A THEORETICAL STUDY OF REACTIVITIES AND SYNTHESIS OF MOLECULES: APPLICATIONS AND EXTENSIONS OF THE THEORY OF ATOMS IN MOLECULES

By

CHENG CHANG, M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

A THEORETICAL STUDY OF THE REACTIVITY AND SYNTHESIS OF MOLECULES

DOCTOR OF PHILOSOPHY (1990) McMASTER UNIVERSITY

(Chemistry) Hamilton, Ontorio

TITLE: A THEORETICAL STUDY OF THE REACTIVITY AND SYNTHESIS OF MOLECULES: APPLICATIONS AND EXTENSIONS OF THE THEORY OF ATOMS IN MOLECULES

AUTHOR: Cheng Chang, B.Sc. (Beijing Teacher's College)
M.Sc. (Beijing Teacher's College)

SUPERVISOR: Professor R.F.W. Bader

NUMBER OF PAGES: xv, 198

ABSTRACT

This thesis applies and extends the theory of atoms in molecules to some general, important and chemically interesting subjects: a means of studying molecules of biochemical size using high-level quantum chemistry, the chemical reactivities of aromatic molecules as well as the nature of π-type hydrogen bonding.

Quantum mechanical principles define an atom in a molecule and its properties as well as the chemical bonds which link the atoms to yield a molecular structure. The resulting theory is called the theory of atoms in molecules. The theory also predicts the chemical reactivity of a molecule through the properties of the Laplacian of its electronic charge distribution, the Laplacian of the charge density also provides the physical basis of the Lewis and VSEPR models. The theory is reviewed in Chapter 1.

Molecules of biochemical interest always cause difficulties for computational chemists because of the size. Semiempirical methods have been used in the past, but must alternatively be replaced by ab initio theory. An attempt to use a nonempirical method with reliable trial functions to solve such problems, as made possible through the application of the tools provided in the theory of atoms in molecules, is presented in Chapter 2, and gives an encouraging result.

The study of the chemical reactivities of aromatic molecules is of general interest to chemists. An enormous amount of research has been done in this area. A new way of using the theory to study the reactivities and energetics of monosubstituted benzenes and their para and meta protonated intermediates is presented in Chapters 3 and 4. The directing ability of a

substituent and its ability to activate or deactivate the phenyl group in electrophilic substitution reaction are related to atomic, bond and molecular properties of the aromatic molecules.

The nature of π-type hydrogen bonding is discussed in Chapter 5. The molecular structures of three C_6H_6HF complexes are determined for the first time.

ACKNOWLEDGMENTS

I wish to take this chance to thank all the members of Dr.Bader's Theoretical Chemistry Laboratory for their help, cooperation and friendship during the time I have been studying here.

This is the first time for me to leave my country to study in a completely different environment -- different language, different culture, different concepts of value, almost everything is different except for the language and concepts of science. It was the friendship of Dr. Marshall Carroll and Mr. James Cheeseman that encouraged me to overcome many unexpected difficulties with regard to both the study and life aspects of my early time here.

I would also like to thank Dr. Preston MacDougall for his help in my understanding of many important concepts of the theory; talking with him about science has always been stimulating.

I would like particularly to thank Dr. Keith Laidig and Mr. Danny Legare. The former contributed to developing, maintaining and converting the programs I used and the computers on which they have been run; a part of this thesis could not be done without his work. The latter contributed his precious time in proof-reading and helping me to correct my English grammar.

I would like to give my special thanks to Dr. Richard Bader, my supervisor. I have really enjoyed the many discussions I have had with him, from which I have learned so much, although at times the discussions were filled with heated arguments and differences of opinion. His enthusiasm in approaching chemistry and his being ultimately responsible for each of the subjects on which we collaborated and for this thesis have

left a deep impression on me.

This thesis is dedicated to my wife, Sufang Wang, my parents and my motherland -- China.

TABLE OF CONTENTS

TABLE OF CONTENTS (con'd)

LIST OF FIGURES

LIST OF FIGURES (con'd)

LIST OF TABLES

LIST OF TABLES (con'd)

LIST OF TABLES (con'd)

LIST OF TABLES (con'd)

Introduction

Science is based upon experiment and observation. Its purpose is to enable us to predict the natural events that occur around us and to understand them as best we can. The purpose of theory is to provide the conceptual framwork in this quest for understanding and predicting observation (Bader 1990).

Chemical atoms, bonds, and the relation between the structures and reactions of chemical compounds serve as the central concepts of chemistry around which a growing number and variety of chemical reactions have been organized, classified and thus partially understood. Many efforts have been made towards understanding how to model these concepts, the most important developments among th · being the Lewis model, resonance theory, molecular orbital theory and ab initio electronic structure theory. Though the Lewis model, resonance theory and molecular orbital theory can successfully explain some aspects of chemical bonding and chemical reactions, they remain as chemical models and lack physical grounds. Ab initio electronic structure theory is derived from quantum mechanics, however, it has little to say about the concepts of atoms and bonds.

The theory of atoms in molecules is derived from fundamental principles of physics and it defines all the important chemical concepts mentioned above. The theory is obtained when the definition of a quantum subsystem is applied to the observed topological properties of a system's distribution of charge. The topology of ρ, the charge density, as displayed in the global properties of its gradient vector field, yields a faithful mapping of the chemical concepts of atoms, bonds and structure, while the topology of the Laplacian of ρ provides a mapping of the electron

pairs of the Lewis model and reveals the reactivities of a total chemical system.

The theory of atoms in molecules predicts the properties of atoms in molecules just as quantum mechanics predicts the properties of the total system, simply because the theory is a natural extension of quantum mechanics which is normally applied to a total system. When a molecule, which is a total system when isolated, is put into an environment it becomes a subsystem. Thus one has to use subsystem quantum mechanics to study it. Therefore the theory of atoms in molecules is also a complement of quantum mechanics for a total system.

This thesis presents an extension and application of the theory of atoms in molecules to the field of biochemistry (Chapter 2), organic (Chapter 3,4 and 5) and inorganic chemistry (Chapter 5).

CHAPTER 1

The Theory of Atoms in Molecules

This Chapter introduces the theory of atoms in molecules (Bader 1985, Bader and Nguyen-Dang 1981, Bader et al 1981). Discussions of the topological properties of the electronic charge distribution and its Laplacian together with the properties of atoms in molecules are emphasized as these ideas are central to the original work presented in this and the following chapters.

Section 1.1 gives the definition of the charge density ρ. While section 1.2 details its topological features which yields the definitions of an atoms in a molecule and of molecular structure. Section 1.3 describes another important part of the theory, the topological features of the Laplacian of the charge distribution. Section 1.4 delineates the atomic properties of interest in this work.

The conventional atomic units ($\hbar = e = m_e = 1$) are used in the equations and results throughout this work unless otherwise noted.

1.1. Molecular Electronic Charge Distribution

The electronic charge density $\rho(\mathbf{r})$,or charge density is a fundamental physical quantity which is experimentally measurable. The electron density can be determined in x-ray diffraction studies of crystals (Stewart 1979) and can be calculated theoretically. In this work, only theoretical charge densities obtained from single-determinant state functions calculated by the Hartree-Fock-Roothaan (HFR) procedure for a fixed configuration of the nuclei are used. The HFR procedure is a single-

determinant self-consistent field method using a finite basis set of atomic orbitals centered on the nuclei (LCAO-SCF method).

The state wave function Ψ determines all of the information that can be known about a quantum system. For a fixed arrangement of the nuclei, the state function, which is denoted by $\Psi(x)$, satisfies the stationary state Schrödinger equation

$$\hat{H}\Psi(x) = E\Psi(x) \qquad [1.1]$$

where x denotes the collection of electronic space and spin coordinates. The charge distribution for a system is related to the state function through the definition

$$\rho(r) = n \sum_{spins} \int d\tau_2 \int d\tau_3 \cdots \int d\tau_n \Psi^*(x)\Psi(x) \qquad [1.2]$$

where n is the number of electrons, r is the set of space coordinates of one electron, x is the collection of electronic space and spin coordinates and $d\tau_i$ denotes the Cartesian coordinates of the ith electron. For the Hartree-Fock-Roothaan SCF procedure, $\rho(r)$ can further be expressed as

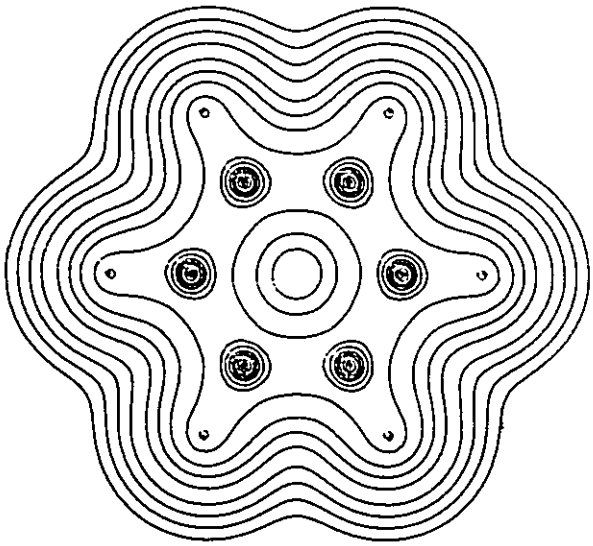
$$\rho(r) = \sum_i n_i \phi_i(r)\phi_i(r) \qquad [1.3]$$

where n_i is the electronic occupation number of the i^{th} real Hartree-Fock spin orbital $\phi_i(r)$. The physical meaning of $\rho(r)$ is n times the probability of finding an electron in a unit volume at the point in space denoted by the position vector $r = ix + jy + kz$.

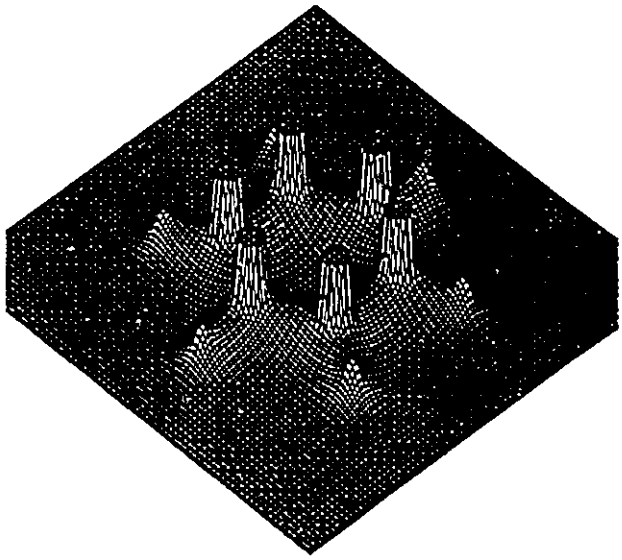
1.2. Topological Properties of the Charge Distribution

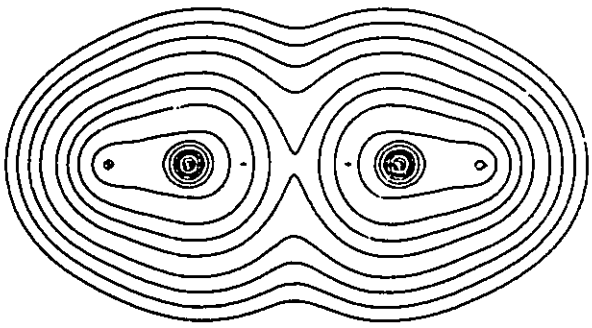
1.2.1 Critical Points

The charge density, $\rho(r)$, is a scalar field defined over three-dimensional space. The topological properties of such a scalar field are conveniently summarized in terms of the number and kind of its critical

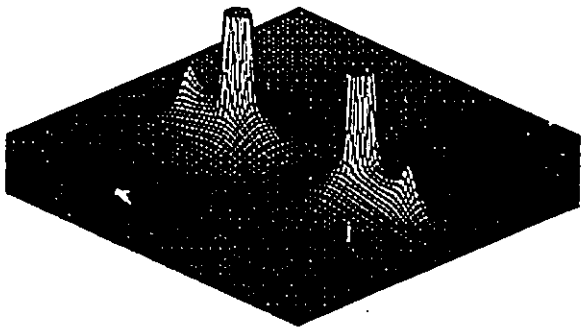


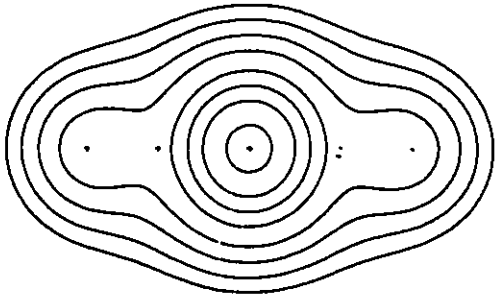
A





B





C

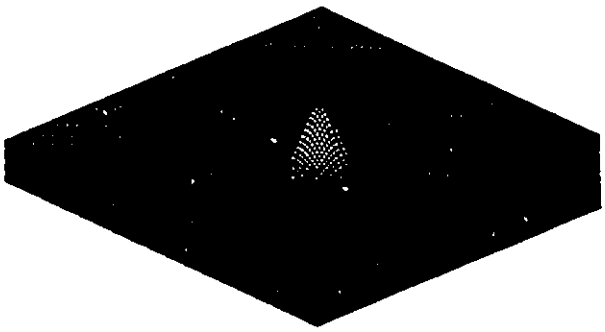


Figure 1.1.

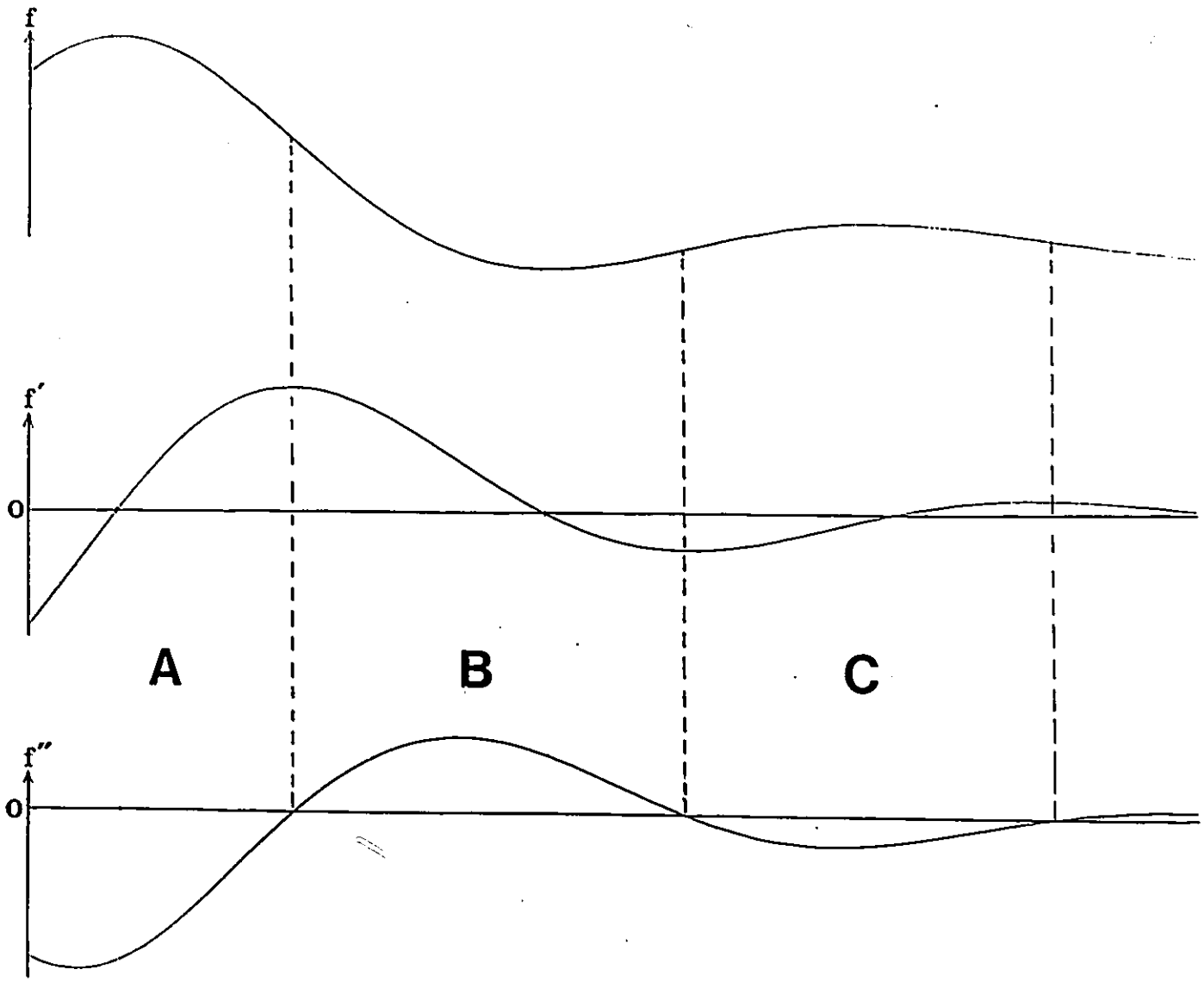
Contour and relief maps of the electronic charge density for three planes of the benzene molecule, C_6H_6. (A) The plane containing the nuclei. (B) The plane perpendicular to the above and containing two pairs of carbon-hydrogen bonds. (C) The plane perpendicular to a carbon-hydrogen bond at its bond critical point. The outermost contour of ρ has a value of 0.001 au and the contours increase in steps of 2×10^n, 4×10^n, and 8×10^n with n beginning at -3 and increasing in steps of unity. A nuclear position is denoted by a cross, and the projected positions of out-of-plane nuclei are indicated by open crosses. This set of contour values is used throughout the thesis unless otherwise noted. The charge densities have been generated from state functions calculated using the RHF/6-31G**//6-31G scheme (Hariharan and Pople 1973 for 6-31** basis; Hehre et al 1972a for 6-31G basis).

points, points where $\rho(r)$ is an extremum. Their characterization which will be given in the following paragraphs.

A contour plot of the charge distribution is a set of lines which connect points of equal charge density. Figure 1.1 displays contour plots of the charge density of benzene in the molecular plane, in a plane which is perpendicular to the molecular one and contains two pair of carbon-hydrogen bonds, and in a plane perpendicular to the C-H bond axis. The corresponding relief plots, which show the relative height of each contour line in the third dimension are also displayed.

From Fig.1.1, one can find three topological features. First of all, in any plane containing nuclei, ρ exhibits the local maxima at the positions of the nuclei (see Fig.1.1A and B). Secondly, the charge density appears as a saddle point between some neighboring nuclei in the plane containing nuclei, Fig.1.1A and B, and shows a maximum in the plane through the same point in the plane perpendicular to the above plane, Fig.1.1C. The third feature is that at the centre of the benzene ring, the charge density displays a minimum in the molecular plane, Fig.1.1A, and appears as a saddle point in the plane perpendicular to the molecular plane, Fig.1.1B. One more topological feature can be found in a molecule possessing a cage structure such as found in tetrahedrane C_4H_4 (Bader et al. 1981), for which ρ exhibits a point which is a minima with respect to all directions.

Each topological feature of $\rho(r)$, whether it be a maximum, a minimum or a saddle, has associated with it a point in space called a critical point, where the first derivatives of $\rho(r)$ vanish. Thus at such a point denoted by the position vector r_c, $\nabla\rho(r_c) = 0$, where $\nabla\rho$ denotes the operation


f
f′
0
A
B
C
f″
0

$$\nabla\rho = i\partial\rho/\partial x + j\partial\rho/\partial y + k\partial\rho/\partial z, \qquad [1.4]$$

or equally we have

$$\partial\rho/\partial r_i = 0 \qquad r_i = x, y, z \qquad [1.5]$$

Whether a function is a maximum or a minimum at an extremum is determined by the sign of its second derivative or curvature at this point. This is shown in Figure 1.2, which displays the behaviour of an arbitrary one-dimensional function, y = f(x), and of its first and second derivatives. From Fig.1.2, it is clear that whenever the function y = f(x) reaches its extremum, y', the first derivative of f(x) has the value zero, as the tangent of f(x) is going to change its sign. Accordingly, the second derivative or the curvature of f(x), y'', is negative for a maximum in f(x) and positive for a minimum in f(x).

It is clear now from Fig.1.1 that a critical point at a nuclear position has three negative curvatures, Fig1.1A and B. The critical point between certain pairs of nuclei has one positive curvature along the internuclear axis, Fig.1.1B, and a negative curvature along each of the two perpendicular axes, Fig.1.1C. The critical point at the centre of the ring has two positive curvatures along the two axes in the molecular plane, Fig.1.1A, and one negative curvature along the axis perpendicular to the plane, Fig.1.1B.

A critical point can be classified according to its rank and signature, written as (rank, signature). The rank of a critical point at r_c equals the number of nonzero eigenvalues (principal curvatures) of the Hessian matrix A where

$$A_{ij} = (\partial^2\rho/(\partial r_i \partial r_j))_{r=r_c} \qquad [1.6]$$

is a real symmetric matrix which can therefore be diagonalized. Associated

Figure 1.2

Plots of a function, $f(x)=1-x*\exp(-x)+(x-2)^2+x^3*\cos(9x)$, and its first and second derivatives.

with the principal curvatures are the corresponding principal axes (the eigenvectors μ_1, μ_2 and μ_3). The signature is the sum of the signs of the eigenvalues. For example, for a nuclear position, the curvatures of the charge density along all the three axes are negative, thus it is of rank 3 and signature -3.

A critical point with its rank less than three, i.e., with at least one zero curvature, is said to be degenerate. Such a critical point is unstable in the sense that a small change in the charge density, as caused by a displacement of the nuclei, causes it to either vanish or to bifurcate into a number of non-degenerate or stable (rank = 3) critical points.

In the charge distributions studied in the present work, only critical points of rank 3 are found. There are four types of critical point of rank 3: (3,-3), (3,-1), (3,+1) and (3,+3).

The (3,-3) critical point shows a maximum in ρ in all three perpendicular directions -- a local maximum of charge concentration, this is the topological behaviour of a nuclear position in the general case (Bader and Nguyen-Dang 1981) or of free electrons in metal clusters (Cao et al 1987, Gatti et al 1987).

There are two negative and one positive curvatures at a (3,-1) critical point and the sum of their signs is -1. The two eigenvectors (μ_1 and μ_2) associated with the negative eigenvalues define a surface on which the critical point is a local maximum, see Figure 1.1 (c). The eigenvector (μ_3) associated with the positive eigenvalue defines a unique line linking the neighboring nuclei along which ρ increases for motion away from the critical point.

The (3,+1) and (3,+3) critical points are called ring and cage

critical point, respectively. The eigenvectors associated with the two positive curvatures of a ring critical point define the ring surface with local minimum in ρ (see Fig.1.1 a), and the remaining eigenvector defines a unique axis perpendicular to the ring surface at the critical point and ρ exhibits a maximum at the point along the axis, (see Fig.1.1 b). The charge density is a minimum with respect to all directions at a cage critical point. The value of ρ at both ring and cage critical points is necessarily smaller than the value of ρ evaluated at the surrounding (3,-1) critical points.

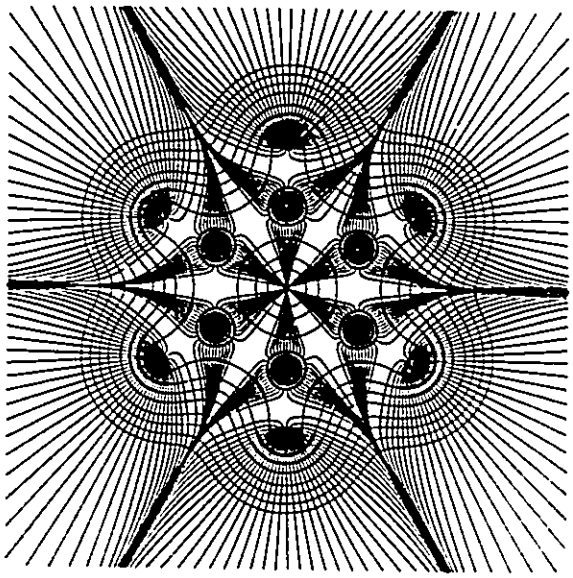
1.2.2 Gradient Vector Field of the Charge Density and the Definition of Atoms in Molecules.

It is instructive to examine the gradient vector field of the charge density (Fig.1.3) to gain more understanding of critical points and derive the concept of atoms in molecules. The gradient vector field of the charge density is represented through a display of the trajectories traced out by the vector $\nabla\rho$ (for its definition, see Eq.[1.4]). A trajectory of $\nabla\rho$, also called a gradient path, starting at some arbitrary point r_0 is obtained by calculating $\nabla\rho(r_0)$, moving a distance Δr away from this point in the direction indicated by the vector $\nabla\rho(r_0)$ and repeating this procedure until the path so generated terminates.

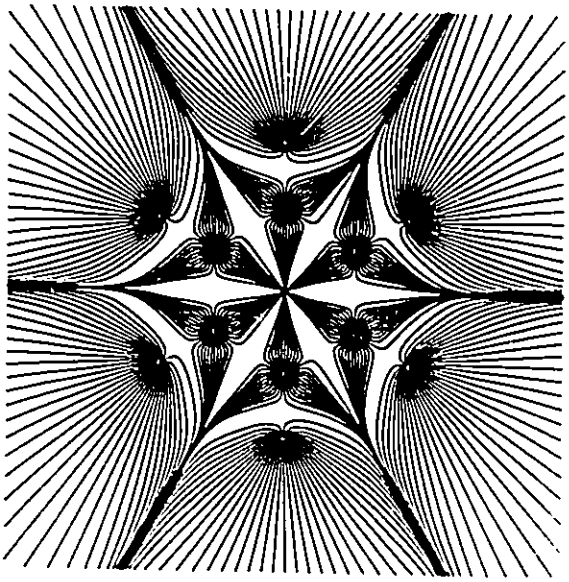
Mathematically, the gradient path is the integral curve of the differential equation

$$dr(s)/ds = \nabla\rho(r(s)) \qquad [1.7]$$

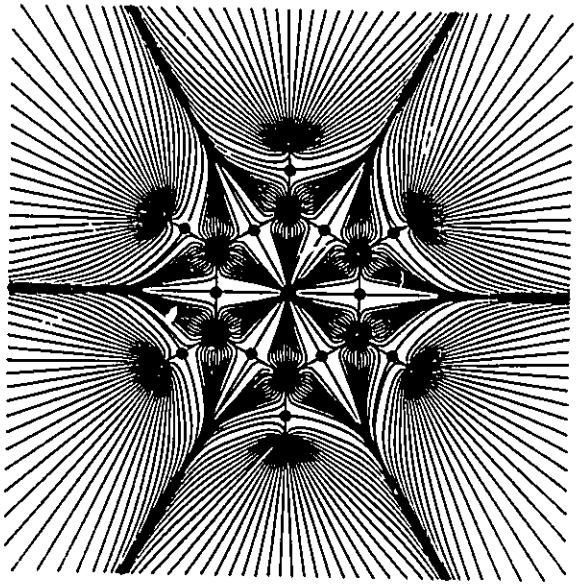
for some initial value $r(0) = r_0$, where s is a path parameter. Thus the points $r(s)$ of gradient path through r_0 are given by



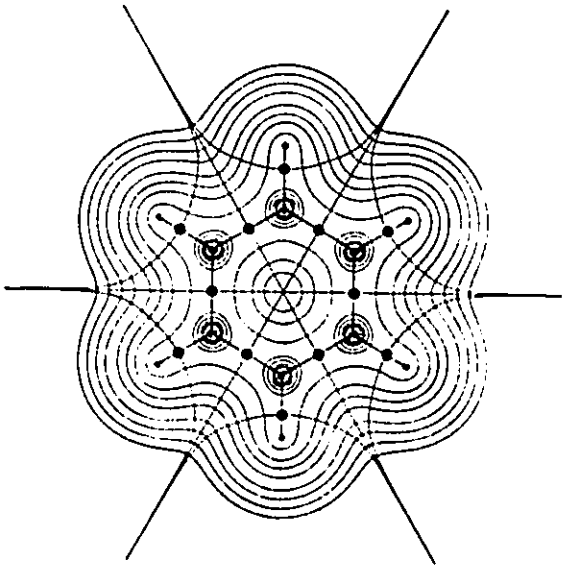
a



b



c



d

$$r(s) = r_o + \int_0^{\infty} \nabla\rho(r(t))dt \qquad [1.8]$$

Each and every such path starts and ends at a critical point or at infinity.

Fig.1.3(a) is a diagram of the gradient paths of the benzene molecule overlaid on the contour map of the charge density; from this we have the following observations:

a) Since the gradient vector of a scalar points in the direction of steepest ascent in the scalar, the trajectories of $\nabla\rho$ are perpendicular to lines of constant density -- the contour lines of ρ.

b) The vector $\nabla\rho$ is tangent to its trajectory at each point r.

c) Trajectories cannot cross since $\nabla\rho(r)$ defines but one direction at each point r.

Figure 1.3 (b) is a diagram of the gradient paths of the benzene molecule. From this figure we can observe some interesting features. First, all the trajectories terminate at the positions of nuclei which we know are the positions of maximum charge density, the (3,-3) critical points. Secondly, if one follows the trajectories between two neighboring nuclei down from the nuclear position, one will walk down from one nucleus towards another, then suddenly turn away. The third observation is that there is a border between each two neighboring nuclei, the trajectories from each nucleus never cross over these borders. Thus, without further mathematical analysis, the display of the gradient paths for a molecule makes visible to the eye the definitions of its atoms. It is clear that each of the atoms is enclosed by several border lines, through which no trajectory passes, which are defined as interatomic surfaces, or is opened to the infinity. In the regions where trajectories change direction, we

Figure 1.3

Maps of the gradient vector field of the charge density for the plane containing the nuclei of the benzene molecule. Each line represents a trajectory of $\nabla\rho(r)$. In (a) Trajectories which terminate at the positions of the nuclei are superimposed on the charge density contour plot. (b) Only trajectories displayed in (a) are shown. Each trajectory is arbitrarily terminated at the surface of a small circle about a nucleus. The set of trajectories which terminate at a given nucleus or attractor define the basin of the attractor. (c) The same as (b), but overlaid with the trajectories associated with the (3,-1) critical points. These trajectories define the boundaries of the atoms and the molecular graph. Bond critical points, which are denoted by dots, are also shown. (d) A superposition of the trajectories associated with the (3,-1) critical points on a contour map of the charge density.

can always find a point at which $\nabla\rho = 0$ and two curvatures of ρ are negative and one positive -- a (3,-1) critical point.

From above, we see that nuclei act as attractors of the gradient vector field of $\rho(r)$. The largest neighborhood containing trajectories such that any trajectory originating in it terminates at a nucleus is called a basin. An atom, free or bound, is defined topologically as the union of an attractor and its associated basin.

Alternatively, the atom can be defined in terms of its boundary. The basin of a single nucleus in an isolated atom spans the entire three dimensional space R^3. For an atom in a molecule the atomic basin is a subspace of R^3. This atom is separated from neighboring atoms by interatomic surfaces, each of which always contains a bond critical point. The interatomic surface S_{AB} consists of all gradient paths which terminate at the (3,-1) critical point. The atomic surface S_A of an atom A is defined as the boundary of the basin. In general this boundary is the union of a number of interatomic surfaces, separating two neighboring basins, and some portions which may be infinitely distant from the attractor. The interatomic surfaces and the surfaces found at infinity are the only surfaces S of R^3 which satisfy the equation

$$\nabla\rho(r)\cdot n(r) = 0 \quad \text{for every } r \in S \qquad [1.9]$$

where n is the unit vector normal to the surface at r, pointed outward. A surface satisfying the condition (Eq.[1.9]) is called a zero flux surface. An atom, isolated or bound, is a region of real space containing a single nucleus and bounded by a zero flux surface. The atomic boundaries displayed in Fig.1.3c represent the intersection of the atomic surfaces with the given plane.

1.2.3 Chemical Bond, Molecular Structure and the Properties of Bond Critical Point

All the gradient paths on the interatomic surface S_{AB} terminate at the (3,-1) critical point between the atoms A and B. However, in the axis perpendicular to the surface S_{AB}, one and only one pair of gradient paths originate at the (3,-1) critical point and terminate at the neighboring nuclei of A and B. This pair of gradient paths forms a line linking the two nuclei, the line generated by μ_3 of the critical point, along which the charge distribution is a maximum with respect to any lateral displacement, is termed an atomic interaction line. For a molecular system at an equilibrium geometry this line is classified as a bond path, and the (3,-1) critical point on this line is called a bond critical point. Two atoms are bonded to each other if and only if a bond path connects the two atoms.

The network of bond paths linking pairs of nuclei is called the molecular graph of the system. A molecular graph summarizes which atoms are bonded to one another, i.e. the molecular structure. The molecular graph is in agreement with chemical expectations. It is important to note that only one bond path links two bonded atoms. For example, the two carbon nuclei in the ethylene molecule are linked by one bond path; "double bonds" do not appear in ρ as two lines linking two C; only a single path is observed. However the "double bond" character can be found in the properties of bond critical point, the bond order and the ellipticity ϵ, which are discussed below.

The concept of molecular structure is different from that of molecular geometry. A molecular structure is a representation of which

atoms are bonded to one another in a molecule. It is a generic property. A molecular geometry gives the internuclear distances and angles in the molecule. Geometry is a non-generic property since any infinitesimal change in a set of nuclear coordinates results in a different geometry. The molecular geometry may be varied without changing the molecular structure but the converse is not true. All molecular geometries having the same molecular graph belong to the same structural region. The theory also describs the mechanisms of structural change (Bader et al 1981).

The bond critical point is important in the theory of atoms in molecules. The value of the charge density evaluated at a bond critical point gives a measure of the degree of accumulation of electronic charge between the given nuclei. It is a general observation for covalent bonds that increasing strength through a similar series of bonds is reflected in a corresponding increase in the amount of charge density at the bond critical point. The bond order has been defined for C-C bonds using the values of ρ_b as

$$n = A \cdot \exp(B - \rho_b) \qquad [1.10]$$

where A and B are constants (Bader et al 1983).

In a bond with cylindrical symmetry such as the bonds in diatomic or linear molecules, the two negative curvatures of ρ at the bond critical point (λ_1 and λ_2) are of equal magnitude. However, if electronic charge is preferentially accumulated in a given plane containing the bond path (as it is for a bond with π-character, see Figure 1.4 for the contour diagram of the C-C bond of benzene as an example), then the rate of falloff in ρ from its maximum value $\rho(r_c)$ in the interatomic surface is less along the axis lying in this plane than along the one perpendicular to it, and the

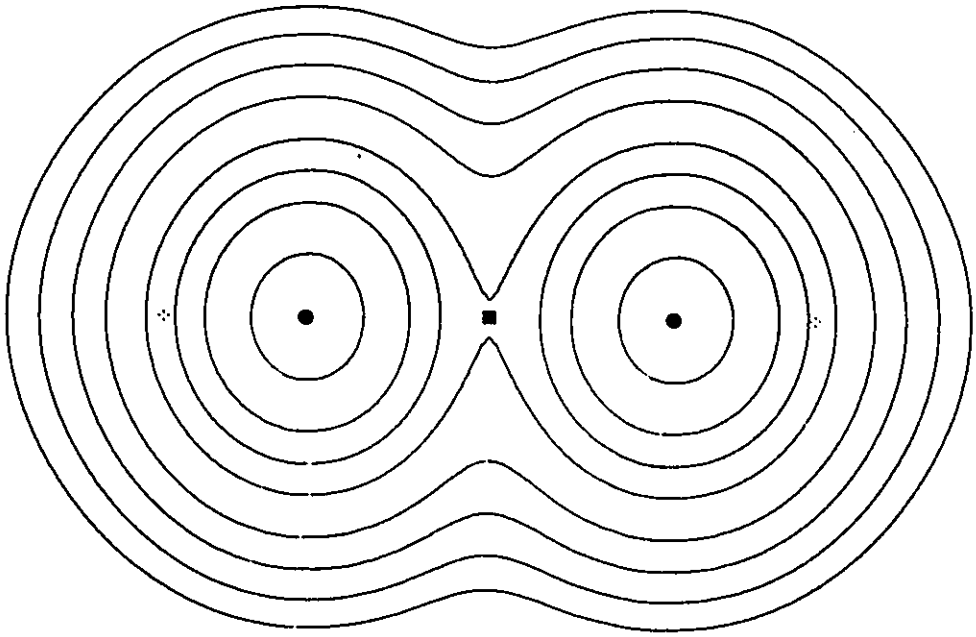


Figure 1.4

Contour plot of the electron density in a plane perpendicular to plane of nuclei through a pair of carbon-carbon bond critical points (denoted in solid dot) of benzene. The central point is the ring critical point (denoted in a solid triangle).

magnitude of the corresponding curvature of ρ is smaller. If λ_2 is taken to be the value of smaller curvature, then the ellipticity of the bond

$$\epsilon = (\lambda_1/\lambda_2 - 1) \qquad [1.11]$$

provides a measure of the extent to which charge is preferentially accumulated in a given plane. The bigger the ellipticity of the bond, the greater the π character. Any bond with cylindrical symmetry has an ellipticity ϵ=0, while the C-C bond of the benzene molecule has ellipticity ϵ=0.231 with the 6-31G**//6-31G (Hariharan and Pople 1973) calculation for comparison. The axis of the softer curvature λ_2, the major axis, determines the relative orientation of this plane within a molecule.

The sign of the value of $\nabla^2\rho$ at each critical point may also be used as the criterion for determining the type of interaction for the two atoms involved. In general,- a negative sign of $\nabla^2\rho$ means a covalent interaction while a positive sign means a closed shell interaction, intereactions which are detailed in section 1.3.1.

1.3. Topology of the Laplacian of the Charge Density

The topology of a molecular charge distribution yields a unified theory of molecular structure, one that defines atoms, chemical bonds, structure and changes in structure. However, the topology of ρ does not give any evidence of a shell structure in an atom or information related to Lewis model for chemical reactions which is of general interest to chemists. These requirements can be achieved by the study of the topology of the Laplacian of the charge density.

The Laplacian of ρ is defined by the equation

$$\nabla^2\rho(r) = \partial^2\rho/\partial x^2 + \partial^2\rho/\partial y^2 + \partial^2\rho/\partial z^2. \qquad [1.12]$$

It is the sum of the three principal curvatures of the Hessian of ρ. AS explained below, it determines where electronic charge is locally concentrated and depleted (Bader 1989, Bader et al 1988, Bader and MacDougall 1985, Bader and Essen 1984, Bader et al 1984). To aid in understanding the concepts of charge "concentration" and "depletion" as determined by the Laplacian of ρ, we examine the behaviour of a function f(x) displayed in Fig.1.5 together with its first and second derivatives. The slope of f(x) is defined as

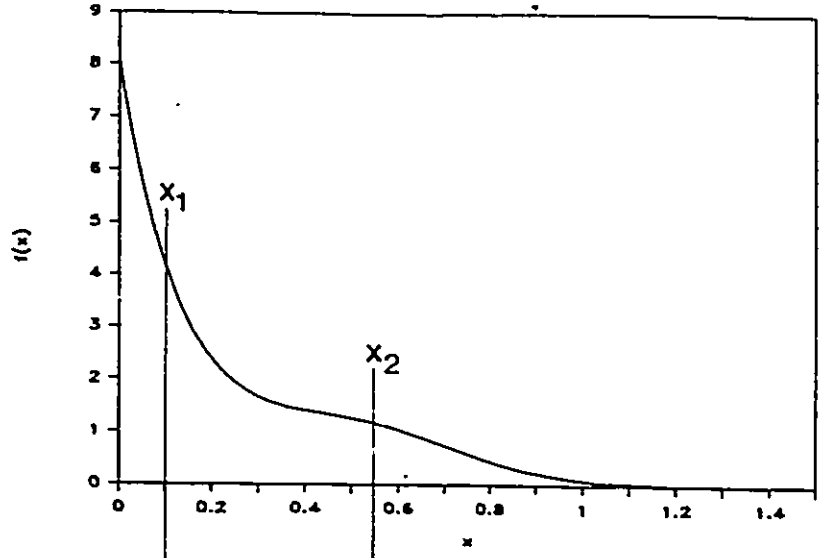
$$df(x)/dx = f'(x) = \{[f(x+\Delta x)-f(x)]/\Delta x\}_{\Delta x \to 0} \qquad [1.13]$$

and its second derivative, at point x, is the limiting difference in the slope of f(x) at the points $x+\Delta x$ and $x-\Delta x$, which bracket the point x:

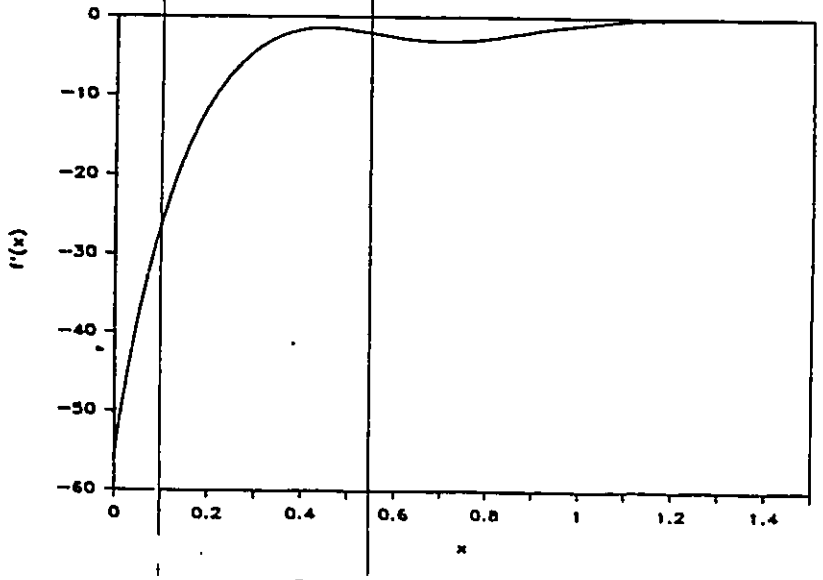
$$d^2f/dx^2 = f''(x) = (\{[f(x+\Delta x)-f(x)]/\Delta x - [f(x)-f(x-\Delta x)]/\Delta x\}/\Delta x)_{\Delta x \to 0} \qquad [1.14]$$

Since f(x) is a steeply rising function in the region of x_1 in Fig.1.5, the magnitude of the slope at $x_1-\Delta x$ is greater than that at $x_1+\Delta x$ and, since both slopes are negative, the curvature of f(x) at x_1 is positive. The function exhibits a shoulder in the region of the point x_2, and the slope at $x_2+\Delta x$ is of greater magnitude than at $x_2-\Delta x$. Thus the curvature of f(x) is negative in the region where f(x) exhibits a shoulder. The function exhibits a point of inflection between x_1 and x_2 where the curvature is zero. From Eq.[1.14] for the curvature, one sees that where the curvature of f(x) is negative, the value of f(x) at x is greater than the average of its values at the neighboring points $x+\Delta x$ and $x-\Delta x$ with the reverse being true where the curvature is positive. It is in this sense that f(x) is said to be concentrated in regions where $d^2f(x)/dx^2$ is less than 0 and depleted where $d^2f(x)/dx^2$ exceeds 0. To graphically emphasize the fact

$$f(x) = 8\exp(-7x) + \exp(-10(x-0.5)^{**}2)$$



First derivative



Second derivative

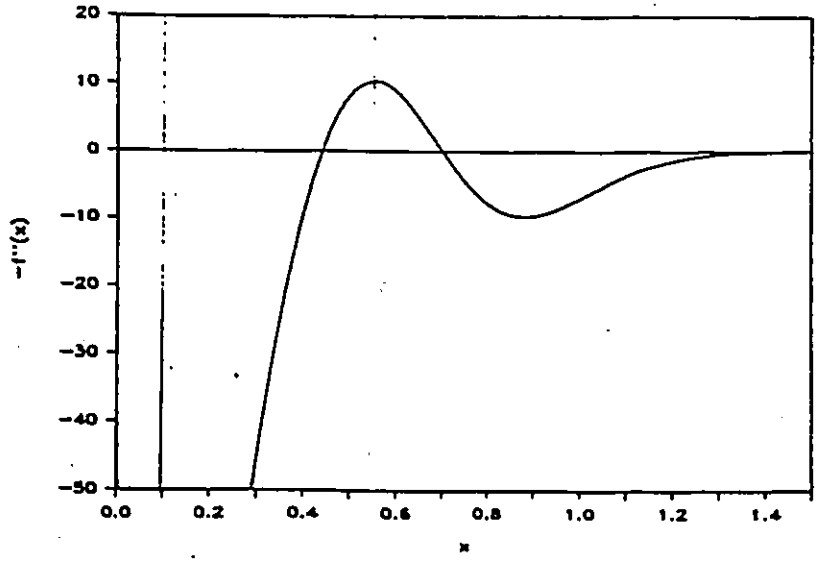


Figure 1.5

Plots of a monotonically decreasing function f(x) and its first and second derivatives. The negative of the second derivative is shown to emphasize that a function is concertrated in a region where its second derivative is negative.

that f(x) is concentrated where its curvature is negative, the negative of the curvature is plotted. Note that the curvature exhibits its largest negative value at x_2 and f(x) is said to be maximally concentrated at that point, the point where the negative of the curvature exhibits a maximum.

The slope of f(x) in Fig.1.5 is always negative. Thus f(x) itself does not exhibit either maxima or minima (ie, points where $f'(x)=0$) aside from the maximum at x=0. We see that the existence of regions where a function is locally concentrated or depleted as determined by the sign of its second derivative does not imply the corresponding existence of maxima or minima in the function itself.

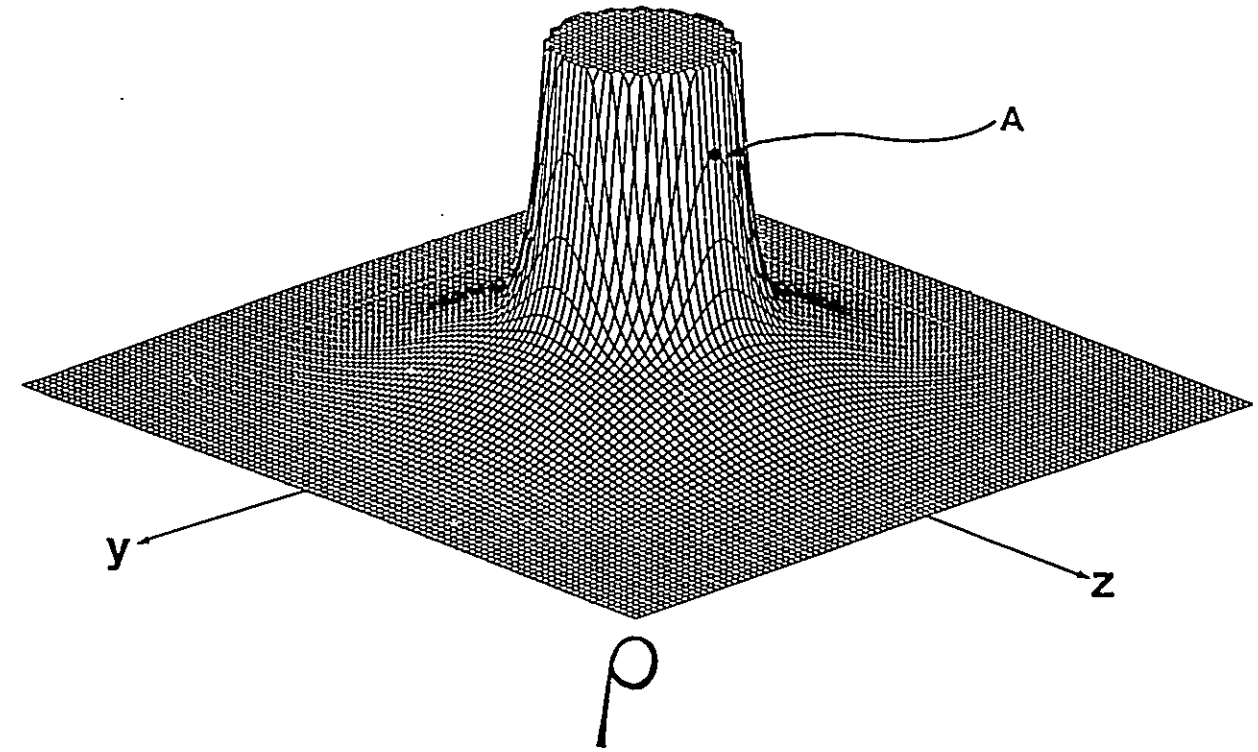
All the foregoing considerations carry over into three dimensions, and in particular it follows from Eq.[1.14] that the value of $\rho(\mathbf{r})$ is greater than the average of its values over an infinitesimal sphere centered on r when the sum of the three curvatures of ρ is negative, that is, when $\nabla^2\rho(\mathbf{r})<0$ and $\rho(\mathbf{r})$ is less than this average when $\nabla^2\rho(\mathbf{r})>0$. The charge density decays exponentially from a nucleus and, in general, the curvature of $\rho(\mathbf{r})$ along a radial line from the nucleus is positive, as for f(x) at x_1 in Fig.1.5, while the two curvatures perpendicular to a radial line are negative. Thus $\nabla^2\rho(\mathbf{r})$ may change sign even in the absence of a shoulder in the charge distribution, the presence of which in any event is relatively rare.

1.3.1 Atomic Shell Structure

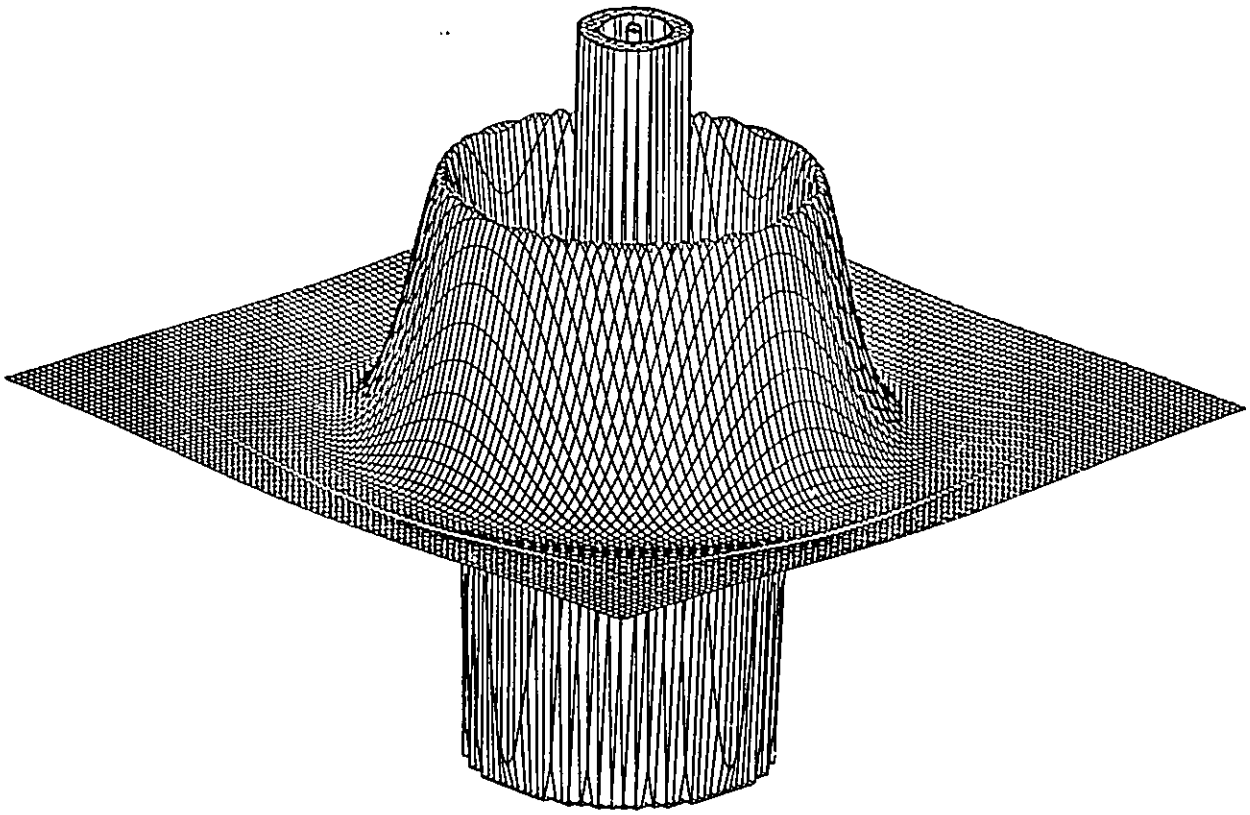
The Laplacian recovers the shell model of electronic structure in an atom by displaying a corresponding number of alternating shells of charge concentration and charge depletion beginning with a region of charge

concentration at the nucleus. For example, Fig.1.6 displays the value of the charge density and of its associated Laplacian distribution as a function of position for an isolated argon atom. Since the charge distribution is spherically symmetric, the same display is obtained for any plane containing the nucleus. The charge density exhibits a single maximum at the position of the nucleus (the maximum value is not shown) and decays monotonically for motion away from the nucleus along any radial line. The charge density does not show the atomic shell structure but its Laplacian does. If we consider a spherical "Laplacian shell" to consist of a region where the Laplacian is negative surrounded by a region where the Laplacian is positive, then the number of Laplacian shells is equal to the number of shells that are partially or fully occupied within the orbital model of electronic structure. For example, argon is a third row atom, thus we see three shells in Fig.1.6, denoted as inner shells and the valence shell, respectively.

