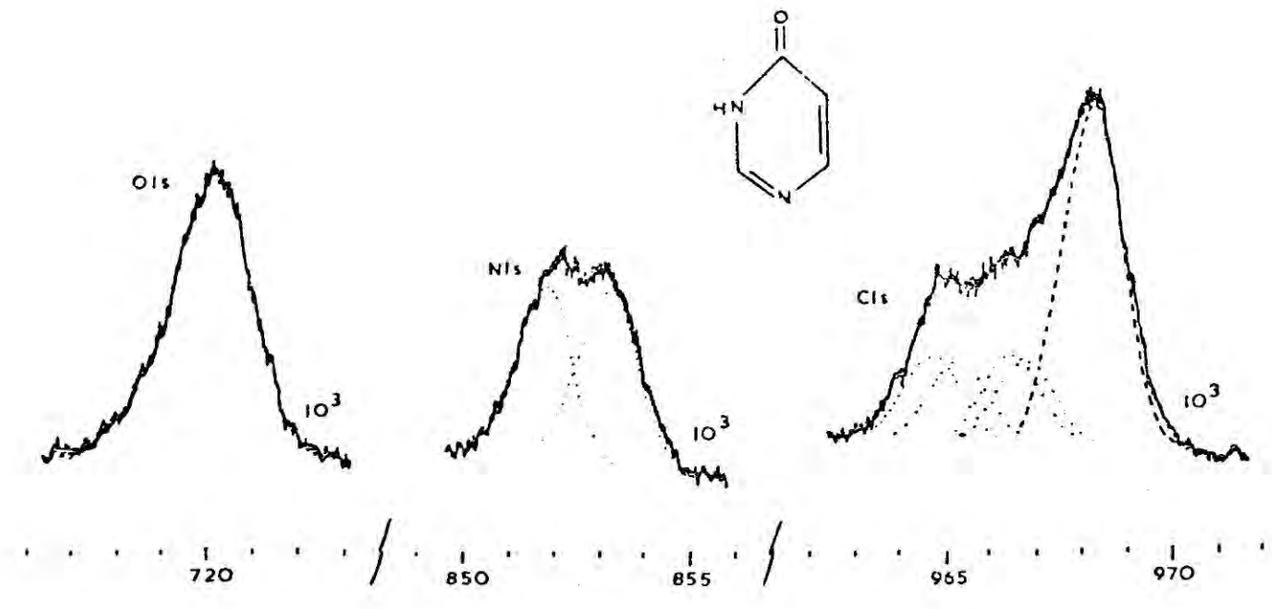

N
N
H
N1s
C1s
850
855
965
970

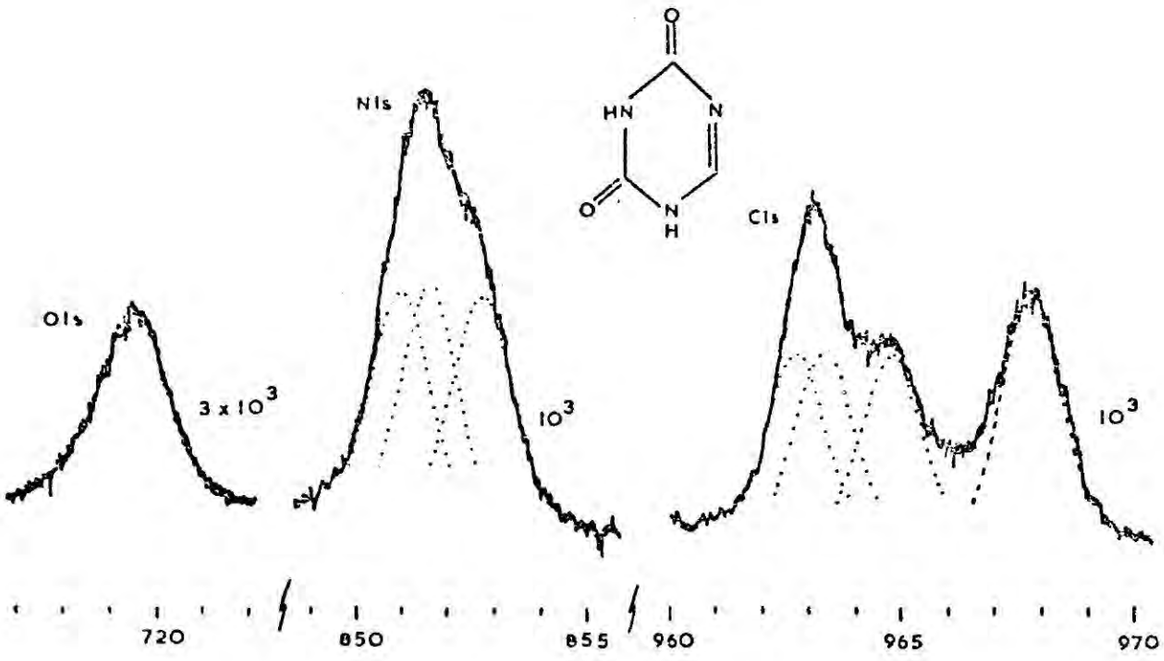

O
N
O1s
N1s
C1s
715
720
850
965
970

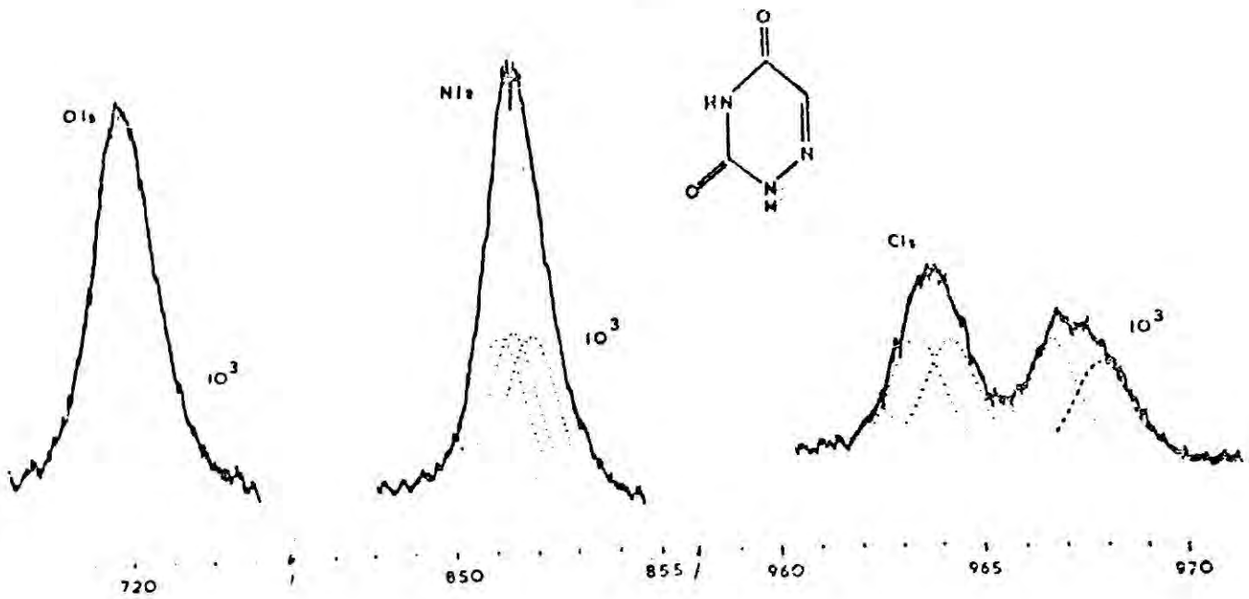

H
S
N1s
C1s
S2s
S2p1/2 2p3/2
855
965
970
1025
1085
1090

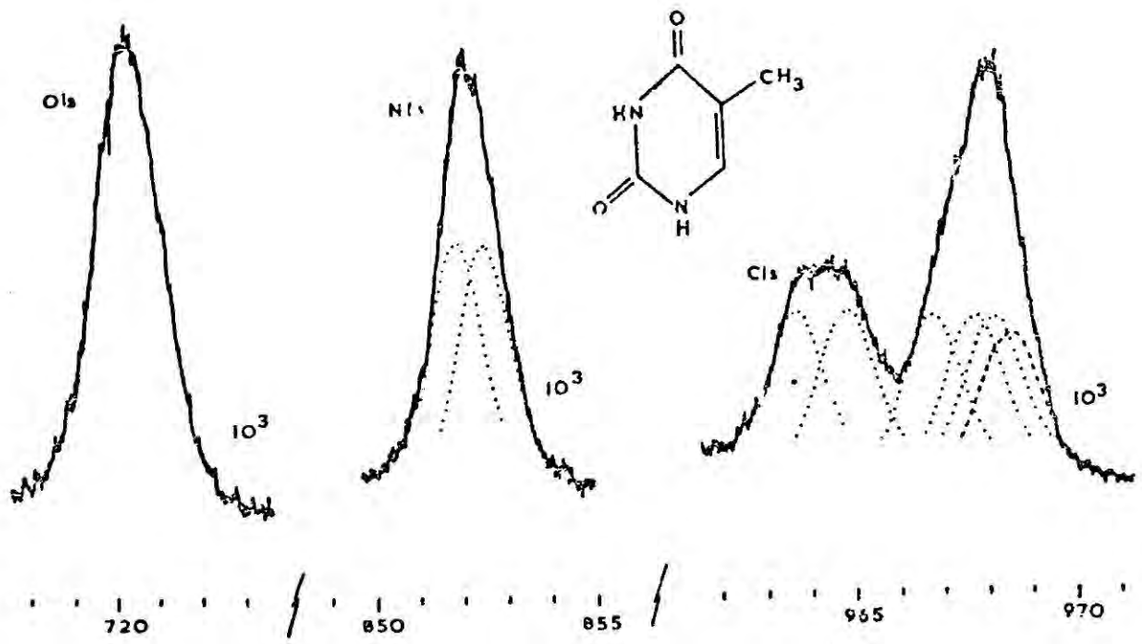
PYRIMIDINES

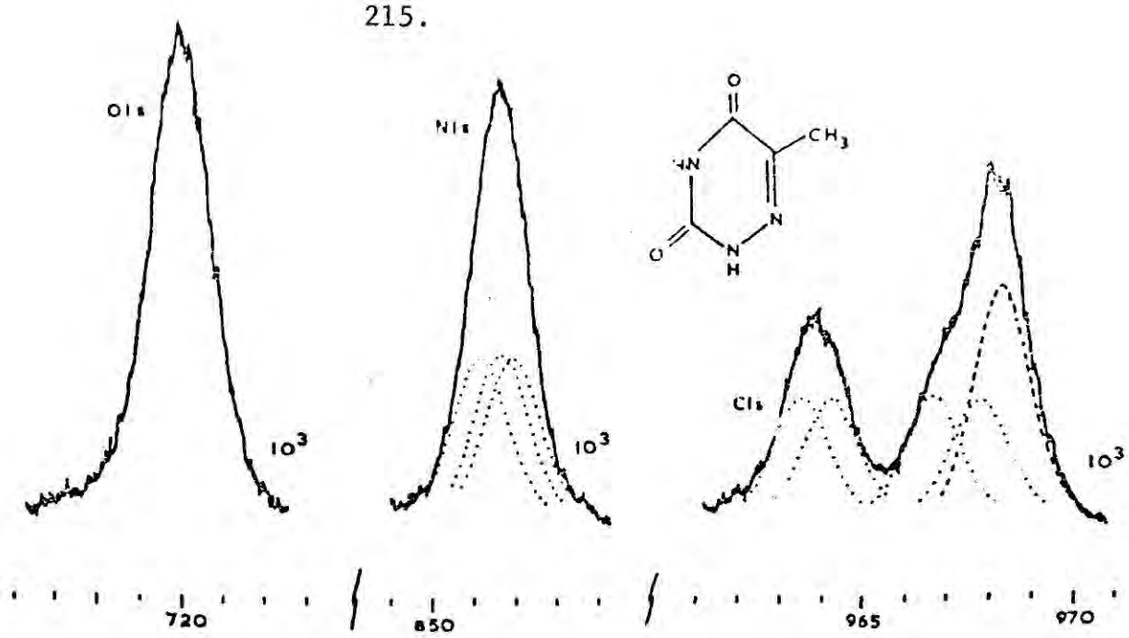


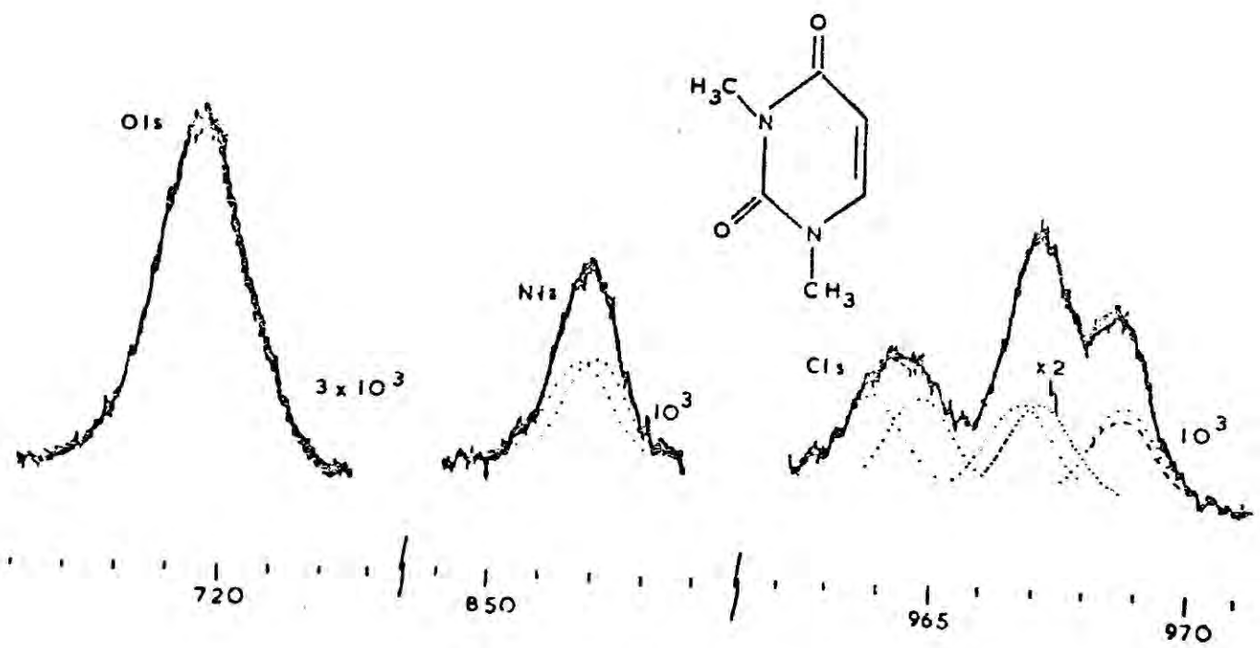


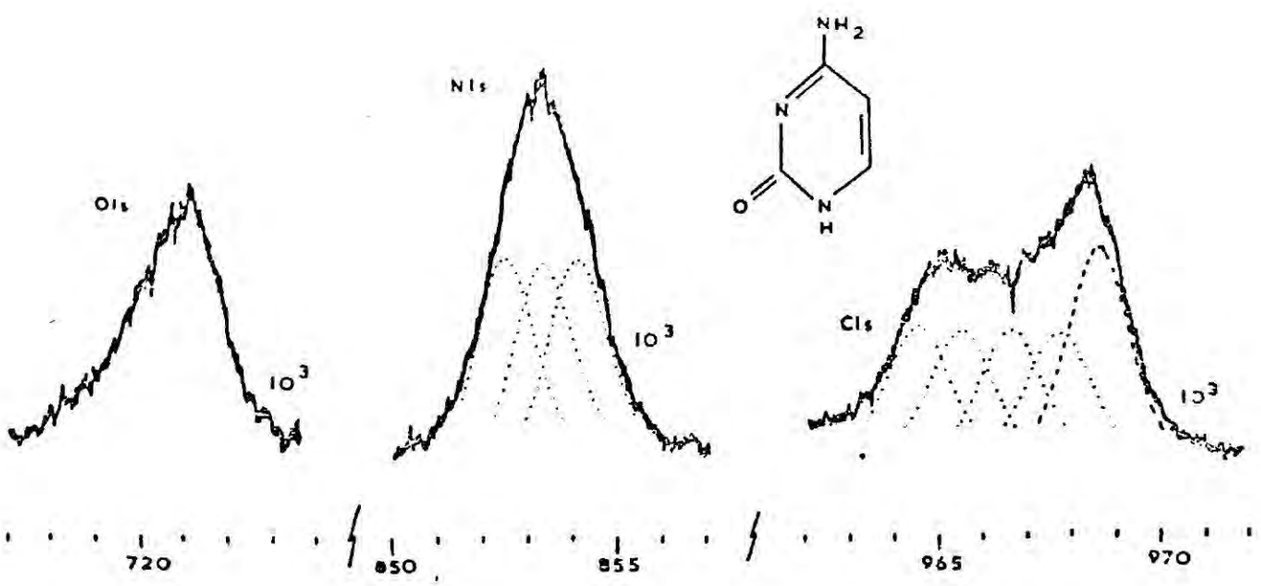

O
HN
N
O
N
H
O1s
3 x 10^3
N1s
10^3
C1s
10^3
720
850
855
960
965
970

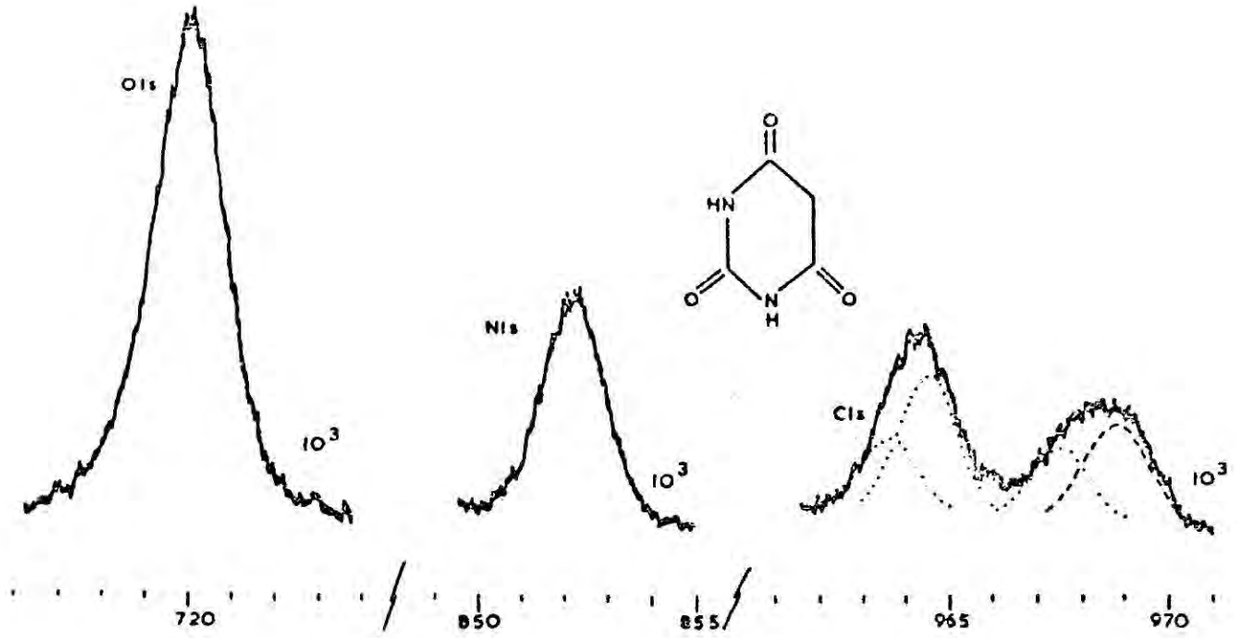

O
HN
N
O
N
H
O1s
10^3
N1s
10^3
C1s
10^3
720
850
855
960
965
970

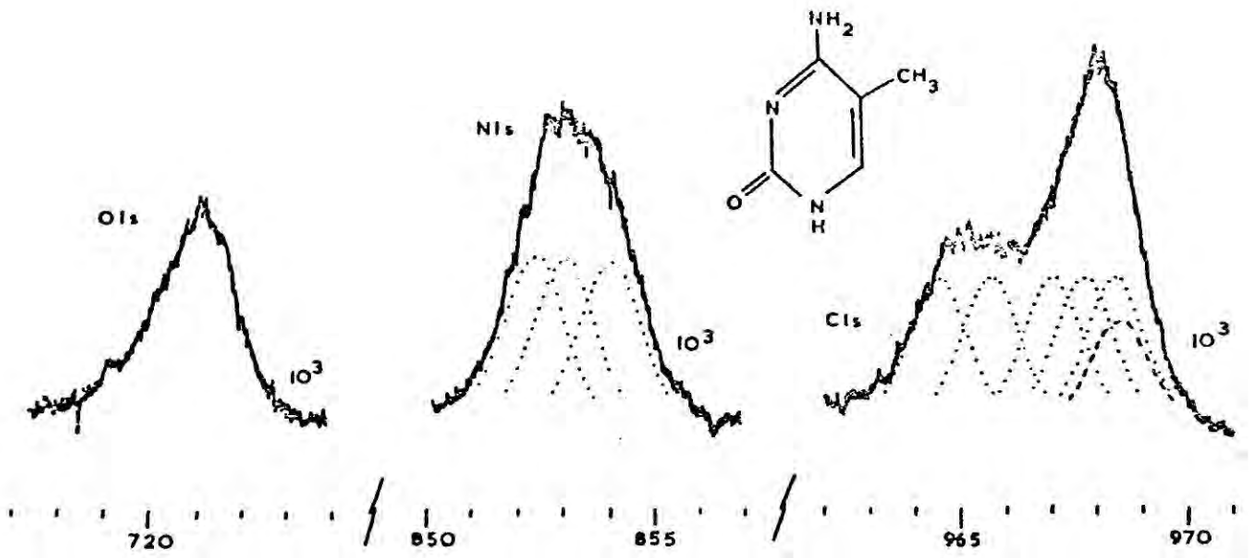