Figure 1.7 displays two kinds of interactions: those of the shared (H_2, B_2, N_2, O_2 and NO) and closed shell (He_2, Ar_2, KF, LiCl and NaCl). In Fig.1.7 the positive values are shown in the solid line while those which are negative in the dashed line. Around each atom of the molecules in the column at the left hand side of Fig.1.7, the number of the identifiable Laplacian shells remains the same as for the isolated atom, in spite of the fact that its outer shells have been distorted. This is the general case of shared interaction, since neither atom loses all of the valence electrons but shares them in the interaction. In the column at the right hand side of Fig.1.7, each atom appears a full electron shell after the interaction. The Laplacian map for Ar_2 shows that more electronic charge



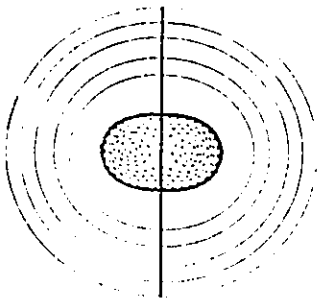
ρ



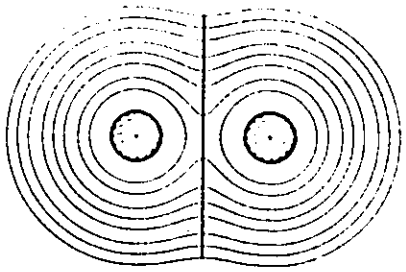
$-\nabla^2\rho$

Figure 1.6

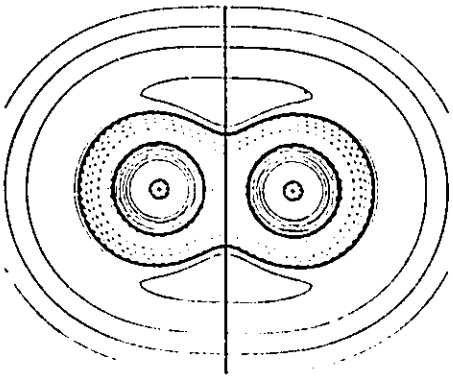
Relief maps of the charge density and the negative of its Laplacian for a plane containing the nucleus of an argon atom. Function values above an arbitrary magnitude are not shown. At the point A, the curvature of ρ along z (the radial curvature) is positive while the curvature along the y axis (a tangential curvature) is negative. In addition to the spike-like maximum in $-\nabla^2\rho$ at the nucleus, there are two shells of charge concentration and three shells of charge depletion, corresponding to the three quantum shells in this atom.



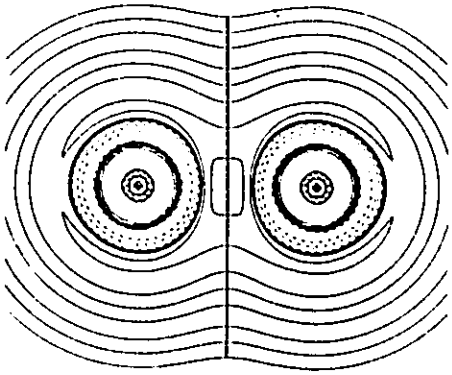
H_2



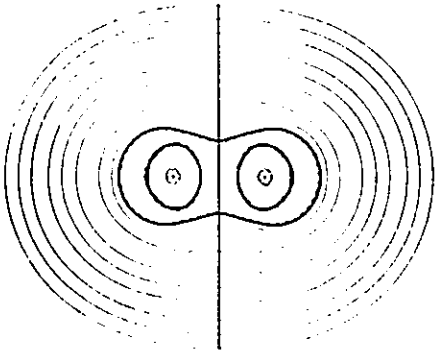
He_2



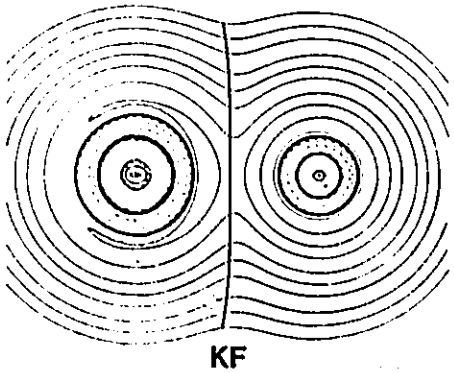
B_2



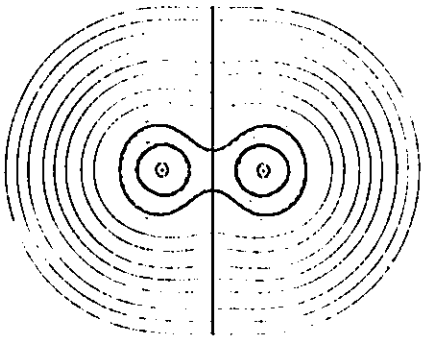
Ar_2



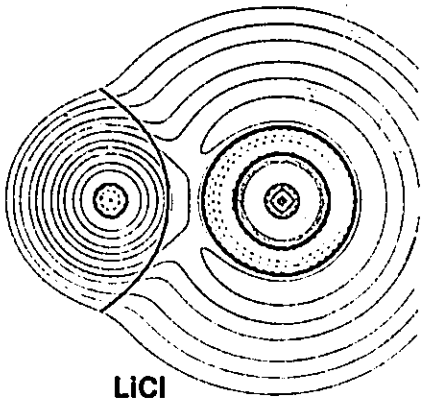
N_2



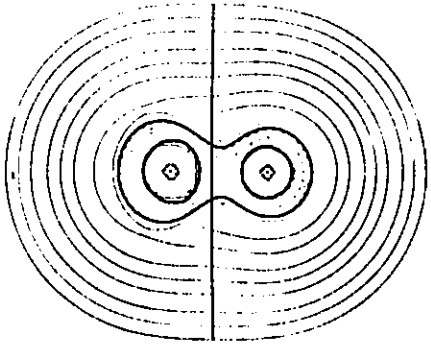
KF



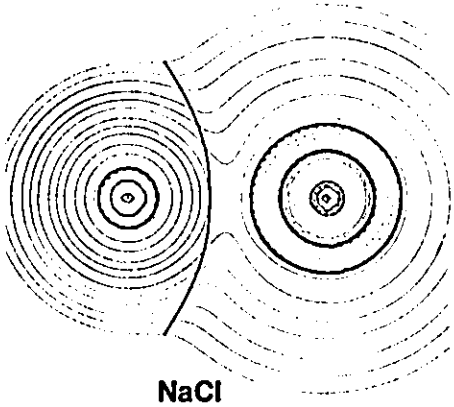
O_2



LiCl



NO



NaCl

Figure 1.7

Contour maps of $\nabla^2\rho$ for molecules with shared (in the column at the left hand side) and closed-shell (in the column at the right hand side) interactions.

is concentrated on the nonbonded than on the bonded side of each argon atom. Charge is depleted from the internuclear region to a greater extent is concentrated on the nonbonded than on the bonded side of each argon than it is from the nonbonded region of each nucleus. Thus the net forces on the nuclei are repulsive in agreement with the unbound nature of Ar_2. A subset of closed-shell interaction which are bound are ionic interactions. For an ionic diatomic molecule to achieve electrostatic equilibrium the charge distributions of both the anion and cation must be polarized in a direction counter to the direction of electron transfer. The cationic nucleus is attracted by the net negative field of the anion and the charge distribution of the cation must polarize away from the anion to balance this attractive force. The anion nucleus is repelled by the net positive field of the cation and the charge distribution of the anion must polarize towards the cation to balance this repulsive force. These features are displayed in the Laplacian distribution of KF, LiCl and NaCl (Fig.1.7).

The type of interaction between two atoms can also be determined from $\nabla^2\rho$ at the corresponding bond critical point. The reasoning is as follows: the charge density is a maximum in the plane perpendicular to the bond path at the critical point, and charge is locally concentrated there since the two curvatures of ρ (λ_1 and λ_2) are negative; the charge density is a local minimum at the critical point along the bond path since the third curvature (λ_3) is positive and charge is locally depleted at r_c with respect to neighboring points along the bond path. Thus the formation of an interatomic surface and a chemical bond is the result of a competition between the perpendicular contractions of ρ, which lead to a concentration or compression of charge between the nuclei towards and along the bond

path, and the parallel expansion of ρ which leads to separate concentration of charge in the basins of the neighboring atoms. When $\nabla^2\rho < 0$ the perpendicular contractions in ρ dominate the interaction and the result is a sharing of electronic charge between the nuclei as is typical of covalent or polar bonds. When $\nabla^2\rho > 0$ the contraction of each atomic density towards its nucleus dominates the interaction, resulting in a depletion of charge at r_c and in the interatomic surface. The latter happens in the interaction between closed shell atoms as found in noble gas repulsive states and in ionic and hydrogen-bonded molecules.

1.3.2 The Valence Shell of Charge Concentration (VSCC)

The values of the Laplacian provide a quantitative measure of the local charge concentrations and depletions as well as their number and positions. Extrema in the Laplacian of ρ are classified by rank and signature in the same way as are critical points in the charge density. The critical points in $\nabla^2\rho$ occur where $\nabla(\nabla^2\rho) = 0$ and their types denoted by the principal curvatures of $\nabla^2\rho$ at the critical point. The topological discussions hereafter will always refer to the negative of the Laplacian, the quantity $-\nabla^2\rho$. Since charge is concentrated where $\nabla^2\rho < 0$, a local maximum in $-\nabla^2\rho$ implies that electronic charge has the maximal trend of concentration at the place where the maximum value of $-\nabla^2\rho$ applies. Thus a local maximum in the $-\nabla^2\rho$ field, a (3,-3) critical point, denotes the presence of a region of local charge concentration. A minimum in the $-\nabla^2\rho$ field, a (3,+3) critical point, denotes the presence of a region of local charge depletion.

The outer quantum shell of an atom is divided into an inner region

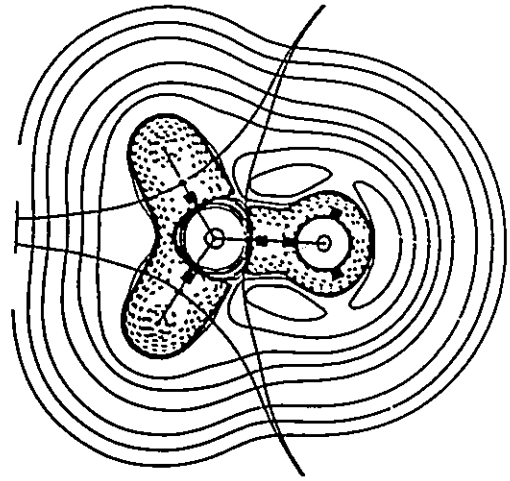
over which $\nabla^2\rho < 0$ and an outer one over which $\nabla^2\rho > 0$. The portion of the shell over which $\nabla^2\rho < 0$, is called the Valence Shell Charge Concentration or VSCC, while the portion, $\nabla^2\rho > 0$, is called the Valence Shell Charge Depletion or VSCD.

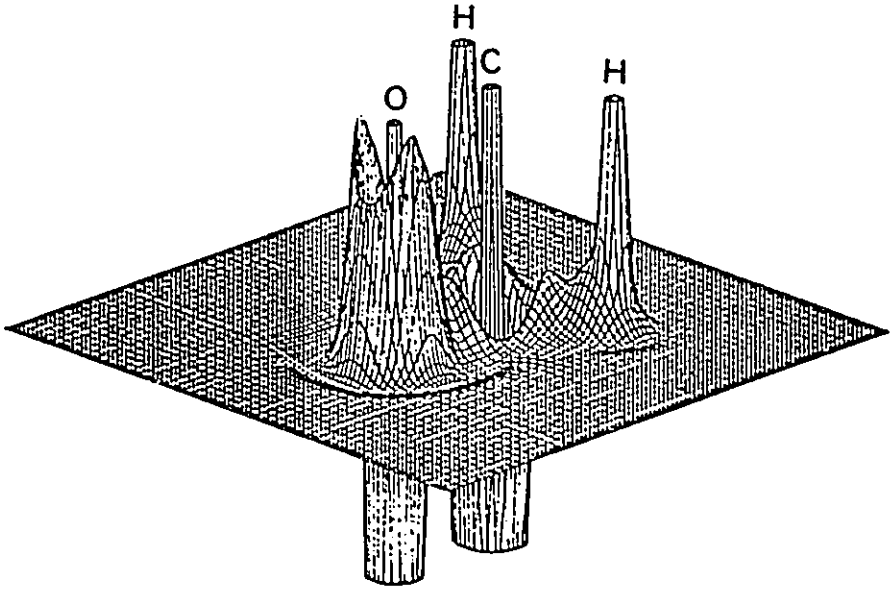
For an isolated atom, each point on the surface of the sphere of maximum concentration in the VSCC always has a negative curvature in the direction perpendicular to the surface since $-\nabla^2\rho$ is a maximum on the surface of the sphere and this surface persists in an atom in a molecule, while the other two curvatures tangent to the surface are zero in an isolated atom and assume either positive or negative values when the atom is in a chemical combination, causing the sphere to be distorted. When the two tangential curvatures are both negative, a local maximum is formed on the surface, a (3,-3) critical point in $-\nabla^2\rho$. When they both assume positive values, a local minimum is formed on the surface and the result is a (3,+1) critical point in $-\nabla^2\rho$. If they assume values of opposite sign, there is a saddle point in the surface and the result is a (3,-1) critical point in $-\nabla^2\rho$. The local maxima that are created within the valence shell of charge concentration provide a one-to-one mapping of the electron pairs of the Lewis model, both bonded and nonbonded.

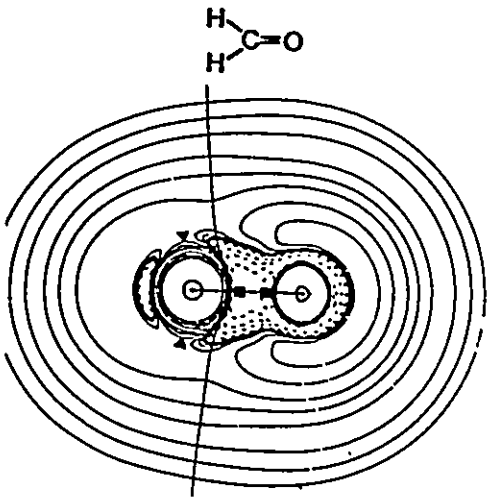
Fig.1.8 displays the contour maps and their corresponding relief maps of the Laplacian of charge density for two planes of the formaldehyde molecule. In the plane containing the nuclei, three bonded maxima in carbon and a bonded maximum and two nonbonded maxima in oxygen are clearly seen. Fig.1.8 also shows holes on the VSCC of carbon atom (at positions of above and bellow molecular plane), corresponding to ring critical points (3,+1) in $\nabla^2\rho$. These (3,+1) critical points define the minima in the

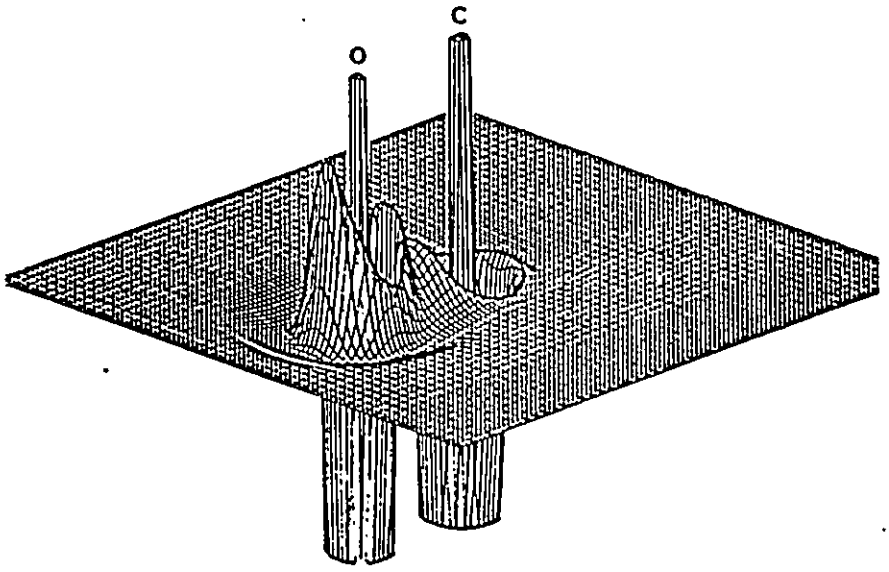
Figure 1.8

Contour and relief maps of $\nabla^2\rho$ for the molecular plane of formaldehyde (top) and a plane perpendicular to this and along the C-O axis (bottom). The dashed (solid) lines denote regions of charge concentration (depletion). The Laplacian is also negative within the region bounded by the innermost solid contour enclosing each C and O nucleus. The top contour diagram shows the positions of the three bonded charge concentrations in the VSCC of carbon and the single bonded and two nonbonded concentrations in the valence shell of oxygen. The zero flux surfaces and bond critical points (black dots) are also shown. The bottom contour diagram shows the two critical points (indicated by solid triangles) which determine the sites of nucleophilic attack at carbon, the (3,+1) critical point is in its VSCC.




H
C
O
H


H
C=O
H


C
O

sphere of charge concentration.

The nonbonded maxima and the local minima of $\nabla^2\rho$ on the VSCC are the most important aspects of the topology of the Laplacian of the charge density, since they are directly related to the chemical reactivity of a molecule.

1.3.3 The Laplacian of Charge Density and Chemical Reactivity

It has been shown in section 1.3.1 that charge is concentrated in regions where the Laplacian is negative and is depleted in regions where the Laplacian is positive. These properties of the Laplacian determine what kind of role each molecule would play in chemical reactions. In other words, the regions of charge concentration are approached by electrophiles and the regions of charge depletion are approached by nucleophiles. Thus when two reactants interact with each other, the local charge depletion of one of the reactants would seek the region of local charge concentration and avoid the region with the larger value of charge depletion of another reactant, and vice versa.

Evidently, the above properties of the Laplacian of charge density reveal the physical basis of Lewis model, i.e. the positions of nonbonded maxima of charge concentration are related to the electron pairs of the Lewis base, and atoms which have local minima of charge concentration in their VSCC are related to the form of the Lewis acid. The positions of the local charge concentration and depletion together with their magnitudes are determined by the positions of the corresponding critical points in the VSCC's of the respective base and acid atoms. This information enables one to predict positions of attack within a molecule and hence the geometries

of approach of the reactants. For example, Figure 1.8 serves as an example of the above features of $\nabla^2\rho$. The two equivalent oxygen nonbonded maxima in the molecular plane would serve as the sites of electrophilic attack, while the position of the minima of the charge concentration at the carbon VSCC in the perpendicular plane would serve as the site of nucleophilic attack.

1.4. Properties of Atom in Molecules

If one inserts an operator F(r), which corresponds to a physical property, between the state functions in Eq.[1.2], one obtains a real space density distribution function for that property. All atomic properties are defined in terms of such density distributions, the atomic average of a property being obtained by the integration of the corresponding density over the region of space, Ω, assigned to the atom, its atomic basin.

1.4.1 Atomic Electron Population

The average electron population of an atom, $N(\Omega)$, is obtained by integrating the electron density over the basin of an atom:

$$N(\Omega) = \int_\Omega dr\rho(r) \qquad [1.15]$$

The net charge $q(\Omega)$ is given by

$$q(\Omega) = Z_\Omega - N(\Omega) \qquad [1.16]$$

where Z_Ω is the nuclear charge.

Where orbitals of pure σ and pure π symmetry can be distinguished, the atomic σ population, $N_\sigma(\Omega)$, and π population, $N_\pi(\Omega)$ can be determined:

$$N_\sigma(\Omega) = \int_\Omega dr\rho_\sigma(r) \qquad [1.17]$$

$$N_\pi(\Omega) = \int_\Omega dr\rho_\pi(r). \qquad [1.18]$$

where

$$\rho_\sigma(r) = \Sigma_{i\sigma}\, \lambda_{i\sigma}\phi_{i\sigma}(r)\phi_{i\sigma}(r) \qquad [1.19]$$

$$\rho_\pi(r) = \Sigma_{i\pi}\, \lambda_{i\pi}\phi_{i\pi}(r)\phi_{i\pi}(r). \qquad [1.20]$$

1.4.2 Atomic Energy

In order to discuss atomic energy, it is necessary to start from the virial theorem for a subsystem. For a stationary state this theorem is

$$V(\Omega) + 2T(\Omega) = -L(\Omega) \qquad [1.21]$$

where $V(\Omega)$ is the average potential energy of the electrons in the atom Ω, $T(\Omega)$ is its average kinetic energy, and $L(\Omega)$ is defined as

$$-L(\Omega) = (1/4)\int_\Omega dr\nabla^2\rho(r) = (1/4)\int dS(\Omega;r)\nabla\rho(r)\cdot n(r) \qquad [1.22]$$

The density associated with $L(\Omega)$ has the property

$$L(r) = K(r) - G(r) = (-1/4)\nabla^2\rho(r) \qquad [1.23]$$

where $K(r)$, termed the "Schrödinger kinetic energy density", is given by

$$K(r) = (-1/4)(\int d\tau'(\psi^*\nabla^2\psi + (\nabla^2\psi^*)\psi)) \qquad [1.24]$$

and $G(r)$, termed the "gradient kinetic energy density", is given by

$$G(r) = (1/2)\int d\tau'\nabla\psi\cdot\nabla\psi \qquad [1.25]$$

where $d\tau' = d\tau_2 d\tau_3 \ldots d\tau_n$ and n is the number of electrons. Since the Lagrangian $L(\Omega)$ vanishes for an atom in a molecule because of the zero flux surface condition (Eq.[1.9]),

$$\int dS(\Omega;r)\nabla\rho(r)\cdot n(r) = 0 \qquad [1.26]$$

Eq.[1.21] reduces to the usual statement of the virial theorem,

$$V(\Omega) + 2T(\Omega) = 0 \qquad [1.27]$$

where $V(\Omega)$ is the average atomic potential energy and $T(\Omega)$ is the average atomic kinetic energy $T(\Omega)$ of the electrons which is defined alternatively as

$$T(\Omega) = \int_\Omega K(r)dr \qquad [1.28]$$

or

$$T(\Omega) = \int_\Omega G(r)dr \qquad [1.29]$$

For Hartree-Fock wavefunctions, $G(\Omega)$ and $K(\Omega)$ may be written as sums over functions of orbitals, that is,

$$G(\Omega) = (1/2)\int_\Omega(\Sigma_i\lambda_i\nabla\phi_i\cdot\nabla\phi_i)dr \qquad [1.30]$$

and

$$K(\Omega) = (-1/2)\int_\Omega(\Sigma_i\lambda_i\phi_i\nabla^2\phi_i)dr. \qquad [1.31]$$

These are the expression calculated to obtain the average electron kinetic energy $T(\Omega)$ of an atom in a molecule using the program PROAIM (Biegler-Konig et al 1982). From Eqs. [1.22] and [1.26], it is seen that $L(\Omega)$ vanishes for a region of space bounded by a surface of zero flux in $\nabla\rho$. Thus the vanishing of their integral is a measure of how well the surface of the atom has been approximated in the numerical procedure. The quatity provides a measure of how accurately the energy of an atom in a molecule has been determined since $E(\Omega) = -K(\Omega) = -G(\Omega)$ only if $L(\Omega)$ vanishes (see Eq.[1.23]). In general, in this thesis an upper limit of 5×10^{-3} au ($\approx$3 kcal/mol) for $L(\Omega)$ was accepted.

The total energy of an atom in a molecule, the atomic energy $E(\Omega)$ is obtained for the average kinetic energy of the atom using the virial relationship, that is

$$E(\Omega) = -T(\Omega) \qquad [1.32]$$

Practically speaking, this equation is not exactly satisfied for a finite basis set calculation (6-31G** for example). The ratio -V/T differs slightly from the correct value of two for an equilibrium geometry. To correct for the error in the virial, each atomic kinetic energy must be multiplied by the factor (V/T + 1) to obtain a set of atomic energies that correctly sum to the total energy of the molecule, although correction is

in general small.

The average atomic potential energy $V(\Omega)$ can broken down as follows:

$$V(\Omega) = V_A(\Omega) + V_R(\Omega) \qquad [1.33]$$

where the quantity, $V_A(\Omega)$, the total attractive potential energy, is the interaction of all the nuclei in the system with the charge density of Ω and $V_R(\Omega)$ is the total repulsive potential energy. Explicitly,

$$V_A(\Omega) = - \int_\Omega (\Sigma_\alpha (Z_\alpha / r_\alpha)) \rho(r) dr. \qquad [1.34]$$

Further,

$$V_R(\Omega) = V_{ee}(\Omega) + V_{NN}(\Omega) \qquad [1.35]$$

where $V_{ee}(\Omega)$ is the repulsive electron-electron contribution to the potential energy and $V_{NN}(\Omega)$ is the repulsive nuclear-nuclear contribution to the potential energy. Since the virial theorem is not exactly satisfied, the potential energies must be corrected also. This is accomplished by multiplying the above potential energy values by the term $2(1+V/T)/(V/T)$.

1.4.3 Atomic First Moment

The first moment (or atomic dipole) of atom Ω, denoted $\mu(\Omega)$, is the atomic average of the electronic position vector $\mathbf{r}_\Omega = \mathbf{r} - \mathbf{X}_\Omega$ where $\mathbf{X}_\Omega$ is the position vector of the nucleus of atom Ω measured from an origin (Bader et al 1987b). Thus,

$$\mu(\Omega) = -\int_\Omega \mathbf{r}_\Omega \rho(r) dr \qquad [1.36]$$

The cartesian components of $\mu(\Omega)$ are calculated as follows:

$$\mu_x(\Omega) = -\int_\Omega x_\Omega \rho(r) dr$$

$$\mu_y(\Omega) = -\int_\Omega y_\Omega \rho(r) dr$$

$$\mu_z(\Omega) = -\int_\Omega z_\Omega \rho(r) dr \qquad [1.37]$$

The atomic first moment measures the distortion of the charge density of Ω from sphericity. It quantitatively describes to what extent the charge density is polarized in a given direction.

The dipole moment of a neutral molecule, μ, may be expressed as a sum over the net charges q(Ω) and first moments μ(Ω) of every atom in the molecule as

$$\mu = \Sigma_\Omega[q(\Omega)X_\Omega + \mu(\Omega)] \qquad [1.38]$$

where X_Ω is the position vector of the nucleus of atom Ω measured from some arbitrary origin. While the individual atomic contributions $q(\Omega)X_\Omega$ depend upon the choice of origin, their sum does not, and thus each molecular dipole moment is uniquely determined by a charge transfer contribution arising from the net charge on the atoms

$$\mu_c = \Sigma_\Omega q(\Omega)X_\Omega \qquad [1.39]$$

and a first moment contribution arising from the polarization of the atomic charge densities

$$\mu_a = \Sigma_\Omega \mu(\Omega). \qquad [1.40]$$

1.4.4 Atomic Quadrupole Moment

The x containing components of the traceless symmetric quadrupole moment tensor Q for atom Ω are given as follows:

$$Q_{xx}(\Omega) = -\int_\Omega (3x^2 - r_\Omega^2)\rho(r)dr$$

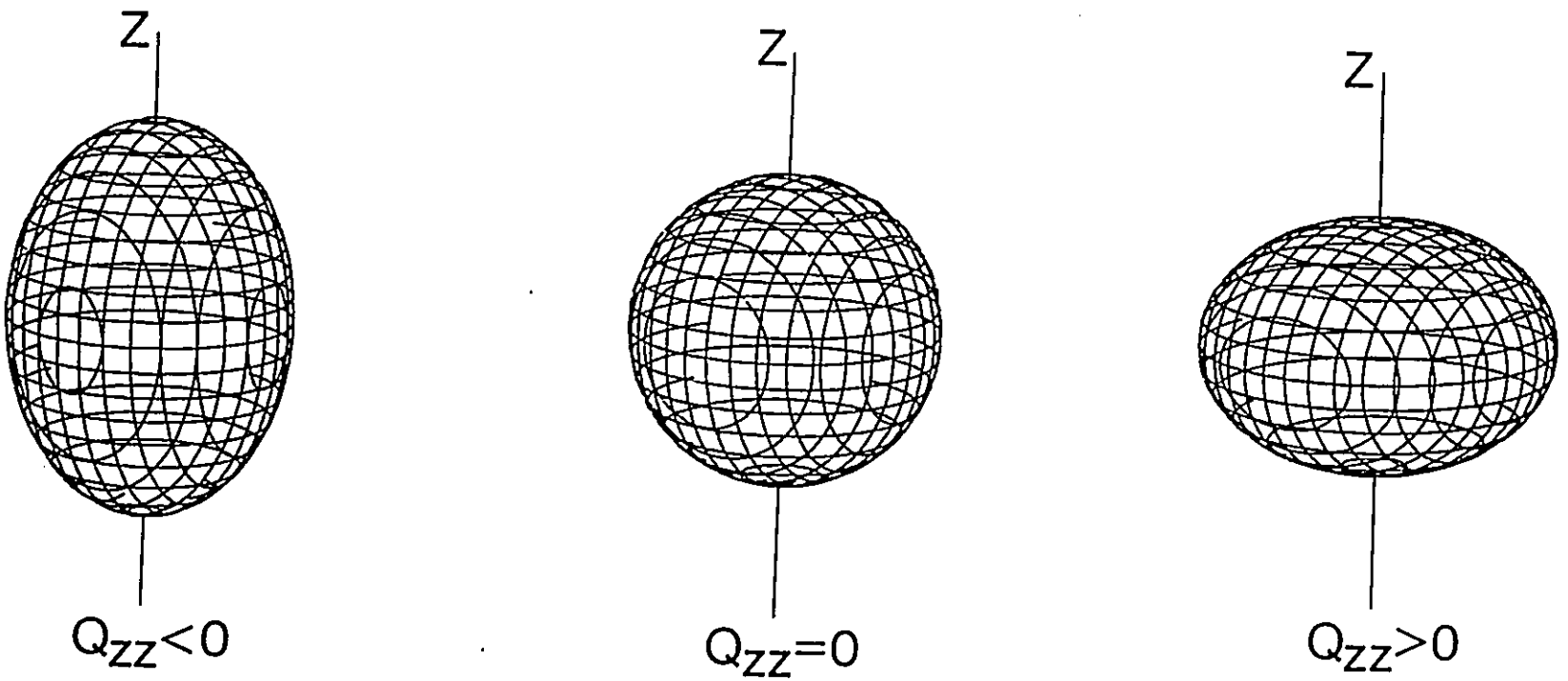
$$Q_{xy}(\Omega) = -\int_\Omega (3xy)\rho(r)dr$$

$$Q_{xz}(\Omega) = -\int_\Omega (3xz)\rho(r)dr \qquad [1.41]$$

Q can always be diagonalized, and so, we examine $Q_{xx}(\Omega)$, $Q_{yy}(\Omega)$ and $Q_{zz}(\Omega)$. These components reflect the preferential accumulation of charge in a given plane. The diagonal elements always sum to zero. For a spherical distribution, $Q_{xx}(\Omega) = Q_{yy}(\Omega) = Q_{zz}(\Omega) = 0$, see Fig. 1.9. If the sphere is

Figure 1.9

Display of quadrupole moment. $Q_{xx}=Q_{yy}$ has been assumed.


Z
Z
Z
$Q_{zz}<0$
$Q_{zz}=0$
$Q_{zz}>0$

flattened at its poles to yield an oblate spheroid, then (with the polar axis along z) $Q_{zz}(\Omega) > 0$ and $Q_{xx}(\Omega) = Q_{yy}(\Omega) = -(1/2)Q_{zz}(\Omega)$ inferring the accumulation of negative charge in the xy plane (see also Fig.1.9).

The atomic quadrupole moment is a property which provides information that both complements and supplements the σ and π populations of the orbital model. It will be detailed in the Chapter 3.

1.4.5 Atomic Volume

An atomic surface is the union of a number of interatomic surfaces, one such surface for each bonded neighbor, and of some portions which may be infinitely distant from the nucleus if the atom is not an interior atom. The latter open portions of the atomic surface are replaced by a particular density envelope in the calculation of an atomic volume. Thus, an atomic volume is defined to be a measure of the region of real space enclosed by the intersection of the atomic surface of zero flux and a particular envelope of the charge density:

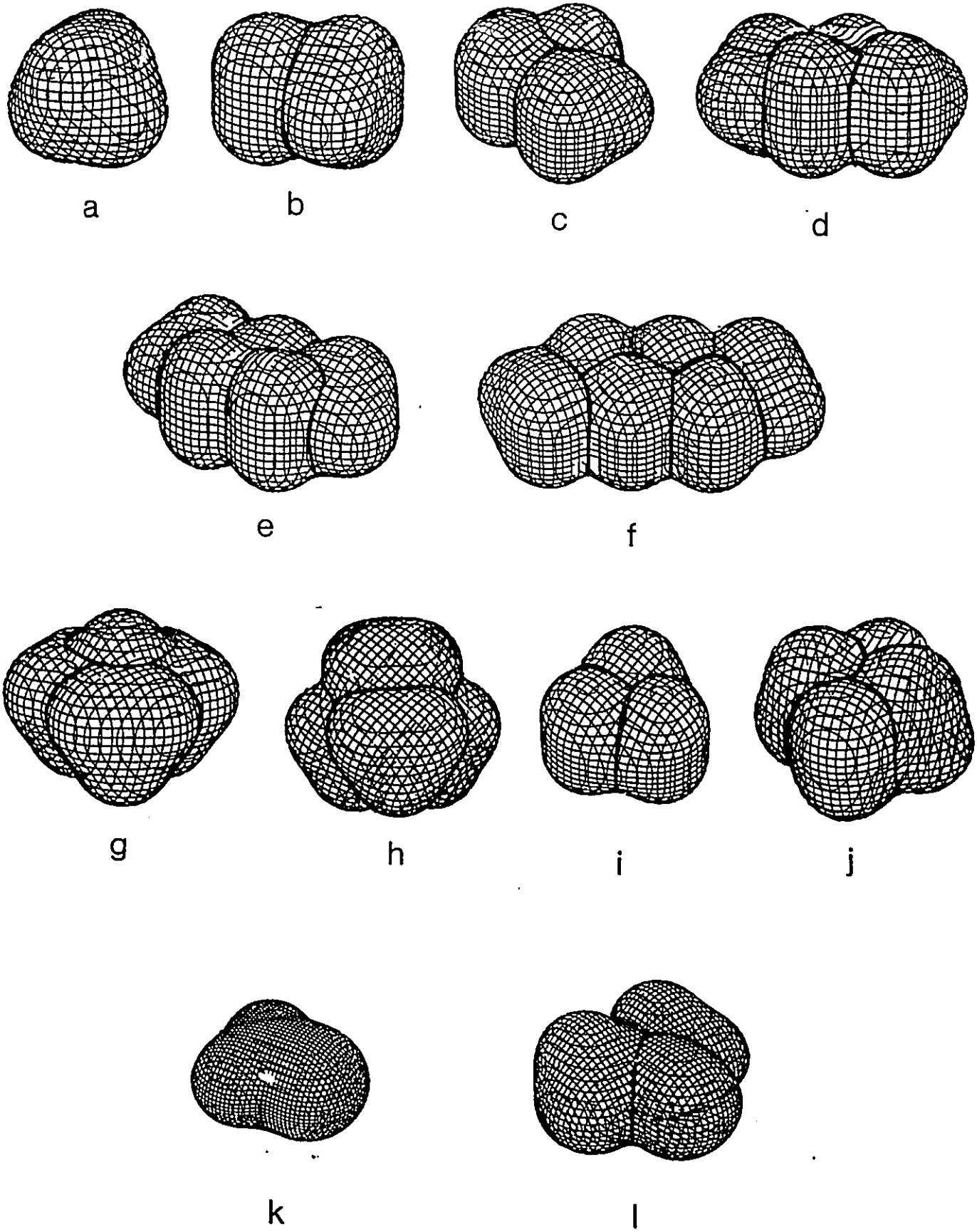
$$v(\Omega) = \int_{\Omega} dr \qquad [1.42]$$

The 0.001 au charge density envelope is chosen because it has previously been shown that the 0.001 au charge density contour for methane and the inert gases gives good agreement with the van der Waals diameters of these molecules as determined by second virial coefficient or viscosity data fitted with a Lennard-Jones 6-12 potential (Bader and Preston 1970). Further, the 0.001 au density envelope contains at least 96% of the electronic charge (Bader et al 1987a).

The shapes of some hydrocarbon molecules, the 0.001 au charge density envelopes, are displayed in Figure 1.10. The lengths, widths and atomic volumes have been calculated (Bader et al 1987a). What is to be

Figure 1.10

Shapes of molecules represented by envelopes of constant electronic charge density. The envelope shown has a value of 0.001 au. The molecules are: (a) to (f), the normal alkanes from methane to hexane, (g) isobutane, (h) neopentane, (i) cyclopropane, (j) cyclobutane, (k) formaldehyde, $H_2C{=}O$, (l) acetone, $(CH_3)_2C{=}O$. The intersections of the "zero flux" interatomic surfaces with the envelope are shown in some cases. They define the methyl, methylene, hydrogen and carbonyl groups. The isobutane molecule (g) for example, exhibits three methyl groups topped by a hydrogen atom.


a
b
c
d
e
f
g
h
i
j
k
l

stressed is that the volumes of methyl and methylene groups in normal hydrocarbons are transferable, as are their charge distributions, populations, dipoles and energies (Bader et al 1987a,b).

All of the average properties of an atom, isolated or in a molecule, are defined by quantum mechanics. The most important consequence of the definition of those properties is that the average value of an observable for the total system <M>, is given by the sum of its atomic contribution M(Ω)

$$\langle M \rangle = \Sigma_{\Omega} M(\Omega) \qquad [1.43]$$

Eq.[1.43] states that each atom makes an additive contribution to the value of every property for a total system. For example, the molecular volume v_M is obtained by summing the volumes of the constituent atoms v(Ω). Similarly, the volume of a methyl group is obtained by summing the volumes of the one carbon and three hydrogens comprising this group. Such a group is bounded by a surface of zero flux in the gradient vector of the charge density and is a quantum subsystem with well defined properties. Eq.[1.43] displays the principle underlying the cornerstone of chemistry - that atoms and functional groupings of atoms make recognizable contributions to the total properties of a system. In practice, we recognize a group and predict its effect upon the static and reactive properties of a system in terms of a set of properties assigned to the group. In the limiting case of a group being essentially the same in two different systems, one obtains a so-called additivity scheme for the total properties, for in this case the atomic contributions as well as being additive in the sense of Eq.[1.43], are transferable between molecules.

Chapter 2

The Theoretical Synthesis of Polypeptides from Amino Acid Fragments

Chapter 1 has provided an introduction to the theory of atoms in molecules. It detailed the definition of an atom in a molecule as well as its properties, the physical concept of the chemical bond, molecular structure, and the chemical reactivity of molecules. The present chapter describes the application of this theory to the problem of building large polypeptides from amino acid fragments.

Many of the problems associated with the prediction and understanding of biochemical processes and of catalysis in general can be grouped under the heading of molecular recognition. Thus, one is concerned with the functional specificity, orientation, temporal evolution and function of the interaction between an active site and its substrate. The active site is a portion of a biological macromolecule or is made up of contributions from a number of portions which are brought into juxtaposition by a particular conformation of the macromolecule. It has recently been demonstrated (Rebek 1987) that processes exhibiting the characteristics of molecular recognition can be displayed by smaller molecules that are synthetically accessible. Even these smaller model systems, however, present the same problem for one wishing to predict their properties from theory in that they are too large to be adequately described by present day computational techniques. This problem is certainly not unique to systems of biological interest and this point brings us to the principal goal of this Chapter: the use of theory to synthesize large molecules from smaller fragments, and to do this in a new way by using groups defined in real space. This approach also offers the

opportunity to construct only that portion of a large molecule that is of interest. Experimental studies have shown that complementarity in size, shape and functionality between the active site and the substrate all play a role in molecular recognition (Rebek 1987). The time . come to solve the problem of how this is to be done and its solution is made ssible by the theory of atoms in molecules, which provides the necessary information for the prediction and understanding of these properties and for determining their role in the formation of the substrate-active site complex.

Section 2.1 shows how the amino acid building blocks are created for the theoretical synthesis of polypeptides. The optimised geometries for all the molecules in this study are also given in section 2.1.

Theory predicts what can be observed and it should be possible to construct a molecule using theory in the same manner that this is done in the laboratory. Chemists do not begin a synthesis starting with separate beakers of nuclei and electrons, the step corresponding to the theoretical description of a molecule starting from the Hamiltonian, but rather from pieces of other molecules. The first step in a parallel theoretical synthesis is the identification of the pieces, the atoms and functional groups of chemistry. Theory defines an atom in a molecule as a region of real space bounded by a surface of zero flux in the gradient vector field of the charge density, as discussed in Chapter 1. The atoms of theory recover the experimentally measurable properties of atoms in molecules (Bader 1986, Wiberg et al 1987, Bader et al 1987a). They are the atoms of chemistry because a) they are the most transferable pieces of a system that one can define in real space and which simultaneously exhaust the space of

a system, b) their properties, as well as being defined by quantum mechanics, are directly determined by their distribution of charge, that is, if an atom has the same form in the real space of two systems, then it contributes the same amounts to the total properties in each system and c) the atomic contributions $M(\Omega)$ to a property M are additive to yield its molecular value $\langle M \rangle$, see Eq.[1.43].

This theory enables one to take advantage of the most important postulate of the atomic hypothesis: that atoms and functional groupings of atoms exhibit characteristic and measurable sets of properties that vary between relatively narrow limits. In certain instances, the methyl and methylene groups of hydrocarbons for example, the groups both experimentally and theoretically, are transferable between systems with essentially no change (Bader 1986, Wiberg et al 1987, Bader et al 1987a) and one can construct molecules from these groups with errors on the same order as the experimental ones. It is based on this property that amino acid fragments can also be used as building blocks in theory. Data are provided in Section 2.2 in order to demonstrate the possibility of the transferability of molecular fragments as defined by theory. It shows that it is possible to construct a molecule from the atoms of theory and predict its properties with acceptably small errors even in cases where one does not approach the limit of perfect transferability.

Section 2.3 demonstrates that the synthesis of a polypeptide can be achieved by linking groups or building blocks which are calculated through the use of relatively large basis sets. Theory determines the properties of the individual atoms, then the properties of the total system are determined by summing the corresponding atomic properties. All of the

properties of the system synthesized from atomic groupings can be predicted from the calculation of small molecules.

2.1 Conformation and Structure

In order for amino acids to be used as building blocks, it is necessary that each fragment be transferable among different chemical environments; in other words, that the geometry, charge distribution, reaction properties, atomic and bond properties, exhibit no significant differences in different polypeptide chains. Thus the first step of the theoretical synthesis is to study the transferability of amino acid fragments, followed by the theoretical synthesis using the standard fragments.

The optimization of the geometries of all molecules in this chapter were carried out with Hartree-Fock SCF method at the 4-31G level (Ditchfield et al 1971), while the 4-31G** basis was employed in the final determination of Ψ and the charge density. It has to be pointed out here that as this is a big project, five different computers (VAX-11/780, VAX-8650, IBM-4381, CRAY X-MP/24 and Multiflow-Trace) and several Gaussian series programs (Gaussian 82, 86 and 88) were involved in the SCF calculations, whichever was available at the time. Two other computers (FPS, Stardent 350(FPS/Stellar)) as well as Multiflow Trace 14/350 computer were employed in the topological analysis and atomic integrations. The neutral forms of the amino acids rather than the Zwitterions found in solution are considered in this exploratory investigation, the purpose of which is to test the transferability of fragments between systems. In any event, the repeating sub-unit of a polypeptide is removed from the change

which is present only at the two ends of the chain of fragments. It has been shown (Schafer et al 1990) that the geometries of amino acids, whether neutral or ionized, and their polypeptide derivatives determined by ab initio methods for gas phase molecules as is done here yield geometries which are either more complete, or more accurate, or more highly resolved than the results of any other method, either computational or experimental.

We start with two amino acids, glycine -- the smallest symmetrical one (C_s symmetry) and alanine -- the smallest unsymmetrical one (C_1 symmetry). The structures of these two amino acids are shown in Fig.2.1, together with the numbering assignments for each atom in the fragment. These numbering assignments are used throughout the study of all the molecules. The comformation studies of glycine and alanine molecules, especially the former, have been done by many authors with various methods (Sham et al 1984, Luke et al 1984, Palla et al 1980, Shipman and Christoffersen 1973 (and references therein)). The conformations shown in Fig.2.1 are the ones with lowest energy for glycine and alanine, respectively. The dihedral angle ψ (where ψ is the dihedral angle of O4C2C1N3) is 0° for glycine and -9.4° for alanine. The angle ω (where ω is the dihedral angle of HO5C2C1) is 180° for glycine and about 170° for alanine. Full geometry optimization had been carried out for both glycine and alanine molecules. As the methyl group on alanine can have two different orientations, with a methyl hydrogens cis or trans to the N3 atom, a related conformation study had also been done. The energy of the trans conformation is the same as the cis one, to within 0.006 kcal/mol, which implies that the methyl group can freely rotate in the alanine molecule. Optimised geometries of glycine and alanine are listed in Table

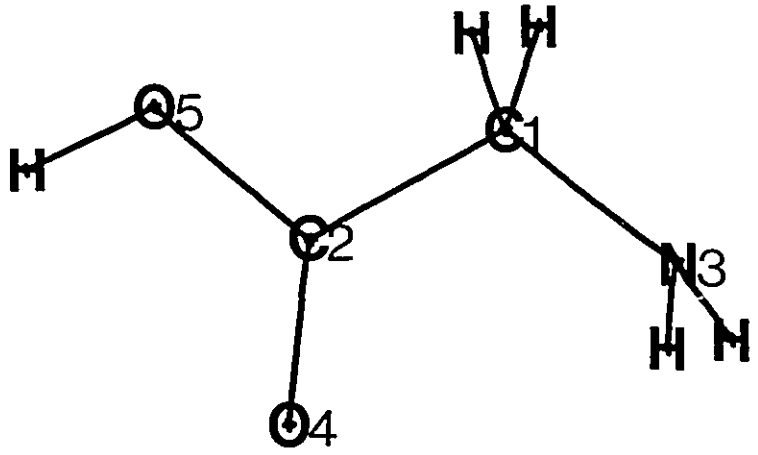
2.1. The structure with the global energy minimum for glycine possesses as C_s symmetry, while that for alanine exhibits no symmetry because of the existence of methyl group (see both Fig.2.1 and Table 2.1). In alanine, the group C1C2O4O5H lies in a plane with the hydrogen atom slightly out of the plane by less than 0.5°, while the plane of C1C2N3 coincides with the plane bisecting the HN3H angle. The dihedral angle about the C1-C2 bond is ϕ=9.4°. The structure of alanine shown in Fig.2.1 is the one with one of three methyl hydrogen trans to N3 atom, yielding a dihedral angle, H_aCC1N3, of -177.5°.

It is noticed that each amino acid can have three different forms when used as a building block, two of them being terminal fragments and one being an internal fragment in a peptide chain. For example, glycine has the form of either an amino-terminal fragment, NH_2CH_2CO-NHR, or a carboxyl-terminal fragment, RCO-$NHCH_2COOH$, when at the end of a polypeptide, and the form RCO-$NHCH_2CO$-NHR′ when presents an internal fragment, where R and R′ are polypeptide chains of arbitrary length (see Fig.2.2). For the sake of convenience, the glycine fragment NH_2CH_2CO-NHR will hereafter be abbreviated as G|NHR, RCO-$NHCH_2COOH$ as RCO|G′, and RCO-$NHCH_2CO$-NHR′ as RCO|G"|NHR′, where the glycyl fragment G ≡ NH_2CH_2CO-, the glycine fragment G′ ≡ -$NHCH_2COOH$, and the glycyl unit G" ≡ -$NHCH_2CO$-, respectively. The vertical bars denote the interatomic surfaces of zero flux in the gradient vector field of the charge density which define the relevant fragments. Thus G, G′ and G" are the glycine fragments to be used in the synthesis of polypeptides.

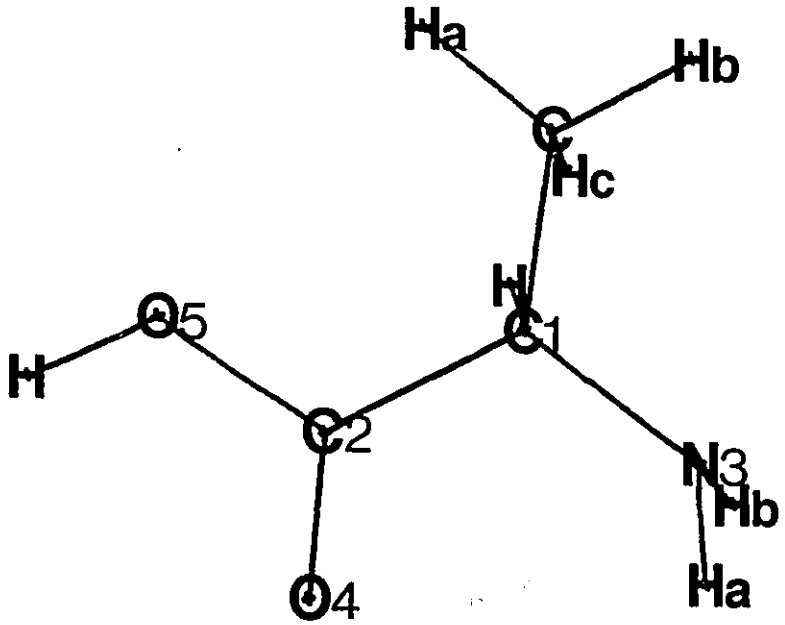
Three glycine fragments and the two terminal alanine fragments have been studied here. Our purpose is to create standard fragments for use as

Figure 2.1.

Molecular graphs of A) glycine and B) alanine molecules.



A



B

Table 2.1. Optimized geometries[a] of glycine and alanine for 4-31G basis set.