O
CH3
HN
O
N
H
O1s
10^3
N1s
10^3
C1s
10^3
720
850
855
965
970


O1s
N1s
C1s
10^3
720
850
965
970


O1s
N1s
C1s
3×10^3
10^3
x2
720
850
965
970


O1s
N1s
C1s
10^3
720
850
855
965
970


O1s
N1s
C1s
10^3
720
850
855
965
970


O1s
N1s
C1s
10^3
720
850
855
965
970

APPENDIX II E.S.C.A. INSTRUMENTATION

An AEI ES100 spectrometer has been employed for the work in this thesis, the essential components of which are shown schematically in Fig. A2.1.

(i) X-ray generator. Fig. A2.2 shows a section drawing of the source. The X-ray tube is of conventional design and consists of a heated cathode at high negative potential and a water cooled anode at ground potential. The anode is a hollow copper tube faced with the required anode material (magnesium in this work giving a photon energy 1253.7 eV). Cooling is essential and is accomplished by means of a water jet directed onto the back of the anode face. Usual operating conditions are 10 - 15 kV at 20 - 50 mA (12 kV and 25 mA respectively in this work). The X-ray tube and sample compartments are separated by a thin window (Al foil, 2/10 thousanths of an inch thickness),through which the X-ray beam passes, thus preventing scattered electrons from the X-ray source penetrating into the sample region.