Glycine

Bond Length		Bond Angle	
C1-C2	1.505	N3C1C2	114.7
C1-N3	1.434	O4C2C1	126.3
C2-O4	1.206	O5C2C1	111.4
C2-O5	1.349	HN3C1	115.9
H-C1	1.082	HC1C2	107.6
H-N3	0.994	HO5C2	114.1
H-O5	0.955	HN3H	110.6
		HC1H	106.8

SCF Energy -282.40095 au for 4-31G basis set
-282.57347 au for 4-31G**//4-31G calculation

Alanine

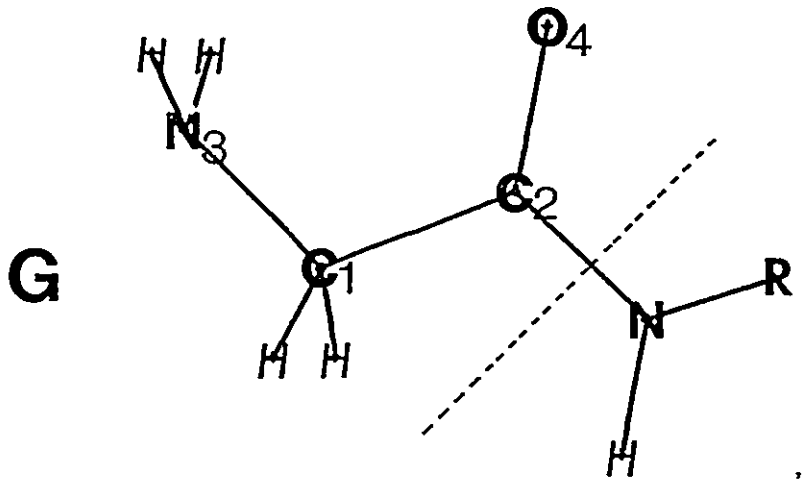
Bond Length		Bond Angle		Dihedral Angle	
C1-C2	1.512	N3C1C2	112.7	O4C2C1N3	-9.4
C1-N3	1.440	O4C2C1	126.2	O5C2C1N3	170.7
C2-O4	1.207	O5C2C1	112.0	HC1C2O4	-128.7
C2-O5	1.350	H_aN3C1	115.1	$H_aN3C1C2$	66.9
H-C1	1.082	H_bN3C1	115.9	$H_bN3C1C2$	-66.7
H_a-N3	0.996	HC1C2	106.6	HO5C2O4	-0.4
H_b-N3	0.995	HO5C2	114.0	CC1C2O4	114.0
H-O5	0.955	HN3H	112.1	H_aCC1N3	-177.5
C-C1	1.533	CC1C2	109.8	H_bCC1N3	61.8
H_a-C	1.079	H_aCC1	111.2	H_cCC1N3	-57.0
H_b-C	1.084	H_bCC1	110.5		
H_c-C	1.080	H_cCC1	108.7		

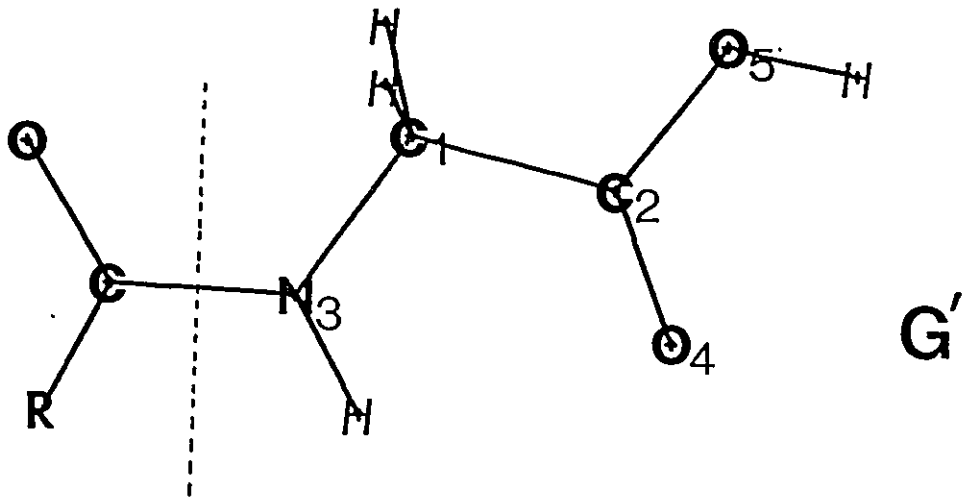
SCF Energy -321.39019 au for 4-31G basis set
-321.57552 au for 4-31G**//4-31G calculation

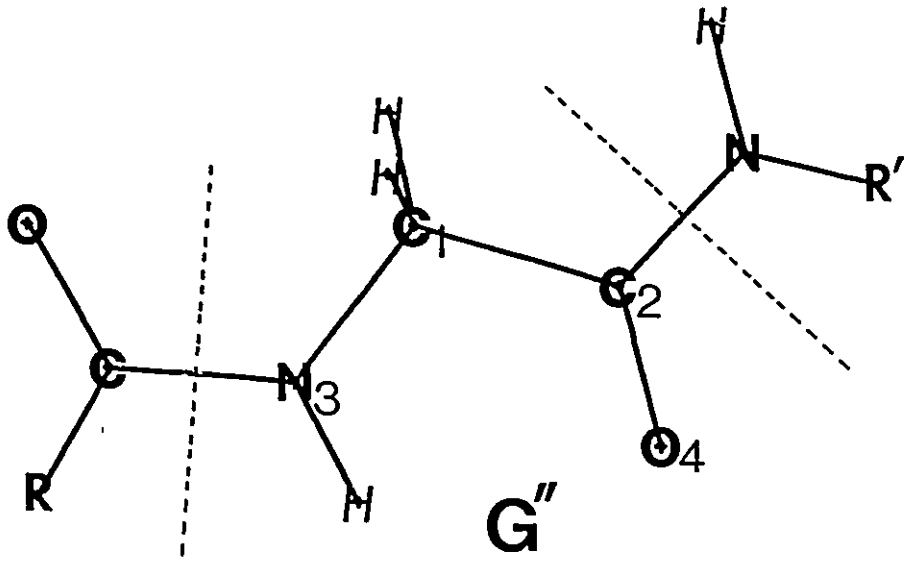
a. Bond lengths are in Å and bond angles in degrees.

Figure 2.2.

Display of three glycine fragments in polypeptides. The dashed lines are schematic representation of the interatomic surfaces which separate the fragments from polypeptide chains. The numbering assignments for the atoms in the studied glycine and alanine fragments are constant throughout this chapter.


G
H H
N3
O4
C2
C1
H H
N
R
H


H
H
O5
H
C1
C2
O
C
N3
O4
R
H
G′

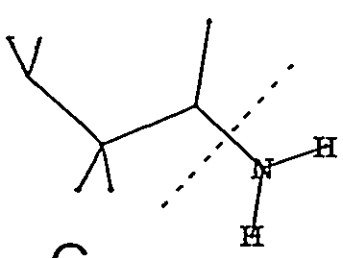

H
H
H
N
R′
C1
C2
O
C
N3
O4
R
H
G″

building blocks by using fragment of amino acid determined in simple molecules with relatively high level basis sets. We build the model molecules which contain the desired fragment by replacing R or R′ in G|NHR, RCO|G′, RCO|G"|NHR′, A|NHR and RCO|A′ with H, where A and A′ are two terminal alanines with the alanyl fragment A ≡ NH_2CHCH_3CO- and the alanine fragment A′ ≡ -$NHCHCH_3COOH$. The study of the transferability of these amino acid fragments is made by comparing the properties of each fragment in the model molecule with its common part in the di- or tri-peptides, i.e., in different chemical environments. Thus comparisons are made of various properties of the glycine fragment G in G|NH_2 (glycylamide), G|G′ (glycylglycine), G|A′ (glycylalanine) and G|G"|G′ (glycylglycylglycine), G′ in COH|G′ (formylglycine), G|G′, A|G′ (alanylglycine) and G|G"|G′, G" in COH|G"|NH_2 (formylglycylamide) and G|G"|G′, A in A|NH_2 (alanylamide) and A|G′, and A′ in COH|A′ and G|A′ (see Fig.2.3).

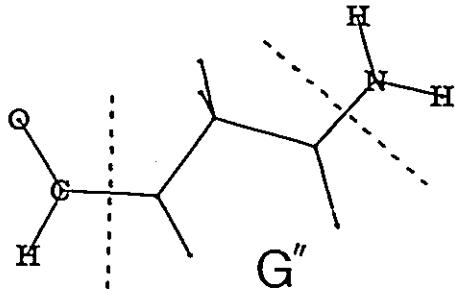
The previous conformation studies (Bonaccorsi et al 1984, Siam et al 1984, Wright and Borkman 1982, Peters and Peters 1980, Palla et al 1980) which employed either standard bond lengths and bond angles or lower level basis sets, were used as references for initial geometries. To obtain detailed geometrical information which has not been studied before, different conformations were calculated and the optimised energies were compared. The two possible geometrical structures with pyramidal and planar NH_2 substituent in G|NH_2 have been checked using GAUSSIAN 82 (Binkley et al 1982) on the VAX-11/780 computer. The latter (-262.61091 au) (see Fig.2.4a) has an energy 18.05 kcal/mol lower than the former (-262.58214 au) at the 4-31G level. This shows that a sp^2 nitrogen atom is preferred in a peptide bond, a result analogous to the equilibrium geometry

Figure 2.3.

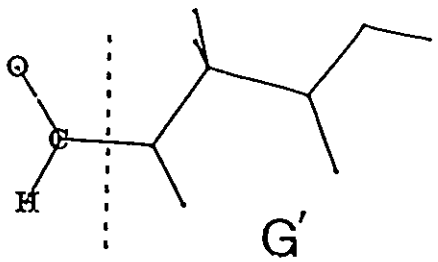
Display of the common fragments of glycine and alanine in different molecules, which are, from left to right and from top to bottom, glycinamide, formylglycinamide, formylglycine, glycylglycine, glycylglycylglycine, glycylalanine, alanylglycine, formylalanine, and alaninamide. The dashed lines represent the interatomic surfaces which define each of the fragments.



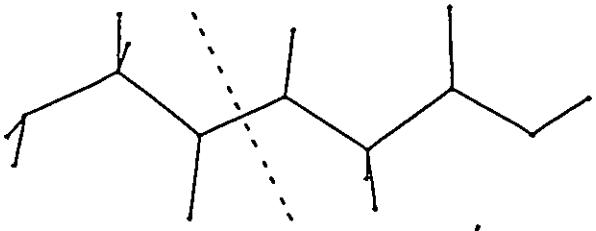
G



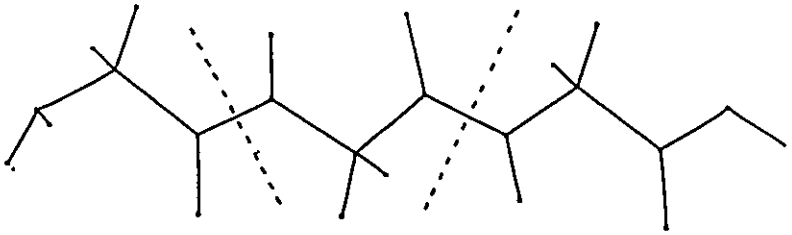
G″



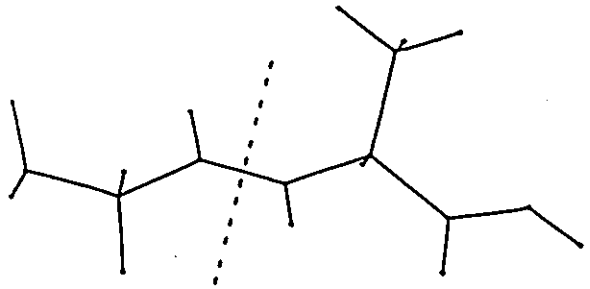
G′



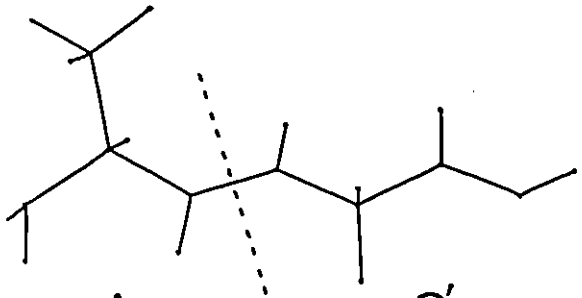
G G′



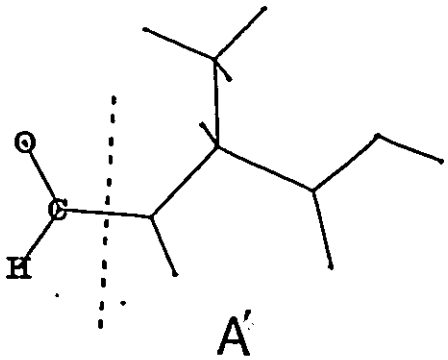
G G″ G′



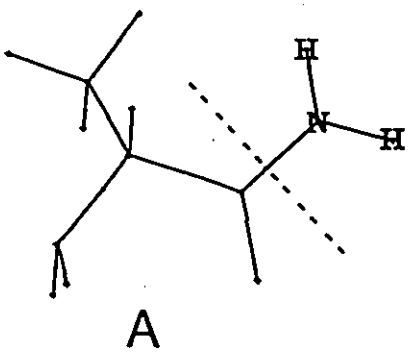
G A′



A G′



A′



A

of formamide (Wiberg and Laidig 1987). Wiberg and Laidig (1987) carried out the calculations with both the 6-31G* and 6-31G** basis set with complete geometry optimization for the formamide molecule and found that the planar structure (sp^2 N) is the one with the minimum energy while two other conformers with a sp^3 nitrogen atom are saddle points. This planar structure for the fragment -CO-NH- were considered through all the other four model molecules. The conformation study of Gly-Gly, Gly-Ala, Ala-Gly and Gly-Gly-Gly by Wright and Borkman (1982) indicated that both Gly-Gly dipeptide and Gly-Gly-Gly tripeptide have a plane of symmetry in the conformation with the lowest energy ($\phi=U=w=180^o$, where ϕ is the dihedral angle of CN3C1C2, U is of N3C1C2N and w is of C1C2NC). The conformation with lowest energy for Gla-Ala and Ala-Gly also has the dihedral angles $\phi=U=w=180^o$. Two conformations of the model molecules containing methyl group with one of its three hydrogens trans or cis to N3 atom have also been calculated. The full optimization results obtained by using GAUSSIAN 86 (Frisch et al 1986) on both VAX-8650 and IBM-4381 show the trans conformation has 0.018 kcal/mol lower energy for $A|NH_2$ than the cis and no energy difference for the two isomers of $COH|A'$. These results further support the above conclusion that the methyl group can freely rotate in an alanine fragment. Thus the conformation with one of three methyl hydrogens trans to N3 atom had been chosen for all the molecules containing methyl group, because it is slightly lower in energy.

All the geometrical parameters for these molecules are allowed to optimize except that the framework of the glycine fragment is forced to stay in a plane, a restriction justified by the results of previous conformational studies by many authors (Bonaccorsi et al 1984, Siam et al

1984, Wright and Borkman 1982, Peters and Peters 1980, Palla et al 1980), in 5 (Fig.2.4a-e) out of 7 molecules which contain the glycine fragment. The assumption that the glycine fragment keeps its planar symmetry in different polypeptides has been supported by the calculations of A|G′ and of G|A′ (see Fig.2.4g and 2.4i), in which the skeleton of the heavy atoms in the glycine and glycyl fragment is not restricted to a planar symmetry. Both calculations result in planar frameworks with deviations from the plane of less than 1^{o}. In practice, calculations are started from small molecules, then the results of these calculations are used as initial guesses of geometries of larger molecules. This technique saves large amounts of CPU time.

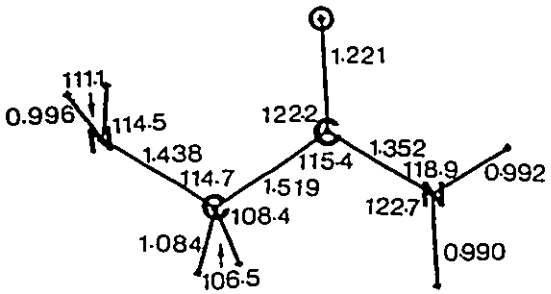
Except for those cases mentioned above, all the optimization calculations on the remaining molecules were carried out using GAUSSIAN 86 for COH|G′, COH|G"|NH_2, A|G′ and G|A′ on VAX-8650, G|G′ on IBM-4381 and G|G"|G′ on CRAY X-MP/24. For the final determination of Ψ, the 4-31G** basis set was employed, and GAUSSIAN 86 has been used for G|NH_2, COH|G′ and COH|G"|NH_2 on VAX-8650 and for A|NH_2, COH|A′ and G|G′ on CRAY, while the direct SCF calculation (Almlof et al 1982, Head-Gordon and Pople 1988) of GAUSSIAN 88 (Frish et al 1988) had to be employed on the Multiflow-Trace for G|A′, A|G′ and G|G"|G′ because of the large size of the integral files.

All the optimised structures, which are depicted by molecular graphs from 4-31G** charge densities, are shown in Fig.2.4 together with the optimised geometrical parameters. From the dihedral angles of molecules containing the alanine fragment, we find that the frames of the fragment -CCONH_2 in A|NH_2, COHNHC- in COH|A′, -CCOOH in COH|A′ and G|A′, NH_2CH_2CONH- in G|A′, and -CCONHCH_2COOH in A|G′ are all close to planar with

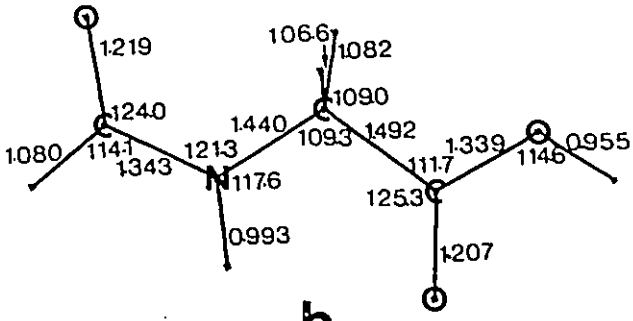
the largest deviation from planarity being less than one degree. The H

Figure 2.4.

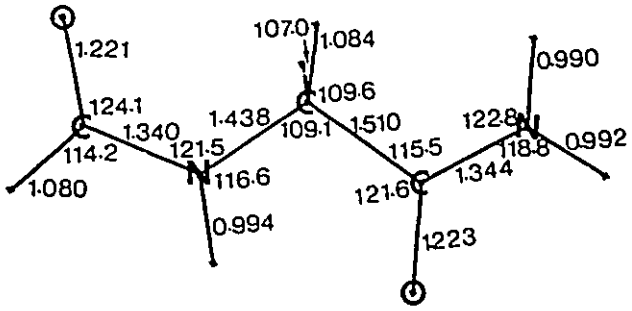
Geometrical parameters of optimised molecules which are (a) glycinamide, (b) formylglycine, (c) formylglycinamide, (d) glycylglycine, (e) glycylglycylglycine, (f) alaninamide, (g) formylalanine, (h) alanylglycine, and (i) glycylalanine. Dihedral angles for the molecules containing alanine fragment are listed under each corresponding molecule. Characters A and G at the top of the list of dihedral angles of dipeptides indicate the fragments to which the numbering assignments of atoms refer. For example, the dihedral angle for O4C2NC under G| is the same as that for OCN3C1 under |A' in glycylalanine, because the numbers of atoms in the former are assigned by referring to the glycine fragment while those in the latter are assigned by referring to the alanine fragment.


1.221
111.1
0.996
114.5
122.2
1.438
115.4
1.352
114.7
118.9
0.992
1.519
122.7
108.4
1.084
0.990
106.5

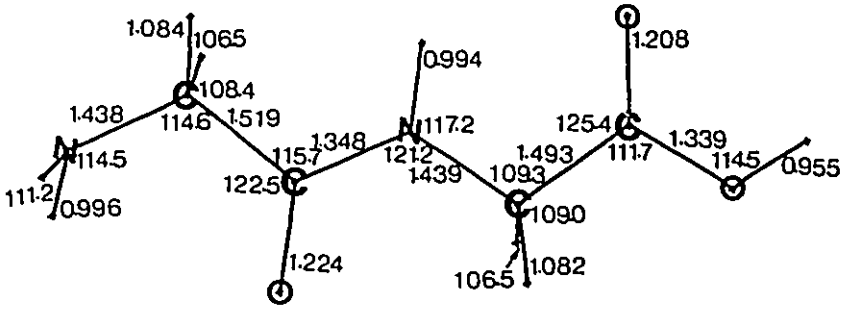
a


1.219
106.6
1.082
124.0
1.090
1.440
1.080
114.1
121.3
109.3
1.492
1.339
0.955
1.343
117.6
111.7
114.6
125.3
0.993
1.207

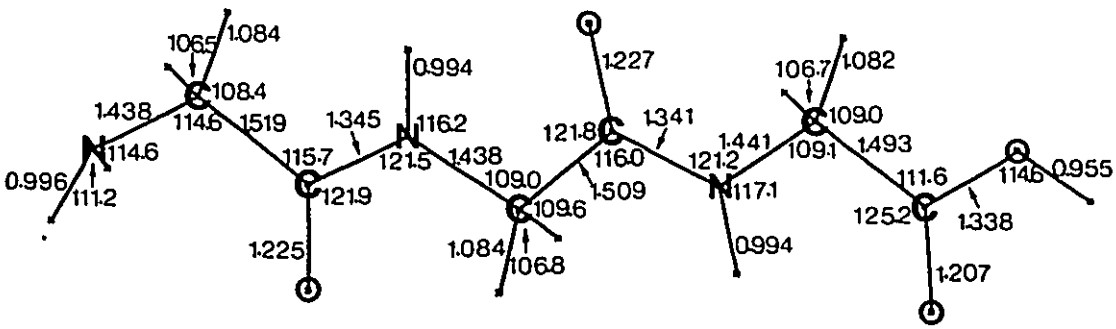
b


1.221
107.0
1.084
0.990
124.1
109.6
1.438
1.340
109.1
1.510
122.8
114.2
121.5
115.5
118.8
0.992
1.080
116.6
1.344
121.6
0.994
1.223

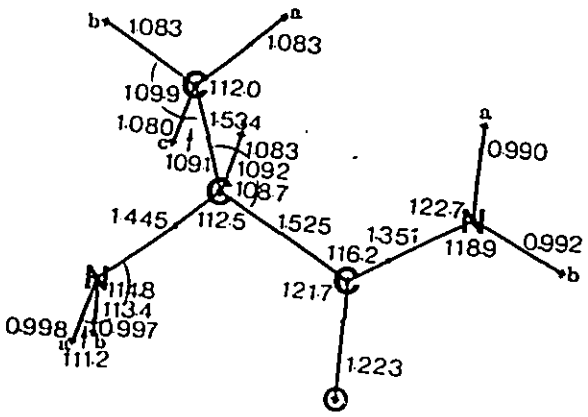
c


1.084
106.5
0.994
1.208
108.4
1.438
114.6
1.519
117.2
125.4
1.348
1.339
114.5
115.7
121.2
1.493
111.7
114.5
0.955
111.2
0.996
122.5
1.439
109.3
109.0
1.224
106.5
1.082

d


106.5
1.084
0.994
1.227
106.7
1.082
108.4
1.438
114.6
1.519
1.345
116.2
121.8
1.341
109.0
1.441
114.6
115.7
121.5
1.438
116.0
121.2
109.1
1.493
0.996
111.2
109.0
1.509
117.1
111.6
114.6
0.955
121.9
109.6
125.2
1.338
1.225
1.084
106.8
0.994
1.207

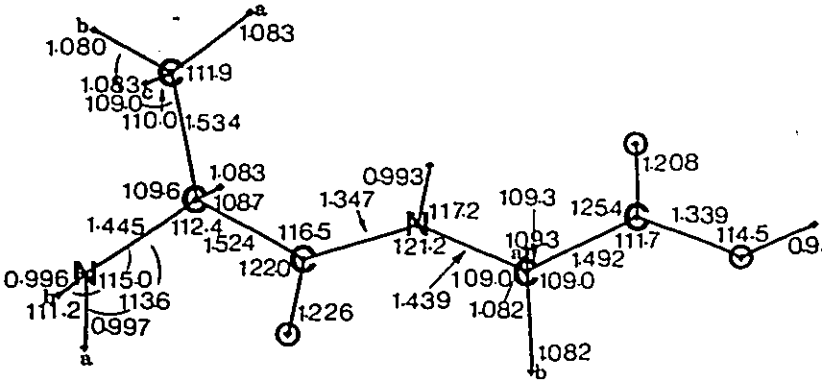
e



O4C2C1N3	-24.1
NC2C1N3	157.7
CC1C2O4	97.7
HC1C2O4	-143.6
H_aN3C1C2	55.8
H_bN3C1C2	-73.5

H_aCC1N3	-178.5
H_bCC1N3	61.2
H_cCC1N3	-57.6
H_aNC2O4	178.8
H_bNC2O4	-1.1

f

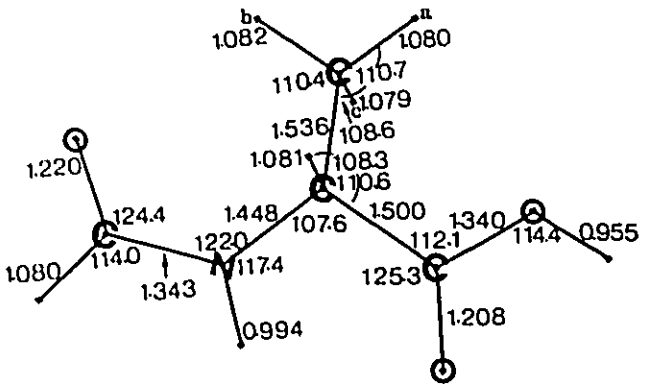


A|	
O4C2C1N3	-25.5
NC2C1N3	156.2
CC1C2O4	96.2
HC1C2O4	-145.1
H_aN3C1C2	55.8
H_bN3C1C2	-74.2

A|	
H_aCC1N3	-177.7
H_bCC1N3	61.8
H_cCC1N3	-56.9
H_aNC2O4	178.9
CNC2O4	-1.0

|G'	
C1N3CO	-1.0
C2C1N3C	180.1
O4C2C1N3	0.1
O5C2C1N3	180.1
HN3C1C2	0.1
HO5C2C1	180.1
H_aC1C2O5	57.9
H_bC1C2O5	-57.9

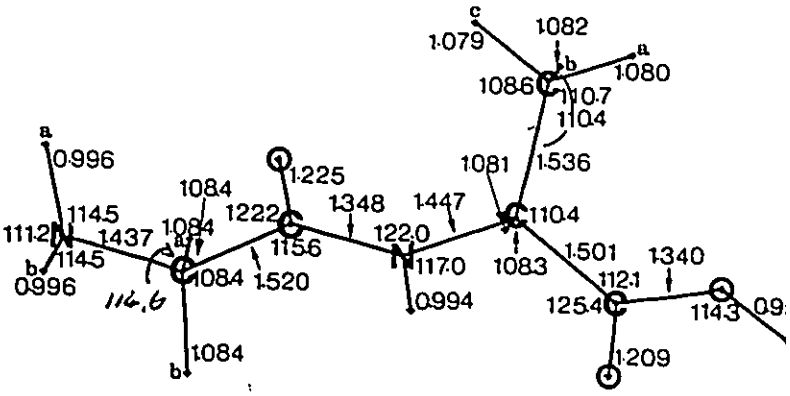
h



O4C2C1N3	-11.0
O5C2C1N3	170.2
CC1C2O4	112.1
HC1C2O4	-128.8
HN3C1C2	16.9
HO5C2O4	-0.3

CN3C1C2	199.2
OCN3C1	-0.8
HCN3C1	179.1
H_aCC1N3	180.4
H_bCC1N3	59.3
H_cCC1N3	-60.0

g



G|	
O4C2NC	-0.7
C1C2NC	179.2
N3C1C2N	180.1
H_aC1C2N	57.9
H_bC1C2N	-57.9
H_aN3C1C2	65.0
H_bN3C1C2	-65.0

|A'	
O4C2C1N3	-11.0
O5C2C1N3	170.2
CC1C2O4	112.1
HC1C2O4	-128.9
HN3C1C2	17.0
HO5C2O4	-0.3

|A'	
CN3C1C2	199.4
OCN3C1	-0.7
CCN3C1	179.2
H_aCC1N3	180.4
H_bCC1N3	59.3
H_cCC1N3	-60.1

i

atoms (not including those in the CH_2 and terminal NH_2) in these fragments also lie in the plane of the heavy atoms with the only exception being the N-H bond in COHNHC- of COH|A' with an out-of-plane deviation angle of ≈2.3 degrees. That the skeletons of the heavy atoms of the glycine fragment in both the dipeptides G|A' and A|G' remain planar (the largest deviation angle is 0.1°) is especially important, since from this fact one can reasonably ignore the unsymmetrical effect of its bonded peptide chains.

2.2. The Transferability of Amino Acid Fragments

Geometry

Table 2.2 lists the comparison of the geometries of glycine fragments G in molecules G|NH_2, G|G', G|A' and G|G"|G', G' in COH|G', G|G', A|G' and G|G"|G', and G" in COH|G"|NH_2 and G|G"|G'. Table 2.3 lists the similar comparisons for the geometry of the alanine fragment A in A|NH_2 and A|G' and A' in COH|A' and G|A'. The atomic radii for the atoms in each peptide bond are also given in these tables for use in building polypeptides. From the tables we can see that the changes in geometry for each glycine or alanine fragment in different chemical environment are very small. Table 2.2 and 2.3 shows that for glycine fragments the average change in bond length is about 0.001 A with a maximum change of 0.004 A, and in bond angle is 0.17° with a maximum change of 0.5°. For alanine fragments the average change in bond length is less than 0.001 A with a maximum change of 0.004 A, the average change in bond angle is 0.10° with a maximum change of 0.4°, and the average change in dihedral angle of alanine fragments is 0.35° with a maximum change of 1.5°. All of the largest changes happen at the site which is closest to the interatomic surface

Table 2.2. Comparison of Bond Lengths (in A) and Bond Angles (in degrees) of Glycine Fragments in Different Molecules.

G Fragment[a]

	G|NH2	G|G´	G|A´	G|G"|G´
C1-C2	1.519	1.519	1.520	1.519
C1-N3	1.438	1.438	1.437	1.438
C2-O4	1.221	1.224	1.225	1.225
C2-N[b]	0.443	0.441	0.441	0.440
C1-H	1.084	1.084	1.084	1.084
N3-H	0.996	0.996	0.996	0.996
N3C1C2	114.7	114.6	114.6	114.6
O4C2C1	122.2	122.5	122.2	122.4
NC2C1	115.4	115.7	115.6	115.7
HN3C1	114.5	114.5	114.5	114.6
HC1C2	108.4	108.4	108.4	108.4
HN3H	111.1	111.2	111.2	111.2
HC1H	106.5	106.5	106.5	106.5

a. G ≡ $NH_2CH_2CO|$
b. This is the atomic radius of carbon C2 at the C2-N peptide bond

G´ Fragment[c]

	COH|G´	G|G´	A|G´	G|G"|G´
C1-C2	1.492	1.493	1.492	1.493
C1-N3	1.440	1.439	1.439	1.441
C2-O4	1.207	1.208	1.208	1.207
C2-O5	1.339	1.339	1.339	1.338
N3-C[d]	0.904	0.907	0.907	0.901
H-C1	1.082	1.082	1.082	1.082
H-N3	0.993	0.994	0.993	0.994
H-O5	0.955	0.955	0.955	0.955
N3C1C2	109.3	109.3	109.3	109.0
O4C2C1	125.3	125.4	125.4	125.2
O5C2C1	111.7	111.7	111.7	111.6
CN3C1	121.3	121.2	121.2	121.2
HN3C1	117.6	117.2	117.2	117.1
HC1C2	109.0	109.0	109.0	109.0
HO5C2	114.6	114.5	114.5	114.6
HC1H	106.6	106.5	106.5	106.7

c. G´ ≡ $|NHCH_2COOH$
d. This is the atomic radius of nidrogen N3 at the N3-C peptide bond

Table 2.2. (con'd)

G" Fragment[e]

	COH|G"|NH2	G|G"|G'
C1-C2	1.510	1.509
C1-N3	1.438	1.438
C2-O4	1.223	1.227
C2-N[f]	0.442	0.440
N3-C[g]	0.902	0.905
C1-H	1.084	1.084
N3-H	0.994	0.994
N3C1C2	109.1	109.0
O4C2C1	121.6	121.8
NC2C1	115.5	116.0
CN3C1	121.5	121.5
HN3C1	116.6	116.2
HC1C2	109.6	109.6
HC1H	107.0	106.8

e. G" ≡ $|NHCH_2CO|$
f. This is the atomic radius of carbon C2 at the C2-N peptide bond
g. This is the atomic radius of nidrogen N3 at the N3-C peptide bond

Table 2.3. Comparison of Bond Lengths (in A) and Bond Angles (in degrees) of Alanine Fragments in Different Molecules.

	A Fragment[a]		A' Fragment[b]		
	A|NH2	A|G'		COH|A'	G|A'
C1-C2	1.525	1.524	C1-C2	1.500	1.501
C1-N3	1.445	1.445	C1-N3	1.448	1.447
C2-O4	1.223	1.226	C2-O4	1.208	1.209
C1-C	1.534	1.534	C2-O5	1.340	1.340
C2-N[c]	0.442	0.440	C1-C	1.536	1.536
H-C1	1.083	1.083	N3-C[d]	0.903	0.907
H_a-N3	0.998	0.997	H-C1	1.081	1.081
H_b-N3	0.997	0.996	H-N3	0.994	0.994
H_a-C	1.083	1.083	H-O5	0.955	0.955
H_b-C	1.083	1.083	H_a-C	1.080	1.080
H_c-C	1.080	1.080	H_b-C	1.082	1.082
N3C1C2	112.5	112.4	H_c-C	1.079	1.079
O4C2C1	121.7	122.0	N3C1C2	107.6	107.6
NC2C1	116.2	116.5	O4C2C1	125.3	125.4
CC1C2	109.2	109.1	O5C2C1	112.1	112.1
H_aN3C1	113.4	113.6	CC1C2	110.6	110.4
H_bN3C1	114.8	115.0	CN3C1	122.0	122.0
HC1C2	108.7	108.7	HN3C1	117.4	117.0
H_aN3H_b	111.2	111.4	HC1C2	108.3	108.3
HC1C	108.9	108.9	HO5C2	114.4	114.3
H_aCC1	112.0	111.9	HC1C	108.7	108.7
H_bCC1	109.9	109.9	H_aCC1	110.7	110.7
H_cCC1	109.1	109.0	H_bCC1	110.4	110.4
O4C2C1N3	-24.1	-25.5	H_cCC1	108.6	108.6
NC2C1N3	157.7	156.3	O4C2C1N3	-11.0	-11.0
CC1C2O4	97.7	96.3	O5C2C1N3	170.2	170.2
H_aN3C1C2	55.8	55.8	CC1C2O4	112.1	112.1
H_bN3C1C2	-73.5	-74.2	CN3C1C2	199.2	199.4
HC1C2O4	-143.6	-145.1	HN3C1C2	16.9	17.0
H_aCC1N3	-178.5	-177.7	HC1C2O4	-128.9	-128.9
H_bCC1N3	61.2	61.8	HO5C2O4	-0.3	-0.3
H_cCC1N3	-57.6	-56.9	H_aCC1N3	180.4	180.4
			H_bCC1N3	59.3	59.3
			H_cCC1N3	-60.0	-60.1

[a]. A ≡ $NH_2CHCH_3CO|$

[b]. A' ≡ $|NHCHCH_3COOH$

c. This is the atomic radius of carbon C2 at the C2-N peptide bond

[d]. This is the atomic radius of nidrogen N3 at the N3-C peptide bond

between each fragment and the substituent group. From the comparison, we can see that geometries of the amino acid fragments are transferable with only small changes. If we replace the fragments in large polypeptides by the corresponding ones in the simple model molecules, the errors which are introduced from the replacement are acceptably small, being less than 0.005 A for bond length and less than 0.5 degrees for bond angle. Such small changes are acceptable since larger errors are introduced by a change in basis set. For instance, for optimised geometries of protonated benzene with 3-21G and 6-31G respectively (see Table AII of Chapter 4), the maximal difference in bond length is 0.006 A and in bond angle is 1.0 degree. The crystal structure data of glycine (Marsh 1958), alanine (Simpson and Marsh 1966), glycylclycine (Biswas et al 1968), glycylalanine (Wang and Paul 1979), alanylglycine (Michel et al 1970) and glycylglycylglycine (Srikrishnan et al 1982) are listed in Appendix of this chapter for comparison.

The Properties of Bond Critical Points

As described in Chapter 1, a bond can be characterized in terms of the local properties of ρ at its critical point: its value ρ_b, its Laplacian $\nabla^2\rho_b$, its ellipticity ϵ and its bonded radius. Tables 2.4 - 2.5 compare the values of ρ, $\nabla^2\rho$, ϵ and bonded radii, for the critical points in glycine and alanine fragments in different chemical environments. All of these properties for each fragment vary by only small amounts. The average change in ρ is 0.001 au with a maximum of 0.005 au at the peptide bond of glycyl complexes. The Laplacian of the charge density is a sensitive quantity, the values of $\nabla^2\rho$ at the carbon-carbon bond critical point of ethane, ethylene, acetylene and benzene for a 6-31G* calculation,

Table 2.4. Comparison of Properties[a] of ρ at Bond Critical Point for Glycine Fragments in Different Molecules.

		G Fragment[b]			
		G|NH2	G|G'	G|A'	G|G"G'
	ρ	0.274	0.274	0.274	0.274
	$\nabla^2\rho$	-0.822	-0.822	-0.821	-0.821
C1-C2	ϵ	0.088	0.089	0.089	0.089
	r_b(C1)	1.349	1.341	1.342	1.344
	r_b(C2)	1.522	1.529	1.530	1.527
	ρ	0.288	0.288	0.288	0.287
	$\nabla^2\rho$	-1.029	-1.028	-1.029	-1.026
C1-N3	ϵ	0.033	0.032	0.031	0.031
	r_b(C1)	1.019	1.019	1.017	1.015
	r_b(N3)	1.699	1.699	1.700	1.703
	ρ	0.413	0.410	0.409	0.409
	$\nabla^2\rho$	0.250	0.239	0.224	0.228
C2-O4	ϵ	0.062	0.057	0.053	0.054
	r_b(C2)	0.761	0.763	0.763	0.763
	r_b(O4)	1.546	1.551	1.552	1.552
	ρ	0.330	0.333	0.333	0.335
	$\nabla^2\rho$	-0.661	-0.621	-0.634	-0.623
C2|N[c]	ϵ	0.021	0.005	0.003	0.000
	r_b(C2)	0.837	0.833	0.833	0.832
	r_b(N)	1.718	1.715	1.713	1.711
	ρ	0.294	0.294	0.294	0.294
	$\nabla^2\rho$	-1.108	-1.112	-1.112	-1.113
C1-H	ϵ	0.038	0.037	0.038	0.038
	r_b(C1)	1.252	1.254	1.254	1.253
	r_b(H)	0.796	0.794	0.794	0.794
	ρ	0.363	0.363	0.363	0.363
	$\nabla^2\rho$	-2.013	-2.012	-2.013	-2.012
N3-H	ϵ	0.062	0.062	0.062	0.062
	r_b(N3)	1.417	1.417	1.416	1.416
	r_b(H)	0.465	0.465	0.465	0.466

Table 2.4 (con'd)

G' Fragment[d]

		COH|G'	G|G'	A|G'	G|G''|G'
	ρ	0.287	0.287	0.287	0.287
	$\nabla^2\rho$	-0.908	-0.907	-0.908	-0.907
C1-C2	ϵ	0.107	0.107	0.107	0.107
	r_b(C1)	1.304	1.300	1.299	1.308
	r_b(C2)	1.516	1.521	1.521	1.514
	ρ	0.277	0.277	0.278	0.275
	$\nabla^2\rho$	-0.794	-0.810	-0.812	-0.760
C1-N3	ϵ	0.063	0.065	0.066	0.063
	r_b(C1)	0.919	0.921	0.921	0.914
	r_b(N3)	1.802	1.799	1.799	1.809
	ρ	0.425	0.425	0.425	0.425
	$\nabla^2\rho$	0.369	0.367	0.366	0.369
C2-O4	ϵ	0.103	0.101	0.101	0.103
	r_b(C2)	0.753	0.753	0.753	0.753
	r_b(O4)	1.529	1.529	1.529	1.529
	ρ	0.305	0.305	0.305	0.306
	$\nabla^2\rho$	-0.013	-0.012	-0.012	-0.013
C2-O5	ϵ	0.115	0.118	0.118	0.113
	r_b(C2)	0.812	0.812	0.812	0.812
	r_b(O5)	1.718	1.719	1.719	1.717
	ρ	0.337	0.333	0.333	0.340
	$\nabla^2\rho$	-0.620	-0.621	-0.601	-0.672
N3|C[e]	ϵ	0.030	0.005	0.008	0.022
	r_b(N3)	1.708	1.715	1.714	1.702
	r_b(C)	0.830	0.833	0.831	0.831
	ρ	0.362	0.362	0.363	0.361
	$\nabla^2\rho$	-2.060	-2.053	-2.058	-2.055
N3-H	ϵ	0.057	0.058	0.058	0.056
	r_b(N3)	1.449	1.447	1.447	1.452
	r_b(H)	0.428	0.431	0.430	0.427
	ρ	0.298	0.298	0.298	0.298
	$\nabla^2\rho$	-1.156	-1.155	-1.154	-1.160
C1-H$_a$	ϵ	0.041	0.040	0.040	0.041
	r_b(C1)	1.278	1.277	1.277	1.278
	r_b(H$_a$)	0.767	0.768	0.768	0.766
	ρ	0.298	0.298	0.298	0.298
	$\nabla^2\rho$	-1.156	-1.155	-1.156	-1.160
C1-H$_b$	ϵ	0.041	0.040	0.040	0.041
	r_b(C1)	1.278	1.277	1.277	1.278
	r_b(H$_b$)	0.767	0.768	0.768	0.766
	ρ	0.373	0.373	0.373	0.373
	$\nabla^2\rho$	-2.443	-2.441	-2.441	-2.445
O5-H	ϵ	0.018	0.018	0.018	0.018
	r_b(O5)	1.473	1.472	1.472	1.473
	r_b(H)	0.333	0.333	0.333	0.332

Table 2.4 (con'd)

		G" fragment[a]	
		COH|G"|NH2 G|G"|G'	
	ρ	0.278	0.279
	$\nabla^2\rho$	-0.848	-0.850
C1-C2	ϵ	0.082	0.084
	r_b(C1)	1.369	1.358
	r_b(C2)	1.484	1.493
	ρ	0.278	0.279
	$\nabla^2\rho$	-0.792	-0.807
C1-N3	ϵ	0.050	0.054
	r_b(C1)	0.916	0.919
	r_b(N3)	1.801	1.798
	ρ	0.412	0.408
	$\nabla^2\rho$	0.206	0.195
C2-O4	ϵ	0.061	0.055
	r_b(C2)	0.763	0.764
	r_b(O4)	1.548	1.554
	ρ	0.337	0.340
	$\nabla^2\rho$	-0.717	-0.672
C2|N[c]	ϵ	0.007	0.022
	r_b(C2)	0.835	0.831
	r_b(N)	1.705	1.702
	ρ	0.339	0.335
	$\nabla^2\rho$	-0.621	-0.623
N3|C[c]	ϵ	0.035	0.000
	r_b(N3)	1.704	1.711
	r_b(C)	0.828	0.832
	ρ	0.296	0.296
	$\nabla^2\rho$	-1.133	-1.135
C1-H	ϵ	0.039	0.038
	r_b(C1)	1.269	1.269
	r_b(H)	0.779	0.779
	ρ	0.361	0.361
	$\nabla^2\rho$	-2.060	-2.053
N3-H	ϵ	0.054	0.056
	r_b(N3)	1.456	1.453
	r_b(H)	0.423	0.426

[a]. All quantities are in au
[b]. G ≡ $NH_2CH_2CO|$
[c]. This is an amide bond
[d]. G' ≡ $|NHCH_2COOH$
[e]. G" ≡ $|NHCH_2CO|$

Table 2.5. Comparison of Properties[a] of ρ at Bond Critical Points for Alanine Fragments in Different Molecules.

A Fragment[b]

		A|NH2	A|G'
	ρ	0.273	0.273
	$\nabla^2\rho$	-0.817	-0.818
C1-C2	ϵ	0.073	0.073
	r_b(C1)	1.346	1.338
	r_b(C2)	1.535	1.542
	ρ	0.285	0.285
	$\nabla^2\rho$	-1.006	-1.006
C1-N3	ϵ	0.041	0.040
	r_b(C1)	1.022	1.020
	r_b(N3)	1.709	1.710
	ρ	0.411	0.408
	$\nabla^2\rho$	0.238	0.228
C2-O4	ϵ	0.057	0.052
	r_b(C2)	0.762	0.763
	r_b(O4)	1.548	1.553
	ρ	0.330	0.333
	$\nabla^2\rho$	-0.644	-0.601
C2|N[c]	ϵ	0.024	0.008
	r_b(C2)	0.835	0.831
	r_b(N)	1.717	1.714
	ρ	0.255	0.255
	$\nabla^2\rho$	-0.703	-0.704
C1-C	ϵ	0.033	0.033
	r_b(C1)	1.460	1.482
	r_b(C)	1.419	1.417
	ρ	0.296	0.297
	$\nabla^2\rho$	-1.119	-1.124
C1-H	ϵ	0.032	0.032
	r_b(C1)	1.242	1.244
	r_b(H)	0.805	0.803

		A|NH2	A|G'
	ρ	0.361	0.361
	$\nabla^2\rho$	-2.008	-2.009
N3-H$_a$	ϵ	0.058	0.059
	r_b(N3)	1.424	1.424
	r_b(H$_a$)	0.461	0.461
	ρ	0.362	0.362
	$\nabla^2\rho$	-1.997	-1.998
N3-H$_b$	ϵ	0.063	0.063
	r_b(N3)	1.412	1.412
	r_b(H$_b$)	0.472	0.472
	ρ	0.289	0.289
	$\nabla^2\rho$	-1.065	-1.068
C-H$_a$	ϵ	0.011	0.011
	r_b(C)	1.235	1.237
	r_b(H$_a$)	0.810	0.809
	ρ	0.290	0.290
	$\nabla^2\rho$	-1.074	-1.072
C-H$_b$	ϵ	0.007	0.007
	r_b(C)	1.246	1.245
	r_b(H$_b$)	0.801	0.802
	ρ	0.293	0.293
	$\nabla^2\rho$	-1.100	-1.101
C-H$_c$	ϵ	0.007	0.007
	r_b(C)	1.248	1.248
	r_b(H$_c$)	0.792	0.792

Table 2.5 (con'd)

A' Fragment[a]

		COH|A'	G|A'			COH|A'	G|A'
	ρ	0.284	0.284		ρ	0.362	0.362
	$\nabla^2\rho$	-0.894	-0.894		$\nabla^2\rho$	-2.060	-2.052
C1-C2	ϵ	0.100	0.101	N3-H	ϵ	0.056	0.057
	r_b(C1)	1.299	1.295		r_b(N3)	1.449	1.446
	r_b(C2)	1.536	1.540		r_b(H)	0.429	0.432
	ρ	0.272	0.273		ρ	0.302	0.302
	$\nabla^2\rho$	-0.753	-0.768		$\nabla^2\rho$	-1.187	-1.189
C1-N3	ϵ	0.056	0.058	C1-H	ϵ	0.034	0.034
	r_b(C1)	0.919	0.921		r_b(C1)	1.280	1.279
	r_b(N3)	1.817	1.814		r_b(H)	0.763	0.763
	ρ	0.424	0.424		ρ	0.373	0.373
	$\nabla^2\rho$	0.368	0.366		$\nabla^2\rho$	-2.440	-2.438
C2-O4	ϵ	0.099	0.098	O5-H	ϵ	0.018	0.018
	r_b(C2)	0.754	0.754		r_b(O5)	1.473	1.472
	r_b(O4)	1.530	1.530		r_b(H)	0.333	0.333
	ρ	0.304	0.304		ρ	0.293	0.292
	$\nabla^2\rho$	0.001	0.002		$\nabla^2\rho$	-1.097	-1.095
C2-O5	ϵ	0.124	0.127	C-H_a	ϵ	0.013	0.013
	r_b(C2)	0.812	0.812		r_b(C)	1.247	1.246
	r_b(O5)	1.720	1.721		r_b(H_a)	0.794	0.795
	ρ	0.253	0.253		ρ	0.290	0.290
	$\nabla^2\rho$	-0.692	-0.692		$\nabla^2\rho$	-1.074	-1.072
C1-C	ϵ	0.036	0.036	C-H_b	ϵ	0.013	0.013
	r_b(C1)	1.517	1.513		r_b(C)	1.242	1.241
	r_b(C)	1.387	1.390		r_b(H_b)	0.803	0.804
	ρ	0.337	0.333		ρ	0.295	0.295
	$\nabla^2\rho$	-0.635	-0.634		$\nabla^2\rho$	-1.123	-1.124
N3|C[c]	ϵ	0.032	0.003	C-H_c	ϵ	0.009	0.009
	r_b(N3)	1.707	1.713		r_b(C)	1.259	1.258
	r_b(C)	0.831	0.833		r_b(H_c)	0.780	0.780

[a]. All quantities are in au
[b]. A ≡ NH_2CHCH_3CO|
[c]. This is an amide bond
[d]. A' ≡ |$NHCHCH_3COOH$

for example, (Bader et al 1983) are -15.94, -28.73, -30.88 and -24.42 A^{-5}, respectively. However, in Tables 2.4 - 2.5, the average change in $\nabla^2\rho$ is 0.006 au with a maximum of 0.052 au for the glycine peptide bonds. The average change in ellipticity ϵ is 0.003 with a maximum of 0.035, compared to the values of 0.45 and 0.23 for ethylene and benzene (Bader et al 1983), as reference values. The average change in bond radius is less than 0.003 au with the maximum of 0.013 au, which is of the nitrogen bond radius at the peptide bond in the G′ fragment. The properties of ρ at a bond critical point, ρ_b, $\nabla^2\rho_b$ and ϵ characterize a chemical bond. The small changes found in the values of these critical point properties for the fragments in the different molecules shows that the bonds in these fragments and hence their total charge distributions are essentially the same. This again demonstrates the transferability of amino acid fragments.

Atomic Properties

Atomic properties are obtained by integrating the corresponding property density over the atomic basin. As discussed in Chapter 1, the Laplacian L(Ω) has been used as a criterion by which the quality of the integration over the given atom is based. If $L(\Omega) < 5\times10^{-3}$, a "good" integration is obtained. The current integration program (Biegler-Konig et al 1982) which we employed can handle most molecules with desirable accuracy, but in the present project it appears to have difficulty in integrating some atoms which have complicated interatomic surfaces. For example, the L(Ω) values of the C2 atom of both G′ and A′ fragment in all molecules are greater than 1.8×10^{-2}. These integration errors introduce difficulties in judging the quality of the integrated results. However, the error in the integrated value can be further determined by summing up

the individual atomic properties comparing with the molecular value. The sum of the atomic populations should equal the number of electrons in the molecule, while the sum of atomic energies should equal the total SCF energy. The error in the sum of the atomic populations ranges from a minimum value of 0.009e for G|NH_2 to a maximum of 0.080e for G|G"|G' (see Tables 2.6 - 2.7). The sums of the atomic energies for each molecule are acceptable (see Tables 2.8 - 2.9), the errors in the sum being 0.0016 au for COH|A' as the smallest one, and 0.0118 au for G|G"|G' as the largest one. These errors are relativly small compared to the total populations of 40e for G|NH_2 and 100e for G|G"|G' and the total energies of 434.2251 au for COH|A' and 695.8204 au for G|G"|G'.

The average change in atomic population is 0.005 e with a maximum of 0.022 e. The maximal differences in a total population for a given fragment in the different complexes are only 0.009e for G, 0.023 for G', 0.017 for G", 0.003 for A and 0.024 for A'. From the Table 2.8 and 2.9 we can see that the average change in an atomic energy is 0.0054 au (3.3 Kcal), and the maximum change is 0.0198 au (12.4 Kcal). The maximal difference in total energy for a given fragment in the different molecules is larger than the error for a total system, being 0.0231 au for G, 0.0583 for G', 0.0048 for G", 0.0390 for A and 0.0185 for A'.

This research group was the first to demonstrate that the size and shape of a molecule, as determined by its nonbonded interactions, could be related to a particular outer contour of the electronic charge density (Bader et al 1967, Bader et al 1987a). The intersection of an atomic surface with a particular envelope of the charge density determines a corresponding atomic volume while the envelope itself determines the van

Table 2.6. Comparison of Atomic Populations N(Ω) for Glycine Fragments in Different Molecules.

G Fragment[a]

	G|NH2	G|G′	G|A′	G|G″|G′
C1	5.337	5.326	5.331	5.332
C2[b]	4.225	4.239	4.237	4.231
N3	8.219	8.221	8.220	8.220
O4	9.373	9.378	9.383	9.381
H(N3)	0.594	0.594	0.595	0.596
H(N3)	0.594	0.594	0.594	0.596
H(C1)	1.056	1.052	1.051	1.052
H(C1)	1.056	1.052	1.052	1.052
$\Sigma\Omega_{eG}$	30.454	30.456	30.463	30.460
$\Sigma\Omega^{c}$	40.009	70.050	78.062	100.080
Total	40.000	70.000	78.000	100.00

G′ Fragment[d]

	COH|G′	G|G′	A|G′	G|G″|G′
C1	5.304	5.297	5.295	5.302
C2	4.169	4.178	4.169	4.173
N3[b]	8.584	8.588	8.590	8.606
O4	9.369	9.371	9.370	9.367
O5	9.328	9.328	9.328	9.329
H(N3)	0.502	0.510	0.509	0.499
H(C1)	0.992	0.995	0.995	0.991
H(C1)	0.992	0.995	0.994	0.991
H(O5)	0.331	0.332	0.332	0.330
$\Sigma\Omega_{eG'}$	39.571	39.594	39.582	39.588
$\Sigma\Omega^{b}$	54.036	70.051	78.040	100.080
Total	54.000	70.000	78.000	100.00

Table 2.6. (con'd)

	G" Fragment[a]	
	COH|G"|NH2 G|G"|G'	
C1	5.333	5.321
C2[b]	4.188	4.199
N3[b]	8.593	8.597
O4	9.371	9.378
H(N3)	0.490	0.499
H(C1)	1.020	1.019
H(C1)	1.020	1.019
$\Sigma_{\Omega \in G''}$	30.015	30.032
Σ_{Ω}[b]	54.023	100.080
Total	54.000	100.00

[a]. G ≡ $NH_2CH_2CO|$
[b]. This is an amide atom
[c]. Ω goes through all the atoms in corresponding molecule
[d]. G' ≡ $|NHCH_2COOH$
[e]. G" ≡ $|NHCH_2CO|$

Table 2.7. Comparison of Atomic Populations N(Ω) for Alanine Fragment in Different Molecules.

A Fragment[a]			A′ Fragment[b]		
	A|NH2	A|G′		COH|A′	G|A′
C1	5.394	5.387	C1	5.378	5.378
C2[c]	4.233	4.248	C2	4.186	4.190
N3	8.226	8.229	N3[c]	8.587	8.591
O4	9.377	9.382	O4	9.372	9.372
C	5.753	5.752	O5	9.330	9.330
H_a(N3)	0.608	0.607	C	5.754	5.755
H_b(N3)	0.583	0.583	H(N3)	0.505	0.513
H(C1)	1.084	1.078	H(O5)	0.332	0.333
H_a(C)	1.085	1.080	H(C1)	0.992	0.993
H_b(C)	1.062	1.064	H_a(C)	1.049	1.051
H_c(C)	1.050	1.048	H_b(C)	1.069	1.071
$\Sigma_{\Omega\in A}$	38.455	38.458	H_c(C)	1.018	1.019
Σ_{Ω}[d]	48.010	78.054	$\Sigma_{\Omega\in A'}$	47.572	47.596
Total	48.000	78.000	Σ_{Ω}[d]	62.044	78.059
			Total	62.000	78.000

[a]. A ≡ NH_2CHCH_3CO|

[b]. A′ ≡ |$NHCHCH_3COOH$

[c]. This is an amide atom

[d]. Ω goes through all the atoms in corresponding molecule

Table 2.8. Comparison of Negative Atomic Energies[a] -E(Ω) for Glycine Fragments in Different Molecules.

G Fragment[b]

	G|NH2	G|G′	G|A′	G|G″|G′
C1	37.4328	37.4324	37.4278	37.4310
C2[c]	36.6251	36.6418	36.6382	36.6406
N3	54.8588	54.8630	54.8584	54.8645
O4	75.5514	75.5497	75.5392	75.5476
H(N3)	0.4637	0.4638	0.4642	0.4645
H(N3)	0.4637	0.4638	0.4638	0.4645
H(C1)	0.6681	0.6662	0.6658	0.6665
H(C1)	0.6681	0.6662	0.6664	0.6665
$\Sigma_{\Omega^{eG}}$	206.7317	206.7469	206.7238	206.7457
$\Sigma_{\Omega^{d}}$	262.7576	489.1894	528.1913	695.8086
SCF	262.7597	489.1964	528.1987	695.8204
Diff	-0.0021	-0.0070	-0.0074	-0.0118

G′ Fragment[e]

	COH|G′	G|G′	A|G′	G|G″|G′
C1	37.4282	37.4176	37.4148	37.4224
C2	36.5514	36.5498	36.5456	36.5437
N3[c]	55.3169	55.3125	55.3069	55.3196
O4	75.6040	75.5947	75.5867	75.5960
O5	75.5624	75.5533	75.5435	75.5548
H(N3)	0.4133	0.4177	0.4176	0.4111
H(C1)	0.6405	0.6422	0.6418	0.6401
H(C1)	0.6405	0.6422	0.6415	0.6401
H(O5)	0.3119	0.3125	0.3124	0.3110
$\Sigma_{\Omega^{eG}}$	282.4691	282.4425	282.4108	282.4388
$\Sigma_{\Omega^{d}}$	395.2127	489.1894	528.1931	695.8086
SCF	395.2148	489.1964	528.1983	695.8204
Diff	-0.0021	-0.0070	-0.0052	-0.0118

Table 2.8. (con'd)

	G" Fragment[d]					
	COH	G"	NH2	G	G"	G'
C1	37.4353			37.4241		
C2[c]	36.5980			36.6114		
N3[c]	55.3260			55.3262		
O4	75.5583			75.5460		
H(N3)	0.4057			0.4108		
H(C1)	0.6528			0.6528		
H(C1)	0.6528			0.6528		
$\Sigma_{\Omega \in G''}$	206.6289			206.6241		
Σ_{Ω}[d]	375.3983			695.8086		
SCF	375.4026			695.8204		
Diff	-0.0043			-0.0118		

a. The value is in au
b. G ≡ $NH_2CH_2CO|$
c. This is an amide atom
d. Ω goes through all the atoms in corresponding molecule
e. G' ≡ $|NHCH_2COOH$
f. G" ≡ $|NHCH_2CO|$

Table 2.9. Comparison of Negative Atomic Energies[a] -E(Ω) for Alanine Fragments in Different Molecules.

	A Fragment[b]		A′ Fragment[c]		
	A|NH2	A|G′		COH|A′	G|A′
C1	37.4539	37.4552	C1	37.4524	37.4421
C2[d]	36.6360	36.6558	C2	36.5682	36.5670
N3	54.8615	54.8725	N3[d]	55.3159	55.3153
O4	75.5342	75.5384	O4	75.5923	75.5870
C	37.6518	37.6585	O5	75.5505	75.5437
H_a(N3)	0.4706	0.4704	C	37.6625	37.6599
H_b(N3)	0.4561	0.4560	H(N3)	0.4149	0.4193
H(C1)	0.6883	0.6861	H(O5)	0.3126	0.3132
H_a(C)	0.6718	0.6698	H(C1)	0.6473	0.6481
H_b(C)	0.6611	0.6621	H_a(C)	0.6585	0.6595
H_c(C)	0.6580	0.6575	H_b(C)	0.6647	0.6657
$\Sigma_{\Omega \in A}$	245.7433	245.7823	H_c(C)	0.6462	0.6467
Σ_{Ω}[e]	301.7598	528.1931	$\Sigma_{\Omega \in A'}$	321.4860	321.4675
SCF	301.7614	528.1983	Σ_{Ω}[e]	434.2136	528.1913
Diff	-0.0016	-0.0052	SCF	434.2251	528.1987
			Diff	-0.0115	-0.0074

[a]. The values are in au
[b]. A ≡ $NH_2CHCH_3CO|$
[c]. A′ ≡ $|NHCHCH_3COOH$
[d]. This is an amide atom
[e]. Ω goes through all the atoms in corresponding molecule

der Waals envelope or shape of the molecule. The resulting shape of a functional group and the volumes of its atoms are readily determined and are transferable to a remarkable degree in many instances, including the amino acid building blocks. This can be clearly seen in Tables 2.10-2.11, which list the comparison of atomic volumes of glycine and alanine fragments in different molecules separately. The changes in atomic volumes for each fragment among various molecules are very small, so these too are transferable.

Chemical Reactivity

It has been demonstrated that the Laplacian distribution function determines not only the sites of electrophilic attack, but also the sites of nucleophilic attack (Bader and Chang 1989a, Carroll et al 1989, Carroll et al 1988, Slee 1986, Bader and MacDougall 1985, Bader et al 1984). These sites are located at the positions of the local nonbonded maxima and local minima respectively, in the negative of the Laplacian distribution.

Each amino acid fragment has both nonbonded maxima (appearing as (3,-3) critical points in $-\nabla^2\rho$), i.e., nonbonded local charge concentrations in the VSCC's of the oxygen and nitrogen atom, which serve as the sites of electrophilic attack, and minima (appearing as (3,+1) critical points in $-\nabla^2\rho$) or local charge depletions in the VSCC's of the carbonyl carbons (see Fig.2.5), which serve as the sites of nucleophilic attack.

Fig.2.5 gives the contour plot of the Laplacian distribution for Gly-Gly dipeptide. The in-plane positions of the critical points which determine the locations of the charge concentrations and depletions and ultimately the direction of approach of a reactant are indicated by

Table 2.10. Comparison of Atomic Volumes (ρ=0.001) V(Ω)[a] for Glycine Fragments in Different Molecules.

G Fragment[b]

	G|NH2	G|G′	G|A′	G|G″|G′
C1	48.74	47.94	48.59	49.08
C2[c]	31.91	32.29	32.15	31.84
N3	110.17	110.61	110.28	110.18
O4	126.89	126.29	125.52	126.42
H(N3)	28.68	28.74	28.77	28.79
H(N3)	28.68	28.74	28.68	28.79
H(C1)	49.21	48.82	48.70	48.85
H(C1)	49.21	48.82	48.82	48.85
$\Sigma_{\Omega \in G}$	473.49	472.25	471.51	472.80

G′ Fragment[d]

	COH|G′	G|G′	A|G′	G|G″|G′
C1	47.38	47.34	46.89	47.05
C2	31.40	31.64	31.73	32.74
N3[c]	101.47	99.82	99.94	103.16
O4	125.16	125.13	125.91	125.75
O5	116.01	115.19	116.04	115.36
H(N3)	22.12	22.44	22.55	21.99
H(C1)	45.15	45.22	45.14	44.95
H(C1)	45.15	45.22	45.11	44.95
H(O5)	16.26	16.34	16.31	16.20
$\Sigma_{\Omega \in A'}$	550.10	548.34	549.62	552.15

G″ Fragment[e]

	COH|G″	NH2 G|G″|G′
C1	47.39	47.09
C2[c]	31.22	31.86
N3[c]	101.18	99.77
O4	125.00	124.35
H(N3)	21.20	21.57
H(C1)	46.80	46.61
H(C1)	46.80	46.61
$\Sigma_{\Omega \in A''}$	419.59	417.86

a. The values are in au.
b. G ≡ $NH_2CH_2CO|$
c. This is an amide atom
c. G′ ≡ $|NHCH_2COOH$
d. G″ ≡ $|NHCH_2CO|$

Table 2.11. Comparison of Atomic Volumes (ρ=0.001) V(Ω)[a] for Alanine Fragments in Different Molecules.

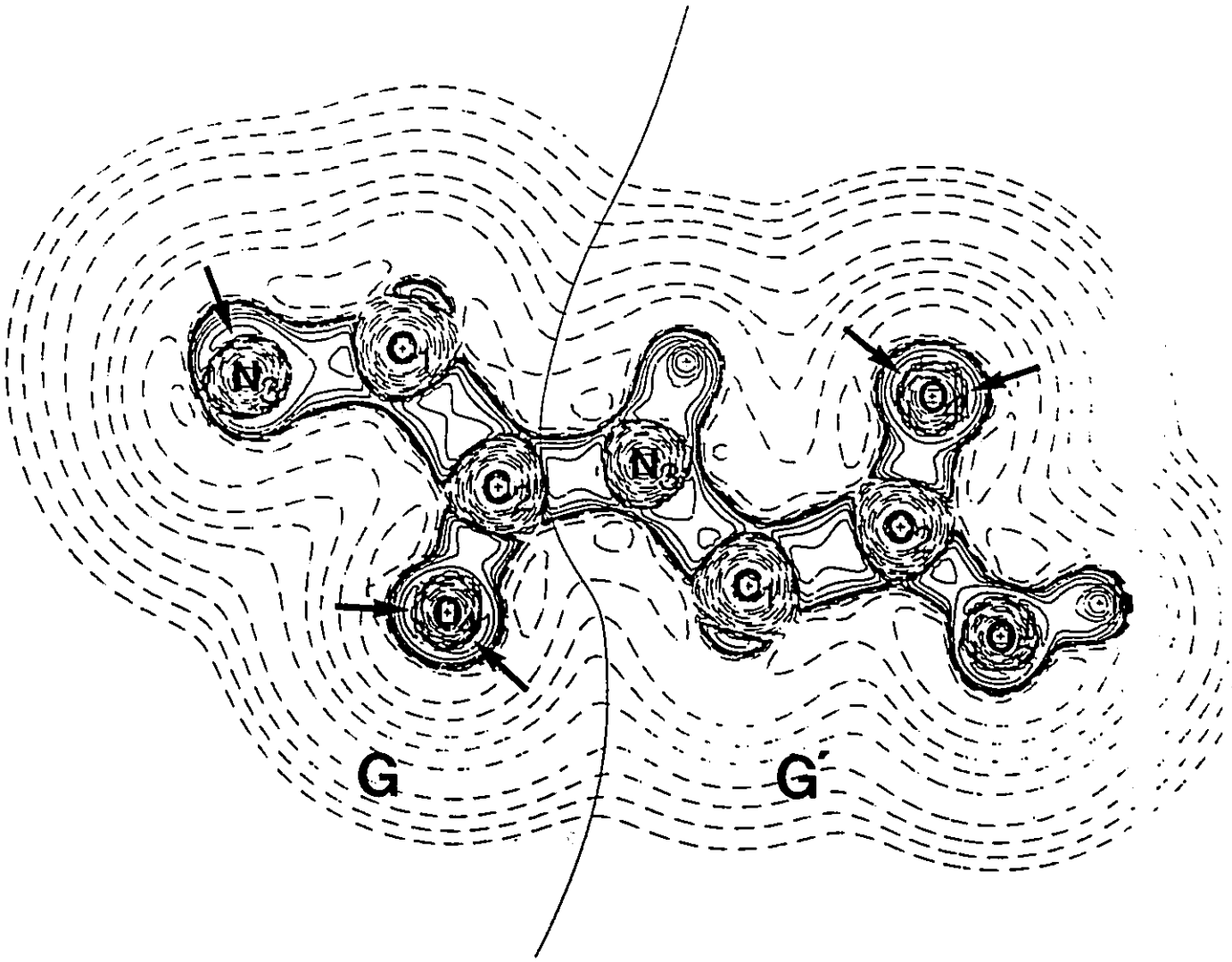
	A Fragment[b]	
	A|NH2	A|G´
C1	42.43	42.24
C2[d]	30.57	30.80
N3	108.74	108.94
O4	126.96	126.39
C	61.78	61.81
H_a(N3)	29.27	29.29
H_b(N3)	28.06	28.03
H(C1)	49.51	48.96
H_a(C)	51.40	51.06
H_b(C)	50.36	50.50
H_c(C)	49.17	49.04
$\Sigma\Omega_{eA}$	628.25	627.06

	A´ Fragment[c]	
	COOH|A´	G|A´
C1	42.22	43.13
C2	30.57	31.46
N3[d]	99.44	97.69
O4	125.25	125.62
O5	114.72	114.61
C	62.48	62.40
H(N3)	22.25	22.60
H(O5)	16.35	16.38
H(C1)	42.98	42.87
H_a(C)	48.07	48.27
H_b(C)	50.95	51.15
H_c(C)	45.72	45.75
$\Sigma\Omega_{eA'}$	701.00	701.93

a. The values are in au.
b. A ≡ $NH_2CHCH_3CO|$
c. A´ ≡ $|NHCHCH_3COOH$
d. This is an amide atom

Figure 2.5.

Contour plot of $\nabla^2\rho$ for glycylglycine dipeptide. The solid line through the peptide bond is the interatomic surface of carbon and nitrogen, which divides the molecule into two different glycine fragments, G and G′. The number on each atom is assigned by referring to each fragment, respectively. The arrows point to positions of the in-plane maximum charge concentrations. The solid (dashed) lines denote regions of charge concentration (depletion), where $\nabla^2\rho < 0$ ($\nabla^2\rho > 0$). The contour values are ±0.002, ±0.004, ±0.008 (the outermost contour has a value of +0.001 au), increasing in powers of 10 to ±20.0 au.


N
O
O
N
O
O
O
O
G
G'

corresponding arrows, while those out-of-plane ones are not shown in the figure. For the carbonyl group in all of the molecules studied in this chapter, there are always two in-plane maxima of charge concentration for the O4 atom and two out-of-plane minima above and below C2 atom, as shown in Fig.1.7 for $H_2C{=}O$ molecule. For the glycyl fragment G, there is an in-plane maximum of charge concentration for the N3 atom. For the glycine fragment G′, there are two out-of-plane maxima for each of N3 and O5 atom, respectively. For the glycyl unit G", there are two out-of-plane maxima above and below N3 atom. The same comments apply to the alanine fragments as well.

The positions as well as the magnitudes of $\nabla^2\rho$ at these extreme points, differ by only small amounts for the fragments in different molecules (see Tables 2.12 - 2.13). The magnitude of $\nabla^2\rho$ at the nonbonded in-plane maxima on N3 in G| and A| and O4 in all fragments change by less than 1% while the out-of-plane maxima on N3 in fragments |G"|, |G′ and A′ change by less than 2%. The magnitudes of the out-of-plane minima on C2 in all the fragments, which determine the positions of nucleophilic attack in the cleavage of the C-N bond, change by less than 0.009 au.

From the above comparisons of the geometrical, bond, atomic and Laplacian properties of five fragments in fourteen different chemical environments, we see that the properties of amino acid fragments do not exhibit significant changes upon transfer between different molecular environment. In other words, amino acid fragments can be regarded as transferable among different complexes; this is the basis of our theoretical synthesis.

Table 2.12. Comparison of Properties at Critical Points in $-\nabla^2\rho$ for Glycine Fragments in Different Complexes.

G Fragment[a]

Posi.	Property	G|NH2	G|G′	G|A′	G|G″|G′
	Type	(3,+1)	(3,+1)	(3,+1)	(3,+1)
C2	R	1.053	1.051	1.052	1.051
	$\nabla^2\rho$	0.101	0.093	0.095	0.092
	Type	(3,-3)	(3,-3)	(3,-3)	(3,-3)
N3	R	0.747	0.747	0.747	0.747
	$\nabla^2\rho$	-3.244	-3.243	-3.243	-3.243
O4	Type	(3,-3)	(3,-3)	(3,-3)	(3,-3)
anti	R	0.648	0.649	0.649	0.649
to N3	$\nabla^2\rho$	-6.196	-6.151	-6.121	-6.138
O4	Type	(3,-3)	(3,-3)	(3,-3)	(3,-3)
syn	R_{O4}	0.648	0.648	0.648	0.648
to N3	$\nabla^2\rho$	-6.212	-6.212	-6.202	-6.201

G′ Fragment[b]

Posi.	Property	COH|G′	G|G′	A|G′	G|G″|G′
	Type	(3,+1)	(3,+1)	(3,+1)	(3,+1)
C2	R	1.050	1.050	1.050	1.049
	$\nabla^2\rho$	0.115	0.115	0.115	0.115
	Type	(3,-3)	(3,-3)	(3,-3)	(3,-3)
N3	R	0.768	0.766	0.767	0.767
	$\nabla^2\rho$	-2.137	-2.181	-2.153	-2.152
O4	Type	(3,-3)	(3,-3)	(3,-3)	(3,-3)
syn	R	0.649	0.649	0.649	0.648
to O5	$\nabla^2\rho$	-6.146	-6.142	-6.142	-6.150
O4	Type	(3,-3)	(3,-1)	(3,-3)	(3,-3)
anti	R	0.649	0.649	0.649	0.649
to O5	$\nabla^2\rho$	-6.125	-6.122	-6.122	-6.119
	Type	(3,-3)	(3,-3)	(3,-3)	(3,-3)
O5	R	0.650	0.650	0.650	0.650
	$\nabla^2\rho$	-5.973	-5.972	-5.972	-5.973

Table 2.12. (con'd)

G" Fragment[c]

Posi.	Property	COH|G"	NH2 G|G"	G'
	Type	(3,+1)	(3,+1)	
C2	R	1.052	1.050	
	$\nabla^2\rho$	0.105	0.096	
	Type	(3,-3)	(3,-3)	
N3	R	0.768	0.767	
	$\nabla^2\rho$	-2.112	-2.160	
O4	Type	(3,-3)	(3,-3)	
syn	R	0.650	0.650	
to N3	$\nabla^2\rho$	-6.072	-6.068	
O4	Type	(3,-3)	(3,-3)	
anti	R	0.649	0.649	
to N3	$\nabla^2\rho$	-6.170	-6.124	

[a]. G ≡ $NH_2CH_2CO|$
[b]. G' ≡ $|NHCH_2COOH$
[c]. G" ≡ $|NHCH_2CO|$

Table 2.13. Comparison of Properties at Critical Points in $-\nabla^2\rho$ for Alanine Fragments in Different Complexes.

G Fragment[a]

Posi.	Property	A|NH2	A|G´
	Type	(3,+1)	(3,+1)
C2	R	1.054	1.052
	$\nabla^2\rho$	0.105	0.097
	Type	(3,-3)	(3,-3)
N3	R	0.748	0.748
	$\nabla^2\rho$	-3.245	-3.231
O4	Type	(3,-3)	(3,-3)
syn	R	0.649	0.649
to N3	$\nabla^2\rho$	-6.167	-6.163
O4	Type	(3,-3)	(3,-3)
anti	R	0.649	0.649
to N3	$\nabla^2\rho$	-6.168	-6.118

A´ Fragment[b]

Posi.	Property	COH|A´	G|A´
	Type	(3,+1)	(3,+1)
C2	R	1.051	1.051
	$\nabla^2\rho$	0.116	0.117
	Type	(3,-3)	(3,-3)
N3	R	0.768	0.767
	$\nabla^2\rho$	-2.123	-2.170
O4	Type	(3,-3)	(3,-3)
syn	R	0.649	0.649
to O5	$\nabla^2\rho$	-6.145	-6.142
O4	Type	(3,-3)	(3,-3)
anti	R	0.649	0.649
to O5	$\nabla^2\rho$	-6.129	-6.125
	Type	(3,-3)	(3,-3)
O5	R	0.650	0.650
	$\nabla^2\rho$	-5.969	-5.968

a. A ≡ $NH_2CHCH_3CO|$

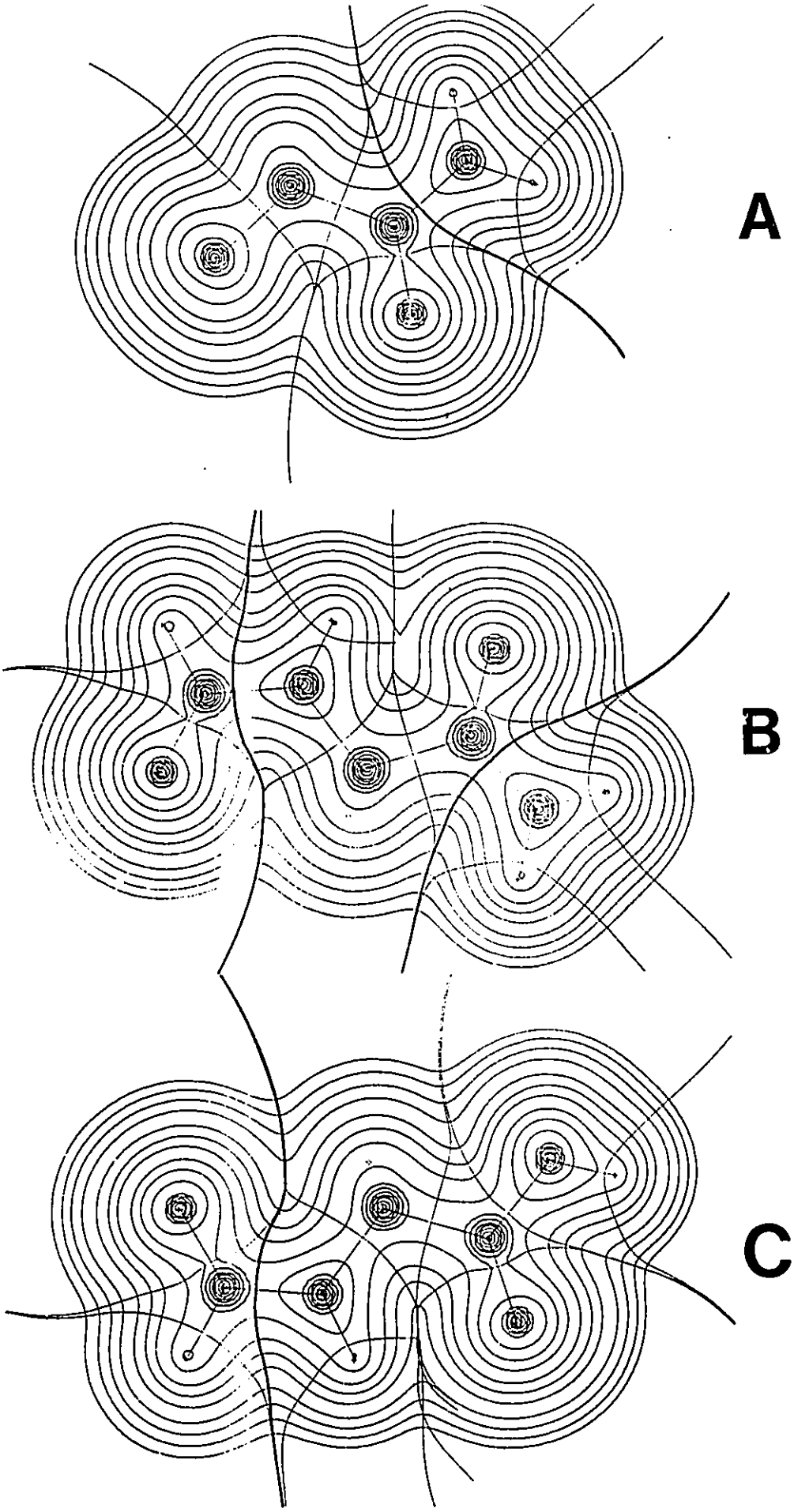
b. A´ ≡ $|NHCHCH_3COOH$

2.3 The Synthesis of Polypeptides From Amino Acid Fragments

The result of section 2.2 shows that if we can create a set of standard amino acid fragments, we may use them to build any size of polypeptide without running the risk of incurring significant change in their properties. We can create the standard fragment, for example the fragment G, by cutting it off from either the simplest model molecule G|NH_2, or a larger molecule like G|G′. An interior repeating residue can be obtained by sandwiching it between the two simplest end residues G and G′ as exemplified by G" in G|G"|G′. The comparisons in the above section have demonstrated that the fragments in simpler molecules exhibit the same properites as those in polypeptides, and we can create the standard fragment by defining it in the simplest model molecule. One can employ relatively large basis sets for the calculation of the model molecules. For example, Figure 2.6 gives the contour maps of the three model molecules which contain glycine fragments respectively. We obtain three different forms of the glycine fragment by cutting them off along the interatomic surface at the peptide bond. Thus we have three standard glycine building blocks and can use them to build arbitrarily long polypeptides containing only glycines, for example Gly-Gly-Gly. Adding the alanine fragments obtained in the same way from model molecules A|NH_2 and COH|A′, we could also build polypeptides containing both glycine and alanine, and so on. All the values of properties in the standard blocks will persist in the larger molecules. It has to be emphasized that our work is only based on the primary structure of molecules, the secondary structure of a polypeptide is to be considered in future studies.

Figure 2.6.

Contour plots of ρ for the model molecules containing A) G|, B) |G"| and C) |G´ fragments, respectively. Zero-flux surfaces and bond paths are also shown. The bold line marks the surface which separate the glycine fragment from the remainder of the molecules, G|NH2 (A), G|G"|G´ (B) and HCO|G´ (C).

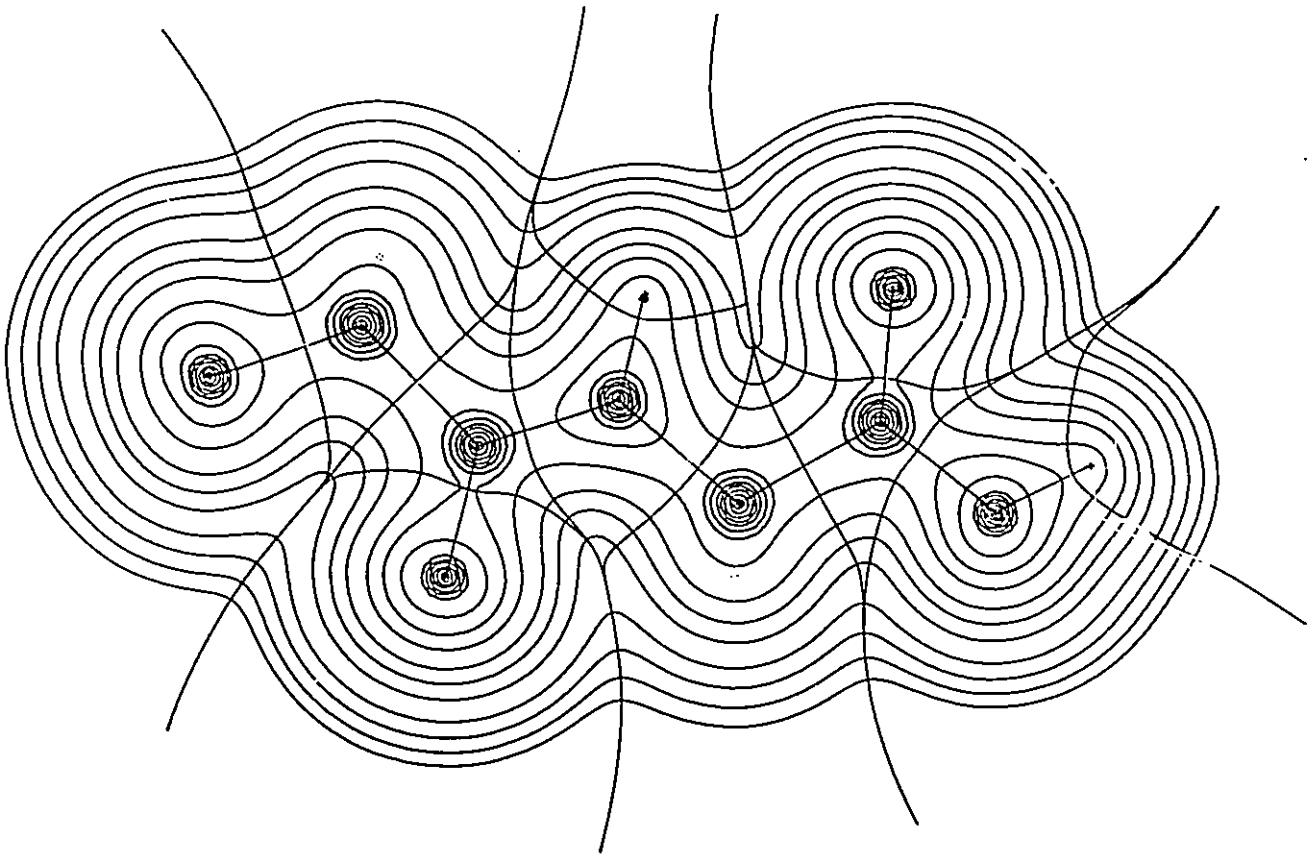

A
B
C

As described in Chapter 1, a bond path is defined by the pair of vectors of $\nabla\rho$ which originate at a bond critical point and terminate at the neighboring nuclei. To link two fragments to creat a polypeptide by forming an amide bond one needs to align, in an antiparalell manner, the gradient vector which is directed at C from the C-N bond critical point of one fragment with the one directed at N from the C-N bond critical point of the other fragment (see Fig.2.7). In this way we can synthesize Gly-Gly from the standard blocks G| and |G', (see Fig.2.7), Gly-Ala from G| and |A', Ala-Gly from A| and |G' and Gly-Gly-Gly from G|, |G"|, and |G'. This construction yields a molecular graph that is superimposable on the calculated one, correctly predicting that the geometrical parameters remain almost unchanged. It has been discussed in Section 2.2 that the errors, which are brought in by using the standard blocks, are less than 0.004 A for bond length, less than 0.5^{o} for bond angles and less than 1.5^{o} for dihedral angles. The predicted length of the C-N peptide bond is given by the sum of the corresponding bonded radii of C and N in the model molecules given in Table 2.4 - 2.5. The C-N peptide bond lengths obtained in this manner are listed in Table 2.14 for synthesized Gly-Gly, Gly-Ala, Aly-Gly and Gly-Gly-Gly and compared with the calculated values of the "real" molecules. The predicted peptide bond lengths differ from the calculated values by less than 0.002 A for four out of five peptide bonds, with the largest difference being 0.005 A at one of the C-N bonds in Gly-Gly-Gly, which is within the range we predicted in the last section for geometrical changes.

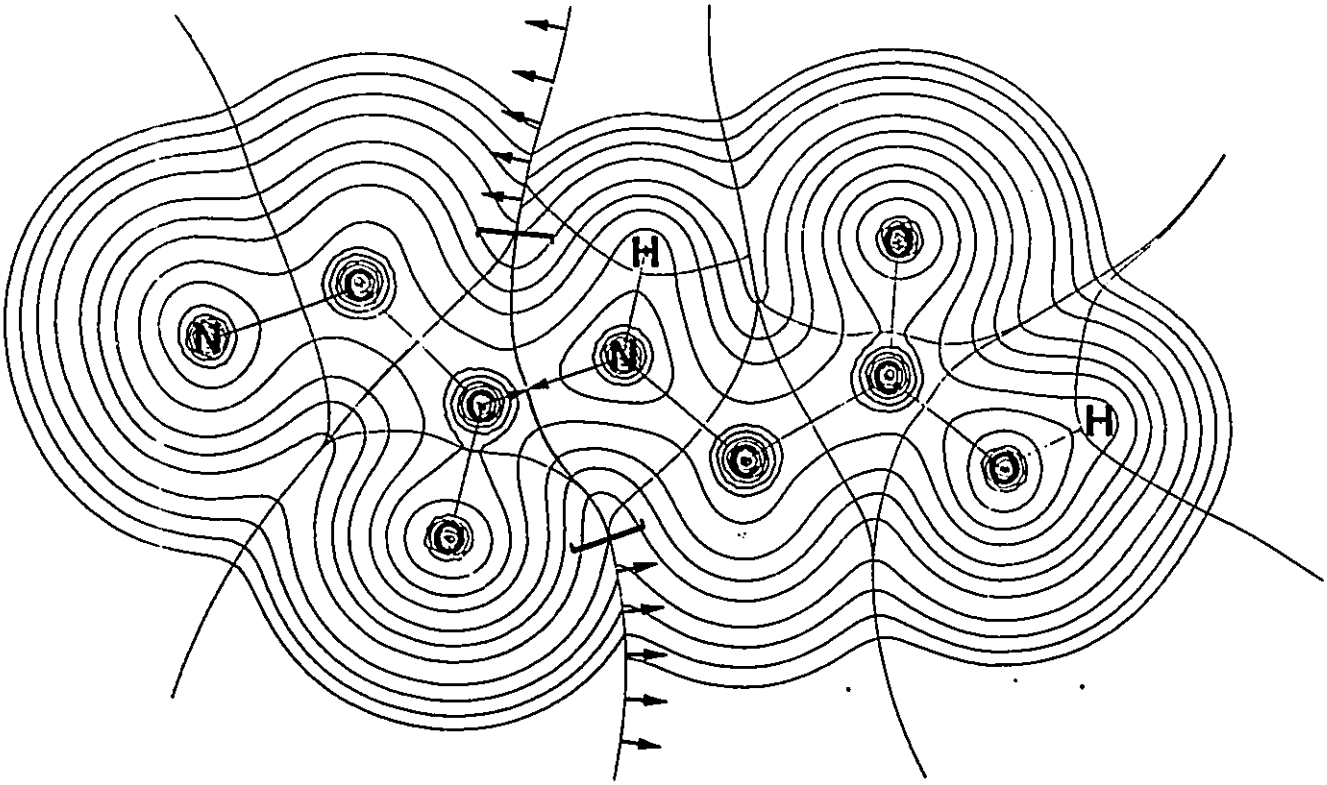
Fig.2.8 displays the contour maps of synthesized Gly-Gly-Gly together with the calculated one for comparison. The contours of the

Figure 2.7.

Contour plots of ρ for the A) calculated and B) synthesized Gly-Gly molecule. The zero-flux surfaces and bond paths are shown. In B), the arrows at the peptide bond critical point denote the linking of the negatives of the two bond path vector which originate at the C-N bond critical point. The interatomic surfaces of the two fragment, G and G′, merge at the range in the brackets, while the arrows point the fragment which the interatomic surface is contributed from.



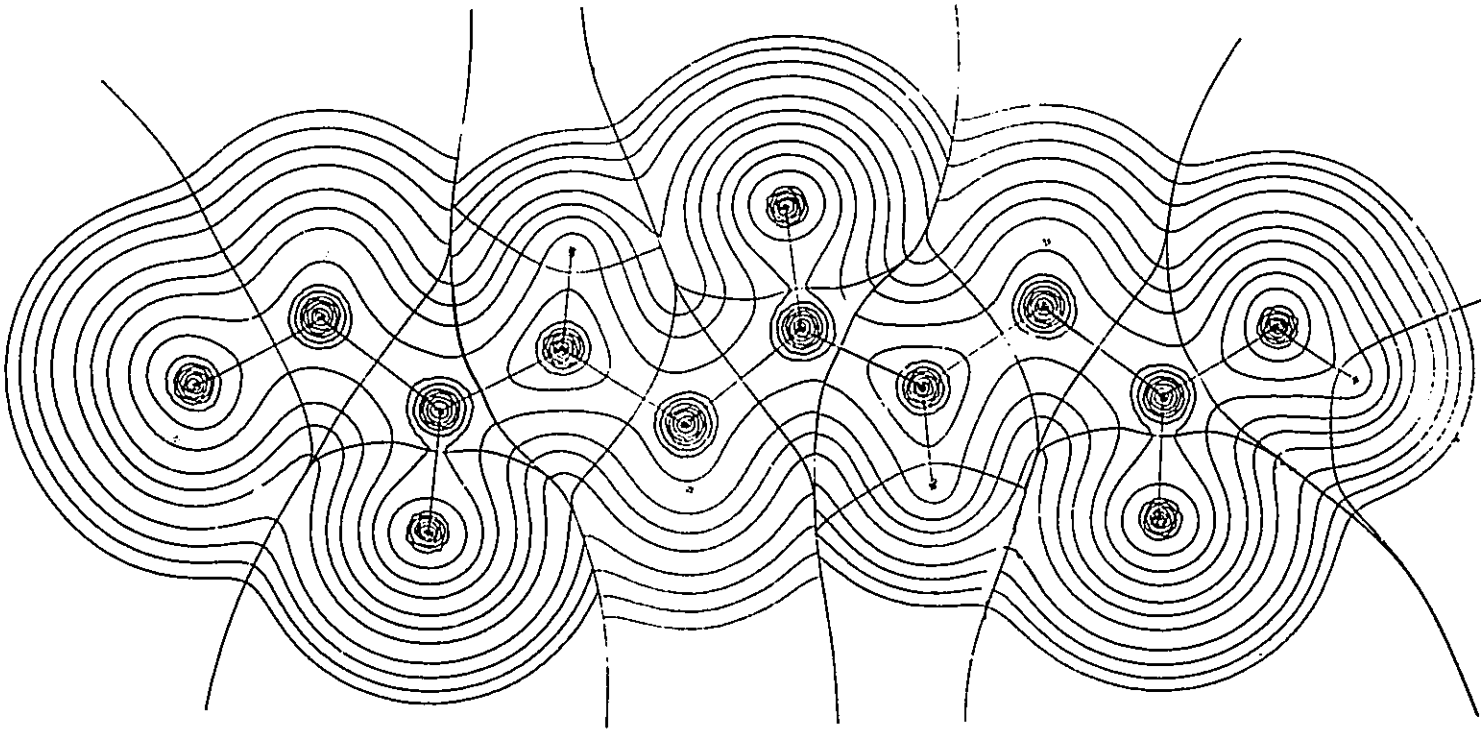
A



B

Figure 2.8.

Contour plots of ρ for the synthesized (top) and calculated (bottom) Gly-Gly-Gly molecule. The zero-flux surfaces and bond paths are also indicated. The synthesized molecule is combined by linking the three different glycine fragments in the manner indicated in Fig.2.7 for Gly-Gly.



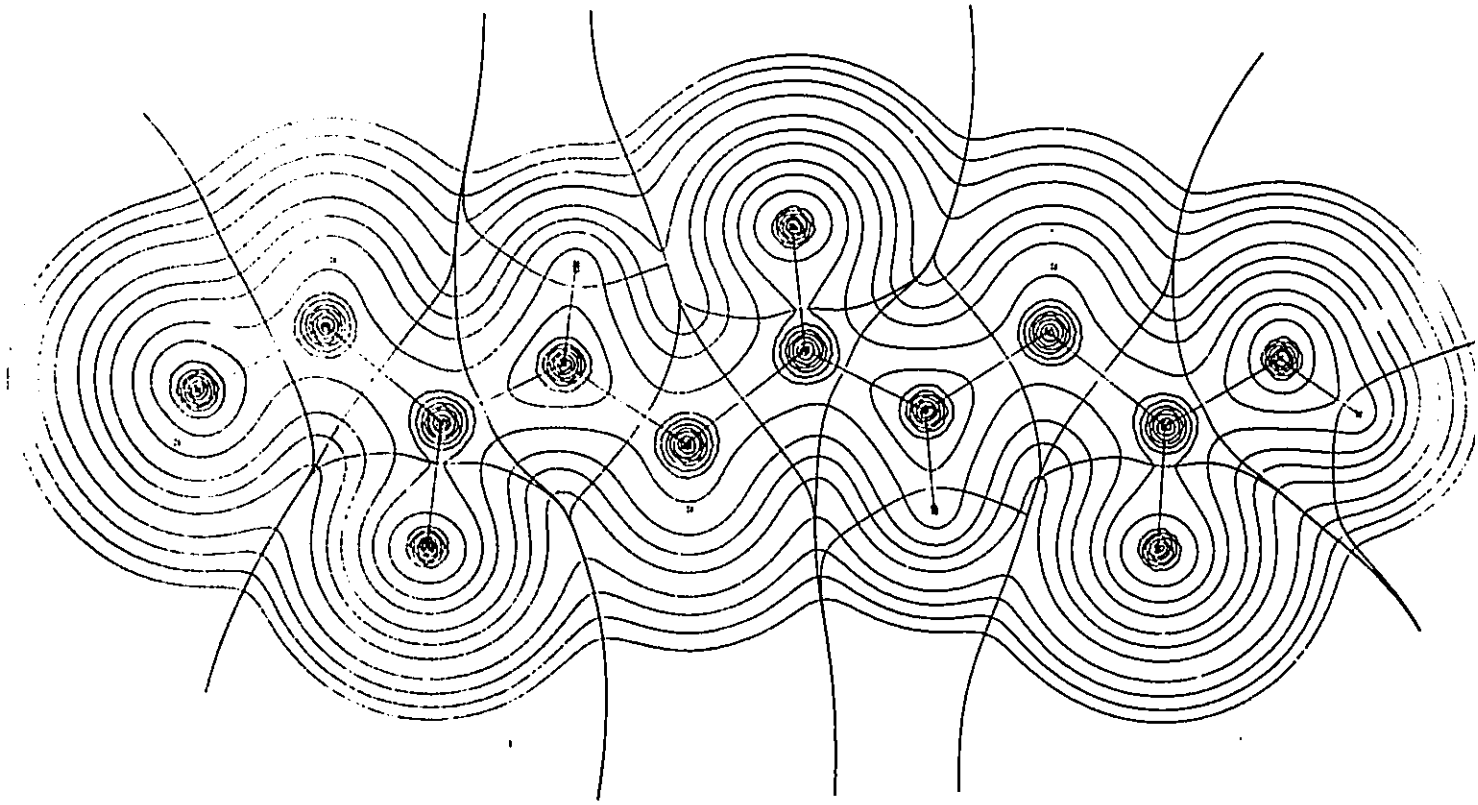


Table 2.14. Comparison of Total Populations and Energies and Peptide Bond Lengths of Synthesized Molecules with the Calculated Ones.

Peptide Bond Lengths (in Å)

	G|G′	G|A′	A|G′	G|G"	G"|G′
Synth.	1.347	1.346	1.346	1.345	1.346
Calc.	1.348	1.348	1.347	1.345	1.341
Diff.	0.001	0.002	0.001	0.000	0.005

Total Populations (in au)

	G|G′	G|A′	A|G′	G|G"|G′
Synth.	70.025	78.026	78.026	100.040
Calc.	70.000	78.000	78.000	100.000
Diff.	0.025	0.026	0.026	0.040

Total Energies (in au)

	G|G′	G|A′	A|G′	G|G"|G′
Synth.	489.2008	528.2177	528.2124	695.8297
Calc(SCF)	489.1934	528.1987	528.1983	695.8204
Diff.	0.0074	0.0190	0.0141	0.0093

adjoining fragments match each other down to the contour line with the value of ρ = 0.02 au. The two interatomic surfaces also merge out to this contour. If one overlaps the contour map of the synthesized Gly-Gly-Gly with the calculated one, one sees that the two maps coincide very well across the whole molecular space except for a small outer regions at the synthesized surface. The interatomic surfaces of C2-C1 in G| and in |G"|, and of N3-C1 in |G"| and in |G', all merge with the C-N surface being retained, as C1 atoms are the least perturbed of the atoms bordering the surface. The C2-O4 surfaces in G| and |G"| fragments and N3-H surfaces in |G"| and |G' fragments also merge with the C-N surface in the region of higher density and it is the outer contours of O4 and H(N) which do not match at the surface due to the disturbance by the replacement of the R group in the real molecule with H in the model molecules. This mismatching of contours is correctable by linking them to the corresponding contours of the C1 atoms. As demonstrated above, mis-matching contours of value less than 0.02 au does not lead to large errors and even the atomic properties of H(N) and O4 are well approximated (see Tables 2.8 - 2.11).

When we use the standard building blocks to build a polypeptide, we only need the data in the first column of Tables 2.2 - 2.13. For example, to build the Gly-Ala dipeptide, we need the data in the first column of Tables which contain G| and |A', respectively. Thus all the errors brought in by using the standard building blocks will be within the range which one has discussed in Section 2.2.

Among the molecular and atomic properties, a most important one is the Laplacian of charge density, because it provides the necessary information for studying chemical reactions. From Tables 2.12 - 2.13, we

can see that not only the position but also the relative ordering of the magnitudes of the critical points is maintained in the synthesized Laplacian distribution . The magnitudes of $\nabla^2\rho$, all in atomic units, for the in-plane nonbonded maxima on O4 in G| of synthesized G|G"|G´ molecule, for example, are 6.212 (syn to NH_2) and 6.196, with corresponding values in the calculated one of 6.201 and 6.138. The magnitudes of the in-plane nonbonded maxima on O4 in |G´ are 6.146 and 6.125 with the former value for the maximum syn to the OH group. In the calculated molecule these values are 6.150 and 6.119, respectively. The nonbonded maxima on O4 are 6.170 and 6.072 in |G"| and 6.124 and 6.068 in the calculated molecule with the latter syn to the NH group. This is the very type of information, information that is so readily read from a Laplacian map, that is of direct use in predicting the relative reactivities at active sites. Thus, Gandour (1981) has observed that carboxylates found at the active sites of enzymes generally employ the more basic syn lone pairs (the syn nonbonded maxima) rather than the less basic anti lone pairs. Rebek et al (1986) have synthesized model systems possessing structures that have enabled experimental studies to be made of the relative basicities of the in-plane syn and anti nonbonded charge concentrations on a carboxyl oxygen in an attempt to learn more about the role of these factors in determining stereoelectronic effects at a carboxylate group in an active site. The availability of Laplacian maps for carboxylate ions and for fragments of the model systems used in these experiments would provide a useful complement to the experimental study of these effects.

The construction of the reactive envelopes of the amino acid fragments and of other biologically important groupings of atoms is

possible and will greatly enhance our ability to predict the specific functional nature of a substrate-enzyme interaction and the relative alignment of the reactants. The accuracy in shape and size as provided by the charge density determined in the present method of synthesis will lead to improved predictions and interpretations when used in the presently employed interpretive models. The reactive surfaces as defined by the Laplacian of the electronic charge distribution provide a new and quantitative tool for the study and prediction of the processes involved in molecular recognition.

The totals of the properties obtained by the summation of the atomic properties in the synthesized molecules is a strict criterion for the validity of the synthesis. Table 2.14 lists a comparison of total populations and energies for the synthesized and calculated molecules. The differences in the total populations and energies between the calculated and synthesized molecules are acceptably small. In fact the differences are in general, less than the errors made in the summation of the integrated populations and energies for the calculated molecules. For example, the difference between the total populations of the synthesized and calculated Gly-Gly molecules is 0.026e compared to an error of 0.050e in the total integrated population of the calculated molecules. Further examples of the difference between the total populations - calculated and synthesized - and of the error in the integrated population of the calculated molecule, using the data in Table 2.7 and 2.14, are respectively: 0.026e and 0.062e for Gly-Ala, 0.027e and 0.040e for Ala-Gly, and 0.041e and 0.080e for Gly-Gly-Gly. The population errors are small, both absolutely and relatively, compared to the total electron population

which range from 70e in a dipeptide to 100e in a tripeptide. The summed total energies of the synthesized molecules are also close to the SCF energies of the calculated molecules, the smallest difference being only 0.0074 au (4.6 kcal/mol) in G|G′ out of a total energy of 489.1934 au (306,974 kcal/mol) and the largest one being 0.0190 au (11.9 kcal/mol) in G|A′ out of a total energy of 528.1987 au (331,450 kcal/mol). For G|G"|G′, the error is 0.0093 au (5.8 kcal/mol) out of a total energy of 695.8204 au (436,634 kcal/mol).