(ii) Electron energy analyser. It is necessary that the analyser should be capable of a resolution in the region of a few parts in 10^4. The ES100 analyser is a hemispherical double focussing analyser based on the principles described by Purcell.[159] The size of the hemispheres is a necessary compromise based upon the cost and ease of construction since larger hemispheres, with their greater resolving power, require accurate machining.and support of

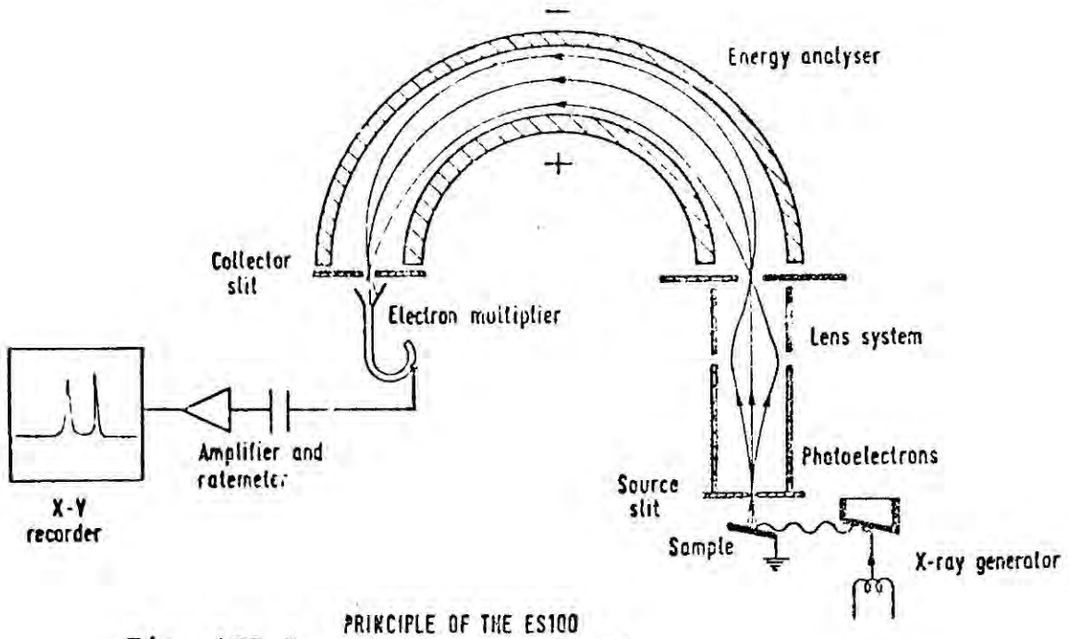


Fig. AII.1. PRINCIPLE OF THE ES100 PHOTOELECTRON SPECTROMETER

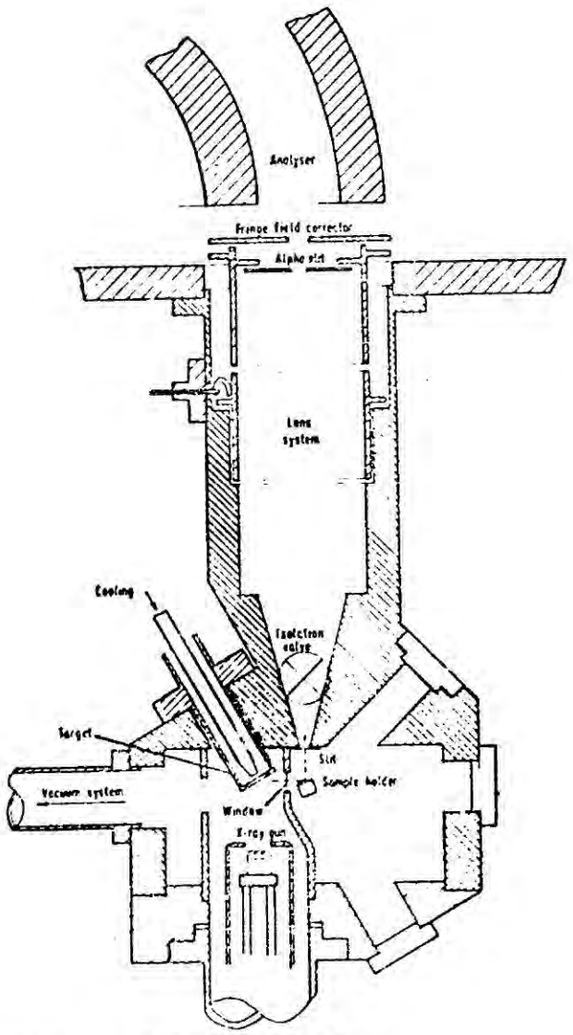


Fig. AII.2. LENS AND SOURCE ARRANGEMENT ES100

the hemispheres without mechanical distortion. Furthermore, as the size of the analyser is increased, so must the efficiency of the pumping system be increased. Before entering the analyser of the ES100, the photoelectrons pass through a retarding lens system which has a twofold purpose:-

a) Allowing more flexible sample handling arrangements by virtue of the removal of the sample region away from the analyser

b) Reduction of the stringency on the resolution requirements of the analyser by means of the retarding potential.[160]

The transmitted electron current, I, of the relevant monoenergetic electrons is given by

$$I = BA\Omega \qquad (A2.1)$$

where B is the brightness of the electron illumination of the entrance slit, in units of current per unit area per unit solid angle, A is the entrance slit area and Ω is the solid angle of the spectrometer angle as viewed from the entrance slit ($A\Omega$ is the spectrometer luminosity). The brightness B is determined by the strength of the X-radiation at the sample, and since it is low, a high luminosity $A\Omega$ is required. The luminosity is given by

$$A\Omega = CR^2 \left(\frac{\Delta E}{E}\right)^2 \qquad (A2.2)$$

If E is reduced by applying a retarding potential to the electrons before they enter the entrance slit, the luminosity may be increased, without affecting ΔE, by increasing the slit dimensions and

acceptance angles. The brightness of the electron beam is proportional to its kinetic energy and is therefore reduced in a retarding field. If Bo is the electron brightness at the sample, where the photoelectrons have energy Eo, and E is the kinetic energy of the electron as it passes through the entrance slit to the analyser, the brightness is given by

$$B = \left(\frac{Bo}{Eo}\right) E \qquad (A2.3)$$

and hence

$$I = BoCR^2\left(\frac{(\Delta E)^2}{E.Eo}\right) \qquad (A2.4)$$

and thus an increase in intensity of the transmitted current I is also obtained.

Electrons of the required kinetic energy may be focussed at the collector slit by either of:-

a) Scanning the retarding potential while keeping a constant potential between the hemispheres.

b) Scanning the retarding potential and the potential between the analyser hemispheres simultaneously keeping a constant ratio between the two.

The ES100 employs the latter method. Magnetic double focussing analysers have been used but, despite their simpler construction, these are bulky since they require Helmholts coils to eliminate stray magnetic fields.

(iii)Detection and Data Acquisition. Photoelectrons passing through the collector slit are detected by a channel electron

multiplier and the pulses obtained are amplified and fed into counting electronics. Spectra may be generated either by continuous or step scans. In the continuous mode of operation the field is increased continuously while the detector signal is monitored by a rate meter. If the signal is sufficiently strong and the signal to background ratio sufficiently high then the spectrum (a graph of counts per second versus kinetic energy of the electrons) is plotted out directly on an X- Y recorder. Alternatively the energy may be incremented in small steps (typically 0.1 eV) and at each setting either a fixed number of counts may be timed or a count can be made for a fixed length of time. By storing this data in a multichannel analyser several scans of the region of interest can be made, thereby averaging any random background fluctuations. The availability of both wide and narrow scan facilities permits both preliminary searches and detailed study of specific regions.