2.4 Conclusion

The predicted properties of the synthesized molecules, obtained by summing the atomic contributions of their fragments, are in acceptable agreement with the corresponding properties of the calculated molecules. This fact, coupled with the above demonstration of the successful prediction of properties of $\nabla^2\rho$ and bond critical point in the synthesis of polypeptides from the atoms of theory, shows that the static and reactive properties of a large molecule that are important in the process of molecular recognition can be predicted from those of its constituent atoms as they appear in smaller systems.

It is to be emphasized that this method of synthesis is made possible by and can be carried out only through the use of the theory of atoms in molecules. The use of a surface other than the zero flux surface for the definition of a functional grouping of atoms in real space will reduce the transferability of chemical information between systems. Equally important, only the atoms of theory have properties which are defined by quantum mechanics and which sum to yield the total properties of

the system. The method also allows one to calculate just that portion of a larger system that is of interest, by replicating the fragment of interest up to its surfaces of zero flux. The surfaces contain all of the information regarding the interaction of the fragment with the remainder of the system. This possibility is of paramount importance in the study of the binding of a substrate to an active site of an enzyme.

The proposed use of the theory of atoms in molecules in the synthesis of large molecules is a new application of the theory. While the use of the Laplacian to predict the course of a generalized Lewis acid-base reaction has been in the process of development for a number of years, its specific application to biological problems through the definition and use of the reactive envelope is new.

Secondary structure resulting from rotation about the C1-N3 and C1-C2 bonds will initially be dealt with in an empirical manner, using the experimental values for the corresponding angles ϕ and ψ. At each stage of the development of this model, however, every attempt will be made to predict secondary structure using the reactive surfaces and van der Waals envelopes to describe the interactions between the residues.

The atomic properties, van der Waals envelope, the Laplacian distribution and the reactive envelope will be determined for each fragment. These fragments, and fragments of substrate molecules, will be used in the modelling of biologically important interactions in real and in model systems. A good example of the latter is the use of the Laplacian distribution function to complement the model studies on the relative basicities of the lone pairs on the carboxylate oxygen atoms (Rebek et al 1986) and to then aid in the extension of these studies to the properties

of the active sites found in an enzyme. Ultimately, the goal is to develop a catalogue of information listing the characteristic properties of each amino acid group that will be made available to workers in the field of molecular recognition.

Appendix 2.1

Crystal Structure Data of Peptides (Zwitterion) Containing Glycine and Alanine Fragments.

Glycine (Marsh 1958)

C1-C2	1.523	C1-N3	1.474	N3C1C2	111.8	O4C2C1	117.4
C2-O4	1.252	C2-O5	1.255	O5C2C1	117.1		

Alanine (Simpson and Marsh 1966)

C1-C2	1.525	C1-N3	1.291	N3C1C2	110.3	O4C2C1	118.3
C2-O4	1.247	C2-O5	1.256	O5C2C1	116.1	CC1C2	111.5
C-C1	1.525						

G\|G′ (Biswas et al 1968)

C1-C2	1.528	N3C1C2	110.3	C1′-C2′	1.516	N3′C1′C2′	112.7
C1-N3	1.497	O4C2C1	120.1	C1′-N3′	1.462	O4′C2′C1′	117.6
C2-O4	1.249	N3′C2C1	116.8	C2′-O4′	1.239	O5′C2′C1′	115.6
C2-N3′	1.328			C2′-O5′	1.262	C2N3′C1′	121.6

G\|A′ (Wang and Paul 1979)

C1-C2	1.527	N3C1C2	108.3	C1′-C2′	1.538	N3′C1′C2′	114.2
C1-N3	1.475	O4C2C1	121.2	C1′-N3′	1.456	O4′C2′C1′	115.9
C2-O4	1.223	N3′C2C1	114.7	C2′-O4′	1.249	O5′C2′C1′	119.9
C2-N3′	1.339			C2′-O5′	1.256	C2N3′C1′	120.6
				C′-C1′	1.524	C′C1′C2′	108.3

A\|G′ (Michel 1970)

C1-C2	1.524	N3C1C2	108.4	C1′-C2′	1.522	N3′C1′C2′	114.2
C1-N3	1.489	O4C2C1	121.6	C1′-N3′	1.443	O4′C2′C1′	119.9
C2-O4	1.215	CC1C2	111.0	C2′-O4′	1.240	O5′C2′C1′	115.8
C-C1	1.514	N3′C2C1	114.3	C2′-O5′	1.263	C2N3′C1′	120.2
C2-N3′	1.334						

G\|G"\|G′ (Srikrishnan et al 1982)

C1-C2	1.508	N3C1C2	109.9	C1′-C2′	1.520	N3′C1′C2′	110.7
C1-N3	1.482	O4C2C1	120.8	C1′-N3′	1.451	O4′C2′C1′	120.3
C2-O4	1.229	N3′C2C1	116.1	C2′-O4′	1.235	O5′C2′C1′	115.6
C2-N3"	1.328	N3"C1"C2"	111.3	C2′-O5′	1.277	C2"N3′C1′	123.1
C1"-C2"	1.510	O4"C2"C1"	122.9	C2"-O4"	1.228	C2N3"C1"	119.4
C1"-N3"	1.447	N3′C2"C1"	113.6	C2"-N3′	1.333		

Chapter 3 Electrophilic Aromatic Substitution

It is the purpose of this Chapter to use the theory of atoms in molecules to explore in a quantitative manner the effects of substituents on the properties and reactivity of the phenyl group. Section 3.1 gives the history and background of this project; Section 3.2 explains the method of calculation; Section 3.3 discusses the molecular structures and bond properties of monosubstituted benzenes; Section 3.4 details atomic properties; and Section 3.5 predicts and explains the susceptibility of substituted benzenes to electrophilic attack. This Chapter is concerned with how the ortho-, meta-, and para-directing abilities of the substituents X and their propensity for activating or deactivating the phenyl group in electrophilic substitution reactions are related to the changes they cause in the charge distribution of the phenyl group for the reactant Ph-X. Chapter 4 relates these changes and those for a number of arenium ion intermediates $(Ph\text{-}X)H^+$ to the energetics of the substituent effect.

3.1 The History and Background of the Study of the Electrophilic Aromatic Substitution

The interpretation of substituent effects on the course of electrophilic substitution reactions in aromatic systems played an important role in the development of models of electron displacement. This history and that of the general development of the electronic models of the inductive and mesomeric (resonance) effects and their kinetic counterparts are described by Ingold (1969). Central to these developments is the concept of assigning particular electronic properties to atoms and

functional groupings of atoms. This leads to classification schemes based on the ability of a given group or substituent to respond to the electrical requirements of a particular reaction as reflected in its effect on the equilibrium or rate constant of the reaction. This idea was quantified by Hammett (1937, 1940) with the introduction of the substituent constant σ in the construction of a linear free energy relationship for the correlation of both equilibrium and rate constant for substituted benzenes. The theory of atoms in molecules enables one to both predict and understand the properties that characterize a functional group.

Even the earliest of models recognized the necessity of distinguishing between polar or inductive and resonance or mesomeric modes of electron displacement. This led Taft to expand on Hammett's linear free energy relationship through the definition of separate substituent constant for the inductive and resonance effects (Taft 1956 and 1960), an idea that has been further generalized by Swain and Lupton (1968). The transcription of these terms into the language of molecular orbital theory identifies the inductive or polar effect and the mesomeric or resonance effect with the transfer of charge density derived from σ and π orbitals, respectively (See for example, Dewar 1949). Thus within molecular orbital theory, the fluorine atom is classified as a σ electron acceptor and a π electron donor corresponding to the Ingold classification of -I and +M. It is shown here that properties ascribed separately to the σ and π sets of orbitals can be recovered in the properties of the total charge density. The use of the charge density is not subject to the limitations of the model in those cases where, because of symmetry, the separation into σ and π sets is not possible.

The properties of the phenyl group Ph, as perturbed by the substituents X=O^-, NH_2, OH, F, CH_3, H, CN, NO_2 and NH_3^+, are studied here. They fall into three groups when gauged relative to H: O-, NH_2, OH, and F are σ acceptors and π donors, CH_3 is both a σ and π donator, while CN, NO_2, and NH_3^+ are both σ and π acceptors. The effects of most of these same substituents on the charge distribution of the carbonyl group in the system HXC=O have been previously reported (Slee et al 1988) where they exhibited the same donor-acceptor properties. Grier and Streitwieser (1982) have also studied the effects of substituents and the integration of the σ and π contributions to these maps over selected regions of space.

The theoretically determined atomic and bond properties are correlated with experiment through Taft's empirical resonance parameter σ_R^o (Taft et al 1963) by using values listed by Exner (1978). This parameter is proportional to the difference in the FNMR chemical shift for meta- and para-substituted fluorobenzenes and thus equals zero for X=H. The ordering of the substituent effects on the atomic and bond properties of phenyl, when similarly referenced, is found to parallel that predicted by the σ_R^o values and the ordering of the substituents in the tables reflects this correspondence.

A knowledge of the electronic charge distribution was the goal of Hammett who in 1940 stated "-σ measures a change in electron density produced by a substituent, the quantity ρ (here it is not charge density-- author) measures the susceptibility of the reaction in question to change in electron density". While many investigations are described as being studies of the electron density, this is in fact not the case. They are instead based upon Mulliken charges, or values of coefficients of basis

functions or overlap integrals. These quantities are not related to nor derived from the density of electronic charge at a point in physical space, which is the definition of electron density. All properties of an atom in a molecule are determined by a direct integration of a corresponding density over a well-defined region of space, the basin of the atom, and for its electron population and multiple moments, the density of interest is the charge density $\rho(r)$.

While the topology of the charge density determines a system's molecular structure, its reactivity is mirrored in the topology of its Laplacian distribution. The Laplacian of the charge density, the quantity $\nabla^2\rho$, has the important property of determining where electronic charge is locally concentrated and locally depleted (see Chapter 1). The local maximum in the concentrations of charge defined in this manner recover the Lewis model of localized electron pairs and the associated VSEPR model of molecular geometry (Bader et al 1984 and 1988) while the course of a generalized Lewis acid-base reaction is predicted by aligning a maximum of charge concentration on the base with a minimum corresponding to a region of charge depletion on the acid (Bader and MacDougall 1985, Carroll et al 1988). It is shown here that the directing ability of a substituent and its overall activating or deactivating character for electrophilic substitution reactions are predicted by the Laplacian distribution of the phenyl group.

3.2. Calculations

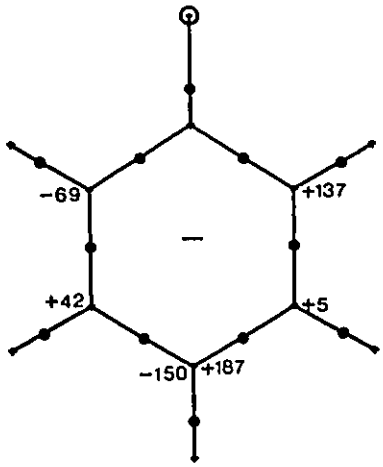
The ab initio calculations were performed on a VAX-8600 computer using the GAUSSIAN82 program (Binkley et al 1982) with the 6-31G** basis

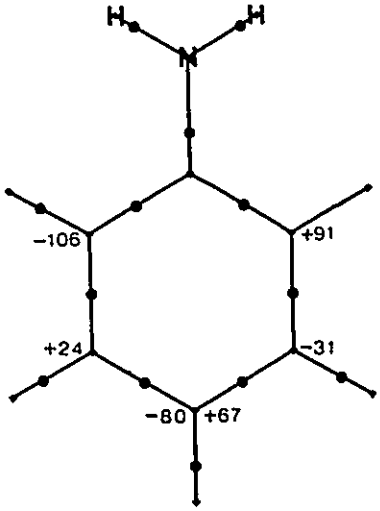
set (Hariharan and Pople 1973) at geometries previously determined by Bock et al. (1985) using the 6-31G basis set (Hehre et al 1972). These authors assumed that the substituent group NH_2, OH and NO_2 in aniline, phenol and nitrobenzene, are in the same plane as the ring, and one of hydrogen atoms of group CH_3 and NH_3^+ in $C_6H_5CH_3$ and $C_6H_5NH_3^+$ lies in the ring plane (see Figure 3.1). Experimentally, the N atom in aniline is pyramidal (Lister et al. 1974, Quack and Stockburger 1972) with the angle between the HNH plane and the ring equal to 40°, while the 6-31G basis predicts the planar form to be most stable. The addition of the polarizing functions reverses this prediction, with the pyramidal form predicted to be more stable by 1.2 kcal/mol. To be consistent, the results reported here are for the planar geometry. The optimised planar geometry, which is not reported by Bock et al. (1985) is given in Table 3.13, and Table 3.14 shows the atomic populations of the pyramidal aniline as a comparison. Aside from these restrictions, a full geometry optimization was carried out by these authors who also include a detailed comparison with experiment. The benzene ring is distorted by substitution, the principal change in ring angle occurring at the ipso carbon, the carbon bearing the substituent X, to a maximum value of 122.9° in Ph-F and to a minimum value of 114.5° in $Ph-O^-$. The remaining angles deviate on the average by less than a degree from 120°. The same two molecules exhibit the largest change in C-C bond length from its value of 1,3883 A in Ph-F to an increase by 0.042 A in $Ph-O^-$. The remaining bond lengths change, in general, by less than 0.006 A.

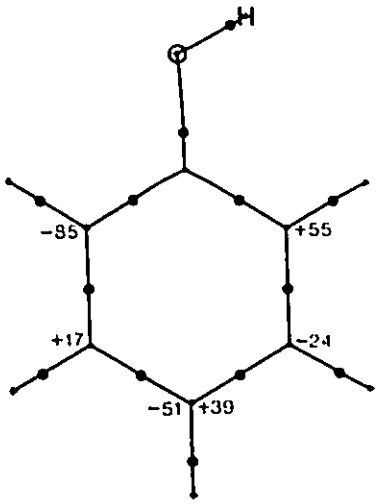
The analysis of the atomic energies, the atomic populations and the Laplacian of charge density was carried out using the AIMPAC series of programs developed in this laboratory (Biegler-Konig et al 1982). All the

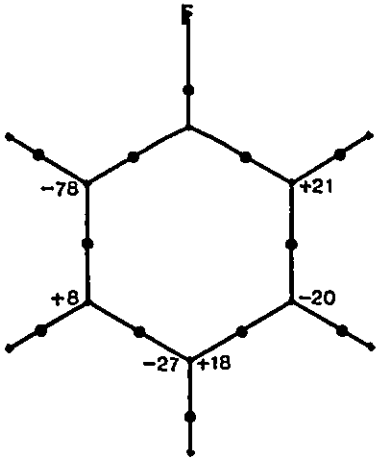
Figure 3.1

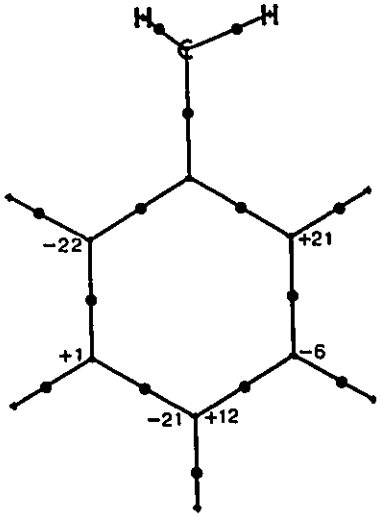
Structure of monosubstituted benzenes as depicted by molecular graphs predicted by theory. The bond critical points are denoted by dots. The numbers on the right-hand side of each structure are the carbon π populations relative to the value in benzene, and those on the left-hand side the corresponding σ populations all multiplied by 1000. An increase in population (+) signifies an increased negative charge at that carbon. In Ph-OH, Ph-CH_3^+ which do not possess a C_2 symmetry axis, the two ortho and the two meta populations are averaged.

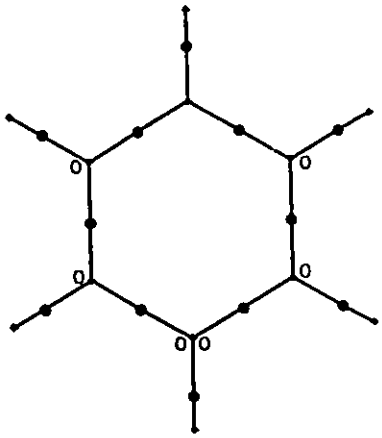

−69
+137
+42
+5
−150
+187

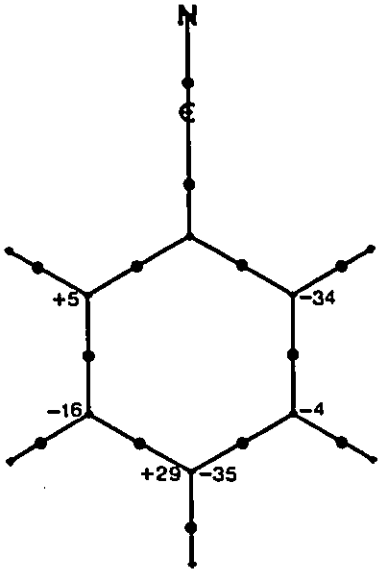

H
N
H
−106
+91
+24
−31
−80
+67

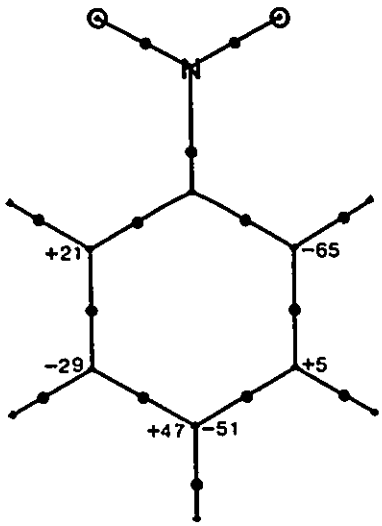

H
−85
+55
+17
−24
−51
+39

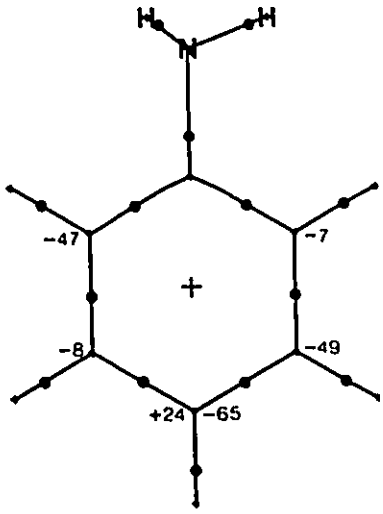

F
−78
+21
+8
−20
−27
+18


H
C
H
−22
+21
+1
−6
−21
+12


0
0
0
0
0
0


N
C
+5
−34
−16
−4
+29
−35


N
+21
−65
−29
+5
+47
−51


H
N
H
−47
−7
−8
−49
+24
−65

atomic integrations satisfy the criterion of $L(\Omega) < 5*10^{-3}$ au with the exception of the C_1 atom in nitrobenzene, which has $L(\Omega)=-7.45*10^{-3}$ au. It follows that the summation of atomic energies and atomic populations for eight out of the nine molecules calculated here agree well with the total properties of corresponding molecules, with the largest differences being 0.0006 au in energy, 0.005 electron in total electronic population and 0.008 electron in total π population (Table 3.4 and 3.5). The atomic energy and electronic population of the C_1 atom in nitrobenzene were obtained by subtraction of the corresponding sum for the remaining atoms from the total property.

3.3 Molecular Structure and Bond Properties of Monosubstituted Benzenes

Structures of studied monosubstituted benzenes, as predicted in terms of the properties of the charge distributions obtained from either theory or experiment, are displayed in Figure 3.1. This diagram, like those given previously for acyclic, cyclic and bicyclic hydrocarbons, including propellanes (Wiberg et al 1987) and for many other kinds of systems, shows that the numerous models used to predict the network of bonds in a molecule may be replaced by a single theory of structure.

The properties of a bond are summarized by the properties of the charge density at its bond critical point. The properties of the bonds in the hydrocarbons referred to above (Wiberg et al 1987) have been classified in this manner. The value of the charge density $\rho(r)$ at the bond critical point, denoted by ρ_b, determines a bond order n which yields values of 1.0, 1.6, 2.0, and 2.9 for ρ_b in the C-C bonds in ethane, benzene, ethylene, and acetylene. The change in the value of $\rho(r)$ for the C-C bonds from its

value in benzene, Table 3.1, parallel the small changes in bond length and in general represent small change in bond order, the largest being associated with the largest changes in bond length which, as noted above, occur for phenoxide ion and fluorobenzene, where n for the bonds to the ipso carbon decrease to 1.4 and increase to 1.7, respectively (an early study (Runtz et al 1977) reported the presence of very slightly inwardly curved bond paths in benzene based upon a density obtained from a minimal STO basis set. The present study, using a much larger basis with polarizing functions, finds no detectable curvature in the C-C bonds of benzene). Also listed in Table 3.1 is the value of $\rho(r)$ at the ring critical point of each of the phenyl groups. This is the minimum value attained by $\rho(r)$ in the ring surface.

It was shown in Chapter 1 that the Laplacian of ρ at a bond critical point, the quantity $\nabla^2\rho_b$, can be used to clarify the bond as being a shared or closed-shell interaction. The curvature of ρ is positive along the bond path at the bond critical point and charge is locally depleted there with respect to points on the path. Its two curvatures perpendicular to this path are negative, as ρ is a maximum at the bond critical point in the interatomic surface which intersects the bond path at this point, charge is locally concentrated at this point in the interatomic surface. The sum of the three curvatures of ρ at this point, the quantity $\nabla^2\rho_b$, is negative for the C-C and C-H bonds, Table 3.2, showing that these interactions are dominated by a contraction of electronic charge toward the internuclear axis. This results in an accumulation of charge between the nuclei which is shared by both nuclei and is the principal source of their bonding, as is typical of covalent interactions (Bader and Essen 1984).

Table 3.1. Values of β_b in Substituted Benzenes in aux10.[a]

Bond\X	O⁻	NH_2	OH	F	CH_3	H	CN	NO_2	NH_3^+
C1-C2	3.089	3.225	3.305	3.369	3.252	3.259	3.229	3.290	3.293
C2-C3	3.273	3.262	3.235	3.246	3.253	3.259	3.281	3.278	3.249
C3-C4	3.186	3.253	3.268	3.253	3.273	3.259	3.266	3.264	3.280
C4-C5	3.186	3.253	3.243	3.253	3.255	3.259	3.266	3.264	3.266
C5-C6	3.273	3.262	3.270	3.246	3.270	3.259	3.281	3.278	3.262
C6-C1	3.089	3.225	3.315	3.369	3.234	3.259	3.229	3.290	3.280
C2-H7	2.895	2.939	2.912	2.963	2.968	2.955	2.989	3.027	2.956
C3-H8	2.900	2.983	2.962	2.972	2.978	2.955	2.971	2.979	3.010
C4-H9	2.863	2.962	2.948	2.960	2.975	2.955	2.976	2.985	3.012
C5-H10	2.900	2.983	2.966	2.972	2.978	2.955	2.971	2.979	3.012
C6-H11	2.895	2.939	2.959	2.963	2.963	2.955	2.989	3.027	2.953
C1-X	3.640	3.094	2.757	2.331	2.633	2.955	2.889	2.708	2.288
Ring	0.193	0.200	0.201	0.202	0.203	0.204	0.202	0.205	0.203

[a]. C1, the ipso carbon, bears the substituent X.

Table 3.2. Values of $\nabla^2\rho_b$ in Substituted Benzenes in au.

Bond\X	O^-	NH_2	OH	F	CH_3	H	CN	NO_2	NH_3^+
C1-C2	-0.939	-1.004	-1.048	-1.106	-0.996	-1.007	-0.991	-1.039	-1.048
C2-C3	-1.004	-1.001	-0.989	-0.997	-1.004	-1.007	-1.021	-1.020	-1.011
C3-C4	-0.961	-1.003	-1.010	-1.002	-1.013	-1.007	-1.014	-1.013	-1.023
C4-C5	-0.961	-1.003	-0.998	-1.002	-1.006	-1.007	-1.014	-1.013	-1.017
C5-C6	-1.004	-1.001	-1.008	-0.997	-1.010	-1.007	-1.021	-1.020	-1.015
C6-C1	-0.939	-1.004	-1.072	-1.106	-0.989	-1.007	-0.991	-1.039	-1.042
C2-H7	-1.079	-1.119	-1.102	-1.156	-1.149	-1.143	-1.179	-1.228	-1.152
C3-H8	-1.095	-1.167	-1.152	-1.161	-1.161	-1.143	-1.159	-1.167	-1.206
C4-H9	-1.049	-1.142	-1.134	-1.147	-1.158	-1.143	-1.165	-1.175	-1.210
C5-H10	-1.095	-1.167	-1.155	-1.161	-1.162	-1.143	-1.159	-1.167	-1.208
C6-H11	-1.079	-1.119	-1.148	-1.156	-1.145	-1.143	-1.179	-1.228	-1.150
C1-X	-0.426	-0.805	-0.148	0.327	-0.720	-1.143	-0.954	-0.487	-0.271

The magnitude of these values parallel those for ρ_b, increasing as the bond is shortened, and more charge is accumulated between the nuclei. The variation in the values of ρ_b and $\nabla^2\rho_b$ of the C-C bonds, like the variation in the bond lengths which they parallel, exhibits no apparent relationship to the directing or activating-deactivating ability of the substituents.

The two negative curvatures of ρ at the bond point, λ_1 and λ_2, serve to define the ellipticity of a bond ϵ, a measure of its π character, as determined by the extent to which charge is preferentially accumulated in a given plane (Bader et al 1983), (for its definition, refer to Eq.[1.11]). In ethylene for example, ϵ=0.4 with the major axis directed in the π plane. The data in Table 3.3 indicate that, in general, the degree of π character of the C-C bond of the phenyl group can increase or decrease independent of geometrical changes. The bond to the ipso carbon are most affected, all exhibiting an increased ellipticity with the exception of phenoxide ion, for which the π character of the neighboring bonds is greatly increased. The alternating polarization of the π density about the phenyl ring is found to be greatest in phenoxide ion and its pattern of C-C bond ellipticities is consistent with a large weighing of the quinone-like resonance structure which localizes the negative charge at the para position, and indeed as noted below, the para position of the phenoxide ion has the largest of all π populations.

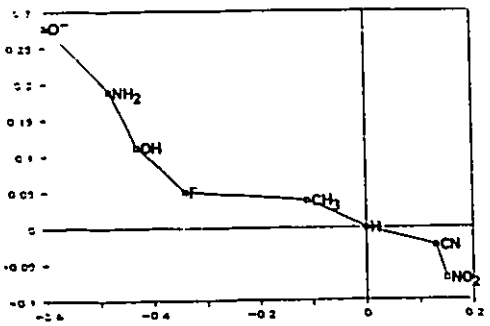
In general, the properties of the C-C bonds do not reflect the differing chemistries of the three groups of substituents, because electronic substitution occurs at a carbon atom and the properties of the individual atoms are blurred in the bond properties. The same is not true of the C-H bond ellipticities which, as illustrated in Figure 3.2, clearly

Table 3.3. Ellipticities at Bond Critical Points of Substituted Benzenes in aux10.

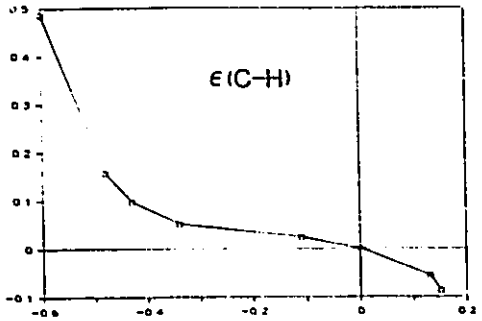
Bond\X	O⁻	NH_2	OH	F	CH_3	H	CN	NO_2	NH_3^+
C1-C2	1.937	2.530	2.868	2.922	2.413	2.308	2.384	2.441	2.940
C2-C3	3.071	2.582	2.410	2.338	2.297	2.308	2.304	2.249	2.176
C3-C4	2.670	2.409	2.440	2.337	2.366	2.308	2.201	2.175	2.110
C4-C5	2.670	2.409	2.306	2.337	2.271	2.308	2.201	2.175	2.033
C5-C6	3.071	2.582	2.472	2.338	2.403	2.308	2.304	2.249	2.249
C6-C1	1.937	2.530	2.676	2.922	2.279	2.308	2.384	2.441	2.847
C2-H7	0.438	0.350	0.293	0.211	0.203	0.161	0.136	0.091	0.159
C3-H8	0.064	0.086	0.117	0.133	0.143	0.161	0.173	0.194	0.134
C4-H9	0.647	0.318	0.259	0.213	0.185	0.161	0.105	0.074	0.093
C5-H10	0.064	0.086	0.109	0.133	0.142	0.161	0.173	0.194	0.127
C6-H11	0.438	0.350	0.252	0.211	0.194	0.161	0.136	0.091	0.168
C1-X	0.125	0.175	0.966	1.724	0.289	0.161	0.588	0.896	0.068

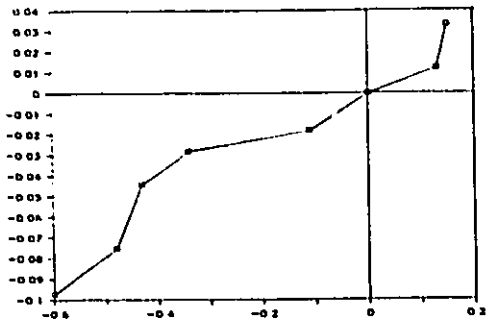
Figure 3.2

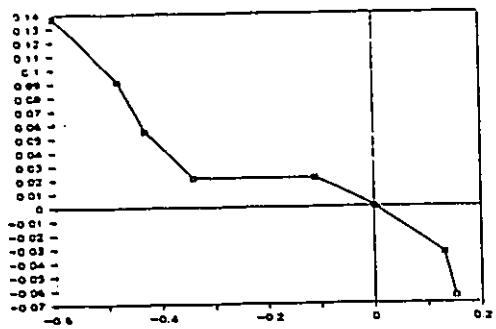
Bond and atomic properties (relative to their values in benzene) for the ortho, para, and meta positions of substituted benzenes plotted versus Taft's σ_R^o resonance parameter. The ordering of the substituents is the same in each diagram. The negative of the relative values in au is plotted for $Q_{zz}(C)$ and for the values of $\nabla^2\rho$ at the secondary charge concentration. For example, $Q_{zz}(C)$ and $\nabla^2\rho$ are negative relative to the values in benzene for the para position in phenoxide ion.

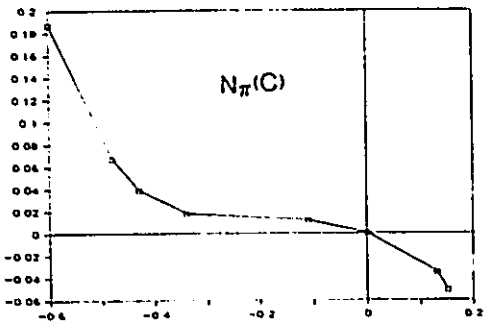


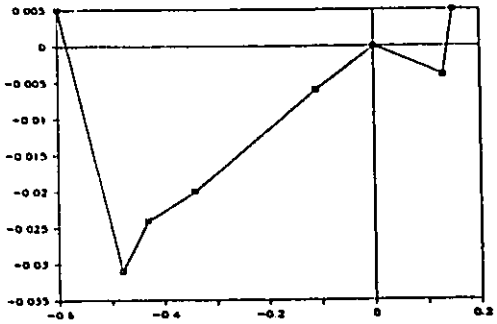
σ_R^o

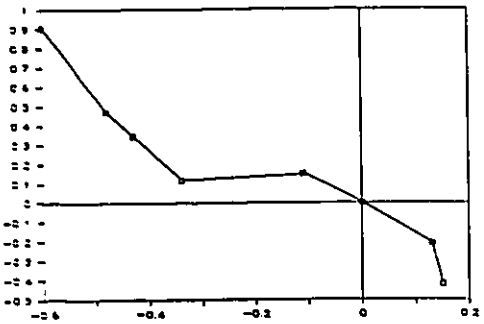


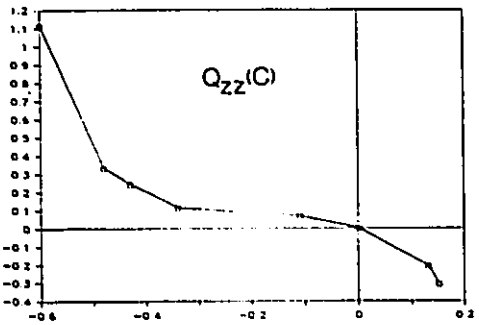


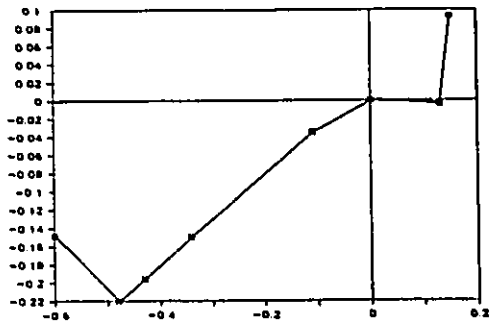


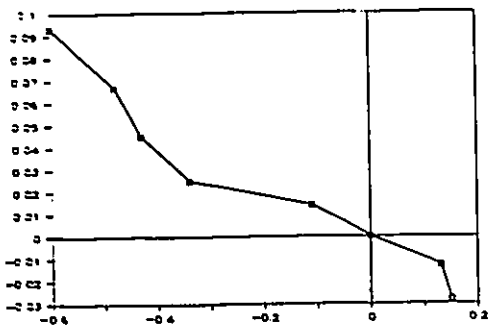


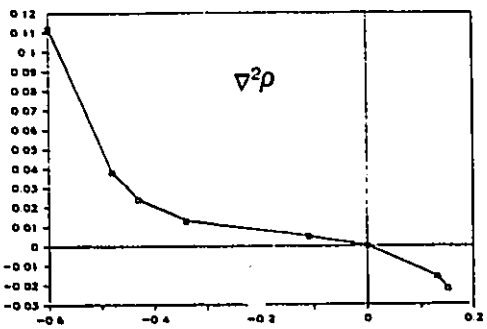


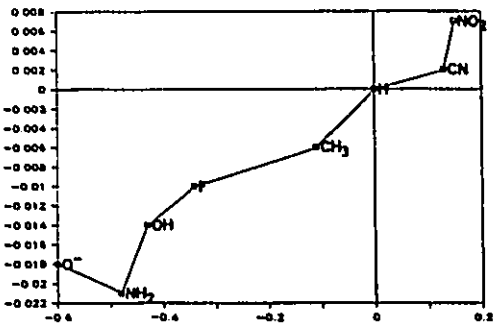












σ_R^o

ortho

para

meta

distinguish between the ortho- and para-directing abilities of the π donor substituents and the meta-directing ability of the π abstractors. (In general, the substituent property relative to its value in benzene is plotted versus the σ_R^o value, since the resonance parameter is itself based upon H as the standard.) The bond path to the proton serves as a probe of the atomic surface of an individual carbon atom. The relative C-H bond ellipticities show that the total charge density at the bond critical point in the atomic surface of ortho and para carbon atoms is increasingly polarized along an axis which is perpendicular to the aromatic ring in the order predicted by the σ_R^o values, starting from values less than that found in benzene for the -M or π-electron-withdrawing groups. Just the reverse behaviour is found for the meta carbon atom where +M or π donating groups cause a decrease in the extent of polarization of the density in the C-H surface perpendicular to the benzene ring. Exner (1978) has stated the opinion shared by others that "a single scale of mesomeric effects is an unrealizable objective". Thus one does not demand or anticipate linear correlations in the plots of the atomic and bond properties versus σ_R^o, but rather monotonically increasing or decreasing relations such as exhibited in Figure 3.2, plots which reflect the corresponding trends in the experimentally determined substituent effects on the course and rate of the substitution reaction.

3.4 Atomic properties of Monosubstituted benzenes

Atomic Populations

An atomic population N(Ω) is determined by the integration of the charge density over the basin of atom Ω, Eq.[1.15] and, since ρ is

expressed as a sum of orbital contributions, the σ and π atomic populations may be separately determined (Wiberg and Wendolowski 1981). The total atomic populations and the atomic π populations $N_\pi(\Omega)$ are given in Table 3.4 and 3.5. Also given in Table 3.4 is the net charge on the phenyl group, q(Ph).

It is well established that the atomic population of carbon relative to that of hydrogen increases along with an increase in the s character of its bonds to H (Bader et al 1987a, Wiberg et al 1987). Thus the net charges on H in ethane, benzene, ethylene and acetylene at the 6-31G**//6-31G* levels are -0.079, -0.045, -0.041, and +0.121e, respectively. As anticipated on the basis of this model the net charges on H in benzene and ethylene are nearly the same. The model applies not just to carbon, as evidenced by the populations on the nitrogen atom in planar (sp^2) and pyramidal (sp^3) aniline being 8.427e and 8.331e, respectively. The basis set is flexible enough to distinguish between σ and π populations on the H atoms, and of the 1.043e on H in benzene, 0.031e comes from orbitals of π symmetry.

All the substituents but methyl withdraw charge from phenyl relative to H. Since H is more electronegative than C in saturated unstrained hydrocarbons, the order of electron-withdrawing abilities is H > CH_3 > CH_2 > CH > C (Wiberg et al 1987) and there is a transfer of charge from methyl to phenyl. The changes in the total electron population of each atom in the phenyl group and of the group itself, together with the changes in their σ and π components, as caused by replacement of H by X, are given in Tables 3.6, 3.7 and 3.8. The ordering of charge withdrawal is the same as that observed in substituted ketones XHC=O, as their

Table 3.4. Atomic Populations N(Ω) of Substituted Benzenes.

Ω \ X	O⁻	NH_2	OH	F	CH_3	H	CN	NO_2	NH_3^+
C1	4.945	5.455	5.460	5.531	6.010	5.957	5.874	5.685[c]	5.842
C2	6.025	5.942	5.945	5.900	5.955	5.957	5.928	5.912	5.900
C3	6.004	5.950	5.948	5.945	5.952	5.957	5.937	5.930	5.900
C4	5.994	5.944	5.945	5.944	5.948	5.957	5.951	5.953	5.916
C5	6.004	5.950	5.952	5.945	5.951	5.957	5.937	5.930	5.901
C6	6.025	5.942	5.910	5.900	5.959	5.957	5.928	5.912	5.906
H7	1.101	1.063	1.052	1.004	1.059	1.043	1.004	0.951	1.005
H8	1.119	1.048	1.037	1.029	1.052	1.043	1.023	1.018	0.978
H9	1.130	1.056	1.044	1.036	1.053	1.043	1.023	1.017	0.977
H10	1.119	1.048	1.036	1.029	1.052	1.043	1.023	1.018	0.978
H11	1.101	1.063	1.015	1.004	1.058	1.043	1.004	0.951	1.002
C					5.754		4.887		
N		8.427					8.482	6.673	8.272
O	9.436		9.292					8.525	
F				9.738					
H[a]		0.557	0.364		1.071	1.043			0.473
H[b]					1.063				0.474
Total	50.003	50.002	50.000	50.005	50.000	42.000	54.001	64.000	49.998
q(ø)[d]	+0.436	+0.541	+0.656	+0.738	-0.049	+0.043	+0.369	+0.723	+0.693

a. In the plane of benzene ring.
b. Out of the plane of benzene ring.
c. Obtained by difference.
d. Net charge on phenyl group.

Table 3.5. Atomic π Populations $N_\pi(Q)$ of Substituted Benzenes.

Q \ X	O^-	NH_2	OH	F	CH_3	H	CN	NO_2	NH_3^+
C1	0.656	0.858	0.906	0.969	0.961	0.969	1.043	1.074	1.134
C2	1.106	1.060	1.040	0.990	0.990	0.969	0.935	0.904	0.959
C3	0.974	0.938	0.946	0.949	0.963	0.969	0.965	0.974	0.922
C4	1.156	1.036	1.008	0.987	0.981	0.969	0.934	0.918	0.904
C5	0.974	0.938	0.944	0.949	0.963	0.969	0.965	0.974	0.919
C6	1.106	1.060	1.008	0.990	0.991	0.969	0.935	0.904	0.966
H7	0.042	0.036	0.034	0.028	0.032	0.031	0.026	0.022	0.028
H8	0.036	0.028	0.026	0.030	0.031	0.031	0.030	0.028	0.026
H9	0.052	0.034	0.032	0.030	0.032	0.031	0.028	0.026	0.025
H10	0.036	0.028	0.030	0.030	0.031	0.031	0.030	0.028	0.026
H11	0.042	0.036	0.030	0.028	0.032	0.031	0.026	0.022	0.028
C					0.894		0.648		
N		1.907					1.432	1.050	1.542
O	1.828		1.982					1.540	
F				2.018					
H^a		0.020	0.010		0.029	0.031			0.008
H^b					0.537				0.258
Total	8.008	7.999	7.998	7.998	8.005	6.000	7.997	10.004	8.002

a. In the plane of benzene ring.
b. Out of the plane of benzene ring.

π-donating/withdrawing ability relative to H. From the values of $\Delta N(Ph)$, $\Delta N_\pi(Ph)$ and $\Delta N_\sigma(Ph)$ in Table 3.6, 3.7 and 3.8, one can see, except for CH_3, that all the substituents, compared with H of benzene, withdraw σ electrons. This is analogous to the suggestion in classical inductive theory that those substituents (except for O^-) are -I effect ones while CH_3 is a +I effect one. All the substituents in front of benzene lose π electrons while those behind gain, which agrees exactly with the assignment of classical mesomeric effect theory (Ingold 1953). The orders of magnitude of electronic transfer for all substituents are also analogous to those of classical theory, respectively. From this point of view, we may relate the σ electron transfer to the inductive effect and π electron transfer to the mesomeric effect.

NH_2, OH and F are three typical -I and +M substituents; their ability to withdraw σ electrons from the phenyl group relative to hydrogen increases in this order, the corresponding values of $\Delta N_\sigma(Ph)$ in Table 3.8 being -0.582, -0.652 and -0.708, respectively and the relative ability of donating π electrons decreases, the values of $\Delta N_\pi(Ph)$ being 0.084, 0.039 and 0.013 in Table 3.7, respectively. CN and NO_2 are two typical -I and -M substituents, withdrawing both σ and π electrons in increasing order from CN to NO_2. CH_3 is the only σ electron donor in this study with the smallest σ electron transfer, 0.058 e, and the second smallest π electron transfer.

Two charged substituents O^- and NH_3^+ make the situation complicated, and it is worth discussing them one by one. O^- is a strong σ electron-attracting substituent by our study, which is different from the conclusion made from classical theory (Ingold 1953), and it is the

Table 3.6. Relative Atomic Populations of Substituted Benzenes.

Q \ X	O^-	NH_2	OH	F	CH_3	H	CN	NO_2	NH_3^+
C1	-1.012	-0.502	-0.497	-0.426	+0.053	0.000	-0.083	-0.260	-0.115
C2	+0.068	-0.015	-0.012	-0.057	-0.002	0.000	-0.029	-0.044	-0.057
C3	+0.047	-0.007	-0.009	-0.012	-0.005	0.000	-0.020	-0.024	-0.057
C4	+0.037	-0.013	-0.012	-0.013	-0.009	0.000	-0.006	-0.004	-0.041
C5	+0.047	-0.007	-0.005	-0.012	-0.006	0.000	-0.020	-0.024	-0.056
C6	+0.068	-0.015	-0.047	-0.057	+0.002	0.000	-0.029	-0.044	-0.051
H7	+0.058	+0.020	+0.009	-0.039	+0.016	0.000	-0.039	-0.092	-0.038
H8	+0.076	+0.005	-0.006	-0.014	+0.009	0.000	-0.020	-0.025	-0.065
H9	+0.087	+0.013	+0.001	-0.007	+0.010	0.000	-0.020	-0.026	-0.066
H10	+0.076	+0.005	-0.007	-0.014	+0.009	0.000	-0.020	-0.025	-0.065
H11	+0.058	+0.020	-0.028	-0.039	+0.015	0.000	-0.039	-0.092	-0.041
ΔN(Ph	-0.393	-0.498	-0.613	-0.695	+0.092	0.000	-0.326	-0.680	-0.650

Table 3.7. Relative Atomic π Populations of Substituted Benzenes.

Q \ X	O^-	NH_2	OH	F	CH_3	H	CN	NO2	NH_3^+
C1	-0.313	-0.111	-0.063	-0.000	-0.008	0.000	+0.074	+0.105	+0.165
C2	+0.137	+0.091	+0.071	+0.021	+0.021	0.000	-0.034	-0.065	-0.010
C3	+0.005	-0.031	-0.023	-0.020	-0.006	0.000	-0.004	+0.005	-0.047
C4	+0.187	+0.067	+0.039	+0.018	+0.012	0.000	-0.035	-0.051	-0.065
C5	+0.005	-0.031	-0.025	-0.020	-0.006	0.000	-0.004	+0.005	-0.050
C6	+0.137	+0.091	+0.039	+0.021	+0.022	0.000	-0.034	-0.065	-0.003
H7	+0.011	+0.005	+0.003	-0.003	+0.001	0.000	-0.005	-0.009	-0.003
H8	+0.005	-0.003	-0.003	-0.001	-0.000	0.000	-0.001	-0.003	-0.005
H9	+0.021	+0.003	+0.001	-0.001	+0.001	0.000	-0.003	-0.005	-0.006
H10	+0.005	-0.003	-0.001	-0.001	-0.000	0.000	-0.001	-0.003	-0.005
H11	+0.011	+0.005	-0.001	-0.003	+0.001	0.000	-0.005	-0.009	-0.003
ΔN_π(Ph	+0.203	+0.084	+0.039	+0.013	+0.034	0.000	-0.049	-0.099	-0.035
ΣC2-C6	+0.471	+0.187	+0.101	+0.020	+0.043	0.000	-0.111	-0.171	-0.175

Table 3.8. Relative Atomic σ Populations of Substituted Benzenes.

Q \ X	O^-	NH_2	OH	F	CH_3	H	CN	NO_2	NH_3^+
C1	-0.699	-0.391	-0.434	-0.426	+0.061	0.000	-0.157	-0.365	-0.280
C2	-0.069	-0.106	-0.083	-0.078	-0.023	0.000	+0.005	+0.021	-0.047
C3	+0.042	+0.024	+0.014	+0.008	+0.001	0.000	-0.016	-0.029	-0.010
C4	-0.150	-0.080	-0.051	-0.027	-0.021	0.000	+0.029	+0.047	+0.024
C5	+0.042	+0.024	+0.020	+0.008	+0.000	0.000	-0.016	-0.029	-0.006
C6	-0.069	-0.106	-0.086	-0.078	-0.020	0.000	+0.005	+0.021	-0.048
H7	+0.047	+0.015	+0.006	-0.036	+0.015	0.000	-0.034	-0.083	-0.035
H8	+0.071	+0.008	-0.003	-0.013	+0.009	0.000	-0.019	-0.022	-0.060
H9	+0.066	+0.010	-0.000	-0.006	+0.009	0.000	-0.017	-0.021	-0.060
H10	+0.071	+0.008	-0.006	-0.013	+0.009	0.000	-0.019	-0.022	-0.060
H11	+0.047	+0.015	-0.027	-0.036	+0.014	0.000	-0.034	-0.083	-0.038
ΔN_σ(Ph	-0.596	-0.582	-0.652	-0.708	+0.058	0.000	-0.277	-0.581	-0.615

strongest π electron donor in this study. Thus according to these two effects, O^- can be classified with the NH_2, OH and F groups, which have the same behaviour. The π electrons on the ring of the phenoxide anion transfer in the alternate way, although the same positive signs appear on C_2-C_6, which is due to the effect of strong π donation from the O^- substituent. NH_3^+ is also an unusual substituent. It is the strongest σ electron withdrawer and the weakest π electron withdrawer if we can put it in the group containing CN and NO_2. These features make it work more like an inductive attractor.

From the above discussion, the groups F, OH, NH_2, and O^- are σ withdrawing (decreasing in this order) and π donating (increasing in this order). Methyl is both σ and π donating, while CN, NO_2 and NH_3^+ are σ and π withdrawing, increasing in the given order for σ withdrawal and with NH_3^+ withdrawing the smallest amount of π electrons. The ipso carbon is strongly perturbed by the substituent. All of the σ-withdrawing substituents, particularly the π-donating set, impart a considerable positive charge to this atom, ranging from +1.06 in phenoxide to 0.47 in fluorobenzene. While q(C1) is somewhat smaller for X=CN and NO_2, the substituent atom bonded to C1 bears a substantial positive charge in these molecules with q(C)=1.11 in Ph-CN and q(N)=+0.33 in Ph-NO_2. The large positive charge on C1 or its bonded substituent atom preclude the possibility of electrophilic attack at the ipso carbon.

It is clear from the data in Table 3.6 that the relative activating-deactivating and directing ability of a substituent is not determined by the total charge of the phenyl group or by the total charges on the atoms: with the sole exception of methyl, both the activating and

deactivating substituents withdraw electronic charge from phenyl and the variation in the relative carbon atom populations exhibit no correlation with the experimental ortho-, meta-, and para-directing abilities of the substituents. The changes in the total populations of the carbon atoms are smaller than the changes in either of their σ or π components, a result of alternating and opposing polarization of these component populations around the ring of atoms, as shown in Figure 3.1. Similar results were obtained in an analysis of Mulliken populations (Bock et al 1985, Hehre et al 1972, 1976).

The F atom withdraws π charge from phenyl and the π population of methyl remains essentially unchanged when acting as a substituent; however, relative to H which withdraws π density, both substituents are π electron donors. While the transfer of π charge is small in these two cases, their perturbation of the charge distribution of the phenyl group is significant. As pointed out by Hehre et al. (1972) and by Libit and Hoffmann (1974), equally as important as the transfer of charge between a substituent and a π system is the accompanying polarization of charge in the π system. The same is true for populations of the atoms of theory. The substituent X in the carbonyls HXC=O, for example (Slee et al 1988), polarizes the C=O bond and it is the π population of the carbonyl oxygen that exhibits the trends in the relative donating or abstracting abilities of X. The same is true for the π populations of the substituted phenyls with the additional feature that the polarization extends over the complete π system thereby leading to an alternation in the π population of the ring carbon atoms, Table 3.7. The pattern of the alternating changes in the π populations relative to benzene shown in Figure 3.1 calls to mind the contributing

resonance structures of valence bond theory which place a negative charge at the ortho or para positions for the π-donating groups, and a positive charge at these same positions for π-withdrawing groups. The π populations of the ortho and para carbons increase (their charges become more negative) and those of the ipso and meta carbons decrease relative to their values in benzene for the π-donating groups. With the exception of O^-, the increase at ortho is greater than at para. The π population of the ipso carbon is increased for the π withdrawing substituents CN, NO_2, and NH_3^+ while that of the ortho and para carbons is decreased. The relative π population of the meta carbon is decreased and increased slightly for the first two of these substituents. In protonated aniline, the relative π populations undergo a progressive decrease in the order ortho, meta, para. The total changes in the π populations of the carbons susceptible to electrophilic attack-all but C1-are also given in Table 3.7 as ΣC2>C6. Activating groups increase these π populations, the values decreasing in the expected order, being smallest for F. As discussed in the energy analysis (see Chapter 4), it appears that, in the gas phase at least, F may slightly activate the phenyl group. The π populations of these same carbon atoms are decreased for the deactivating substituents and in the anticipated order.

The plots of the relative ortho, meta, and para π populations versus the resonance parameter σ_R^o, Figure 3.2, exhibit the anticipated monotonic behaviour with respect to their directing ability with the exception of O^- and CN in the meta position. These plots of relative π populations, with the same two exceptions, show a remarkable similarity to the corresponding plots for the relative ellipticities of the C-H bonds.

The changes in the ellipticities of the C-H bonds at their bond critical point provide a measure of the changes in the average π populations of the carbon atoms; the ellipticity of the carbon density increase with an increase in π population of the carbon atom.

Hehre et al. (1976) have correlated the Mulliken π charges on the carbon atoms in substituted benzenes, obtained in STO-3G calculations, with σ_R^o and with ^1H, ^{13}C, and ^{19}F NMR substituent shifts. They find the total π charge transferred to or from phenyl is generally well correlated directly with σ_R^o. Local π charges at the meta and para carbons show in addition to σ_R^o effects, a significant dependence on the polar parameter σ_I.

With the exception of methyl, all the substituents withdraw σ charge from the ipso carbon relative to H. The relative σ populations of the ortho, meta, and para carbons exhibit an alternating loss and gain of charge which is the opposite of that found for the relative π populations. Thus the polarization of the σ charge of the benzene ring is the opposite to that required to account for the relative directing ability of the substituents. In three cases, O^-, CN, and NO_2, the change in the σ population is greater for the para than for the ortho. Thus, there do not appear to be different mechanisms underlying the σ and π polarizations as envisaged in the classical inductive and resonance models, respectively. Rather, the two components of the charge distribution are mutually polarizing and only the magnitudes of the changes in the total charges at each carbon exhibit an approximate falloff with distance from the point of substitution.

Atomic Quadrupole Moment

The quadrupole polarization of an atomic density is another property of the total density which provides information that both complements and supplements the σ and π population of the orbital model. If z is the axis of a coordinate system centered at the nucleus, then the corresponding component of the quadrupole moment of the charge density for atom Ω is

$$Q_{zz}(\Omega) = -e \int_\Omega \rho(3z^2 - r^2)dr \qquad [3.1]$$

with corresponding definitions for the other diagonal components, [1.43]. A set of principal axes can always be found such that the off-diagonal elements vanish. Each of the diagonal moments of the density has the form of the familiar d_{z2} orbital. A negative value for Q_{zz} and positive values for Q_{xx} and Q_{yy} implies that the sphere is distorted in such a way that electronic charge is removed from the xy plane and concentrated along the z axis (see Fig.1.8). The carbon atoms in ethylene and benzene exhibit large negative values for $Q_{zz}(C)$ (with z perpendicular to the plane of the nuclei in both molecules) equal to -3.38 and -3.34 au, respectively, and an sp^2 carbon atom with a single π electron has a quadrupole moment of approximately -3.4 au. With z taken as the internuclear axis in acetylene, $Q_{zz}(C)$ in this molecule is large and positive, equal to +4.14 au corresponding to a torus-like concentration of π density about the z axis as reflected in the negative values for $Q_{xx}(C) = Q_{yy}(C) = -2.07$ au. In the planar methyl cation with a nearly vacant p_π orbital, $Q_{zz}(C) = +1.22$ au and the sphere is flattened to give an oblate spheroid (Bader 1986). This orbital vacancy is partly filled in tert-butyl cation by inductive and hyperconjugative electron release from the methyl groups, and the moment $Q_{zz}(C)$ is correspondingly reduced to +0.43 au in this molecule.

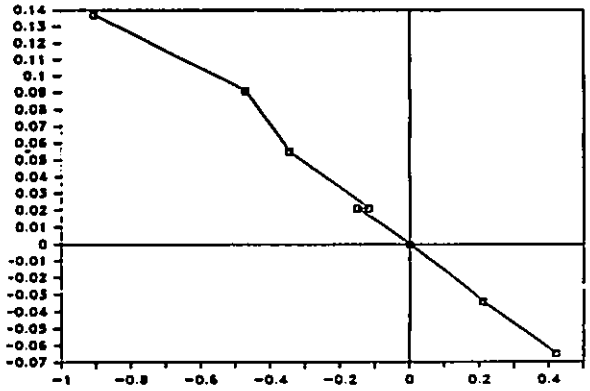
The parallelism between the quadrupole moment and π population is further demonstrated in Figure 3.3 which shows plots of the quadrupole moments of carbon atoms at the ortho, para, and meta positions of monosubstituted benzenes relative to their values in benzene (Table 3.9) versus their relative π populations. Figure 3.2 shows plots of the relative quadrupole moments of the ortho, para, and meta carbon atoms versus σ_R^o. These latter correlations bear a remarkable similarity to those previously given for the relative C-H bond ellipticities and π populations. The departure of the meta π population in phenoxide ion from the linear relationship with $Q_{zz}(C)$ apparent in Figure 3.3, results in a better correlation of the quadrupole moment with σ_R^o, as is also the case of the same position in cyanobenzene. The correlation in Figure 3.2 demonstrate that the quadrupole polarization of the charge density of a carbon atom in a monosubstituted benzene molecule parallels its relative reactivity toward electrophilic attack.

3.5 Susceptibility to Electrophilic Attack as Determined by the Laplacian of ρ

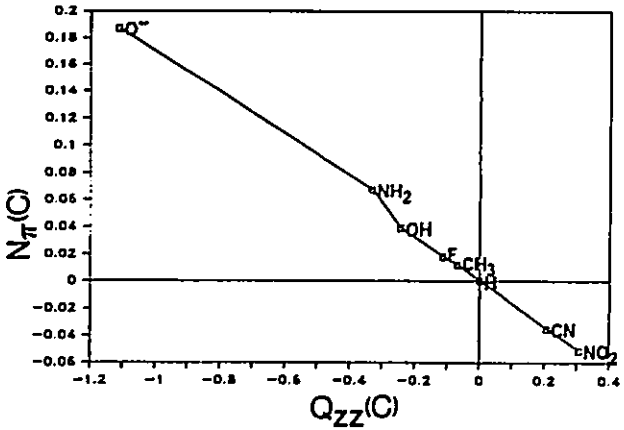
It has been shown in section 3.4 that the total charge on a carbon atom exhibits a poor correlation with the relative reactivity of a carbon atom in a monosubstituted benzene toward electrophilic attack, and neither does the σ population. The π electron distribution, however, exhibits such a contribution. The Laplacian of the charge density, $\nabla^2\rho$, which need not separate σ and π electrons in order to obtain more information, does show where charges are concentrated and depleted naturally. It provides rich information about the chemical reactivity of any system, while the π

Figure 3.3

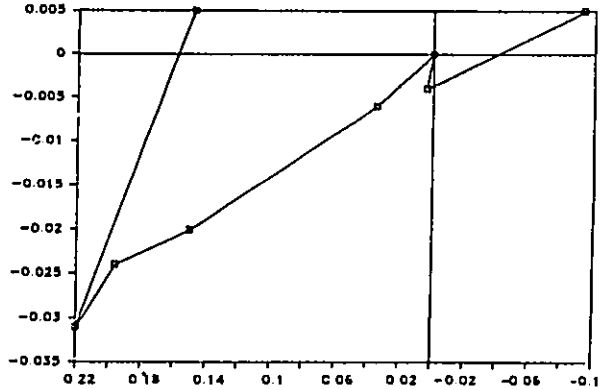
The relative π population of carbon plotted versus its relative quadruple moment $Q_{zz}(C)$ at the ortho, para, and meta positions of substituted benzenes.



ortho



para



meta

Table 3.9. Relative Quadrupole Moments[a] (Q_{zz}) of Substituted Benzenes.

Q \X	O⁻	NH_2	OH	F	CH_3	H	CN	NO_2	NH_3^+
C1	+1.3354	+0.5926	+0.5170	+0.3199	+0.3915	0.0000	-0.1455	-0.3144	-0.4547
C2	-0.9036	-0.4716	-0.4152	-0.1190	-0.1683	0.0000	+0.2100	+0.4203	+0.1097
C3	+0.1483	+0.2203	+0.1941	+0.1497	+0.0272	0.0000	+0.0040	-0.0933	+0.2324
C4	-1.1120	-0.3321	-0.2439	-0.1125	-0.0692	0.0000	+0.2083	+0.3070	+0.3554
C5	+0.1483	+0.2203	+0.1994	+0.1497	+0.0427	0.0000	+0.0040	-0.0933	+0.2608
C6	-0.9036	-0.4716	-0.2744	-0.1190	-0.1328	0.0000	+0.2100	+0.4203	+0.0995

a. The value of Q_{zz}(C) for benzene is -3.3401 au.

electron population can only be used in conjugated systems.

It is now well documented that the relative orientation of acid and base molecules in a reaction or in a crystal structure can be predicted by aligning the centre of charge concentration on the base with the centre of charge depletion on the acid, as defined by the Laplacian of their charge distribution (Bader et al 1984 and 1988, Bader and MacDougall 1985, Caroll et al 1988,see also Chapter 1). In previous examples, the base possessed an unshared electron pair in the Lewis sense. A phenyl group does not possess any Lewis unshared pairs to serve as centers of electrophilic attack and the present Chapter demonstrates how the definition of a base can be generalized to include such molecules in terms of a property of the Laplacian distribution.

It has been mentioned in Chapter 1 that there is a shell of charge concentration and one of charge depletion associated with each quantum shell of an atom. Each shell of charge concentration for a free atom possesses a sphere over which the electronic charge density is uniformly and maximally concentrated. Upon chemical combination, local maxima, minima and saddle points in $-\nabla^2\rho$ are created on the surface of this sphere within the outer or valence shell charge concentration (VSCC), the maxima corresponding in number and relative location to the localized electron pairs as assumed in VSEPR model (Bader et al 1984 and 1988). The structure exhibited by the Laplacian of ρ in terms of its valence shell critical point is characteristic for a given atom with a given number of bonded neighbors. In methane for example, the VSCC of the carbon atom exhibits four bonded maxima, i.e., four local concentrations of electronic charge, one along each of the C-H bond paths. Each maximum is linked to

the other three by trajectories which originate at intervening saddle points on the surface of the sphere of charge concentration. (These lines are topologically identical with the bond paths linking the nuclear maxima in ρ in the definition of a molecular graph.) This network of lines, called an atomic graph, partitions the surface into four segments and the basic structure is that of a tetrahedron with curved faces, Figure 3.4. In the centre of each face there is a local minimum in the surface of the VSCC. (A face critical point is topologically equivalent to a ring critical point of a molecular graph.) As an example of a system with a nonbonded electron pair, the atomic graph for the nitrogen atom in ammonia is topologically equivalent to that for the methane carbon, but one of the maxima now represents a nonbonded electron pair. It is of greater magnitude and large spatial extent than are the three bonded maxima and the angle it forms with a bonded maximum is increased above the tetrahedral value, to 112.3°.

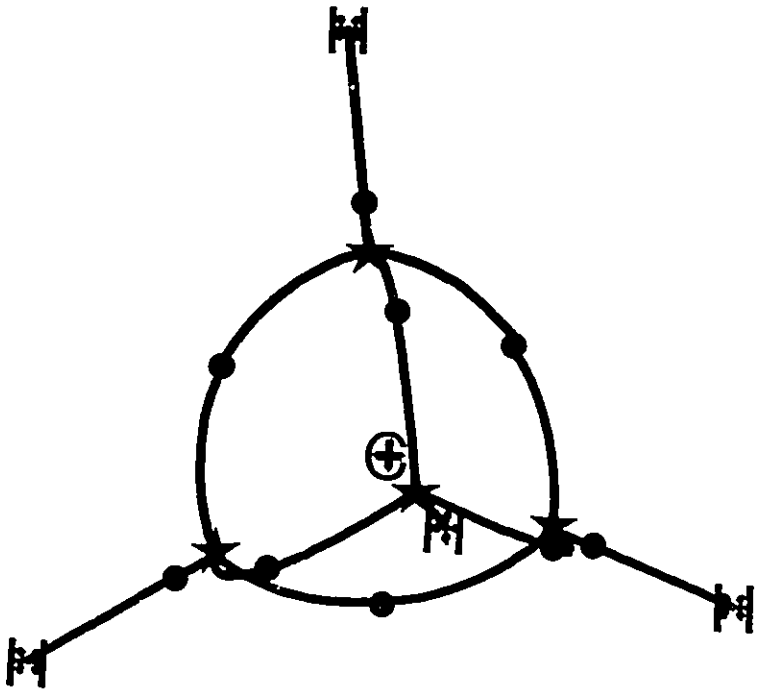
The atomic graph for a carbon atom in benzene has three bonded maxima, Figure 3.4. The maximum on the C-H bond path is linked by pairs of trajectories emanating from intervening saddle points, to each of the maxima on a C-C bond path. These latter two bonded maxima are linked not by one but by two sets of trajectories emanating from saddle points located on either side of the plane of the benzene ring. The sphere of charge concentration is

partitioned into three curved faces, with a minimum or ring critical point located within each face. The Laplacian distributions for benzene and phenoxide ion are illustrated in Figure 3.5 in the form of relief plots.

The bonded maxima dominate the VSCC of carbon, with values in the

Figure 3.4

Atomic graphs depicting the structure of the VSCC of a carbon atom in methane (top) and benzene. Local maxima are denoted by stars and the intervening saddle points by dots. In methane the four maxima are at the corners of a tetrahedron. The two lower maxima for C in benzene are linked to maxima on neighboring carbons which are not shown. The three maxima lie in the plane of the ring and of the two saddles linking the two lower maxima; one is above and the other below the plane of the ring.



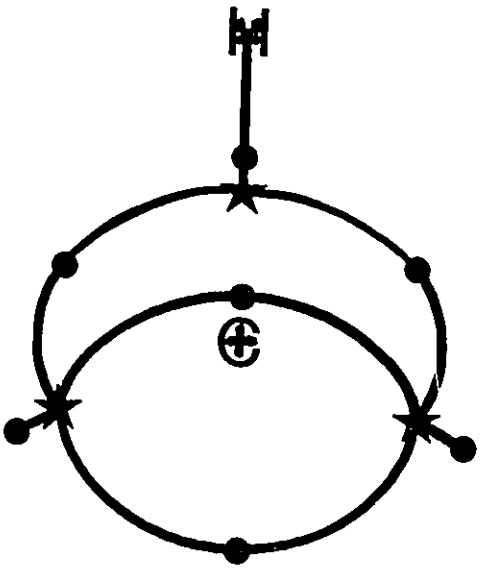
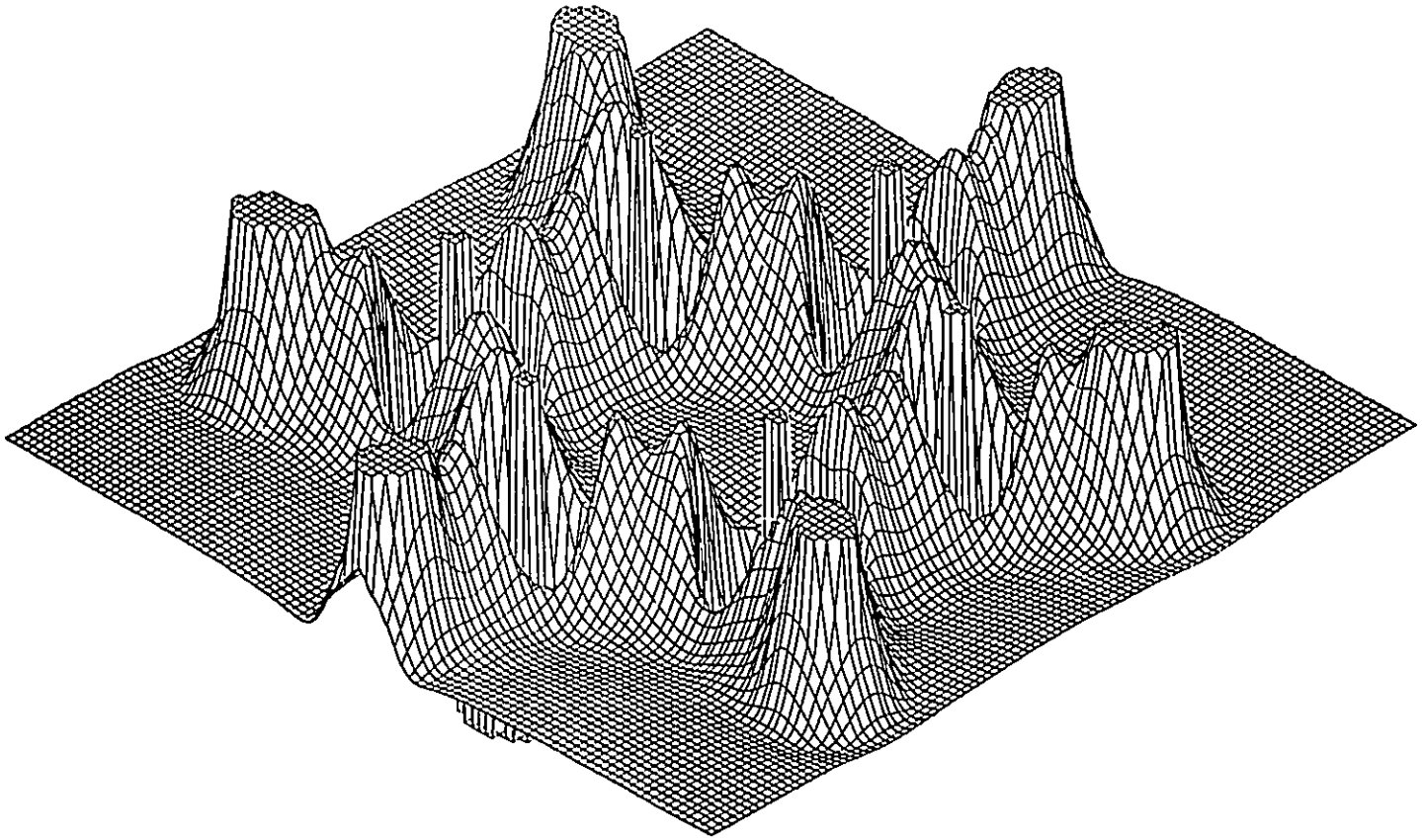
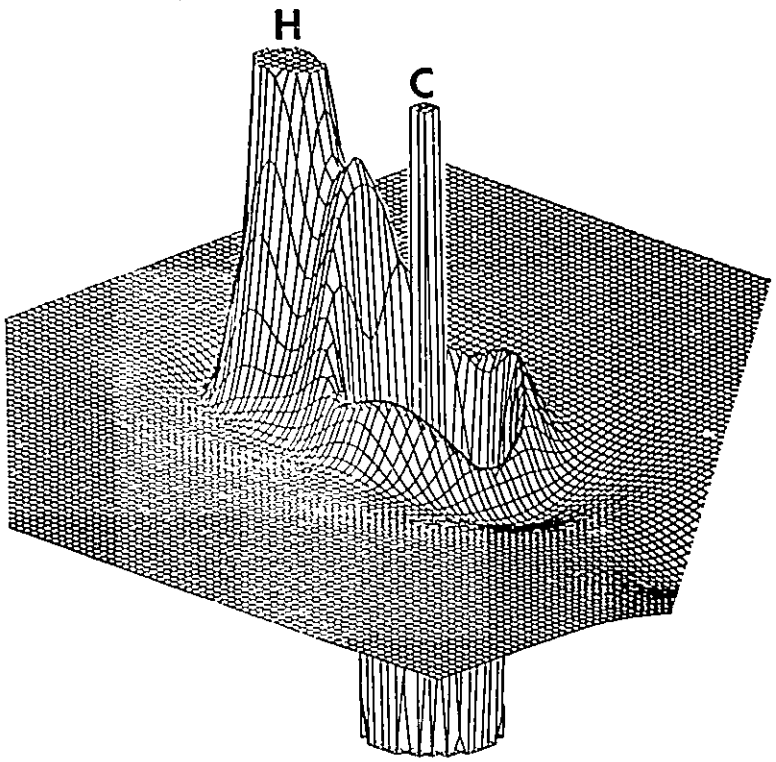


Figure 3.5

Relief maps of the negative of the Laplacian of ρ. The upper diagram is in the plane of the nuclei for benzene. Note the large shared concentration of electronic charge between each pair of neighboring carbon atoms and carbon and hydrogen atoms. The inner quantum shell of carbon consists of a spikelike charge concentration at the nucleus, surrounded by a region of charge depletion. The valence shell of charge concentration in this plane exhibits three bonded charge concentrations as indicated in the atomic graph, Fig.3.4. The lower diagram is for the para carbon atom in phenoxide ion in the symmetry plane along C_2 axis and perpendicular to the ring plane. The VSCC of carbon exhibits the bonded maximum with hydrogen and the two secondary charge concentrations which appear as two-dimensional maxima in this plane.




H
C

range of -1.2 to -1.4 au depending on the substituent. These are followed by the saddle points linking the bonded maxima on the C-C bond paths, where $\nabla^2\rho$ has values in the range from -0.25 to -0.11 au. Note that the saddle points appear as maxima in the relief plots of the negative of Laplacian when viewed in a plane containing the two negative curvatures. A true maximum appears as much in any plane containing the critical point, as exemplified by the bonded maxima in Figure 3.5. The magnitudes of $\nabla^2\rho$ at the saddle points linking the C and H bonded maxima are approximately half as great as those linking the carbon maxima. The value of the Laplacian at a ring critical point is small in magnitude and is of either sign depending on whether or not the sum of its two positive curvatures in the ring surface exceeds the magnitude of its radial curvature.

Next to the bonded charge concentrations, the Laplacian attains its most negative values at the out-of-plane saddle points and these points locate the greatest concentrations of electronic charge that are not involved with the bonding of nuclei, Figure 3.5. Because of this and their relative location, it is reasonable to propose that these secondary charge concentration will serve as the sites for electrophilic attack of a phenyl group carbon atom and this conjecture is borne out by the data presented in Table 3.10 and displayed in Figure 3.2. These data give the values of $\nabla^2\rho$ at the out-of-plane saddle points -- the secondary charge concentrations-- relative to their values in benzene. They exhibit the pattern of alternating values around the ring necessary to account for the relative directing abilities of the substituents and, as is evident from Figure 3.2, the intensity of their effect parallels the experimentally determined resonance parameter σ_R^o. The group to the left of benzene, the ortho- and

Table 3.10. Relative Values of $\nabla^2\rho_b$ in Substituted Benzenes.

At secondary concentration in VSCC of carbon atoms[a].

Ω \ X	O⁻	NH₂	OH	F	CH₃	H	CN	NO₂	NH₃⁺
C1		+0.020	-0.044	-0.124	+0.008	0.000	-0.054	-0.124	-0.191
C2	-0.093	-0.067	-0.054	-0.025	-0.012	0.000	+0.013	+0.028	-0.003
C3	+0.018	+0.021	+0.014	+0.010	+0.006	0.000	-0.002	-0.007	+0.013
C4	-0.112	-0.038	-0.024	-0.013	-0.005	0.000	+0.016	+0.022	+0.021
C5	+0.018	+0.021	+0.015	+0.010	+0.006	0.000	-0.002	-0.007	+0.014
C6	-0.093	-0.067	-0.036	-0.025	-0.015	0.000	+0.013	+0.028	-0.006
ΣC2-C6	-0.262	-0.130	-0.085	-0.043	-0.020	0.000	+0.038	+0.064	+0.039

At image points in VSCD of carbon atoms[b].

Ω \ X	O⁻	NH₂	OH	F	CH₃	H	CN	NO₂	NH₃⁺
C1	+0.035	+0.016	+0.006	-0.004	+0.005	0.000	-0.004	-0.014	-0.015
C2	-0.027	-0.015	-0.011	-0.003	-0.002	0.000	+0.005	+0.010	+0.005
C3	+0.003	+0.007	+0.005	+0.003	+0.003	0.000	-0.000	-0.002	+0.007
C4	-0.029	-0.010	-0.006	-0.003	-0.000	0.000	+0.005	+0.007	+0.009
C5	+0.003	+0.007	+0.005	+0.003	+0.003	0.000	-0.000	-0.002	+0.007
C6	-0.027	-0.015	-0.007	-0.003	-0.002	0.000	+0.005	+0.010	+0.004

[a]. The value for benzene is -0.140 au.
[b]. The value for benzene is 0.087 au.

para-directing group, have negative values at ortho and para carbon atoms, and positive values at meta ones, showing it is easier for them to react with an electrophile at ortho and para positions and harder at the meta position relative to benzene. The group to the right of benzene exhibits the exactly opposite behaviou. The difference in the value of $\nabla^2\rho$ at the (3,-1) critical point compared to benzene shows that the reactivity increases or decrease towards electrophilic substitution. If the value is more negative than that of benzene, this position will be activated, and if more positive, it will be deactivated. If we sum up the relative values of $\nabla^2\rho$ on the ring atoms except for the ipso carbon (see Table 3.10), this sum will show the activating or deactivating ability of a substituent to the phenyl group. If the sign of the sum is negative, the total system is activated; and if positive, is deactivated. Also the magnitude of the sum shows the extent of activating or deactivating ability of each substituent, i.e., the more negative the sum, the more active will be the phenyl group toward an electrophilic attack; and the more positive the more deactivated.

The secondary charge concentration also increases at the ipso carbon for the -M substituents, as does its relative π population in these cases. as pointed out earlier, however, the large positive charge on this carbon and on the neighboring substituent atom precludes the possibility of electrophilic attack at this carbon. The even larger positive charge found on C1 in phenol and in fluorobenzene nullifies the increase in secondary charge concentration found for this atom in these molecules as well.

The secondary charge concentrations are displaced off a nuclear-centered axis perpendicular to the ring plane, toward the centre of the ring. The position of the secondary charge concentrations determine the

initial direction of approach of an electrophile, and the greater this angle the less accessible is the charge concentration to electrophilic attack. In benzene this angle equals 32.7° and its value for the para carbon atom undergoes a continuous decrease for the para-directing groups in parallel with their activating ability, Table 3.11. These same groups cause the angle for the secondary charge concentrations at the meta carbons to increase slightly and become less accessible to attack. The meta-directing, deactivating substituents CN and NO_2 cause the angle at the para carbon atom to increase while that at the meta position is left essentially unchanged from its value in benzene. When we relate the values of $\nabla^2\rho$ of para and meta carbon atoms to their off-axis angles, the reason why these angles correlate with the activating ability of the corresponding carbon atoms become clear. Actually an off-axis angle is a reflection of the magnitude of charge concentration. For each of the ring atoms, its two secondary charge concentrations above and below the ring plane, together with the carbon nucleus, form a complementary angle of the two off-axis angles. From Table 3.12, a nice correlation between complementary angles and $\nabla^2\rho$ values at the secondary charge concentration is observed. The more negative the value of $\nabla^2\rho$, the larger the complementary angle, then the smaller the off-axis angles. While factors other than the position of the secondary charge concentration may determine the direction of attack, the present results indicate that the approach of the electrophile toward a ring carbon atom should be shifted off axis toward the ring centre. This has been found to be true and will be detailed in Chapter 5.

The valence shell of charge depletion (where $\nabla^2\rho > 0$) of a carbon atom in benzene possesses critical points which spatially behave as the

Table 3.11. Off-axis Angles of Reactive Centres in the Laplacian of Substituted Benzenes.

Position\X	O^-	NH_2	OH	F	CH_3	H	CN	NO_2	NH_3^+
Meta	34.2	36.1	35.3	34.7	33.7	32.7	32.7	32.1	36.5
Para	18.6	27.2	29.2	31.1	31.9	32.7	35.8	37.5	38.2

Table 3.12. Comparison of Complementary Angles[a] of Secondary Charge Concentrations With Their Relative $\nabla^2\rho$ values.

	O^-	NH_2	OH	F	CH_3	H	CN	NO_2	NH_3^+
Meta Angle	111.6	107.8	109.4	110.6	112.6	114.6	114.6	115.8	107.0
$\nabla^2\rho$	+0.018	+0.021	+0.014	+0.010	+0.006	0.000	-0.002	-0.007	+0.013
Para Angle	142.8	125.6	121.6	117.8	116.2	114.6	108.4	105.0	103.6
$\nabla^2\rho$	-0.112	-0.038	-0.024	-0.013	-0.005	0.000	+0.016	+0.022	+0.021

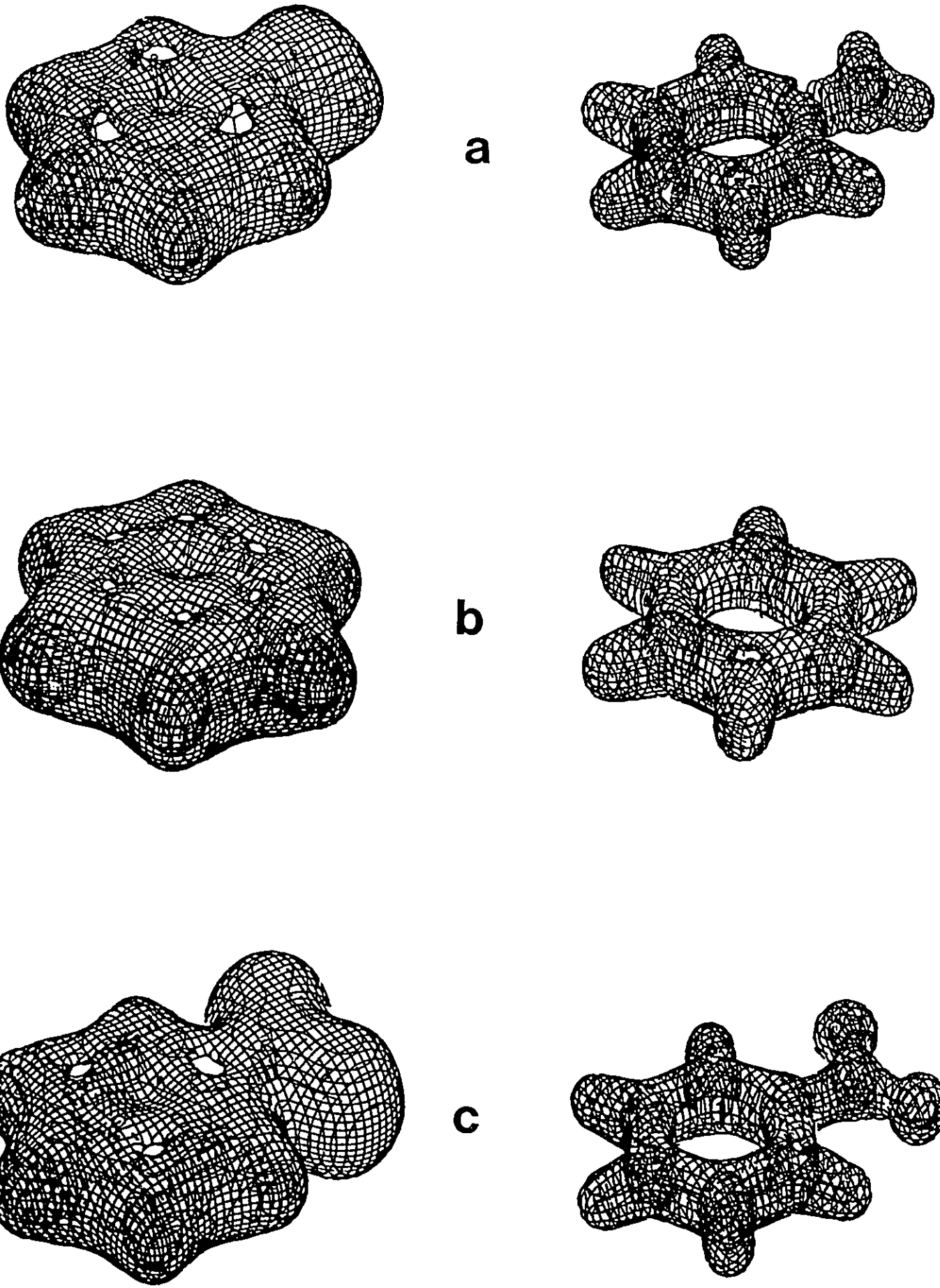
[a]. A complementary angle is the angle formed by two secondary concentraions above and below benzene ring plane with the corresponding nucleus.

images of the points of secondary charge concentration. They represent the points of least removal of negative charge in the shell of charge depletion which must be traversed by an electrophile in its approach to the secondary charge concentrations. From Table 3.10, we can see that the critical values in charge depletion shell change much more slowly than those in a charge concentration shell with different substituents. Figure 3.6 gives both charge concentration envelops (right), $V^2\rho=0$, and charge depletion one (left), $V^2\rho=0.09$ au, for aniline, benzene and nitrobenzene molecules. The $\nabla^2\rho=0$ envelope clearly shows how charges concentrate in the molecular space (see the envelopes at the right hand side of Fig.3.6). But here it does not give enough information about how the distribution changes with different substituents, while the charge depletion envelope does give some interesting observations. At the activated positions, ortho and para to NH_2 and meta to NO_2 (see the left hand side plots of Fig.3.6), the original small openings present in the benzene molecule becomes larger. While at the deactivated positions, the original openings have disappeared. Such pictures graphically illustrates the ortho-, meta- and para-directing ability of a substituent. The values of the Laplacian of ρ at the image points, relative to their value in benzene, are given in Table 3.10. There is again an alternation in values around the ring which parallels the directing ability of the substituent: the values of $\nabla^2\rho$ are less positive, and the hole in the envelope of charge depletion is greater at the ortho and para positions and smaller at the meta positions for the activating group with just the opposite behaviour found for the deactivating groups CN and NO_2.

The substituent NH_3^+ is anomalous in most of its atomic properties

Figure 3.6

Three-dimensional displays of the Laplacian distribution for substituted benzenes: (a) aniline, (b) benzene, (c) nitrobenzene. The diagrams on left-hand side are for an envelope of value +0.09 au and they show the presence of the "holes" in the valence shell of charge depletion (VSCD) which give access to the centers of secondary charge concentration. The diagrams on right-hand side are for the zero envelope which separates the shell of charge concentration from the shell of charge depletion in the valence shell of each molecule.


a
b
c

just as it is experimentally, since it is found to direct para as much as, or a little more than, it directs meta, even though it is strongly deactivating (March 1977). This is the only substituent which bears a large positive charge, +0.31e, and as a consequence, the property of the carbon atoms of the phenyl group, both relative and total, do not exhibit alternating values as found for the other substituent, but rather there is a monotonic change with distance from the ipso carbon. This fact, coupled with the presence of the net charge itself and the steric requirements resulting from its solvation, can account for the observed directing effects of NH_3^+. A similar conclusion was reached by Chandra and Coulson (1965).

3.8 Conclusion

It has been shown that the properties of the charge density of the phenyl group in its isolated reactant state reflects the observed directing and activating-deactivating affects of a substituent X in an electrophilic substitution reaction. These effects, while not evident in the total electron populations or net charges of the atoms or groups involved, are reflected in the π contributions to the carbon atom populations, and in related properties of the total charge density, the ellipticity of the C-H bond, and the quadrupole moment of the carbon atom charge densities. The net positive charge found on the ipso carbon atom of the phenyl group for the -I substituents is responsible for precluding electrophilic attack at this position of the ring. The same chemistry is predicted by the Laplacian of the charge density when the base sites in the Lewis model of an acid-base reaction are identified with the secondary concentrations of

electronic charge in the VSCC's of the ring carbon atoms. It is the concentration of electronic charge as defined by the Laplacian of ρ, rather than the presence of net negative charge, that determines the site of electrophilic attack. The Laplacian of ρ is related to the local statement of the virial theorem (Bader and Nguyen-dang 1981; Bader 1988) and as a consequence, its predictions regarding reactivity are determined by the full quantum potential, as opposed to electrostatic potential maps which are based only on the classical potential as determined by the charges of a system. It also follows from this theory that an acid-base reaction corresponds to the combination of a region with excess electronic potential energy which is negative with a region of excess electronic kinetic energy which is positive, where the excesses are measured relative to their average virial ratio of 2:1. This provides an energetic basis for the Laplacian model and, when coupled with the very specific definition of the sites and intensity of reactivity provided by this distribution function, one has a useful predictive model for the electrophilic substitution reaction based on a property of the reactant charge distribution.

It is important to realize, as pointed out some time ago by Coulson et al. (1952) that π density is not spatially separate from σ density, and the example used was an early study of the charge density of benzene. Thus, one should not view the correlation of benzene chemistry with the carbon π populations as implying that an approaching electrophile "sees" the π charges as distinct from the σ distribution. As demonstrated here, an increasing π population does correlate directly with a quadrupole polarization of the atomic density, a physical moment of the density. It is also true that π density is in general less tightly bound than is σ

density and this, coupled with the presence of a recognizable moment, offers some understanding as to why the reactivity is reflected in the π density. A full discussion of the energetics of these aspects of benzene chemistry is given in Chapter 4 where the theory of atoms in molecules is used to relate the substituent effect on the energy to the energies and charge distributions of the individual atoms in both the reactants and the arenium ion intermediates.

Table 3.13. The Optimised Geometry of Aniline at 6-31G[a]

Bond	length	Bond	length	Angle		Angle	
C1-C2	1.3970	C2-H7	1.0742	C2-C1-C6	118.6156	H8-C3-C4	119.9269
C2-C3	1.3847	C3-H8	1.0737	C1-C2-C3	120.4062	H9-C4-C3	120.6224
C3-C4	1.3879	C4-H9	1.0723	C2-C3-C4	120.9084	H9-C4-C5	120.6224
C4-C5	1.3879	C5-H10	1.0737	C3-C4-C5	118.7551	H10-C5-C4	119.9269
C5-C6	1.3847	C6-H11	1.0742	C4-C5-C6	120.9084	H11-C6-C1	119.5378
C6-C1	1.3970	C1-N	1.3805	C5-C6-C1	120.4062	N-C1-C2	120.6922
		N-H	0.9890	H7-C2-C1	119.5378	H-N-C1	121.0649

a. Bond lengths in angstroms and angles in degrees.

Table 3.14. Atomic Populations of Pyramidal Aniline[a]

C1	5.502	C6	5.944	H10	1.046
C2	5.944	N	8.331	H11	1.059
C3	5.950	H7	1.059	H13	0.586
C4	5.944	H8	1.046		
C5	5.950	H9	1.053		

a. Compare Table 3.4.

Chapter 4 Energetics of Protonated Monosubstituted Benzenes

It has been shown in Chapter 3 that the directing ability of a substituent and its ability to activate or deactivate a phenyl group in electrophilic aromatic substitution reactions are the apparent result of the properties of the charge distributions of the reactant molecules. The classical, experimentally based ordering of the substituents (Ingold 1969, Hammett 1937), as reflected in the empirical resonance parameter σ_R° of Taft (1956 and 1960), is recovered by the properties of atomic charge distributions referred to in Chapter 3, as is the orbital classification of their σ-π donor-acceptor properties. The three groups of substituents discussed in Chapter 3 with respect to their effect on the charge distribution of benzene are: the σ abstractor-π donor substituents listed in order of decreasing σ_R° values and increasing π-donating ability F, OH, NH_2, O^-, the σ-π donating CH_3 group and the σ-π abstracting substituents NH_3^+, CN and NO_2, their π electron-withdrawing ability increasing in the order indicated.

This chapter is a study of the energetics of the aromatic substitution reaction. Together with the protonated benzene, the corresponding arenium ion intermediates both meta and para to the substituent, have been investigated for the substituents OH, F and CN. The focus is on the effect of a substituent X on the energy of the phenyl group Ph, and on the energy of formation of the protonated arenium ion intermediate, $[Ph\text{-}X]H^+$. These effects are studied in terms of the energetics for the reaction

$$Ph\text{-}X + [Ph\text{-}H]H^+ \rightarrow Ph\text{-}H + [Ph\text{-}X]H^+ \qquad \Delta E_1 \qquad [4.1]$$

for protonation both meta and para to the substituent. Section 4.1

briefly describes the method of calculation. An atomic view of the aromatic electrophilic substitution reaction is obtained by relating the changes in energy caused by substitution of X for H to the changes in the energies of the individual atoms and functional groups, which is studied in Section 4.2. The ability of the theory to define the properties of groupings of atoms within a molecule is used to compare the bond and atomic properties of a fragment of the pentadienyl cation with those of the corresponding fragment in the benzenium ion and in this way obtain a concise, chemical description of the nature of this intermediate, which is discussed in the section 4.3.

4.1 Calculations

The calculations for the reactant molecules were performed using the 6-31G** basis set (Hariharan and Pople 1973) at geometries previously optimised by Bock et al (1985) using the 6-31G basis set. Further details are given in Chapter 3. The optimised geometries and corresponding energies of the protonated intermediates have been determined for three basis sets, 3-21G, 4-31G and 6-31G. A fourth calculation using the 6-31G* set was also performed for protonated benzene to determine the effect of polarizing functions. The results of these calculations are tabulated in the Appendix. Further calculations using the 6-31G** basis at the 6-31G optimised geometries were performed and it is the 6-31G**//6-31G results for the reactants and intermediates that are tabulated and discussed in the text of this chapter. Protonated benzene was calculated to possess C_{2v} symmetry. A similar symmetry was assumed for the intermediates $C_6H_6FH^+$ and $C_6H_6CNH^+$ with the protonation at the para position, while all the other

intermediates were assumed with C_s symmetry. Hence all atoms other than the two hydrogens bonded to the protonated carbon atom are assumed to lie in a plane. All other geometrical parameters however, have been optimised. Previous calculations on protonated benzene using the STO-3G and 4-31G basis sets with some geometrical constraints have been performed by Hehre and Pople (1972). Ermler et al (1976) used contracted forms of a (11s,7p,1d) basis set for carbon and of a (7s,1p) basis set for hydrogen in a calculation of the properties of benzenium ion in which all C-C and C-H bond lengths other than those to the ipso carbon were held fixed at the values in benzene in a C_{2v} geometry. Sordo et al (1979) have carried out MINDO/3 calculations over a portion of the energy surface to compare the properties of both the π- and σ-complexes while Gleghorn and McConkey (1975) have used the MINDO/2′ method to investigate regions of the surface for protonated benzene and toluene.

The usual fixed-nucleus total energies E obtained in SCF calculations are listed in Table 4.1 for the 6-31G**//6-31G results; the energies of the monosubstituted benzenes and the arenium cation intermediates are reported. Also listed are the value of γ, defined as V/T, which is used as a correction factor in the calculation of atomic energies. Because of small errors in the virial theorem arising from the use of a finite basis set, the use of T(Ω) to obtain E(Ω), Eq.[1.32], requires that the atomic kinetic energies be multiplied by the factor $(1 + \gamma)$. For an equilibrium geometry $(1 + \gamma)$ should equal -1. Empirically, this error is found to decrease when the basis used to calculate E is augmented beyond that used in the determination of the energies, as is done here. The correction factor and total energies are essentially the same

Table 4.1 Energies and Virial Ratios for Substituted Benzenes, Arenium Intermediates and Pentadienyl Cation in au.

Substituted Benzenes

X	-E	γ	X	-E	γ	X	-E	γ
O^-	304.97635	-2.00116	F	329.56208	-2.00159	CN	322.44725	-2.00149
NH_2	285.74544	-2.00119	CH_3	269.75318	-2.00107	NO_2	434.18073	-2.00226
OH	305.57254	-2.00159	H	230.71382	-2.00108	NH_3^+	286.10658	-2.00159

Arenium Cation Intermediates

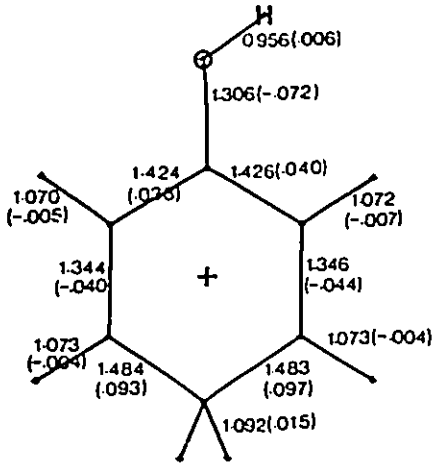
	Para		Meta				
X	-E	γ	-E	γ	X	-E	γ
OH	305.91284	-2.00155	305.87963	-2.00148	H	231.02698	-2.00093
F	329.87696	-2.00162	329.85871	-2.00154			
CN	322.73228	-2.00139	322.73457	-2.00138			

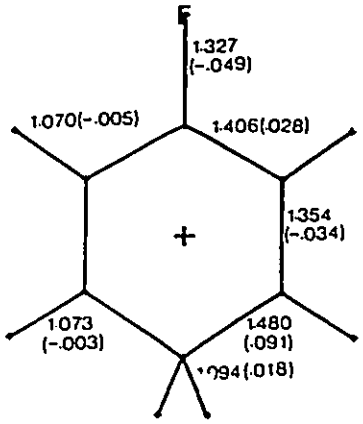
Pentadienyl Cation

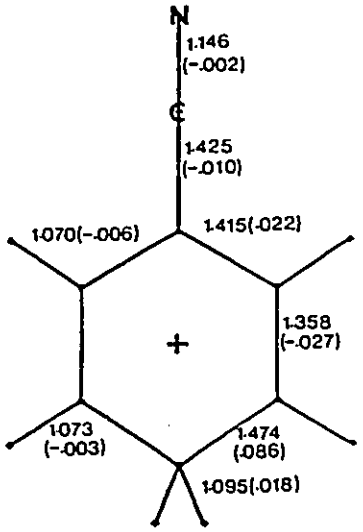
E = -193.12815 =2.00098

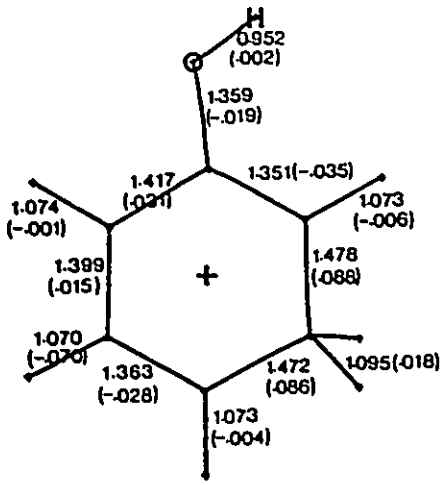
Figure 4.1.

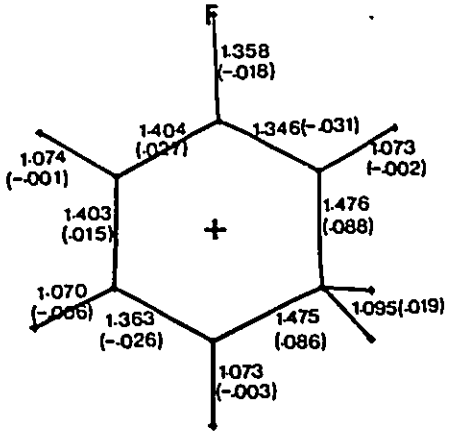
Molecular graphs as predicted by theory for meta- and para-protonated monosubstituted benzenes and for the benzenium ion. Also shown is the structure for the pentadienyl cation. The bond lengths are indicated and the values in parentheses are the changes in bond lengths relative to their values in Ph-X.

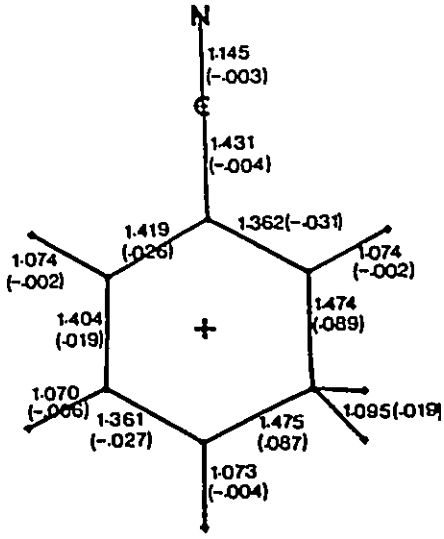

H
0·956(·006)
O
1·306(−·072)
1·424
(·026)
1·426(·040)
1·070
(−·005)
1·072
(−·007)
1·344
(−·040)
1·346
(−·044)
1·073
(−·004)
1·073(−·004)
1·484
(·093)
1·483
(·097)
1·092(·015)

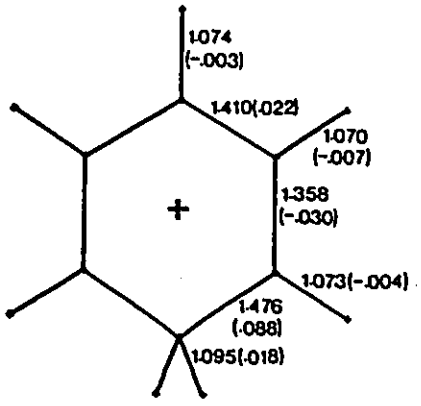

F
1·327
(−·049)
1·070(−·005)
1·406(·028)
1·354
(−·034)
1·073
(−·003)
1·480
(·091)
1·094(·018)

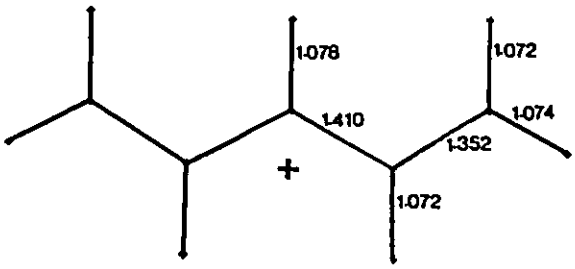

N
1·146
(−·002)
C
1·425
(−·010)
1·070(−·006)
1·415(·022)
1·358
(−·027)
1·073
(−·003)
1·474
(·086)
1·095(·018)


H
0·952
(·002)
O
1·359
(−·019)
1·417
(·021)
1·351(−·035)
1·074
(−·001)
1·073
(−·006)
1·399
(·015)
1·478
(·088)
1·070
(−·070)
1·095(·018)
1·363
(−·028)
1·472
(·086)
1·073
(−·004)


F
1·358
(−·018)
1·404
(·027)
1·346(−·031)
1·074
(−·001)
1·073
(−·002)
1·403
(·015)
1·476
(·088)
1·070
(−·066)
1·095(·019)
1·363
(−·026)
1·475
(·086)
1·073
(−·003)


N
1·145
(−·003)
C
1·431
(−·004)
1·419
(·026)
1·362(−·031)
1·074
(−·002)
1·074
(−·002)
1·404
(·019)
1·474
(·089)
1·070
(−·066)
1·095(·019)
1·361
(−·027)
1·475
(·087)
1·073
(−·004)


1·074
(−·003)
1·410(·022)
1·070
(−·007)
1·358
(−·030)
1·073(−·004)
1·476
(·088)
1·095(·018)


1·078
1·072
1·410
1·074
1·352
1·072

for Ph-X and its protonated intermediate and there is no difficulty in comparing the associated changes in atomic energies. The same trends in values are obtained with the uncorrected values of T(Ω).

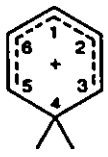
The geometries of the intermediates and the principal changes in bond lengths relative to Ph-X are summarized in Fig.4.1. The nature of the changes relative to the site of protonation is similar for all of the molecules. The C-C bond lengths to the protonated carbon increase by 0.086 to 0.091 A while the associated angle decreases to 115 - 116°. Its H-C-H angles vary over only one degree, from 103 to 104° and the associated C-H bond lengths increase over the unprotonated values by 0.015 to 0.019 A. The length of a C-C single bond to the central methylene group in pentane, the standard methylene group, and of its C-H bonds using the 6-31G* basis are 1.528 and 1.090 A, respectively. These lengths and the associated CCC and HCH bond angles of 113.1 and 106.3°, differ significantly from the corresponding values for the CH_2 fragment in the benzenium ion using the same basis set (Table A2). The C-C bonds, once removed from the site of protonation, decrease in length by 0.03 to 0.04 A while those twice removed are lengthened by a smaller amount. Also shown in Figure 4.1 is the molecular graph along with the geometrical parameters for the minimum energy configuration of the pentadienyl cation determined in a 6-31G**//6-31G calculation. The alternate lengthening and shortening of pairs of C-C bonds around the benzene ring, starting at the site of protonation in [Ph-H]H+, yields a fragment whose C-C and C-H bond lengths closely resemble those of this ion. The same fragment is present in all of the intermediates and the dominant changes in the geometry of the phenyl group in Ph-X are determined by the location of a given bond relative to the site

of protonation, independent of the nature and location of the substituent X.