(iv) Sample handling.

A) Involatile non-metallic solid samples. The most straight forward method of sample preparation for an involatile solid is to mount it on the spectrometer probe by means of double sided adhesive tape. This has the dissadvantage that the sample is not in electrical contact with the spectrometer and hence sample charging may occur. However, the binding energies do not change with time and the C_{1s} binding energy from the adhesive tape may be used for calibration purposes. An improvement on this method is to mount a small quantity of sample on electrically conducting adhesive

tape.[161] The layer of pump oil which forms on the sample surface may be used as a reference.[47,161] When possible, a more satisfactory method is the deposition of a thin layer of the sample onto a conducting backing (e.g. gold) by evaporation from a solution in a suitable volatile solvent. If the layer is sufficiently thin, such that the core levels of the conducting backing are observable, the photo-conductivity induced throughout the sample ensures that the sample takes up the same potential as the sample plate and eliminates sample charging effects. Furthermore, the $4f_{7/2}$ signal from the gold backing (binding energy 84.0 eV) may be used as reference. Other methods of solid sample handling include

a) Pressing a disc of a powder sample and mounting this on the probe. This generally improves count rates compared with powder samples and the slight deposit of hydrocarbon on the surface may be used as a reference (see Chapter 5).

b) A powder sample may be pressed into a wire gauze on the probe.

c) Samples in the form of foils or sheets can be clipped directly onto the probe.

Other methods of calibrating for charging effects include:-

a) Making an intimate mixture of a powder sample with a reference powder (which may then be pressed into a disc).

b) An internal element may sometimes be used as a reference if its chemical environment is known not to change from sample to sample.

c) A conducting surface may be deposited on the sample and will take up the same potential as the sample surface (e.g. the gold decoration technique).

B) Liquid samples. These are introduced by injection through a septum plug into a heatable (25 to 150° C) evacuated reservoir shaft. The vapour diffuses through a metrosil leak and condenses on the tip of the cooled (typically -80 to -150° C) sample probe. Thus the sample surface is in a state of continual renewal, thereby preventing surface contamination by the residual spectrometer atmosphere and also diminishing difficulties from X-ray damage. While the layer of condensed sample is thin, and the Au $4f_{7/2}$ peak is observable, no sample charging occurs. However, with a high rate of condensation over an extended period sample charging may occur resulting in changes in binding energy and peak width.[162]

C) Volatile solids. These are generally studied by sublimation of the sample from a capillary tube, which may be heated, and subsequent condensation onto a cooled probe. Very volatile solids may be injected into the reservoir shaft to reduce the rate of condensation. The considerations of sample charging apply just as with condensed liquids.

D) Gases. Gases may be studied by condensation onto a cooled probe although several electron spectrometers have facilities for the study of samples in the gas phase thereby offering several advantages[163] including lack of sample charging and ease of theoretical interpretation of results.

APPENDIX III E.S.C.A. SPECTRUM DECONVOLUTION

In this work, partially resolved peaks have been resolved into their individual components by use of a Du Pont 310 curve resolver (an analogue computer). For detailed deconvolutions a prior knowledge of line shapes (usually gaussian) and peak widths for a particular level is required, obtained from a study of similar compounds with well resolved peaks, which have been studied under the same experimental conditions. Peak areas for a given level are proportional to the number of atoms in a particular environment after due allowance has been made for any shake-up, shake-off and CI satellites which may arise, and thus if the molecular formula is known then individual binding energies may be determined by setting peak widths and areas and varying peak positions to obtain a fit to the experimental spectrum. In the compounds studied in this work, the satellite structure due to shake up processes were of negligible proportions. The thus deconvoluted peaks may then be assigned by means of the methods described in the text. The curve resolver also possesses an integration facility by means of which the areas under peaks may be determined.

APPENDIX IV CO-ORDINATES FOR TYPICAL SPECIES (a.u.)

Atom	x	y	z
(i) Bridged species			
X = H	0.0	0.0	2.53 (minimum)
X = F	0.0	0.0	2.83 (minimum)
X = Cl	0.0	0.0	3.77 (minimum)
C(1)	-1.392519	0.0	0.0
C(2)	1.392519	0.0	0.0
H(1)	-2.416765	-1.774045	0.0
H(2)	-2.416765	1.774045	0.0
H(3)	2.416765	1.774045	0.0
H(4)	2.416765	-1.774045	0.0
(ii) Ethyl Cation and 2-Haloethyl Cations			
X = H	-0.689478	-1.947023	0.0
X = F	-0.83898	-2.369205	0.0
X = Cl	-1.110229	-3.135191	0.0
C(1)	0.0	0.0	0.0
C(2)	2.963833	0.0	0.0
H(1)	-0.689478	0.972474	-1.686771
H(2)	-0.689478	0.972474	1.686771
H(3)	3.988078	0.0	1.774046
H(4)	3.988078	0.0	-1.774046
(iii) 1-Haloethyl Cations			
X = F	4.211069	-2.160275	0.0
X = Cl	4.626815	-2.880369	0.0
C(1)	0.0	0.0	0.0
C(2)	2.963833	0.0	0.0
H(1)	-0.689478	0.972474	-1.686771
H(2)	-0.689478	-1.947023	0.0
H(3)	-0.689478	0.972474	1.686771
H(4)	3.988078	1.774046	0.0
(iv) Ethane and Haloethanes			
X = H	-0.689478	-1.947023	0.0
X = F	-0.83898	-2.369205	0.0
X = Cl	-1.110229	-3.135191	0.0
C(1)	0.0	0.0	0.0
C(2)	2.963833	0.0	0.0
H(1)	-0.689478	0.972474	-1.686771
H(2)	-0.689478	0.972474	1.686771
H(3)	3.653311	-0.972474	-1.686771
H(4)	3.653311	1.947023	0.0
H(5)	3.653311	-0.972474	1.686771

Atom	x	y	z

(v) Ethylene and Haloetheylenes

X = H	3.547064	-1.759315	0.0
X = F	3.796512	-2.199554	0.0
X = Cl	4.18958	-2.880369	0.0
C(1)	0.0	0.0	0.0
C(2)	2.526598	0.0	0.0
H(1)	-1.015742	-1.759315	0.0
H(2)	-1.015742	1.759315	0.0
H(3)	3.547064	1.759315	0.0

(vi) Reaction Co-ordinate Bridge-Protonated Ethylene → Ethyl Cation

Point (i)

C(1)	-2.818367	0.0	0.0
C(2)	0.0	0.0	0.0
H(1)	-3.857495	-1.757891	-0.200722
H(2)	-3.857495	1.757891	-0.200722
H(3)	1.015742	1.759315	0.0
H(4)	1.015742	-1.759315	0.0
H(5)	-1.845559	0.0	2.416265

Point (ii)

C(1)	-2.854734	0.0	0.0
C(2)	0.0	0.0	0.0
H(1)	-3.868341	-1.741083	-0.406759
H(2)	-3.868341	1.741083	-0.406759
H(3)	1.015742	1.759315	0.0
H(4)	1.015742	-1.759315	0.0
H(5)	-2.300118	0.0	2.30253

Point (iii)

C(1)	-2.891101	0.0	0.0
C(2)	0.0	0.0	0.0
H(1)	-3.837397	-1.723624	-0.609805
H(2)	-3.837397	1.723624	-0.609805
H(3)	1.015742	1.759315	0.0
H(4)	1.015742	-1.759315	0.0
H(5)	-2.754677	0.0	2.188796

Point (iv)

C(1)	-2.927467	0.0	0.0
C(2)	0.0	0.0	0.0
H(1)	-3.765	-1.705518	-0.801235
H(2)	-3.765	1.705518	-0.801235
H(3)	1.015742	1.759315	0.0
H(4)	1.015742	-1.759315	0.0
H(5)	-3.209236	0.0	2.075061

Atom	x	y	z

(vii) Reaction Co-ordinate 2-Fluoroethyl Cation → Bridge-Protonated Fluoroethylene

Point (i)

F	-1.115955	0.824032	2.095863
C(1)	0.0	0.0	0.0
C(2)	2.904235	0.0	0.0
H(1)	-0.91458	0.675334	-1.717663
H(2)	0.004521	-2.14135	0.0
H(3)	3.928481	0.606759	1.667057
H(4)	3.928481	-0.606759	-1.667057

Point (ii)

F	-1.256251	0.413552	2.13725
C(1)	0.0	0.0	0.0
C(2)	2.844636	0.0	0.0
H(1)	-1.026725	0.337993	-1.74676
H(2)	0.69852	-2.335675	0.0
H(3)	3.868882	0.30806	1.747094
H(4)	3.868882	-0.30806	1.747094

(viii) Reaction Co-ordinate Bridge-Protonated Fluoroethylene → 1-Fluoroethyl Cation

Point (i)

F	-1.256685	-1.088321	1.885027
C(1)	0.0	0.0	0.0
C(2)	2.844636	0.0	0.0
H(1)	-1.024246	0.887022	1.885027
H(2)	2.146116	-2.335675	0.0
H(3)	3.871361	0.337993	1.74676
H(4)	3.871361	0.337993	-1.74676

Point (ii)

F	-1.256685	-1.885027	1 088321
C(1)	0.0	0.0	0.0
C(2)	2.904235	0.0	0.0
H(1)	-1.024244	1.536367	-0.887022
H(2)	2.899714	-2.14135	0.0
H(3)	3.818815	0.675334	1.717663
H(4)	3.818815	0.675334	-1.717663

Atom	x	y	z

(ix) Reaction Co-ordinate 2-Chloroethyl Cation → Bridge-Protonated Chloroethylene

Point (i)

Cl	1.528808	0.98518	2.784636
C(1)	0.0	0.0	0.0
C(2)	2.892315	0.0	0.0
H(1)	0.143321	-2.180214	0.0
H(2)	-0.946298	0.609804	-1.723626
H(3)	3.916561	1.042758	-1.435233
H(4)	3.916561	-1.042758	1.435233

Point (ii)

Cl	-1.68435	0.325352	2.849411
C(1)	0.0	0.0	0.0
C(2)	2.820797	0.0	0.0
H(1)	0.97612	-2.413404	0.0
H(2)	-1.039129	0.20072	-1.757892
H(3)	3.845043	0.368845	-1.735278
H(4)	3.845043	-0.368845	1.735278

(x) Reaction Co-ordinate Bridge-Protonated Chloroethylene → 1-Chloroethyl Cation

Point (i)

Cl	-1.662982	-0.890083	2.739393
C(1)	0.0	0.0	0.0
C(2)	2.820797	0.0	0.0
H(1)	1.844677	-2.413405	0.0
H(2)	-1.024244	0.54821	-1.687217
H(3)	3.859926	0.20072	-1.757892
H(4)	3.859926	0.20072	1.757892

Point (ii)

Cl	-1.662982	-2.330267	1.693038
C(1)	0.0	0.0	0.0
C(2)	2.892315	0.0	0.0
H(1)	2.748994	-2.180214	0.0
H(2)	-1.024244	1.435233	-1.042758
H(3)	3.838613	0.609804	-1.723626
H(4)	3.838613	0.609804	1.723626

APPENDIX V GAUSSIAN EXPONENTS AND CONTRACTION COEFFICIENTS AND SLATER EXPONENTS

(i) Exponents and Contraction coefficients for hydrogen, carbon, fluorine and chlorine.

Exponents				Contraction Coefficients			
Hydrogen	0.450038	D	01	Hydrogen	0.7048	D	01
S	0.681277	D	00		0.40789	D	00
	0.151374	D	00		0.64767	D	00
Carbon	0.9947	D	03	Carbon	0.72	D	-02
	0.16	D	03		0.473	D	-01
	0.3991	D	02		0.1819	D	00
S	0.1182	D	02		0.4474	D	00
	0.3698	D	01		0.4438	D	00
	0.6026	D	00		0.434	D	00
	0.1817	D	00		0.6859	D	00
	0.4279	D	01		0.1093	D	00
P	0.8699	D	00		0.4597	D	00
	0.2036	D	00		0.6302	D	00
Fluorine	0.2723	D	04	Fluorine	0.59	D	-02
	0.4164	D	03		0.42	D	-01
	0.9773	D	02		0.1792	D	00
S	0.2787	D	02		0.4544	D	00
	0.8712	D	01		0.4436	D	00
	0.1396	D	01		0.5064	D	00
	0.4209	D	00		0.6190	D	00
	0.1053	D	02		0.127	D	00
P	0.2188	D	01		0.4784	D	00
	0.4785	D	00		0.6129	D	00
Chlorine	0.286563	D	05	Chlorine	0.1558	D	-02
	0.4299	D	04		0.11941	D	-01
	0.976335	D	03		0.59635	D	-01
	0.274415	D	03		0.208871	D	00
S	0.890063	D	02		0.444010	D	00
	0.312371	D	02		0.388159	D	00
	0.776951	D	01		0.350098	D	00
	0.307933	D	01		0.730327	D	00
	0.651038	D	00		0.390248	D	00
	0.240798	D	00		0.782082	D	00
	0.150436	D	03		0.278870	D	-01
	0.347101	D	02		0.173468	D	00
P	0.104071	D	02		0.469717	D	00
	0.337330	D	01		0.485035	D	00
	0.748495	D	00		0.485239	D	00
	0.207855	D	00		0.648858	D	00

(ii) For STO-3G calculations, the basis set consists of a least squares fit of Slater type orbitals to three gaussian functions. The Slater exponents, evaluated by application of Burns rules[126] are

H	1s	1.200				
C	1s	5.6727	2s	1.6083	2p	1.5679
Cl	1s	16.5239	2s	5.7152	3s	2.3561
	2p	6.4966	3p	2.0387	3d	1.8000