4.2. Molecular Structures and Atomic Properties of Arenium Ion Intermediates

The molecular structures of the protonated monosubstituted benzenes shown in Fig.4.1 and used to summarize their molecular geometries, are predicted by quantum mechanics (see Chapter 1).

The substitution reaction of Ph-X with the electrophile Y^+ is assumed to proceed via an intermediate arenium ion $[Ph\text{-}X]Y^+$ (Brouwer et al 1970, Lowry and Richardson 1981, Streitwieser and Heathcock 1985). These are stable σ-bonded (as opposed to π-bonded) complexes which can be isolated in certain instances (Olah and Kuhn 1958). The prototype benzenium ion has been prepared in solution and identified by its NMR spectrum (Olah et al 1978). With the assumed relationship between ^{13}C chemical shifts and the π density at a carbon nucleus (Spiesecke and Schneider 1961, O'Brien et al 1975), this spectrum has been used to support the basic model of the benzenium ion as a pentadienyl cation, bridged by a methylene group and generally represented as in **a**.



a

The bond and atomic properties of the protonated benzenes are used to provide a precise picture of the structure and bonding in the benzenium ion and of the manner in which it is perturbed by a substituent.

The properties of the charge density at a bond critical point, its value ρ_b, the value of its Laplacian $\nabla^2\rho_b$ and the bond ellipticity ϵ, serve to characterize the bond and are listed in Table 4.2 for the intermediates considered here. The value of ρ_b can be used to define a bond order and the ρ_b values for ethane, benzene, ethylene and acetylene fix their respective C-C bond orders at 1.0, 1.6, 2.0 and 2.9 (For the definition of bond order, see Eq.[1.12]. For 6-31G**//6-31G densities, A=6.55 and B=0.2513). The bond ellipticity ϵ measures the extent to which charge density is preferentially accumulated in a given plane (see Chap.1). In ethylene, ϵ=0.45 and in benzene, ϵ=0.23 with the major axis being coincident with the axis of the π orbitals in both molecules. The increase in ellipticity of the C-C bond from benzene to ethylene is primarily a result of an increased contraction in the density toward the bond path, the magnitude of λ_1 increasing from 0.70 to 0.81 au, rather than from 0.57 to 0.56 au. (The magnitudes of both curvatures are increased over their value in ethane where they degenerate to the value -0.48 au.) It is by this mechanism that more charge is accumulated along the bond path, reducing the positive curvature from 0.27 to 0.20 au, thereby decreasing $\nabla^2\rho$ and increasing the bond order from 1.6 to 2.0.

The C-C bond orders of the substituted benzenes differ little from the value of 1.6, the small changes that are present reflecting the changes in equilibrium bond lengths (see Chap.3). This situation changes for the protonated intermediates, as seen from Table 4.2 which lists the n values for the C-C bonds in the intermediates. The C-C bonds to the protonated carbon have reduced bond orders ranging from 1.20 to 1.25 and their ellipticities, while reduced in value to $\approx$ 0.04 to 0.05, are significantly

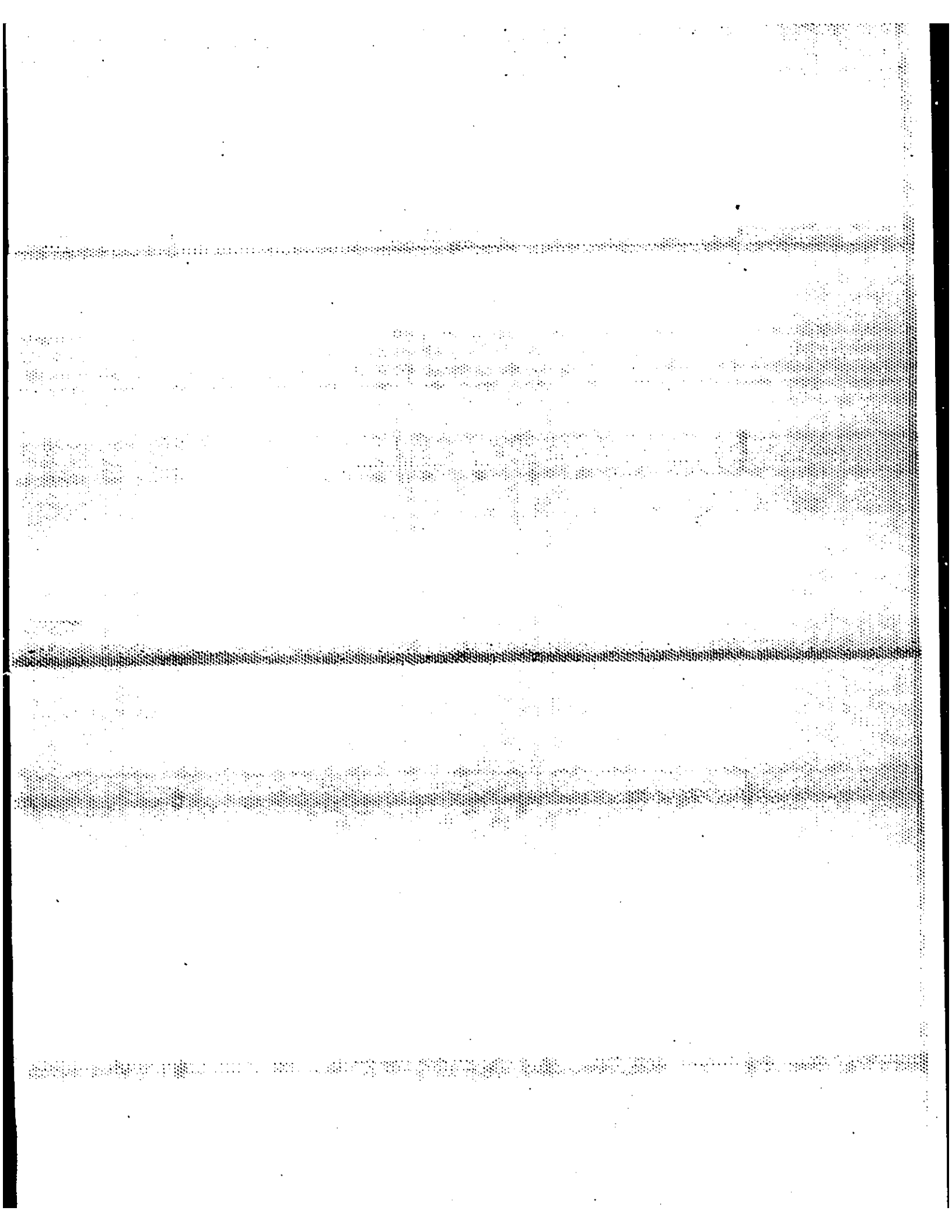


Table 4.2. Bond and Ring Critical Point Data for Arenium Ion Intermediates.

ρ_b (in aux10)

	Para				Meta			
	OH	F	H	CN	OH	F	H	CN
C1-C2	3.112	3.227	3.174	3.128	3.526	3.585	3.457	3.443
C2-C3	3.510	3.470	3.457	3.460	2.796	2.812	2.832	2.839
C3-C4	2.789	2.809	2.832	2.841	2.856	2.839	2.832	2.843
C4-C5	2.786	2.809	2.832	2.841	3.430	3.430	3.457	3.445
C5-C6	3.521	3.470	3.457	3.460	3.223	3.192	3.174	3.201
C6-C1	3.130	3.227	3.174	3.128	3.212	3.285	3.174	3.137
C2-H2	2.996	3.030	3.039	3.050	3.037	3.065	3.082	3.094
C3-H3	3.076	3.087	3.082	3.087	2.779[a]	2.777[a]	2.779[a]	2.778[a]
C4-H4	2.820[a]	2.796[a]	2.779[a]	2.[illegible]5[a]	3.087	3.090	3.082	3.091
C5-H5	3.075	3.087	3.082	3.087	3.043	3.047	3.039	3.044
C6-H6	3.031	3.030	3.039	3.050	3.106	3.099	3.110	3.117
C1-X	3.263	2.737	3.110	3.043	2.944	2.524	3.039	2.943
Ring	0.189	0.192	0.195	0.193	0.194	0.194	0.195	0.194

Bond order n

	para				meta			
	OH	F	H	CN	OH	F	H	CN
C1-C2	1.48	1.60	1.54	1.50	1.94	2.02	1.86	1.84
C2-C3	1.92	1.87	1.86	1.86	1.20	1.22	1.23	1.24
C3-C4	1.20	1.21	1.23	1.24	1.25	1.24	1.23	1.24
C4-C5	1.20	1.21	1.23	1.24	1.82	1.82	1.86	1.84
C5-C6	1.94	1.87	1.86	1.86	1.59	1.56	1.54	1.57
C6-C1	1.50	1.60	1.54	1.50	1.58	1.66	1.54	1.50
Σn	9.24	9.36	9.26	9.20	9.38	9.52	9.26	9.23
Σn(PhX)	9.86	9.98	9.60	9.78	9.86	9.98	9.60	9.78

$\nabla^2\rho_b$ (in au)

	para				meta			
	OH	F	H	CN	OH	F	H	CN
C1-C2	-1.015	-1.120	-1.008	-0.999	-1.151	-1.200	-1.131	-1.105
C2-C3	-1.137	-1.123	-1.131	-1.128	-0.797	-0.811	-0.823	-0.830
C3-C4	-0.804	-0.812	-0.823	-0.828	-0.836	-0.826	-0.823	-0.829
C4-C5	-0.800	-0.812	-0.823	-0.828	-1.120	-1.116	-1.131	-1.130
C5-C6	-1.144	-1.123	-1.131	-1.128	-1.036	-1.023	-1.008	-1.030
C6-C1	-1.049	-1.120	-1.008	-0.999	-1.034	-1.067	-1.008	-0.971
C2-H2	-1.192	-1.246	-1.237	-1.257	-1.246	-1.302	-1.300	-1.327
C3-H3	-1.291	-1.310	-1.300	-1.311	-1.026[a]	-1.030[a]	-1.026[a]	-1.031[a]
C4-H4	-1.050[a]	-1.040[a]	-1.026[a]	-1.027[a]	-1.307	-1.313	-1.300	-1.317
C5-H5	-1.288	-1.310	-1.300	-1.311	-1.244	-1.251	-1.237	-1.249
C6-H6	-1.237	-1.246	-1.237	-1.257	-1.348	-1.347	-1.335	-1.358
C1-X	+0.008	+0.288	-1.335	-1.001	-0.277	+0.224	-1.237	-0.935
Ring	+0.151	+0.154	+0.156	+0.154	+0.155	+0.155	+0.156	+0.155

Table 4.2. (con'd)

	Ellipticities x 10							
	para				meta			
	OH	F	H	CN	OH	F	H	CN
C1-C2	1.556	1.834	1.412	1.596	3.500	3.449	2.773	2.896
C2-C3	3.324	2.964	2.773	2.739	0.563	0.493	0.427	0.454
C3-C4	0.385	0.419	0.427	0.436	0.477	0.467	0.427	0.426
C4-C5	0.380	0.419	0.427	0.436	2.553	2.616	2.773	2.640
C5-C6	3.360	2.964	2.773	2.739	1.599	1.542	1.412	1.534
C6-C1	1.516	1.834	1.412	1.596	1.543	1.847	1.412	1.472
C2-H2	0.333	0.275	0.264	0.271	0.004	0.073	0.136	0.130
C3-H3	0.123	0.129	0.136	0.132	0.310[a]	0.325[a]	0.313[a]	0.327[a]
C4-H4	0.278[a]	0.309[a]	0.313[a]	0.322[a]	0.150	0.140	0.136	0.152
C5-H5	0.102	0.129	0.136	0.132	0.255	0.260	0.264	0.264
C6-H6	0.297	0.275	0.264	0.271	0.217	0.219	0.256	0.226
C1-X	1.983	3.923	0.256	0.041	0.097	0.463	0.264	0.639

[a]. Critical point of a methylenic C-H bond.

different from zero. In agreement with their calculated lengths, these bonds are not reduced to bonds of order one. Their values of $\nabla^2\rho_b$ are also more negative than for a normal single bond ($\nabla^2\rho_b$ = -0.66 au) indicating that charge is concentrated along the bond path and in the interatomic surface to a greater degree than in a C-C bond of order one. The changes in the properties of the remaining C-C bonds parallel their geometry changes which, as noted above for the benzenium ion, result in a geometry close to that for the pentadienyl cation. The properties of the C-C and C-H bonds in this fragment of the benzenium ion are remarkably similar to those for the pentadienyl cation as can be seen from the comparison of values given in Table 4.3. The same delocalization of π charge that lessens the bond order of the terminal bonds of the pentadienyl cation and increases that of its interior bonds is operative over the corresponding carbon framework in the benzenium ion intermediate. The results in Tables 4.2 and 4.3 show that, as found for the bond lengths in the intermediates, Fig.4.1, the properties of a bond are determined primarily by its position relative to the protonated carbon and show only small variations with substituent and with whether the protonation is meta or para to the substituent. Thus for example, bonds C1-C2 and C1-C6 for para substitution are equivalent to bonds C6-C1 and C6-C5 for meta substitution and, with the exception of substitution meta to F, their bond orders lie within ±0.06 of the value for benzenium ion. Thus the pentadienyl fragment is recognizable in terms of its bond properties in all of the intermediates. With the exception of the net charge of the carbon bearing the substituent, the same observation holds for the atomic properties as shown below.

The delocalization of electronic charge in benzene and pentadienyl

cation that is assumed responsible for their extra stability, is reflected in the sums of their C-C bond orders, 9.6 and 6.9 respectively, which exceed the values of 9.0 and 6.0 calculated for a structure with localized bonds. The C-C bond orders of the protonated intermediates, denoted by Σn in Table 4.2, sum to less than their values in Ph-X by amounts ranging from 0.6 in para protonated phenol to 0.4 in the benzenium ion. The sum of the bond orders of just the pentadienyl fragment is, however, sustained or increased over the value for the parent ion, ranging from a value of 6.8 in meta protonated cyanobenzene to a maximum value of 7.1 in meta protonated fluorobenzene.

The NMR investigation of the benzenium ion (Olah et al 1978) was also consistent with a planar ring structure and provided no support for the presence of a C3-C5 bridge to give a homocyclopentadienyl ring system. The presence of a homoallylic bond is made evident in a system's molecular graph as determined by the topology of the charge density (Cremer et al 1983, Gatti et al 1985) and the present results, in agreement with experiment, show that no such bond is present or incipient in either the benzenium cation or its substituted congeners.

The net charges on the atoms in the intermediates, Table 4.4, indicate that just under one-third of the positive charge resides on the CH_2 group. The proportioning of the remaining positive charge around the pentadienyl fragment of the benzenium ion is similar to that in the isolated cation, Table 4.3, with the largest positive charges appearing on the pair of atoms meta to the point of protonation, C2 and C6. C3 and C5, the atoms linked to the protonated carbon which bears a substantial positive charge, possess smaller net charges than their counterparts in the

Table 4.3. Comparison of Bond and Atomic Properties of Benzenium and Pentadienyl Cations.

Bond Properties[a]

Bond	Cation	n	ϵ	$\nabla^2\rho_b$	Bond	Cation	ρ_b	ϵ	$\nabla^2\rho_b$
C1-C2	$C_6H_7^+$	1.54	0.141	-1.008	C1-H	$C_6H_7^+$	0.311	0.026	-1.335
	$C_5H_7^+$	1.54	0.136	-1.029		$C_5H_7^+$	0.308	0.029	-1.296
C2-C3	$C_6H_7^+$	1.86	0.277	-1.131	C2-H	$C_6H_7^+$	0.304	0.026	-1.237
	$C_5H_7^+$	1.89	0.282	-1.152		$C_5H_7^+$	0.302	0.027	-1.219
					C3-H[b]	$C_6H_7^+$	0.308	0.014	-1.300
						$C_5H_7^+$	0.307	0.023	-1.300

a. The numbering of the atoms is as shown in a. Values in au.
b. These are the averages of two inequivalent hydrogens in $C_6H_7^+$.

Atomic Properties

	q(C1)	q(C2)	q(C3)	q(H1)	q(H2)	q(H3)	Q_{zz}(C1)	Q_{zz}(C2)	Q_{zz}(C3)[d]
$C_6H_7^+$	+0.019	+0.149	+0.018	+0.090	+0.054	+0.079	-1.466	-3.641	-1.829
$C_5H_7^+$	+0.010	+0.164	+0.093	+0.067	+0.043	+0.081	-1.308	-3.754	-1.730
$\Delta E(\Omega)$[c]	+7.1	-7.2	-45.0	+5.9	+2.9	-4.1			

c. $\Delta E(\Omega) = E(\Omega)$ for $C_6H_7^+$ - $E(\Omega)$ for $C_5H_7^+$ in kcal/mole.
d. Q_{zz}(C) in au.

Table 4.4 Atomic Charges in Arenium Ion Intermediates.

	q(Q) Para				q(Q) Meta			
Q \ X	OH	F	H	CN	OH	F	H	CN
C1	+0.701	+0.561	+0.019	+0.059	+0.708	+0.647	+0.149	+0.186
C2	+0.132	+0.202	+0.149	+0.168	+0.062	+0.096	+0.018	+0.061
C3	+0.040	+0.036	+0.018	+0.034	+0.135	+0.129	+0.126	+0.128[a]
C4	+0.156	+0.138	+0.126	+0.127[a]	+0.021	+0.025	+0.018	+0.021
C5	+0.032	+0.036	+0.018	+0.034	+0.160	+0.164	+0.149	+0.168
C6	+0.165	+0.202	+0.149	+0.168	+0.038	+0.053	+0.019	+0.044
H2	+0.038	+0.090	+0.054	+0.078	+0.060	+0.113	+0.079	+0.110
H3	+0.073	+0.090	+0.079	+0.092	+0.081	+0.094	+0.082	+0.097[b]
H4	+0.061	+0.082	+0.082	+0.092[b]	+0.084	+0.091	+0.079	+0.095
H5	+0.072	+0.090	+0.079	+0.092	+0.080	+0.068	+0.054	+0.070
H6	+0.073	+0.090	+0.054	+0.078	+0.121	+0.129	+0.090	+0.120
C				+1.146				+1.150
N				-1.257				-1.348
O	-1.309				-1.294			
F		-0.703				-0.708		
H	+0.700		+0.090		+0.676		+0.054	
CH_2	+0.278	+0.302	+0.290	+0.311	+0.297	+0.317	+0.290	+0.322
X	-0.609	-0.703	+0.090	-0.111	-0.618	-0.708	+0.054	-0.198

[a]. Protonated carbon.
[b]. Out-of-plane hydrogen atoms.

pentadienyl cation. With the exception of the charge on the atom bearing the substituent X, the same pattern of alternation in the magnitude of the atomic charges relative to the site of protonation is found for all the substituted intermediates, for both meta and para protonation.

The ^{13}C NMR chemical shifts for benzenium ion in solution (Olah et al 1978) would indicate that the pair of atoms C2 and C6 have greater π densities than do C1 and the pair C3 and C5 in agreement with the Hückel model which assigns 1.0 π charge each to C2 and C6 and 2/3 to each of the other three atoms. The atomic π charges for pentadienyl are in agreement with the Hückel values, equalling 1.00 for C2 and C6, 0.61 for C1 and 0.63 for C3 and C5. Thus on the basis of the Huckel model, C2 and C6 should possess the smallest positive charges in both cations. However, net charges assigned on the basis of the Hückel model err because they ignore the counter polarization of the σ density and C2 and C6 actually possess the largest net charge. The same observations hold for the allyl cation (Slee and MacDougall 1988) where the central atom bears the largest net positive charge in spite of the fact that it has a unit π population as predicted by the Hückel model.

It was demonstrated in Chapter 3 that the quadrupolar polarization of a carbon atom perpendicular to the plane of the phenyl ring, as measured by $Q_{zz}(Q)$, parallels its π population. The more negative the value of $Q_{zz}(Q)$, the greater the extent to which electronic charge is removed from the plane of the ring and concentrated along a z axis passing through a carbon nucleus and perpendicular to the plane of the nuclei. The changes in the quadrupolar polarizations of the ring carbon atoms are found to mimic the alternation in the π populations of the ring carbon atoms induced

by the π donor and abstractor substituents shown in Table 3.7. This is an important observation, for it relates the π model to a property of the total charge distribution. Thus, although π density is not defined in the arenium ion intermediates, one can still determine the quadrupolar polarizations of the carbon atom densities.

As anticipated on the basis of the Hückel model and the calculated π populations, carbons C2 and C6 possess the most negative values of $Q_{zz}(C)$ in the pentadienyl cation, Table 4.3, and the relative magnitude of this moment parallels the π populations of the carbon atoms in this cation. The values of $Q_{zz}(C)$ for the corresponding atoms in the benzenium ion mimic those for the pentadienyl cation and thus the carbon atom densities in the pentadienyl fragment in the benzenium ion exhibit a pattern of π-like quadrupolar polarizations very similar to that found in the parent cation. The data in Table 4.5 show that this same pentadienyl pattern of atomic quadrupolar polarizations is found for all the intermediates, with the nature of the substituent and its location relative to the site of protonation having only a secondary effect on the magnitude of the polarization. The changes in $Q_{zz}(C)$ relative to their values in Ph-X are also given in Table 4.5. The moment for the protonated carbon undergoes the largest decrease in magnitude. The other atoms exhibit changes which reflect the change in the pattern of delocalization from that of a substituted benzene to that for a pentadienyl fragment, the moments for the atoms ortho and para to the protonated carbon decreasing in magnitude and those meta to it increasing in magnitude.

The decrease in the quadrupolar polarization of the carbon atom charge distributions of ≈ 6.7 au on protonation is substantial and

Table 4.5 Quadrupole Moments $Q_{zz}(\Omega)$ of Arenium Ion Intermediates in au.

		Para				Meta			
	Ω \ X	OH	F	H	CN	OH	F	H	CN
	C1	-1.167	-1.372	-1.466	-1.694	-3.236	-3.445	-3.641	-3.750
	C2	-3.909	-3.688	-3.641	-3.585	-2.483	-2.001	-1.829	-1.675
	C3	-1.893	-1.801	-1.829	-1.783	-0.795	-0.852	-0.804	-0.878[a]
T	C4	-0.734	-0.815	-0.804	-0.846[a]	-1.743	-1.753	-1.829	-1.703
O	C5	-1.992	-1.801	-1.829	-1.783	-3.599	-3.585	-3.641	-3.619
T	C6	-3.743	-3.688	-3.641	-3.585	-1.559	-1.535	-1.466	-1.409
A	C				-1.717				-1.265
L	N				0.467				0.041
	O	-0.503				-0.635			
	F		-0.224				-0.214		
	H	+0.026		-0.279		+0.029		-0.271	
	C1	1.656	1.648	1.874	1.792	-0.413	-0.425	-0.301	-0.264
R[b]	C2	-0.154	-0.229	-0.301	-0.455	1.272	1.458	1.511	1.455
E	C3	1.253	1.389	1.511	1.553	2.351	2.338	2.536	2.458[a]
L	C4	2.850	2.638	2.536	2.286[a]	1.841	1.700	1.511	1.429
A	C5	1.149	1.389	1.511	1.553	-0.458	-0.395	-0.301	-0.283
T	C6	-0.128	-0.229	-0.301	-0.455	2.056	1.924	1.874	1.721
I	C				-0.576				-0.124
V	N				0.420				-0.006
E	O	0.221				0.089			
	F		-0.006				0.004		
	H	-0.003		0.042		0.000		0.050	

[a]. Protonated carbon.
[b]. Values relative to those in Ph-X.

represents about one-third of the magnitude of the sum of the $Q_{zz}(C)$ values for the unprotonated species. However, the sum of the $Q_{zz}(C)$ values for the pentadienyl group in the intermediates is, with one exception, slightly larger than the value of 12.28 au found for the pentadienyl cation itself. If one uses the value of $Q_{zz}(C)$ as a measure of the π-like population of a carbon atom, then the pentadienyl fragments in the intermediates have nearly the same pattern of π populations as does the parent ion. This density is similarly delocalized over the C-C bonds as is seen by comparing the values of bond ellipticities for the intermediates in Table 4.2 with the values for corresponding bonds in pentadienyl in Table 4.3. There are only a few differences of significance, for phenol and fluorobenzene.

The atomic and bond properties show that a perturbed pentadienyl fragment is present in each of the arenium ion intermediates, with the largest single difference from the parent cation being the net charge on the carbon bearing the substituent atom X. Table 4.6 lists the changes in the atomic populations of the atoms in Ph-X on protonation. The pattern of changes in the properties of the ring atoms with respect to the site of protonation, are remarkably similar for all the intermediates, for both para and meta substitution. The replacement of the delocalized electronic structure of the phenyl ring with that of the pentadienyl group upon protonation of a substituted benzene, is the single most important driving force in determining the changes in the atomic properties of the benzene ring and the most important factor in determining the change in the properties of a given carbon atom is its position relative to the protonated carbon atom rather than to the carbon bearing the substituent. For example, with the single exception of $|\Delta N(C)|$ for C1 exceeding that for

Table 4.6 Atomic Populations of Arenium Ion Intermediates [Ph-X]H+ Relative to Their Values in Ph-X.

	Para				Meta			
Q \ X	OH	F	H	CN	OH	F	H	CN
C1	-0.161	-0.092	0.024	0.067	-0.168	-0.178	-0.106	-0.060
C2	-0.077	-0.102	-0.106	-0.096	-0.007	0.004	0.025	0.011
C3	0.012	0.018	0.025	0.029	-0.083	-0.074	-0.083	-0.065[a]
C4	-0.101	-0.082	-0.083	-0.078[a]	0.034	0.031	0.025	0.028
C5	0.016	0.019	0.025	0.029	-0.112	-0.109	-0.106	-0.105
C6	-0.075	-0.102	-0.106	-0.096	0.052	0.047	0.024	0.028
H2	-0.090	-0.094	-0.097	-0.082	-0.112	-0.117	-0.122	-0.114
H3	-0.110	-0.119	-0.122	-0.115	-0.118	-0.123	-0.125	-0.120[b]
H3					0.919	0.906	0.918	0.903[c]
H4	-0.105	-0.118	-0.125	-0.115[b]	-0.128	-0.127	-0.122	-0.118
H4	0.939	0.918	0.918	0.908[c]				
H5	-0.108	-0.119	-0.122	-0.115	-0.096	-0.097	-0.097	-0.093
H6	-0.088	-0.094	-0.097	-0.082	-0.136	-0.133	-0.133	-0.124
C				-0.033				-0.037
N				-0.225				-0.134
O	0.017				0.002			
F		-0.035				-0.030		
H	-0.064		-0.133		-0.040		-0.097	

[a]. Protonated carbon.
[b]. Out-of-plane hydrogen, change relative to H of Ph-X.
[c]. Out-of-plane hydrogen, change relative to proton.

C2 in para protonated phenol, all the carbon atoms meta to the site of protonation undergo the largest loss in charge of the atoms in this fragment, corresponding to atoms C2 and C5 bearing the largest positive charge in the pentadienyl fragment, Table 4.3. The perturbations of the carbon atom π densities caused by the replacement of H by X are reflected in the values of $Q_{zz}(C)$ (see Chap.1 and 3), resulting in an alternating pattern of increases and decreases which mimic the charge alternation associated with the classical resonance structures for ortho-, para- and meta-directing substituents. The changes in the values of $Q_{zz}(C)$ caused by protonation are, however, of larger magnitude and the final pattern of quadrupolar polarizations is that associated with a pentadienyl fragment, erasing the pattern created by the substituent.

4.3. Energies of Formation of Protonated Monosubstituted Benzenes

The energy changes for reaction Eq.[4.1], ΔE_1, are listed in Table 4.7 together with the individual energies of protonation, ΔE_p. The calculated protonation energy of benzene is in fair agreement with the experimental value of 183 kcal/mole (Haney and Franklin 1969). The intermediate formed from para protonation of phenol is substantially more stable than protonated benzene while para-protonated fluorobenzene is, within the calculational uncertainties, found to have the same stability as the benzenium ion itself. These results are also in agreement with experiment, as are the calculated values of ΔE_1 for the same two substituents, the corresponding measured values of $\Delta H^o_1(600K)$ being -13.4 and -1.3 kcal/mol for X=OH and F, respectively (Lau and Kebarle 1976). Para-protonated cyanobenzene is less stable than protonated benzene. To

Table 4.7 Energies of Formation and Group Contributions for Arenium Ions in kcal.

	Para				Meta			
	OH	F	H	CN	OH	F	H	CN
ΔE_1	-17.0	-1.1	0.0	+17.6	+3.8	+10.4	0.0	+16.2
ΔE_p	-213.5	-197.6	-196.5	-178.9	-192.7	-186.1	-196.5	-180.3
$\Delta E(X)$	-41.2	-19.0	+32.7	+58.9	+3.4	+5.4	+25.1	+46.4
$\Delta E(CH_2)$	-327.1	-322.5	-313.9	-312.4	-313.3	-311.9	-313.9	-315.4
$\Delta E(\cap^a)$	+154.8	+143.9	+84.7	+74.6	+117.2	+120.4	+92.3	+88.7
$\Delta N(X)$	-0.047	-0.035	-0.133	-0.258	-0.038	-0.030	-0.097	-0.171
$\Delta N(CH_2)$	+0.733	+0.718	+0.710	+0.715	+0.718	+0.709	+0.710	+0.718
$\Delta N(\cap^a)$	-0.681	-0.684	-0.576	-0.461	-0.673	-0.679	-0.612	-0.547

a. $\cap$ denotes the pentadienyl fragment less the substituent group X.

the extent that the transition state for the electrophilic substitution reaction resembles the protonated intermediate (Lowry and Richardson 1981, Streitwieser and Heathcock 1985), the relative energies of the protonated intermediates should reflect the relative activating-deactivating abilities of the substituents. The theoretical results agree with the experimental findings with the possible exception of that for the F substituent. The intermediate formed from para protonation of phenol is substantially more stable than protonated benzene while para protonated fluorobenzene is, within the calculational uncertainties, found to have the same stability as the benzenium ion itself. Para protonated cyanobenzene is less stable than protonated benzene. The finding that the meta protonated intermediates of all the substituents are less stable than protonated benzene - the instabilities increasing in the order OH, F, CN - is in agreement with the reaction energy profiles given by Streitwieser and Heathcock (1985). The para-directing groups OH and F are the most discriminating between para and meta substitution with relatively large energy differences between the two possibilities. The CN group destabilizes the phenyl group upon both para and meta protonation, slightly more so in the former case than in the latter.

Assuming that the energies of the intermediates provide a relative measure of the energies of activation, the present results predict the hydroxyl group to be para directing and activating relative to benzene, F to be para directing but not activating, while the cyano group is deactivating and meta directing. The directing and activating-deactivating ability of a substituent X is generally accounted for in terms of the placement of the positive charge relative to X in the resonance structures

that can be drawn for the transition state or intermediate (Lowry and Richardson 1981, Streitwieser and Heathcock 1985). Substitution ortho or para to X places the positive charge at the meta positions or at the ipso carbon while meta substitution places it at the ortho or para positions. Thus ortho- and para-directing substituents stabilize the transition state relative to benzene since the positive charge in one of the structures is placed at the ipso carbon, adjacent to a π donating group. These same structures destabilize a meta directing substituent and there are no stabilizing structures for meta substitution. Reference to Table 4.8 which gives the net charges on C1 and X1 in Ph-X, where X1 is the substituent atom bonded to C1 in Ph-X, shows that a developing positive charge at C1 would be adjacent to a large negative charge on X1 for the ortho-para directors O^-, NH_2, OH and F with the magnitude of the charge decreasing in the same order, and to a large positive charge on X1 for the meta directors. The carbon atom of the methyl group, which is both σ and π donating, bears a relatively small positive charge, but only in this case does the ipso carbon, the site of the developing positive charge, bear a net negative charge. It should also be noted however, that C1 itself, where the resonance structure places the positive charge, bears a large positive charge for X = O^- to F and a smaller positive charge for X = CN and NO_2.

The changes in the energies of the atoms in Ph-X upon protonation are given in Table 4.9, together with the corresponding values for the proton. These energy changes are to be discussed in conjunction with the corresponding population changes given in Table 4.6. The basin of the newly attached proton attracts $\approx$ 0.9e and this charge is lost primarily by

Table 4.8. Charges and Atomic First Moments of C1 and Bonded Substituent Atom X1 in Ph-X and [Ph-X]H+.

		O-	NH_2	OH	F	CH_3	H	CN	NO_2
Ph	μ(X1)[a]	+0.465	+0.199	+0.180	+0.369	+0.032	-0.098	+1.236	+0.940
↓	q(X1)	-1.436	-1.429	-1.292	-0.738	+0.246	-0.043	+1.113	+0.327
X	μ(C1)	+0.783	+0.775	+0.800	+0.800	-0.046	-0.006	+0.341	+0.641
	q(C1)	+1.055	+0.545	+0.540	+0.469	-0.010	+0.043	+0.126	+0.323
Para	Δμ(X1)[b]			+0.126	+0.118		+0.005	-0.183	
	Δμ(C1)[b]			+0.201	+0.204		+0.073	+0.033	
Meta	Δμ(X1)[b]			+0.028	+0.043		+0.004	-0.160	
	Δμ(C1)[b]			-0.055	-0.066		-0.171	-0.185	

a. The negative end of μ(X1) and of μ(C1) are directed towards the ring centre in all cases but for X=CH_3 where μ(C1) is directed towards methyl and for X=H where both μ(C1) and μ(H) are directed away from the ring.
b. Δμ(Ω) = μ(Ω) of [Ph-X]H+ - μ(Ω) of Ph-X. μ and Δμ in au.

Table 4.9. Atomic Energies of Arenium Ion Intermediates [Ph-X]H$^+$ Relative to Their Values in Ph-X.

	Para				Meta			
Q \ X	OH	F	H	CN	OH	F	H	CN
C1	0.1140	0.0619	-0.0478	-0.0467	0.0973	0.1025	0.0497	0.0331
C2	0.0242	0.0425	0.0497	0.0502	-0.0215	-0.0336	-0.0475	-0.0373
C3	-0.0406	-0.0469	-0.0475	-0.0484	0.0169	0.0094	0.0121	0.0037[a]
C4	0.0184	0.0045	0.0121	0.0133[a]	-0.0623	-0.0570	-0.0475	-0.0457
C5	-0.0390	-0.0469	-0.0475	-0.0484	0.0532	0.0497	0.0497	0.0500
C6	0.0451	0.0425	0.0497	0.0502	-0.0688	-0.0629	-0.0478	-0.0438
H2	0.0357	0.0398	0.0400	0.0343	0.0440	0.0489	0.0492	0.0474
H3	0.0440	0.0484	0.0492	0.0466	0.0665	0.0696	0.0693	0.0683[b]
H3					-0.5826	-0.5761	-0.5817	-0.5747[c]
H4	0.0557	0.0649	0.0693	0.0660[b]	0.0511	0.0513	0.0492	0.0482
H4	-0.5954	-0.5833	-0.5817	-0.5771[c]				
H5	0.0435	0.0484	0.0492	0.0466	0.0404	0.0408	0.0400	0.0388
H6	0.0362	0.0398	0.0400	0.0343	0.0539	0.0534	0.0521	0.0503
C				-0.0025				0.0165
N				0.0963				0.0575
O	-0.1100				-0.0220			
F		-0.0303				0.0086		
H	0.0444		0.0521		0.0274		0.0400	

[a]. Protonated carbon.
[b]. Out-of-plane hydrogen, change relative to the H of Ph-X.
[c]. Out-of-plane hydrogen, change relative to proton.

the atoms of the phenyl group. In all cases the carbon atoms which lose charge, the protonated carbon and those meta to it, are destabilized relative to the unprotonated molecule, while those in positions ortho and para to the site of protonation gain charge, and are stabilized. The ipso carbon for para protonation of phenol and fluorobenzene loses electronic charge and is destabilized. These are the only exceptions to the general result that the pattern of changes in properties of the ring carbons are similar and are determined solely by their position relative to the protonated carbon, independent of the nature and position of X. The pattern of energy loss and gain of the ring atoms linked to the protonated carbon is, like the other atomic properties, determined by the pattern of properties found in the pentadienyl cation. The smallest energy change is for the carbon atom which is protonated. It is only slightly destabilized in spite of relatively large loses in charge because of the formation of a bond to the proton. With the exception of C1 for X = OH and F, stabilization for the remaining carbon atoms is associated with a decrease in the extent of their quadrupolar polarization, while destabilization accompanies an increase in the magnitude of this moment, (cp. Table 4.5). All of the hydrogen atoms of Ph-X lose charge in the formation of [Ph-X]H$^+$ and all are destabilized. The added proton, which attracts over 0.9e to its basin, is stabilized by an amount greater than the energy of a free H atom, -0.5 au. Table 4.7 lists the changes in energies of the substituent X, $\Delta E(X)$, and of the C-H group which is protonated, $\Delta E(CH_2)$. The remaining group, which consists of the pentadienyl fragment without the substituent X, including X = H is labelled Ω and its energy change is $\Delta E(\Omega)$. The energy change $\Delta E(CH_2)$ is for the reaction of a fragment of the phenyl group

$$C\text{-}H|_{Ph\text{-}X} + H^{+} \rightarrow CH_2|_{[Ph\text{-}X]H^{+}} \qquad \Delta E(CH_2) \qquad [4.2]$$

a quantity whose determination the theory of atoms in molecules makes possible. Only in para protonation of phenol and fluorobenzene is the substituent stabilized. The relative values of $\Delta E(X)$ parallel the ordering of the values for ΔE_p and ΔE_1. The energies of reaction [4.2] show that the main source of stabilization of the intermediate comes from charge transferred to the proton and the associated formation of the C-H bond. The relative values of this energy change also parallel the ordering of ΔE_p and ΔE_1. The population and energy changes for reaction [4.2] show relatively small variations with X for meta protonation, $\Delta N(CH_2)$ and $\Delta E(CH_2)$ equalling 0.714 ± 0.005e and -314 ± 2 kcal/mole, respectively, while for para protonation the corresponding values are 0.72 ± 0.01e and-319 ± 8 kcal/mole. The charge transferred to CH_2 to bind the proton is taken primarily from the Π group. The energy of this group is increased for all substituents and its magnitude is in the reverse order to that for ΔE_1 and ΔE_p.

The ipso carbon atom undergoes the largest perturbation when H is replaced by X to form Ph-X and the change in its properties are important to understanding the values of ΔE_1, Table 4.8. There is a relatively large negative charge on X1 and a substantial positive charge on C1 for the -I, +M groups. Both atomic densities are polarized towards the ring centre, in a direction counter to the charge transfer. Atomic densities generally polarize in a direction counter to the direction of a substantial interatomic charge transfer in response to the electric field generated by the transfer of charge. The polarization of X1 in CN and NO_2 is large because of the significant transfer of charge within these substituents,

q(N)=-1.48 in CN and q(O)=-0.53 in NO_2. The moment of C1 is reduced in these molecules as X1 bears a positive charge, but its density is still polarized away from X1 because of the net negative charge on the X group and the large polarization of X1. There is charge transfer between the atoms of the -I,-M groups with X1 bearing the positive charge and C1 bears a smaller positive charge in these cases. The atoms are again polarized towards the ring centre, with the X1 atoms exhibiting particularly large moments. On para protonation, the polarizations of the +M groups OH and F and of C1 towards the ring are increased while for all other cases the polarizations are either decreased or increased by only a small amount. The atomic moment of N in CN changes very little when cyanobenzene is protonated, decreasing by 0.023 au for para protonation and remaining essentially unchanged for meta protonation.

The para protonation of phenol results in the largest transfer of charge to the CH_2 group and to its greatest stabilization. This same reaction yields the smallest net destabilization of the combined X and ∩ groups because of the stabilization of the oxygen atom of the OH group. One can draw a resonance structure for a +M group wherein the substituent bears a positive charge and a negative charge is placed ortho or para to it. In actual fact the loss of charge from OH or F is very small in forming their intermediates, smaller than for the -M groups H and CN. There is, however, a significant increase in the polarization of these two atoms towards the ipso carbon, Table 4.8, and F, like OH, is stabilized in the protonation reaction. While the CH_2 group in para protonated fluorobenzene is ≈ 9 kcal/mole more stable than in the benzenium ion, the net destabilization of the remainder of the system is greater by 8

kcal/mole and the two molecules are equally susceptible to para protonation.

The transfer of charge caused by the para protonation of cyanobenzene results in the smallest stabilization of the CH_2 group and in the largest destabilization of the remainder of the molecule. The largest contributing factor to the destabilization of either of the protonated cyanobenzenes relative to the other molecules is the destabilization of the cyano group itself. The cyano group donates a larger fraction of the charge required to form the bond to the proton than do any of the other substituents, the charge coming primarily from the N atom. This decreases the demand for charge from the ∩ group, but the net effect is the largest destabilization of the ∩-X group. One notes that this behaviour of the cyano group is in direct contradiction with the resonance structure for a -M group which localizes a negative charge on a terminal atom of X and places a positive charge on a ortho or para carbon. The +M groups such as OH and F and the -M groups such as CN respond in opposite ways to the demand for charge, both in contradiction to their dominant resonance forms in the reactant benzene: the +M groups donate almost no charge to the ring, but increase their polarization towards the ring carbon while the -M groups donate significant amounts of charge to the ring and have their polarization towards the ring carbon decreased, Table 4.8. The single largest difference in energies contributing to the differing behaviour of phenol and cyanobenzene towards para protonation is the 100 kcal/mole separating the stabilization of the OH group and the destabilization of the CN group.

4.4. Conclusions

The pattern of charge alternation, including the changes in the quadrupolar polarizations of the ring carbons which parallel their "π" populations, is a function not of the substituent, but of the position of the carbon relative to the site of protonation, a position which then determines its identity in the pentadienyl fragment, the overriding factor in determining its properties. The ortho- and para-directing groups lower the energy of formation of the protonated intermediate by increasing the transfer of electronic charge from the ∩ group to the newly formed C-H bond. The increase in energy of the ∩ group which accompanies this charge transfer is mediated by the increase in the stability of the substituent group itself. Meta directing groups lose charge to the ring and are destabilized. They are destabilized by an amount exceeding the decrease in the destabilization of the ∩ group resulting from the reduction in its loss of charge. The properties of the charge distribution of a benzene ring do determine the initial approach of an electrophile. Cheney et al. (1988) and calculations performed by us have shown that the hydrogen of an acid such as HF approaches the periphery of the benzene ring along an axis tilted toward the ring axis as predicted by the Laplacian of the charge distribution, to form a weak π complex. This will be discussed in detail in Chapter 5. The transformation of this complex into the protonated intermediate involves the charge rearrangements described above and these are different from the pattern of charges found in the reactant molecules which determine the initial site of attack, and they involve much greater changes in the energy.

A4.1 Optimised Geometries[a] of Protonated Benzene for Various Basis Sets.

	3-21G	6-31G	6-31G*
r(C1C2)	1.4066	1.4097	1.4099
r(C2C3)	1.3520	1.3579	1.3529
r(C3C4)	1.4788	1.4763	1.4776
r(C1H1)	1.0732	1.0739	1.0761
r(C2H2)	1.0694	1.0700	1.0726
r(C3H3)	1.0726	1.0733	1.0751
r(C4H4)	1.0958	1.0946	1.0937
(C6C1C2)	123.10	123.02	123.66
(C1C2C3)	119.04	118.97	118.53
(C2C3C4)	121.64	121.64	121.78
(C3C4C5)	115.53	115.76	115.73
(H1C1C2)	118.45	118.49	118.17
(H2C2C1)	119.64	119.74	119.79
(H3C3C4)	117.63	117.72	117.54
(H4C4H4)	103.82	102.77	103.38
-Energy	229.7222	230.9300	231.0147

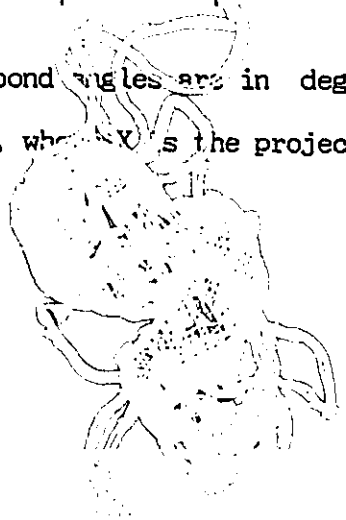
[a]. Bond lengths are in Å, bond angles are in degrees, and energies are in au.

A4.2 Optimised Geometries[a] of Protonated Fluorobenzene for Various Basis Sets.

	para			meta		
	3-21G	4-31G	6-31G	3-21G	4-31G	6-31G
r(C1C2)	1.4036	1.4018	1.4056	1.3411	1.3417	1.3463
r(C2C3)	1.3453	1.3486	1.3536	1.4761	1.4728	1.4755
r(C3C4)	1.4832	1.4777	1.4802	1.4757	1.4721	1.4749
r(C4C5)	1.4832	1.4777	1.4802	1.3566	1.3582	1.3630
r(C5C6)	1.3453	1.3486	1.3536	1.3987	1.3991	1.4034
r(C6C1)	1.4036	1.4018	1.4056	1.4010	1.3997	1.4036
r(C2H2)	1.0690	1.0684	1.0697	1.0716	1.0712	1.0726
r(C3H3)	1.0724	1.0716	1.0729	1.0966	1.0947	1.0952
r(C4H4)	1.0942	1.0932	1.0937	1.0725	1.0719	1.0731
r(C5H5)	1.0724	1.0716	1.0729	1.0692	1.0684	1.0696
r(C6H6	1.0690	1.0684	1.0697	1.0729	1.0727	1.0740
r(C1F)	1.3127	1.3206	1.3266	1.3447	1.3530	1.3580
(C2C1C6)	124.1496	125.0302	125.2792	120.4967	121.4869	121.6861
(C1C2C3)	118.4185	117.7635	117.5498	120.6351	119.9738	119.8560
(C2C3C4)	121.9492	121.9444	121.9870	115.8844	116.0362	116.0710
(C3C4C5)	115.1150	115.5540	115.6472	121.4822	121.7594	121.8511
(C4C5C6)	121.9492	121.9444	121.9870	118.9756	118.9247	118.8860
(C5C6C1)	118.4185	117.7635	117.5498	122.5260	121.8190	121.6498
(H2C2C1)	118.7165	119.0999	119.2218	120.1952	120.4895	120.5148
(H3C3C4)	117.7140	117.7514	117.7757	b121.9613	b121.8536	b121.7878
(H4C4C5)	b122.4425	b122.2230	b122.1764	120.7799	120.6013	120.5474
(H5C5C4)	117.7140	117.7514	117.7757	121.3957	121.3918	121.4181
(H6C6C1)	118.7165	119.0999	119.2218	117.5691	117.8387	117.9425
(H4C4H4)	104.2692	103.1512	103.0706			
(H3C3H3)				103.6468	102.6716	102.6147
(FC1C2)	117.9252	117.4849	117.3604	121.6429	120.9952	120.9303
-Energy	328.04291	329.40791	329.74655	328.02776	329.39535	329.73421

a. Bond lengths are in A, bond angles are in degrees, and energies are in au.

b. This is the angle X-C-C, where X is the projection of two hydrogens on the symmetry plane.



A4.3 Optimised Geometries[a] of Protonated Phenol for Various Basis Sets.

	para			meta		
	3-21G	4-31G	6-31G	3-21G	4-31G	6-31G
r(C1C2)	1.4250	1.4222	1.4257	1.3457	1.3464	1.3507
r(C2C3)	1.3380	1.3410	1.3459	1.4794	1.4752	1.4778
r(C3C4)	1.4870	1.4809	1.4833	1.4726	1.4688	1.4716
r(C4C5)	1.4883	1.4819	1.4843	1.3577	1.3585	1.3633
r(C5C6)	1.3359	1.3392	1.3442	1.3941	1.3950	1.3991
r(C6C1)	1.4244	1.4201	1.4235	1.4163	1.4131	1.4172
r(C2H2)	1.0711	1.0707	1.0720	1.0725	1.0721	1.0734
r(C3H3)	1.0725	1.0717	1.0731	1.0961	1.0943	1.0948
r(C4H4)	1.0922	1.0911	1.0917	1.0726	1.0719	1.0732
r(C5H5)	1.0721	1.0714	1.0728	1.0693	1.0686	1.0698
r(C6H6)	1.0684	1.0682	1.0696	1.0726	1.0724	1.0739
r(C1O)	1.3016	1.3015	1.3061	1.3587	1.3557	1.3592
r(OH)	0.9714	0.9566	0.9559	0.9662	0.9525	0.9517
(C2C1C6)	121.4842	122.1342	122.3142	118.5157	119.1994	119.3230
(C1C2C3)	119.5327	119.0566	118.8930	121.5779	121.1310	121.0527
(C2C3C4)	122.5045	122.4380	122.4709	116.0593	116.2109	116.2529
(C3C4C5)	114.3646	114.7759	114.8537	121.2356	121.3835	121.4553
(C4C5C6)	122.5087	122.4391	122.4608	118.9544	118.9346	118.8818
(C5C6C1)	119.6053	119.1562	119.0074	123.6571	123.1406	123.0343
(H2C2C1)	119.2600	119.4306	119.5421	121.2442	121.3317	121.3136
(H3C3C4)	117.4651	117.5469	117.5882	121.3231	121.2087	121.1668[b]
(H4C4C5)	122.9525	122.8227	122.7764[b]	120.8304	120.7409	120.6717
(H5C5C4)	117.3724	117.4516	117.4916	121.3908	121.3845	121.3869
(H6C6C1)	117.3393	117.7086	117.8177	116.1100	116.5382	116.5835
(H4C4H4)	104.9939	103.8318	103.7574			
(H3C3H3)				103.8600	102.8343	102.8042
(OC1C2)	122.5132	122.0792	122.0007	127.6796	126.9737	126.9703
(HOC1)[c]	119.1064	120.3644	120.2721	116.5598	118.1976	118.2374
-Energy	304.18746	305.45211	305.77235	304.15619	305.42392	305.74446

a. Bond lengths are in Å, bond angles are in degrees, and energies are in au.
b. This is the angle X-C-C, where X is the projection of two hydrogens on the symmetry plane.
c. The H-O bond is on the protonated side of the ring.