REFERENCES

1 M. Born Z. Physik. 37, 863; 38, 803 (1926)

2 L. Salem XXIIIrd Int. Congr. of Pure and Applied Chem., Butterworths, London (1971), vol. 1, p. 197.

3 J. A. Horsley, Y. Jean, C. Moser, L. Salem, R. Stephens and J. S. Wright, J. Amer. Chem. Soc. 94, 279 (1972)

4 P. J. Hay, W. J. Hunt and W. A. Goddard III ibid 94, 638 (1972)

5 E. Schrödinger Ann. Physik 79, 361, 489, 734; 81, 109 (1926) Phys. Rev. 28, 1049 (1926)

6 M. Born and J. R. Oppenheimer, Ann. Physik 84, 457 (1927)

7 W. Pauli Phys. Rev. 58, 716 (1940)

8 J. C. Slater ibid. 35, 509 (1930); 34, 1293 (1929)

9 C. E. Eckart ibid. 36, 878 (1930)

10 D. R. Hartree Proc. Camb. Phil. Soc., 24, 89 (1928)

11 V. Fock Z. Physik, 61, 126 (1930)

12 J. C. Slater Phys. Rev. 35, 210 (1930)

13 T. A. Koopmans Physica 1, 104 (1933)

14 C. C. J. Roothaan Rev. Mod. Phys. 23, 69 (1951)

15 J. C. Slater Phys. Rev. 36, 57 (1930)

16 S. F. Boys Proc. Roy. Soc. (London) A200, 542 (1950)

17 E. Clementi and D. R. Davis J. Comp. Phys 1, 223 (1966)

18 S. Huzinaga J. Chem. Phys. 42, 1293 (1965).

19 T. H. Dunning ibid. 53, 2823 (1970)

20 W. J. Hehre, R. F. Stewart and J. A. Pople, ibid. 51, 2657 (1969)

21 E. Clementi, J. Mehland and W. Von Niessen, ibid. 54 508 (1971)

22 H. Preuss — Z. Naturforsch 11a, 823 (1956); Int. J. Quant. Chem 2, 651 (1968)

23 J. L. Whitten — J. Chem. Phys 44, 359 (1966)

24 S. Shih, R. J. Buenker, S. D. Peyerimhoff and B. Wirsam — Theo. Chim. Acta. 18, 277 (1970)

25 R. Ditchfield, W. J. Hehre and J. A. Pople, J. Chem. Phys. 54, 724 (1971)

26 E. Fermi — Rend. Lincei 6, 602 (1927)

27 L. H. Thomas — Proc. Camb. Phil. Soc. 23, 542 (1927)

28 R. Ditchfield, W. J. Hehre, J. A. Pople and L. Radom, Chem. Phys. Lett. 5, 13 (1970)

29 J. O. Hirschfelder J. Chem. Phys. 33, 1762 (1960)

30 J. O. Hirschfelder and C. A. Coulson, ibid. 36, 941 (1962)

31 P. O. Löwdin — Advan. Chem. Phys. 2, 207 (1959)

32 E. Clementi — J. Chem. Phys 38, 2248 (1963)

33 C. Hollister and O. Sinanoğlu, J. Amer. Chem. Soc. 83, 13 (1966)

34 L. C. Snyder and H. Basch, ibid. 91, 2189 (1969)

35 E. A. Hylleraas — Z. Physik 48, 469 (1928)

36 E. Clementi and H. Popkie, J. Amer. Chem. Soc. 94, 4057 (1972)

37 W. A. Goddard III and R. C. Ladner, ibid. 93, 6750 (1971)

38 E. A. Hylleraas and J. Midtal, Phys. Rev. 103, 829 (1956)

39 A. Speiser — 'Theorie der Gruppen von endlicher Ordnung', Springer, Berlin, 1927.

40 E. Wigner — 'Gruppentheorie', Vieweg, Braunschweig, 1931

41 J. A. Pople, D. P. Santry and G. A. Segal, J. Chem. Phys. 43, S129 (1965)

42 J. A. Pople and G. A. Segal, ibid. 43, S136 (1965)

43 J. A. Pople, D. L. Beveridge and P. A. Dobosh, ibid. 47, 2026 (1967)

44 I. G. Csizmadia, M. G. Harrison, J. W. Moscowitz, S. S. Seung, D. T. Sutcliffe and M. P. Barnett, 'The Polyatom System' Q.C.P.E. No. 47A.

45 E. Clementi and D. R. Davis, J. Comput. Phys 2, 223 (1967)

46 R. S. Mulliken, J. Chem. Phys. 23, 1833, 1841, 2338 (1955)

47 K. Siegbahn, C. Nordling, A. Fahlman, R. Nordberg, K. Hamrin, J. Hedman, G. Johansson, T. Bergmark, S. E. Karlsson, I. Lidgren and B. Lindberg, 'ESCA Atomic, Molecular and Solid State Structure Studied by Means of Electron Spectroscopy' Almquist and Wiksells, Uppsala (1967)

48 D. W. Turner, C. Baker, A. D. Baker and C. R. Brundle, 'Molecular Photoelectron Spectroscopy' John Wiley and Sons Ltd., (1970)

49 S. Hagstrom, C. Nordling and K. Siegbahn, Z. Physik, 178, 439 (1964)

50 D. S. Urch Quart. Rev. 25, 343 (1971)

51 A. Faessler Angew. Chem (Int. Ed) 11, 34 (1972)

52 P. Auger J. Phys. Radium, 6, 205 (1925)

53 P. Auger Compt. Rend, 180, 65 (1925)

54 J. J. Lander Phys. Rev. 91, 1382 (1953)

55 C. C. Chang Surface Sci. 25, 53 (1971)

56 K. Siegbahn, C. Nordling, G. Johansson, J. Hedman, P. F. Heden, K. Hamrin, U. Gelius, T. Bergmark, L. O. Werme, R. Manne and Y. Baer, 'ESCA Applied to Free Molecules' North Holland (1969)

57 J. S. Levinger Phys. Rev. 90, 11 (1953)

58 M. O. Krause, T. A. Karlson and R. D. Dismukes, ibid. 170, 37 (1968)

59 E. Sokolowski Arkiv Fysik, 15, 11 (1959)

60 C. Nordling ibid, 15, 397 (1959)

61 W. G. Richards Int. J. Mass Spec. Ion Phys. 2, 419 (1969)

62 R. J. Gillespie and R. S. Nyholm, Quart. Rev. 11, 339 (1957)

63 M. E. Schwartz Chem. Phys. Lett. 6, 631; 7, 78 (1970)

64 L. G. Parrat Rev. Mod. Phys. 31, 616 (1959)

65 C. S. Fadley, S. B. M. Hagstrom, M. P. Klein and D. A. Shirley, J. Chem. Phys. 48, 3779 (1968)

66 T. D. Thomas ibid. 52, 1373 (1970)

67 D. W. Davis, J. M. Hollander, D. A. Shirley and T. D. Thomas, ibid. 52, 3295 (1970)

68 J. A. Pople and G. A. Segal, ibid. 44 3289 (1966)

69 M. L. Poutsma J. Amer. Chem. Soc. 87, 2172 (1965)

70 R. C. Fahey and C. Schubert, ibid. 87, 5172 (1965)

71 W. H. Mueller and P. E. Butler, ibid. 88, 2866 (1966)

72 M. J. S. Dewar and F. C. Fahey, ibid. 85, 3645 (1963)

73 R. F. Merrit ibid. 89, 609 (1967)

74 F. C. Fahey and H. J. Schneider, ibid. 90, 4429 (1968)

75 A. Hassmer and C. C. Heathcock, Tetrahedron Lett. 1125 (1964)

76 B. Capone Quart. Rev. (London) 18, 45 (1964)

77 H. C. Brown Chem. Eng. News, 45 87 (Feb. 13, 1967)

78 G. A. Olah and J. M. Bollinger, J. Amer. Chem. Soc. 89, 4744 (1967)

79 W. J. Hehre and J. A. Pople, Tetrahedron Lett. 1959 (1970)

80 J. I. Braumann and L. K. Blair, J. Amer. Chem. Soc. 90, 6561 (1968)

81 L. Radom and J. A. Pople, ibid. 92, 4786 (1970)

82 W. A. Latham, W. J. Hehre and J. A. Pople, ibid. 93 808 (1971)

83 L. C. Allen and H. Basch, ibid. 93 6373 (1971)

84 R. E. Christofferson, D. W. Genson and G. M. Maggiova, J. Chem. Phys. 54, 239 (1971)

85 E. Clementi and W. Von Niessen, ibid. 54, 521 (1971)

86 A. Veillard Theo. Chim. Acta 18, 21 (1970)

87 B. Levy and M. C. Morreau, J. Chem. Phys. 54, 3316 (1971)

88 D. R. Whitman and C. J. Hornback, ibid. 51, 398 (1969)

89 W. H. Fink and L. C. Allen, ibid. 46, 2261, 2276 (1967)

90 L. C. Allen Chem. Phys. Lett. 2, 597 (1968)

91 J.-M. Lehn 'Proceedings of the Seminar on Computational Problems in Quantum Chemistry' Strassburg (1969)

92 J.-M. Lehn 'Conformational Analysis', Ed. G. Chiurdoglu, Acad. Press, 129 (1971)

93 R. Hoffmann J. Chem. Phys. 40, 2480 (1964)

94 R. Sustmann, J. E. Williams, M. J. S. Dewar, L. C. Allen and P. von R. Schleyer, J. Amer. Chem. Soc. 91, 5350 (1969)

95 G. V. Pfeiffer and J. G. Jewett, ibid. 92, 2143 (1970)

96 J. E. Williams, V. Buss, L. C. Allen, P. von R. Schleyer, W. A. Latham, W. J. Hehre and J. A. Pople, ibid. 92, 2141 (1970)

97 P. C. Hariharan, W. A. Latham and J. A. Pople, Chem. Phys. Lett. 14, 385 (1972)

98 R. Breslow 'Organic Reaction Mechanisms' Benjamin, New York (1966) p. 114

99 D. M. Brouwer, C. MacLean and E. L. Mackor, Disc. Farad. Soc. 39, 121 (1964)

100 U. Löhle and C. H. Ottinger, J. Chem. Phys. 51, 3097 (1969)

101 O. Chalvet, R. Daudel, I. Jano and F. Peradejordi, Modern Quantum Chemistry, Part 2. Academic Press, New York (1965)

102 M. L. Brogli and M. I. Jano, G. R. Acad. Sci. Paris (5 July 1965) t. 261, 103

103 G. J. Hojtink, E. DeBoer, P. M. Van Der Meij and W. P. Weijland, Rec. Trav. Chim. 75, 487 (1956)

104 K. B. Harvey and G. B. Porter, 'Introduction to Physical Inorganic Chemistry', Addison Wesley Pub. Co Inc., (1963) p. 326

105 L. C. Snyder Robert A. Welch Found. Bull. 29, (1971)

106 F. A. Cotton and G. Wilkinson 'Advanced Inorganic Chemistry', Wiley 3rd Edit.

107 V. I. Vednedev, L. V. Gurvich, V. N. Kondrat'yev, V. A. Medvedev, and Ye L. Frankevich, 'Bond energies, Ionisation Potentials and Electron Affinities'. Edward Arnold (1966)

108 T. L. Cottrell 'The Strength of Chemical Bonds', Butterworths, London (1954)

109 G. S. Hammond J. Amer. Chem. Soc. 77, 334 (1955)

110 A. Maccoll and P. J. Thomas, 'Progress in Reaction Kinetics', Ed. G. Porter, Pergamon Press (London), 4, 119 (1967)

111 Chem. Soc. Spec. Publication No. 18 (1965) S17s

112 R. W. Kilb, C. C. Lin and E. B. Wilson Jr., J. Chem. Phys. 26, 1695 (1957)

113 R. B. Davidson and L. C. Allen, J. Chem. Phys. 54, 2828 (1971)

114 L. Radom, J. A. Pople and P. von R. Schleyer, J. Amer. Chem. Soc. 94, 5935 (1972)

115 E. S. Gould 'Mechanism and Structure in Organic Chemistry', Holt, Rinehart-Winston (1959)

116 R. H. Martin, R. W. Taft and F. W. Lampe, J. Amer. Chem. Soc. 88, 1353 (1966)

117 T. B. McMahon, B. J. Blint, D. P. Ridge and J. L. Beauchamp, ibid. 94, 8934 (1972)

118 R. L. Flurry, D. Breen and D. L. Howland, Theo. Chim. Acta 20, 371 (1971)

119 G. A. Olah and J. Lukas, J. Amer. Chem. Soc. 89, 4739 (1967)

120 G. A. Olah and J. M. Bollinger, J. Amer. Chem. Soc. 90, 947 (1968)

121 B. Roos and P. Siegbahn, 'Proceedings of the Seminar on Computational Problems in Quantum Chemistry', Strassburg (1969)

122 S. Weiss and G. E. Leroi, J. Chem. Phys. 48, 962 (1968)

123 G. Sage and W. Klemperer, ibid. 39, 371 (1963)

124 R. H. Schwendeman and G. D. Jacobs, ibid. 36, 1245 (1962)

125 L. Radom, J. A. Pople, V. Buss and P. von R. Schleyer, J. Amer. Chem. Soc. 92, 6987 (1970)

126 G. Burns J. Chem. Phys. 41, 1521 (1964)

127 J. A. Pople Chem. Phys. Lett. 14, 385 (1972)

128 G. L. Cunningham Jr., A. W. Boyd, R. J. Meyers, W. D. Gwinn and W. I. Le Van, J. Chem. Phys. 19, 676 (1951)

129 D. T. Clark 5th Int. Symp. on Organic Sulphur Chem. Lund (1972)

130 H. O. House 'Modern Synthetic Reactions', Benjamin, New York (1965)

131 D. T. Clark, D. M. J. Lilley and M. Barber, Paper 38, Oxford Conference on Photoionisation and Photoelectron Spectroscopy. Sept. 1970

132 R. M. Acheson 'An introduction to the chemistry of heterocyclic compounds' 2nd Edit. Wiley

133 M. W. Austin, J. R. Blackborow, J. H. Ridd and B. V. Smith, J. Chem. Soc. 1051 (1965)

134 J. D. Vaughan, D. G. Lambert and V. L. Vaughan, J. Amer. Chem. Soc. 86, 2857 (1964)

135 A. Grimison and J. H. Ridd, Proc. Chem. Soc. 256 (1958)

136 C. H. Eugster and P. Bosshard, Helv. Chim. Acta 46, 815 (1963)

137 E. Clementi J. Chem. Phys. 46, 4731 (1967)

138 D. T. Clark and D. A. Armstrong, J. Chem. Soc.(D) 319 (1970)

139 U. Gelius, B. Roos and P. Siegbahn, Chem. Phys. Lett. 4, 471 (1970)

140 U. Gelius, B. Roos and P. Siegbahn, Theo. Chim. Acta 27, 171 (1973)

141 G. Berthier, L. Prand and J. Serre, 'Proc. Int. Symp. Quantum Aspects of Heterocyclic compounds in Chemistry and Biochemistry' (1969), Israel Acad. Sci. and Humanities, p. 40

142 U. Gelius, C. J. Allan, G. Johansson, P. Siegbahn, D. A. Allison and K. Siegbahn, Physica Scripta 3, 237 (1971)

143 D. T. Clark, D. Kilkast and W. K. R. Musgrave, J. Chem. Soc. (D) 516 (1971)

144 D. T. Clark and D. Kilkast, J. Chem. Soc.(B) 2243 (1971)

145 H. R. Mahler and E. H. Cordes, 'Biological Chemistry' Harper Int., 2nd Edit.

146 I. H. Krakoff 'Drills Pharmacology' Ed. J. J. DiPalma p. 1587

147 G. S. Parry Acta Cryst. 7, 313 (1954)

148 K. Hoogsteen ibid. 16, 28 (1963)

149 G. A. Jeffrey, S. Ghose and J. O. Warwicker, ibid. 14, 881 (1961)

150 D. T. Clark, R. D. Chambers, D. Kilkast and W. K. R. Musgrave, J. Chem. Soc. Farad II, 68, 309 (1972)

151 D. T. Clark, D. Kilkast and D. B. Adams, Farad. Disc. Chem. Soc. 54, 182 (1972)

152 B. Mely and A. Pullman, Theo. Chim. Acta 13, 278 (1969)

153 D. T. Clark and I. W. Scanlan, to be published.

154 J. Almlöf, B. Roos, U. Wahlgren and H. Johansen, J. Electron Spectrosc. 2, 51 (1973)

155 J. P. Kokko, J. H. Goldstein and L. Mandell, J. Amer. Chem. Soc. 83, 2909 (1961)

156 R. E. Marsh, R. Bierstedt and E. L. Eichhorn, Acta. Cryst. 15, 310 (1962)

157 G. A. Jeffrey and Y. Kinoshita, ibid. 16, 20 (1963)

158 R. A. Y. Jones, A. R. Katritsky and J. M. Lagowski, Chem. and Ind. (London) 27, 870 (1960)

159 E. M. Purcell Phys. Rev. 54, 818 (1938)

160 J. C. Helmer Appl. Phys. Lett. 13, 268 (1968)

161 D. N. Hendrickson, J. M. Hollander and W. L. Jolly, Inorg. Chem. 8, 2642 (1969)

162 D. T. Clark, D. Kilkast, D. B. Adams and W. K. R. Musgrave, J. Electron Spectrosc. 1, 227 (1972/73)

163 U. Gelius and K. Siegbahn, Farad. Disc. No. 54 (1972)