A4.4 Optimised Geometries[a] of Protonated Cyanobenzene for Various Basis Sets.

	para			meta		
	3-21G	4-31G	6-31G	3-21G	4-31G	6-31G
r(C1C2)	1.4122	1.4113	1.4148	1.3557	1.3573	1.3618
r(C2C3)	1.3517	1.3530	1.3577	1.4766	1.4718	1.4740
r(C3C4)	1.4773	1.4720	1.4742	1.4772	1.4720	1.4747
r(C4C5)	1.4773	1.4720	1.4742	1.3545	1.3558	1.3605
r(C5C6)	1.3517	1.3530	1.3577	1.4009	1.3997	1.4038
r(C6C1)	1.4122	1.4113	1.4148	1.4152	1.4149	1.4187
r(C2H2)	1.0694	1.0687	1.0699	1.0729	1.0724	1.0736
r(C3H3)	1.0726	1.0718	1.0731	1.0965	1.0947	1.0952
r(C4H4)	1.0964	1.0946	1.0952	1.0728	1.0721	1.0734
r(C5H5)	1.0726	1.0718	1.0731	1.0694	1.0686	1.0699
r(C6H6)	1.0694	1.0687	1.0699	1.0736	1.0731	1.0744
r(C1C7)	1.4163	1.4194	1.4247	1.4257	1.4271	1.4313
r(C7N)	1.1399	1.1416	1.1461	1.1383	1.1404	1.1449
(C2C1C6)	122.8586	122.6080	122.7398	118.9762	118.8552	118.8696
(C1C2C3)	118.9331	119.0079	118.8923	121.4768	121.4858	121.4952
(C2C3C4)	121.9303	121.9059	121.8980	115.7601	115.9603	115.9884
(C3C4C5)	115.4146	115.5644	115.6796	121.6151	121.5873	121.6304
(C4C5C6)	121.9303	121.9059	121.8980	119.1948	119.1540	119.1095
(C5C6C1)	118.9331	119.0079	118.8923	122.9770	122.9574	122.9069
(H2C2C1)	119.3023	119.4025	119.4677	120.4269	120.4406	120.3878
(H3C3C4)	117.7901	117.8393	117.8898	122.2834	122.2490	122.1989[b]
(H4C4C5)	122.2927	122.2178	122.1602[b]	120.7248	120.6879	120.6399
(H5C5C4)	117.7901	117.8393	117.8898	121.3852	121.3305	121.3356
(H6C6C1)	119.3023	119.4025	119.4677	117.9957	118.0867	118.1270
(H4C4H4)	103.7198	102.6294	102.5786			
(H3C3H3)				103.6710	102.6221	102.5600
(C7C1C2)	118.5707	118.6960	118.6301	119.5019	121.5133	121.4879
(NC7C1)	180.0000	180.0000	180.0000	180.0418	180.1184	180.1560[c]
-Energy	320.91534	322.25032	322.59164	320.91671	322.25214	322.59348

a. Bond lengths are in A, bond angles are in degrees, and energies are in au.
b. This is the angle X-C-C, where X is the projection of two hydrogens on the symmetry plane.
c. The N is bent away from the side of protonation.

A4.5 Optimised Geometry[a] of Pentadienyl Cation With 6-31G Basis Set.

r(C1C2)	1.4099	(C2C1C6)	124.38
r(C2C3)	1.3516	(C1C2C3)	120.56
r(C1H6)	1.0775	(H7C2C1)	118.93
r(C2H7)	1.0719	(H8C3C2)	121.34
r(C3H8)	1.0724	(H9C3C2)	122.17
r(C3H9)	1.0744		

a. Bond lengths are in A and bond angles are in degrees.

CHAPTER 5

HYDROGEN-BONDED COMPLEXES OF UNSATURATED HYDROCARBONS

An enormous amount of research has been directed towards understanding the phenomenon of hydrogen bonding because of the importance of hydrogen bonds in nature (reviews include Dyke 1984, Schuster et al 1976, Joesten and Schaad 1974, Vinogradov and Linnell 1971, Pimentel and McClellan 1960). One of my former colleagues, Dr.Marshall Carroll worked on typical hydrogen bond complexes, with HF as acid in the base-acid complexes, by using the theory of atoms in molecules, and this subject became part of his Ph.D. thesis (1989). It has been recognized for many years that hydrogen bonds also exist with the π electrons of unsaturated hydrocarbons acting as the proton acceptors. To distinguish this type of hydrogen bond from the classical one is one of the aims of this chapter. Furthermore, the investigation of the nature of the bonding between an unsaturated hydrocarbon and a Lewis acid is an important aspect of chemical reactivity, since the attack of a π-system by an electrophilic reagent is the first step in a number of organic reaction mechanisms. For the dual purposes just mentioned, three complexes of C_6H_6-HF at energy minima on the potential surface of the benzene and hydrogen fluoride reaction and one for each of C_2H_2-HF and C_2H_4-HF have been studied in this chapter.

In Chapters 3 and 4, the directing ability of a substituent and its ability in activating/deactivating a phenyl group in electrophilic aromatic substitution reactions (Bader and Chang 1989a) and the activation energy of electrophilic substitution for meta- and para-protonated substituted benzenes (Bader and Chang 1989b) have been investigated. As a consequence, the study of electrophiles reacting with aromatic molecules is another aim

of the present chapter.

Section 5.1 shows that the hydrogen-bond complexes of unsaturated hydrocarbons with hydrogen fluoride appear tc exhibit special structures. A detailed comparison of the properties of classical (σ-type) and unsaturated hydrocarbon (π-type) hydrogen-bond complexes is given in Section 5.2.

5.1 Geometries and Structures

For many years, both experimental (Read and Flygare 1982, Shea and Flygare 1982, Baiocchi et al 1983, Andrews et al 1985) and theoretical (Pople et al 1982, Frisch et al 1983, Sapse and Jain 1984, Bredas and Street 1988, Cheney et al 1988) efforts have been devoted to the study of the hydrogen bond between unsaturated hydrocarbons and hydrogen fluoride. The T-shaped structures of C_{2v} symmetry with the hydrogen of HF pointing at the centre of the carbon-carbon bond of ethylene and acetylene were obtained from both gas phase rotational spectra (Read and Flygare 1982, Shea and Flygare 1982) and ab initio calculations at the Hartree-Fock level with the 6-31G* basis by Pople et al. (1982) and Frisch et al. (1983). For the C_6H_6-HF complex, a C_{6v} symmetrical average structure with the proton of HF pointing toward the centre of the benzene ring has been obtained by Baiocchi et al. (1983) from the molecular beam electric resonance technique, although its large amplitude motions make it difficult to determine an equilibrium structure. Fixing the geometries of both isolated benzene and HF and performing the point-by-point optimization by varying the distance between the H of HF and the centre of the benzene ring or the C-C bond centre with the 6-31G* basis set, Sapse and Jain (1984) compared

the energies of two C_6H_6-HF complexes, with C_{6v} and C_s symmetry, respectively, and found the former is 0.2 kcal/mol more stable than the latter. In the Hartree-Fock calculation with the 3-21G basis set, Bredas and Street (1988) optimized the intermolecular distance while keeping the molecular geometry frozen for three C_6H_6-HF complexes with HF sitting on the top of the centre of the ring, the C-C bond centre and one of the carbon atoms; their MP2/6-31G*//3-21G calculations also confirmed the experimental result.

However, some questions are raised by these conclusions: how do these π-type complexes form hydrogen bonds? Is the hydrogen of HF really bonded to the centre of the benzene ring in the C_{6v} C_6H_6-HF complex in spite of the fact that it is not the region of charge concentration (see Chapter 3)? What properties, molecular or bond or atomic, decide π-type hydrogen bond strength? The theory of atoms in molecules provides tools for solving these problems.

All the calculations in the present study were carried out with the Hartree-Fock-Roothaan procedure on both the VAX-8650 and IBM-4381 computer with the Gaussian 86 program (Frisch ed al 1986).

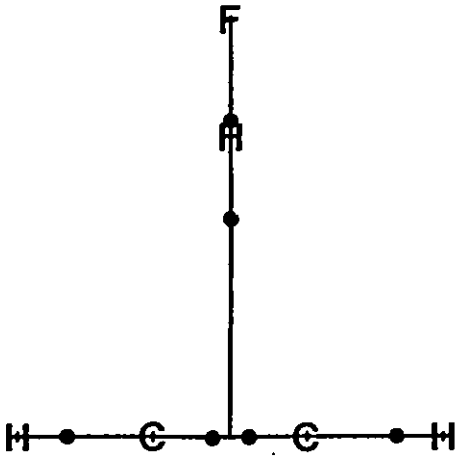
The π-type hydrogen bonded complexes, C_2H_2-HF, C_2H_4-HF and C_6H_6-HF, structures A to E in Figure 5.1, and the monomers, C_2H_2, C_2H_4, C_6H_6 and HF as well, have been fully optimized with the 6-31G basis set, while the 6-31G** basis set was employed for the final calculations at the equilibrium geometries. To our knowledge, this is the first time that the C_6H_6-HF complexes have been fully optimized with a basis set of such high level. Three minima were obtained for C_6H_6-HF complexes, the acidic hydrogen pointing toward the centre of one of the carbon-carbon bonds (abbreviated

as $C_6H_6HF(C)$), one of the carbon atoms ($C_6H_6HF(D)$), and the centre of the ring ($C_6H_6HF(E)$) of benzene separately (see C, D and E in Figure 5.1). Only the T-shaped geometry was obtained for C_2H_2-HF and C_2H_4-HF; the attempt to find an unsymmetrical minimum energy geometry wherein H is linked to a carbon atom failed, and only the symmetrical T-shaped geometries were found for both of these complexes. The results confirm the conclusion made by Sapse and Jain (1984) that there is no bonding between the H of HF and the carbons of acetylene and ethylene.

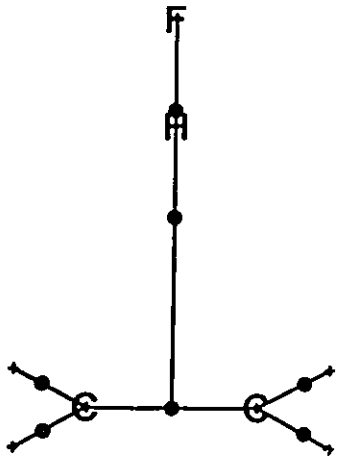
Of three π-type hydrogen bond complexes of C_6H_6-HF, E has the lowest energy (see Table 5.1), though the energy differences between the three structures are so small that one concludes that the three geometries co-exist in the gas phase. Baiocchi et al. (1983) point out that according to experimental results, it is possible that the C_6H_6-HF complex has an asymmetric structure because the C_{6v} symmetric structure which they observed is only an average one; "Rotation of the HF axis about the (C_{6v}) symmetry axis (of benzene) produces the desired symmetric average structure". The small energy difference between these three complexes reveals that the potential surface at the distance of forming the hydrogen bond is very flat. An interesting observation is that when one attempts to optimize the geometry of the complex starting from a geometry where HF is slightly displaced off the C_{6v} symmetric axis, the HF shifts to the top of a carbon atom finally, forming the structure D. This again demonstrates the flat nature of the potential energy surface between benzene and hydrogen fluoride.

The optimized geometries for all the five molecules are listed in Table 5.2. The hydrogens of ethylene bend slightly away from HF, while

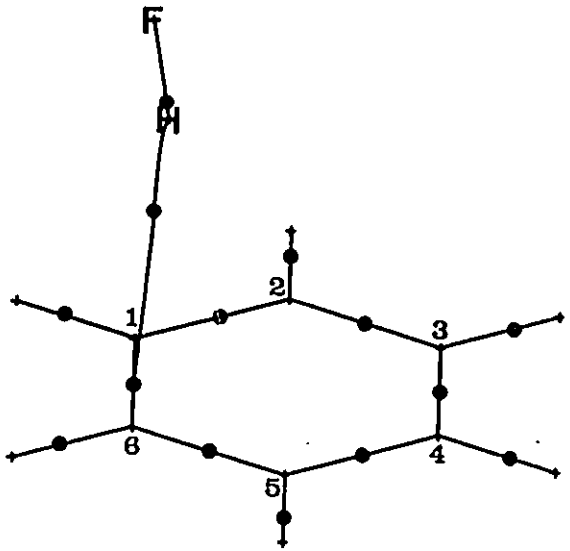
Figure 5.1. Molecular graphs of C_2H_2HF A, C_2H_4HF B, C_6H_6HF(B) C, C_6H_6HF(A) D, and C_6H_6HF(C) E. These graphs were determined from the charge densities of the complexes. The solid circles denote the positions of bond critical points of ρ. The hydrogen fluoride bond critical point is close to the proton. Refer to Table 5. 2 for the values of geometrical parameters.



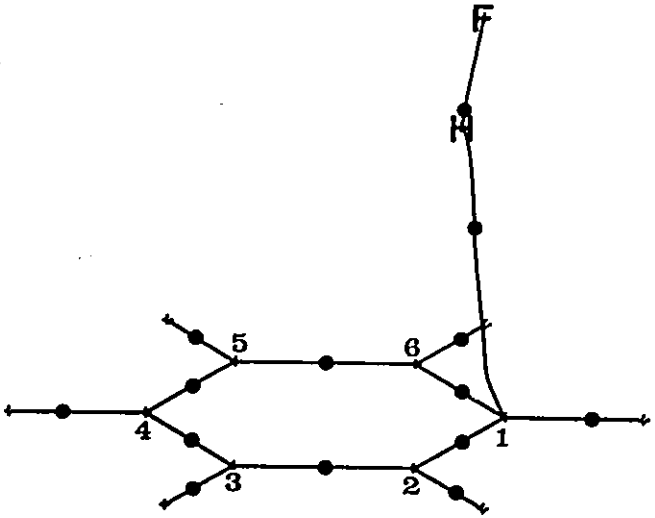
A



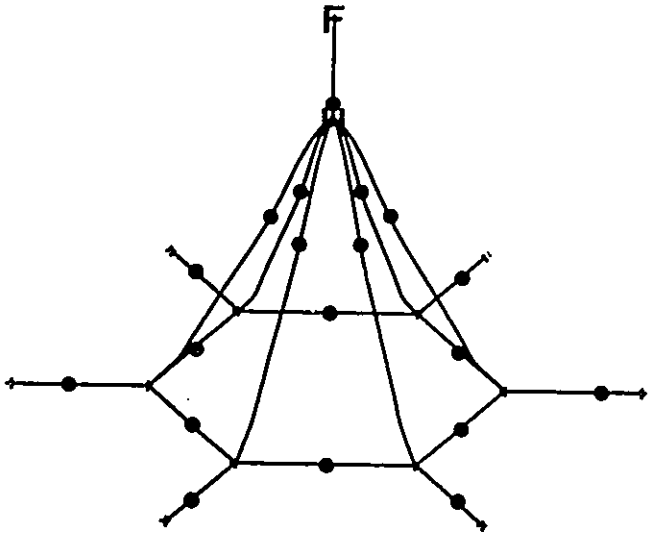
B



C



D



E

Table 5.1. The SCF Energies E_{SCF} and Stablization Energies E_s for the 6-31G**//6-31G Calculation.

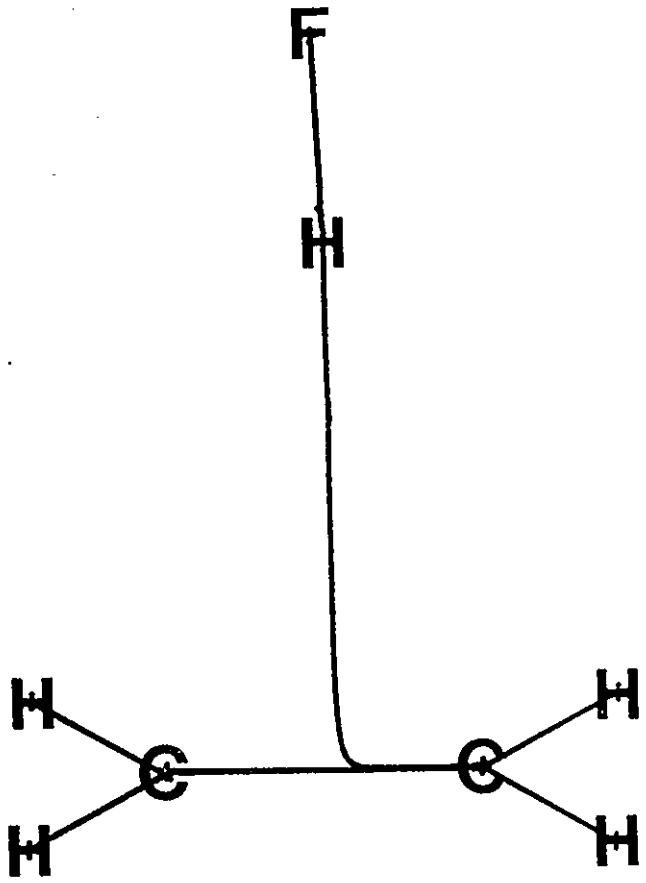
E_{SCF}(Complexes)		E_{SCF}(Monomers)		E_s	
C_2H_2-HF	-176.83887	C_2H_2	-76.82166	-0.00603 au	-3.784 Kcal
C_2H_4-HF	-178.05627	C_2H_4	-78.03877	-0.00632 au	-3.966 Kcal
C_6H_6HF(D)	-330.73052	C6H6	-230.71382	-0.00552 au	-3.464 Kcal
C_6H_6HF(C)	-330.73023	HF	-100.01118	-0.00523 au	-3.282 Kcal
C_6H_6HF(E)	-330.73059			-0.00559 au	-3.508 Kcal

those of acetylene do not, which is different from the result of Pople (1982) using a smaller basis set. It agrees with results determined by Sapse (1984). The nature of the geometric changes of the base on forming the hydrogen bond is similar for all the complexes: all the bond lengths, carbon-carbon and carbon-hydrogen, which are close to the site of protonation, are increased, although only by a very small amount. This trend decreases as atoms are removed farther from the site of protonation as found in C and D for example. Similar to that of C_2H_4-HF, the hydrogen bonded to the carbon atom which forms the hydrogen bond with HF in the complex C_6H_6HF(D) bends down about 0.5 degree from HF. The hydrogens in C_6H_6HF(C) and C_6H_6HF(E) remain in the plane of the benzene ring.

The molecular graphs for the structures of the hydrogen bonded complexes **A**, **B**, **C** and **E** are illustrated in Fig.5.1, they exhibit some special features. In structures B and C the bond path linking the proton of HF terminates at a carbon-carbon bond critical point rather than at a carbon nucleus. They are called conflict structures (Bader et al 1981). Any motion which displaces the proton away from the symmetry axis causes the bond path to shift from the bond critical point to either of the carbon nuclei, see Fig.5.2. Hence the name conflict structure -- the two carbon nuclei are in conflict for the terminus of the line of maximum electron charge density originating at the proton. The structures B and C represent metastable balances in this conflict. In structure E, the proton is simutaneously linked to the six carbon atoms of the benzene ring rather than to the centre of ring, which would again be an unstable conflict structure.

For the 6-31G** basis set employed in this study, there is a

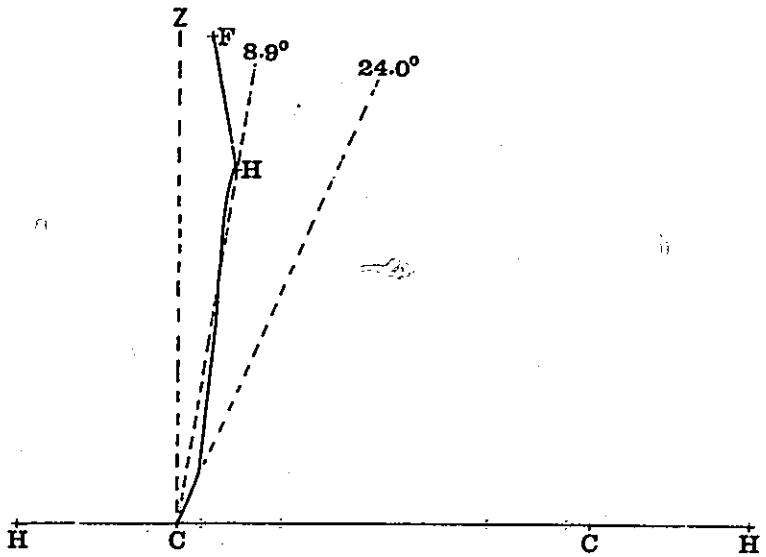
Figure 5.2 The molecular graph of C_2H_4-HF complexe with the angle between the axis of HF and the main axis of C_{2v} symmetry of C_2H_4 being 2^o. The hydrgon bond path which terminates at carbon-carbon critical point in the equilibrium geometry of the complexe has been replaced by one which terminates at a carbon nucleus.

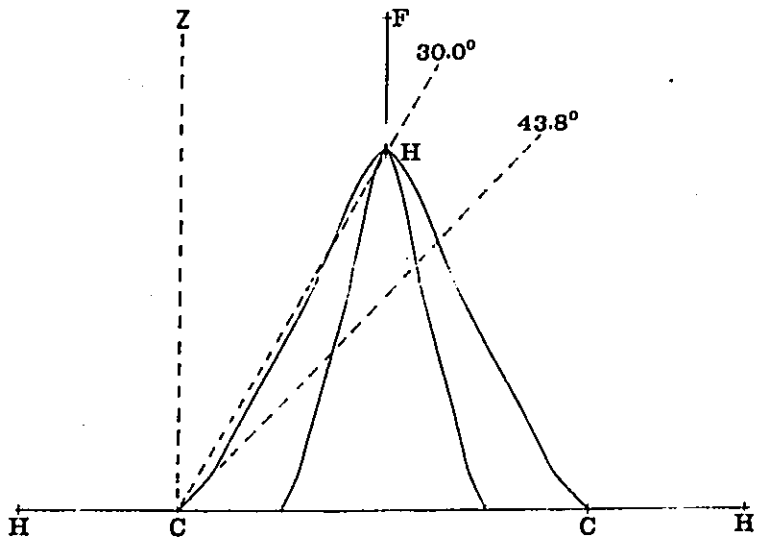

F
H
H
C
C
H
H
H
H

maximum of charge concentration at the carbon-carbon bond of acetylene. When large STO basis sets including d and f polarization functions are employed, this maximum in acetylene disappears and is replaced by a bond critical point (Bader et al 1983). This is due to the fact that there is strong charge concentration along the acetylene carbon-carbon bond; also the curvature of ρ between two carbon nuclei is very small so that even the basis sets such as the 6-31G** cannot give an accurate description. Thus this maximum is an artifact since the value of ρ at the charge maximum is only 0.003 au greater than the value of ρ at the neighboring critical point. Therefor, it is acceptable to treat the non-nuclear attractor as a bond critical point. The atomic properties of the pseudoatom (Cao et al 1987) produced by this incorrect description are averaged into those of the carbon atoms. Thus the C_2H_2-HF complex is also a conflict structure (see A in Fig.5.1), the same as C_2H_4-HF and C_6H_6HF(C).

It has been mentioned in Chapter 3 that next to the bonded charge concentrations, the Laplacian of charge density in Ph-H attains its most negative values at the out-of-plane saddle points; these secondary charge concentrations, which are displaced off a nuclear centered axis perpendicular to the benzene ring plane and toward the centre of the ring, determine the initial direction of approach of an electrophile. This prediction is borne out in the present study. From D of Fig.5.1, it is clear that the approaching HF does not sit at the top of C1 but shifts slightly toward the centre of the ring. Furthermore, from the diagram in the left hand side of Figure 5.3, which displays the symmetry plane of D containing the HF and associated carbons of Fig.5.1, one can see that the hydrogen bond path does not form a straight line between the hydrogen and

Figure 5.3 The cross section map through C1, C4, and HF for both C_6H_6HF(A), LHS, and C_6H_6HF(C), RHS. The line marked Z is the carbon nucleus centerd axis. The lines with the value of degrees on go through either hydrogen atom or the turning point of the hydrogen bond, displaying the off-axis extent. On the RHS picture, the projections of the other four hydrogen bond paths are also shown.


Z
F
8.9°
24.0°
H
H
C
C
H


Z
F
30.0°
43.8°
H
H
C
C
H

carbon, which has the angle off the nuclear centered axis of 8.9°, but forms a line bent toward the centre of the ring. The hydrogen bond path, starting from the hydrogen of HF, points away from the carbon atom and turns towards the carbon nucleus after reaching a point which is close to the position of the secondary charge concentration of the carbon. The off-axis angle of the turning point is 24.0°, while the off-axis angle of the secondary charge concentration of carbon in isolated benzene is 32.7°. (the distance of the turning point from the carbon nucleus C1 is 0.40 Å and that of the secondary charge concentration of carbon from its nucleus in isolated benzene is 0.52 Å). The C_6H_6HF(E) complex is similar (see the right picture of Fig.5.3). The off-axis angle of the straight line between a carbon and the hydrogen of HF is 30.0°, which is similar to the off-axis angle of secondary charge concentration of carbon in isolated benzene. But as the off-axis angle of the secondary charge concentration has been shifted from 32.7° in the free benzene molecule to 40.1° here, the hydrogen bond paths are also bent down towards the centre of the benzene ring, the turning point forming an angle of 43.8°, which is almost the same as that to the secondary charge concentrations of the carbons.

5.2. Comparisons of the Properties of π-Type and σ-Type Hydrogen Bond Complexes

The hydrogen bond lengths (see Table 5.2) are defined as the distance between the proton of HF and the other terminus of the hydrogen bond, either the critical point of a carbon-carbon bond or carbon nucleus. Except for that of C_6H_6HF(E), the hydrogen bond lengths of the unsaturated hydrocarbons are longer than those of classical hydrogen bonds formed by

Table 5.2. Geometries[a] of C_2H_2-HF, C_2H_4-HF and C6H6-HF π Complexes for the 6-31G Basis Set.

	C_2H_2HF	C_2H_4HF	$C_6H_6HF(D)$	$C_6H_6HF(C)$	$C_6H_6HF(E)$
C1-C2	1.1960	1.3264	1.3912	1.3904	1.3899
C2-C3			1.3891	1.3884	1.3899
C3-C4			1.3887	1.3890	1.3899
C4-C5			1.3887	1.3884	1.3899
C5-C6			1.3891	1.3904	1.3899
C6-C1			1.3912	1.3914	1.3899
C1-H1	1.0540	1.0741	1.0732	1.0731	1.0726
C2-H2			1.0729	1.0727	1.0726
C3-H3			1.0727	1.0727	1.0726
C4-H4			1.0728	1.0727	1.0726
C5-H5			1.0727	1.0727	1.0726
C6-H6			1.0729	1.0731	1.0726
H-X[b]	2.3574	2.3581	2.4310	2.4114	2.4119
H-F	0.9235	0.9248	0.9233	0.9232	0.9219
C6C1C2			119.96	119.96	120.00
C1C2C3			119.97	119.98	120.00
C2C3C4			120.00	120.06	120.00
C3C4C5			120.10	120.06	120.00
C4C5C6			120.00	119.98	120.00
C5C6C1			119.97	119.96	120.00
H1C1C2	180.00	121.88	120.02	119.95	120.00
H2C2C1			119.96	119.95	120.00
H3C3C4			120.04	119.98	120.00
H4C4C3			119.95	119.98	120.00
H4C4C5			119.95	119.96	120.00
H5C5C4			120.04	120.07	120.00
H6C6C1			119.96	123.09	120.00
H1C1C4			179.49[c]	180.00	180.00
HX[b]Ring			81.16	85.50	90.00
FHX[b]			160.71[d]	166.74[c]	180.00
HCCH(F)	90.00	90.67			

Monomers

	C_2H_2	C_2H_4	C_6H_6	HF
C-C	1.1942	1.3220	1.3887	
C-H	1.0531	1.0740	1.0773	
H-F				0.9209
HCC	180.00	121.93	120.00	

[a]. Bond lengths are in angstroms and bond angles are in degrees.
[b]. X is the carbon-carbon bond centre for C_2H_2-HF, C_2H_4-HF and C_6H_6-HF(C), the carbon atom for C_6H_6-HF(D), and the ring centre of benzene for C_6H_6-HF(E).
[c]. H bends away from the side of HF.
[d]. F bends away from the benzene ring.

second row atoms, which are less than 2.30 A, and shorter than those of classical ones formed by third row atoms, which are more than 2.54 A in general (Carroll et al 1988).

Table 5.1 lists the SCF energies of the complexes of unsaturated hydrocarbons with hydrogen fluoride as well as of their isomers. The stabilization energies E_s

$$E_s = E_{SCF}(complex) - E_{SCF}(hydrocarbon) - E_{SCF}(HF) \qquad [5.1]$$

are in the range of 3-4 kcal, which show that the strengths of these π-type complexes are close to the average, ≈ 5 kcal/mol, of those classical σ-type hydrogen bonded complexes, (the E_s of σ-type H-bond complexes ranging from 1.5 kcal for NN-HF to 11.8 kcal for H_3N-HF with 6-31g**//6-31G** calculations (Carroll et al 1988)). The conclusion that the benzene-hydrogen fluoride complexes are slightly less stable than those which hydrogen fluoride forms with acetylene and ethylene agrees with the comparison of the result of Bredas and Street (1988) with those of Pople et al. (1982). The order of the stabilization energy for these three complexes in which the hydrogen of HF is bonded to the C-C bond critical point is: **B > A > C** (see Table 5.3). From Table 5.3, one can see that the stabilization energies do not correlate with the properties of the bond critical point of the C-C bond of acetylene, ethylene and benzene isomers, nor with the hydrogen bond lengths of their complexes. However, the stablization energy can be correlated with both the value of the charge density at the hydrogen bond critical point and with the degree of penetration of the van der Waals envelope of the reactant (Carroll and Bader 1988). Boyd and Choi (1986, 1985) studied the hydrogen bond complexes between both nitriles (Boyd and Choi 1985) and hydrogen fluoride

Table 5.3. Comparison of Stablization Energy E_s, Penetration Radius d_p, Properties of C-C Bond and H-bond Critical Point and H-bond Length for Complexes A (C_2H_2HF), B (C_2H_4HF) and C ($C_6H_6HF(C)$), (See Fig.5.1).

	C	A	B
E_s	-3.3 kcal	-3.8 kcal	-4.0 kcal
d_p(HF->B)	0.84 au	0.89 au	0.93 au
d_p(B->HF)	0.45 au	0.52 au	0.53 au
Σd_p	1.29 au	1.41 au	1.47 au
H Bond ρ_b	0.0095 au	0.0113 au	0.0122 au
H Bond $\nabla^2\rho_b$	0.027 au	0.031 au	0.028 au
C-C Bond ρ_b	0.3236 au	0.4113 au	0.3578 au
C-C Bond $\nabla^2\rho_b$	-0.991 au	-1.237 au	-1.155 au
$R_{H\text{-}bond}$	2.411 A	2.357 A	2.358 A

and nitriles and hydrogen chloride (Boyd and Choi 1986), and found a strong correlation between the stablization energies and the properties of the hydrogen bond critical points. Table 5.3 also shows this correlation of stablization energy with the value of charge density at the hydrogen bond critical point. The penetration of the van der Waals envelopes of one reactants by the other is shown in Figure 5.4. The outer contour of value 0.001 au of molecules yields molecular sizes and atomic diameters in good agreement with gas phase van der Waals radii (Bader et al 1987a, Bader et al 1971, Bader and Preston 1970). Thus the 0.001 au envelope of the charge density is also called van der Waals envelope. The extent of the penetration is the measurement of how deep a reactant inserts into the van der Waals envelope of another reactant when their reaction reaches an energy minimum. The greater the penetration, the stronger the interaction. The interatomic surface, which is formed by trajectories of $\nabla\rho$ originating from the hydrogen bond critical point (the solid dot in Figure 5.4), is the border between the two molecules. The distance, d_p in Fig.5.4, of the bond critical point to the outermost contour line (ρ = 0.001) of the isolated unsaturated hydrocarbon, is a measure of the depth to which hydrogen penetrates the van der Waals envelope of unsaturated hydrocarbon molecules. The same measurement applies to the penetration of the unsaturated hydrocarbon into the van der Waals envelope of HF. The depths of penetration of both HF into the van der Waals envelope of base and base into the van der Waals envelope of HF are listed in Table 5.3. The total depth of penetration is simply the sum of the two penetrations, which are also listed in Table 5.3. The sum of d_p is 1.47 au for C_2H_4-HF, 1.41 au for C_2H_2-HF and 1.29 au for C_6H_6-HF, which is in the same order as the

stabilization energy E_B. The same order is also evidenced in the depth of penetration of one reactant into the van der Waals envelope of the other one.

From Fig.5.4, one can also see the polarization of the charge distribution of ethylene caused by the approach of HF. As ethylene acts as a base, an electron donor, its charge distribution moves slightly from a symmetric one in the free molecule to an asymmetric one toward the acid HF. The closer its position to HF, the bigger the shift of the charge.

Such charge transfer from hydrocarbon molecules tc HF can also be found from the change in the atomic populations when compared with those of the isolated molecules (see Table 5.4). Table 5.4a lists the atomic populations for both monomers and dimers, while Table 5.4b lists the change of these properties from monomers to dimers. Since the integration program currently employed has difficulty in giving reliable integrated results for a conflict structure, and for the bridging hydrogen atom in C_6H_6HF(E), the properties of the associated hydrogen and carbon atoms are obtained by difference after the properties of the remaining atoms are determined. As an electron acceptor, the HF molecule gains electrons in all five π-type hydrogen bonded complexes, 0.030 e for C_2H_4-HF, 0.027 for C_2H_2-HF, 0.021 for both C_6H_6HF(D) and C_6H_6-HF(C), and 0.015 for C_6H_6-HF(E). The populations of the hydrogen of HF in A, B and C remain mainly constant, while those in D and E lose a slight amount of charge. The gain in the population of HF stems mainly from the contribution of fluorine, the most electronegative atom. This case is exactly the same as that in the σ-type hydrogen bonded complexes (Carroll et al 1988). As a result of the polarization effect caused by HF, the charges on hydrocarbons in these

Figure 5.4. Contour map of C_2H_4 (dashed contours), overlaid on that for C_2H_4HF (solid contours). The line through the hydrogen bond critical point, denoted by a dot, defines the interatomic surface separating the two molecules which have formed a hydrogen bonded complex. The depth of the penetration of HF into the shell of C_2H_4 is displayed as d_p.

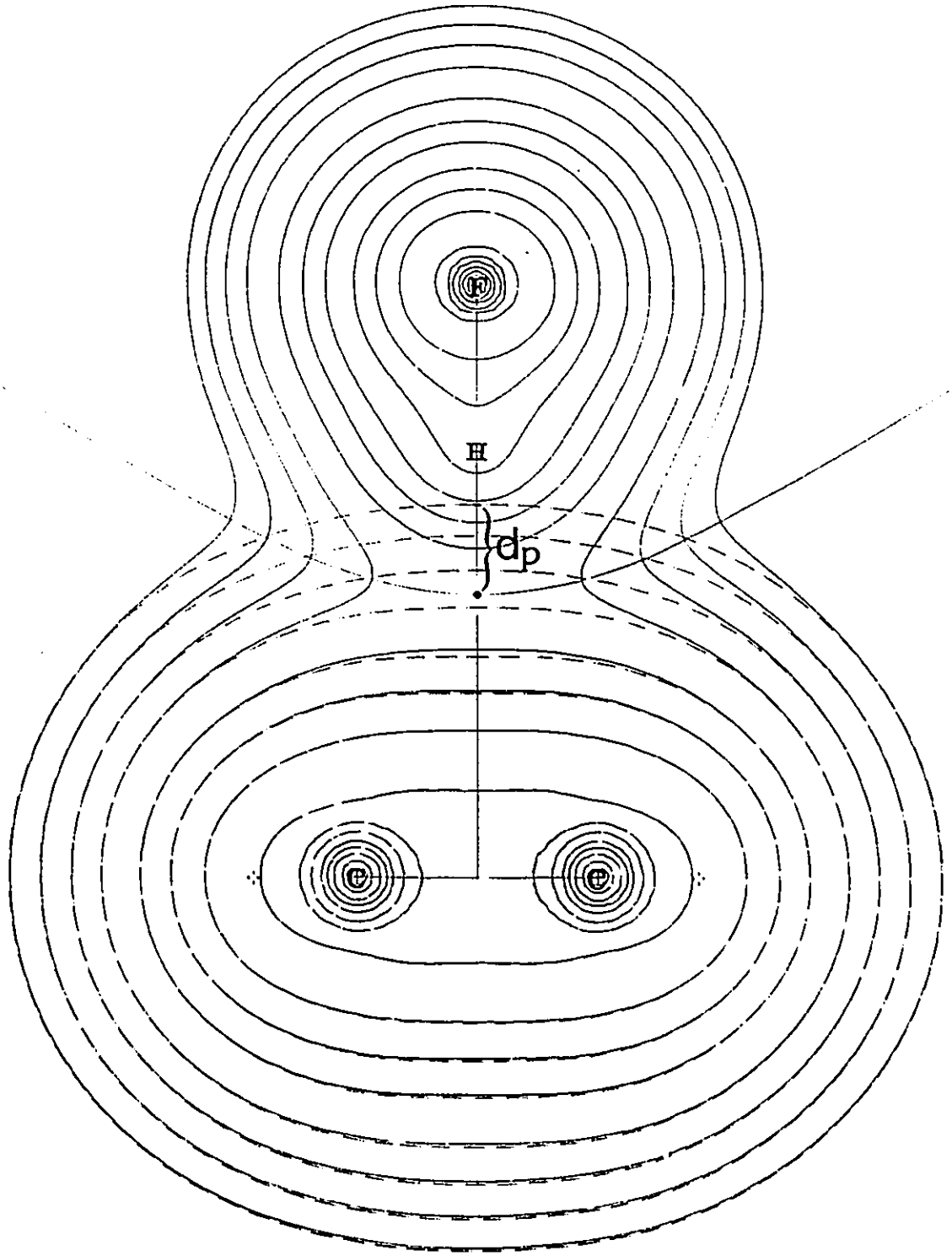

H
H
d_p
C
C

Table 5.4a. Atomic Population of π-type Hydrogen Bonded Complexes.

	C_2H_2-HF	C_2H_4-HF	C_6H_6-HF(D)	C_6H_6-HF(C)	C_6H_6-HF(E)
C1	6.132	5.941[a]	5.987	5.981[a]	5.965
C2	6.132	5.941[a]	5.967	5.958	5.965
C3			5.953	5.952	5.965
C4			5.952	5.952	5.965
C5			5.953	5.958	5.965
C6			5.967	5.981[a]	5.965
H(C1)	0.856	1.022	1.031	1.031	1.033
H(C2)	0.856	1.022	1.032	1.033	1.033
H(C3)			1.034	1.035	1.033
H(C4)			1.035	1.035	1.033
H(C5)			1.034	1.033	1.033
H(C6)			1.032	1.031	1.033
H(F)	0.248	0.248	0.245	0.247	0.241[a]
F	9.779	9.782	9.776	9.774	9.773

Monomers

	C_2H_2	C_2H_4	C_6H_6	HF
C	6.121	5.917	5.957	
H	0.880	1.042	1.043	0.249
F				9.751

Table 5.4b The Relative Atomic Population of π-type Hydrogen Bonded Complexes to their Isomers.

	C_2H_2-HF	C_2H_4-HF	C_6H_6-HF(D)	C_6H_6-HF(C)	C_6H_6-HF(E)
C1	0.011	0.024	0.030	0.024	0.008
C2	0.011	0.024	0.010	0.001	0.008
C3			-0.004	-0.005	0.008
C4			-0.005	-0.005	0.008
C5			-0.004	0.001	0.008
C6			0.010	0.024	0.008
H(C1)	-0.024	-0.020	-0.012	-0.012	-0.010
H(C2)	-0.024	-0.020	-0.011	-0.010	-0.010
H(C3)			-0.009	-0.008	-0.010
H(C4)			-0.008	-0.008	-0.010
H(C5)			-0.009	-0.010	-0.010
H(C6)			-0.011	-0.012	-0.010
H(F)	-0.001	-0.001	-0.004	-0.002	-0.007
F	0.028	0.031	0.025	0.023	0.022
Δq_{acid}	0.027	0.030	0.021	0.021	0.015
Δq_{base}	-0.026	-0.032	-0.023	-0.020	-0.014

[a]. This property is obtained by subtracting from the total value of the property from the SCF calculation.

complexes flow slightly towards the carbon atoms which are closer to the approaching HF; this kind of change in charge distribution is reflected in the population difference of the benzene carbon atoms for complexes C and D, Table 5.4b. The largest charge increase is for the atoms which are nearest to HF, decreasing in amount as the carbon is removed further from the reacting site and those furthest removed lose charge. All the hydrogens on the hydrocarbons lose charge, the amounts varying from 0.024 e in B to 0.008 in C and D. The charges gained by HF are mostly from the contribution of the hydrogens of the hydrocarbons, and the six hydrogens on the benzene ring share the loss of charge, which explains why each of them loses less charge than those in ethylene and acetylene. An interesting observation is that the hydrogens on the benzene ring in all the three complexes lose the same total amount of charge, an average amount of 0.01e.

The changes in atomic energies (see Table 5.5), are different from the changes in the atomic charges. In each complex, HF, which gains charge, is destabilized, compared to the monomer, and both the H and F atoms increase their energies except for the fluorine atom in A. All the carbon atoms, no matter whether they gain or loss charge, are stabilized; the closer to the reaction centre, the greater their stabilization. However, the behaviour of the hydrogens in the hydrocarbons are exactly opposite to that of the carbons; the closer they are to the reaction centre, the more destabilized they become. The behavior of HF with regard to the changes in energy is different from that in the classical hydrogen bonded complexes wherein the base contains a second-row atom, but is the same when the base contains a third-row atom.

Table 5.6 lists properties of ρ at bond and ring critical points

Table 5.5 Atomic Energies of π-type Hydrogen Bonded Complexes.

	C_2H_2-HF	C_2H_4-HF	C_6H_6HF(D)	C_6H_6HF(C)	C_6H_6HF(E)
C1	-37.8562	-37.7612[a]	-37.8278	-37.8244[a]	-37.8142
C2	-37.8562	-37.7612[a]	-37.8151	-37.8117	-37.8142
C3			-37.8085	-37.8104	-37.8142
C4			-37.8095	-37.8104	-37.8142
C5			-37.8085	-37.8117	-37.8142
C6			-37.8185	-37.8244[a]	-37.8142
H(C1)	-0.5605	-0.6408	-0.6481	-0.6484	-0.6494
H(C2)	-0.5605	-0.6408	-0.6488	-0.6496	-0.6494
H(C3)			-0.6502	-0.6504	-0.6494
H(C4)			-0.6505	-0.6504	-0.6494
H(C5)			-0.6502	-0.6496	-0.6494
H(C6)			-0.6488	-0.6484	-0.6494
H(F)	-0.2487	-0.2491	-0.2495	-0.2493	-0.2505[a]
F	-99.7562	-99.7216	-99.6996	-99.6912	-99.6985
ΔE_{acid}	0.0063	0.0405	0.0621	0.0707	0.0622
ΔE_{base}	-0.0110	-0.0468	-0.0661	-0.0748	-0.0666

Monomers

	C_2H_2	C_2H_4	C_6H_6	HF
C	-37.8391	-37.7214	-37.8015	
H	-0.5721	-0.6490	-0.6510	-0.2588
F				-99.7524

[a]. The property is obtained by substracting from the total property of SCF calculation

of complexes A-E as well as the corresponding monomers. All the changes of the properties caused by forming hydrogen bond are small. It is noted that an unusual large ellipticity appears at the hydrogen bond critical point of $C_6H_6HF(E)$ complex in which six carbon nuclei share the bonds with hydrogen, this large number results from a particular small soft curvature, λ_2 (see Table 5.6).

From these comparisons, it can be seen that many properties of the hydrogen bonded complexes of unsaturated hydrocarbons are similar to those of the classical complexes. Both the amount of charge transfered from the base to acid and the stabilization energies in the π-type hydrogen bond complexes are in the middle range of the σ-type ones. The hydrogen bond lengths in the former case lie between those of the hydrogen bonds formed from second-row and third-row atoms in the latter case, with $C_6H_6HF(E)$ being the exception. The energy change for the acid in the π-type H-bonded complexes is more like those bonded with the third-row atoms in the σ-type H-bonded complexes. The largest difference is in the structures of the resulting hydrogen bonded complexes. The hydrogen bonds in σ-type complexes occur between a hydrogen and a single base atom. In π-type complexes, however, hydrogen can be bonded to a carbon-carbon bond critical point or to several carbon atoms, as well as to a carbon atom.

Table 5.6. Properties of ρ at Bond and Ring Critical Points of C_2H_2-HF, C_2H_4-HF and C6H6-HF Complexes and Monomers for the 6-31G**//6-31G Basis Set.

Complexes

Bond	Complex	ρ_b	$\nabla^2\rho_b$	ϵ[a]	λ_1	λ_2	λ_3
	A	0.411	-1.237	0.017y	-0.586	-0.576	-0.074
	B	0.358	-1.155	0.427z	-0.801	-0.561	0.207
C1-C2	C	0.326	-1.007	0.228z	-0.705	-0.574	0.271
	D	0.324	-0.993	0.230z	-0.698	-0.240	0.272
	E	0.325	-1.000	0.229z	-0.701	-0.571	0.272
	C	0.325	-0.999	0.228z	-0.700	-0.570	0.272
C2-C3	D	0.326	-1.005	0.228z	-0.703	-0.573	0.271
	C	0.324	-0.991	0.233z	-0.697	-0.566	0.272
C3-C4	D	0.326	-1.007	0.226z	-0.704	-0.574	0.271
C1-C6	C	0.326	-1.007	0.225z	-0.704	-0.574	0.271
	A	0.305	-1.316	0.000	-0.851	-0.851	0.385
	B	0.296	-1.153	0.012z	-0.784	-0.774	0.405
C1-H1	C	0.299	-1.175	0.015z	-0.797	-0.786	0.408
	D	0.299	-1.169	0.018z	-0.795	-0.782	0.408
	E	0.299	-1.175	0.015z	-0.798	-0.786	0.409
	C	0.299	-1.175	0.015z	-0.798	-0.786	0.409
C2-H2	D	0.299	-1.174	0.015z	-0.797	-0.786	0.409
	C	0.299	-1.175	0.015z	-0.796	-0.783	0.409
C3-H3	D	0.299	-1.174	0.015z	-0.797	-0.785	0.408
C4-H4	D	0.299	-1.174	0.014z	-0.797	-0.786	0.408
	A	0.367	-3.200	0.000	-2.512	-2.512	1.824
	B	0.365	-3.173	0.000	-2.496	-2.496	1.819
F-H	C	0.360	-3.039	0.000	-2.427	-2.427	1.815
	D	0.367	-3.192	0.000	-2.510	-2.510	1.828
	E	0.369	-3.215	0.000	-2.525	-2.525	1.835
	A	0.011	0.031	0.344x	-0.011	-0.008	0.050
	B	0.012	0.028	0.526x	-0.012	-0.008	0.048
H-Bond	C	0.009	0.027	1.302x	-0.009	-0.004	0.040
	D	0.010	0.028	0.472x	-0.009	-0.006	0.044
	E	0.005	0.021	51.063x	-0.003	-0.000(1)	0.024
	C	0.021	0.170		-0.014	0.092	0.092
Ring(ø)	D	0.021	0.170		-0.014	0.092	0.092
	E	0.021	0.170		-0.014	0.092	0.092
Ring(H-Bond)		0.005	0.021		-0.003	0.000(1)	0.024

[a]. Z is the direction of approaching HF for all complexes, x is the direction paralell to C-C bond of C_2H_2HF and C_2H_4HF and perpendicular to the line linking the critical point and the centre of the benzene ring for benzene complexes.

Table 5.6 (con'd)

Monomers

Bond	Complex	ρ_b	$\nabla^2\rho_b$	ϵ	λ_1	λ_2	λ_3
	C_2H_2	0.412	-1.236	0.000	-0.578	-0.578	-0.080
C-C	C_2H_4	0.360	-1.169	0.447	-0.809	-0.559	0.200
	C_6H_6	0.326	-1.007	0.231	-0.705	-0.573	0.271
	C_2H_2	0.305	-1.316	0.000	-0.851	-0.851	0.385
C-H	C_2H_4	0.296	-1.153	0.012	-0.784	-0.774	0.405
	C_6H_6	0.296	-1.143	0.016	-0.781	-0.769	0.407
F-H	FH	0.373	-3.100	0.000	-2.495	-2.495	1.890
Ring(ø)	C_6H_6	0.020	0.170		-0.014	0.092	0.092

REFERENCES

Almlof,J.,Korsell,K.,Faegri,K.,Jr.(1982).J.Comp.Chem.,3,385.

Andrews,L.,Johnson,G.L.,Davis,S.R.(1985).J.Phys.Chem.89,1706.

Bader,R.F.W.(1990).Atoms in Molecules: A Quantum Theory, Oxford University Press: London.

Bader,R.F.W. and Chang,C.(1989a).J.Phys.Chem.93,5095.

Bader,R.F.W. and Chang,C.(1989b).J.Phys.chem.93,2946.

Bader,R.F.W.(1988).Pure.Appl.Chem.60,145.

Bader,R.F.W.,Gillespie,R.J.,MacDougall,P.J.(1988).J.Am.Chem.Soc.110,7329.

Bader,R.F.W.,Carroll,M.T.,Cheeseman,J.R.,Chang,C.(1987a).J.Am.Chem.Soc.109, 7968.

Bader,R.F.W.,Larouche,A.,Gatti,C.,Carroll,M.T.,MacDougall,P.J.,Wiberg,K.B. (1987b).J.Chem.Phys.87,1142.

Bader,R.F.W.(1986).Can.J.Chem.64,1036.

Bader,R.F.W.(1985).Acc.Chem.Res.18,9.

Bader,R.F.W. and MacDougall,P.J.(1985).J.Am.Chem.Soc.107,6788.

Bader,R.F.W. and Essen,H.(1984).J.Chem.Phys.80,1943.

Bader,R.F.W.,MacDougall,P.J.,Lau,C.D.H.(1984).J.Am.Chem.Soc.106,1594.

Bader,R.F.W.,Slee,T.S.,Cremer,D.,Kraka,E.(1983).J.Am.Chem.Soc.105,5061.

Bader,R.F.W.,Tang,T.H.,Tal Y.,Biegler-König,W.(1982).J.Am.Chem.Soc.104,940.

Bader,R.F.W. and Nguyen-Dang,T.T.(1981).Adv.Quant.Chem.14,63.

Bader,R.F.W.,Nguyen-Dang,T.T.,Tal,Y.(1981).Rep.Prog.Phys.44,893.

Bader,R.F.W. and Beddall,P.M.(1972).J.Chem.Phys.56,3320.

Bader,R.F.W.,Beddall,P.M.,Cade,P.E.(1971).J.Am.Chem.Soc.93,3095.

Bader,R.F.W. and Preston,H.J.T.(1970).Theo.Chim.Acta 17,384.

Bader,R.F.W.,Henneker,W.H.,Cade,P.E.J.(1967).J.Chem.Phys.46,3341.

Baiocchi,F.A.,Williams,J.H.,Klemperer,W.(1983).J.Phys.Chem.87,2079.

Biegler-König,F.W.,Bader,R.F.W.,Tang,T.H.(1982).J.Comp.Chem.13,317.

Binkley,J.S.,Frisch,M.J.,DeFress,D.J.,Raghavachari,K.,Whiteside,R.A., Schlegel,H.B.,Fluder,E.M.,Pople,J.A.,(1982) Gaussian 82, Carnegie-Mellon University, Pittsburgh.

Biswas,A.B.,Hughes,E.W.,Sharma,B.D.,Wilson,J.N.(1968).Acta Cryst.B24,40.

Bock,C.W.,Trachtman,M.,George,P.(1985).J.Mol.Struct.(THEOCHEM),122,155. The same authors have studied the effects of the inclusion of polarization functions on the predicted geometries of Ph-H, Ph-F and Ph-CN and found them to be minor,(1985).J.Comput.Chem.6,592.

Bonaccorsi,R.,Palla,P.,Tomasi,J.(1984).J.Am.Chem.Soc.106,1945.

Boyd,R.J. and Choi,S.C.(1986).Chem.Phys.Lett.129,62.

Boyd,R.J. and Choi,S.C.(1985).Chem.Phys.Lett.120,80.

Bredas,J.L. and Street,G.B.(1988).J.Am.Chem.Soc.110,7001.

Brouwer,D.M.,Mackor,E.L.,McLean,C.(1970). In Carbonium Ions;Olah,G.A., von R.Schleyer,P.,Eds.;Wiley; New York,Vol.2,p837.

Cao,W.L.,Catti,C.,MacDougall,P.J.,Bader,R.F.W.(1987).Chem.Phys.Lett.141, 380.

Carroll,M.T.(1989).Ph.D. Thesis, McMaster University.

Carroll,M.T.,Cheeseman,J.R.,Osman,R.,Weinstein,H.(1989).J.Phys.Chem.93, 5120.

Carroll,M.T. and Bader,R.F.W.(1988).Mol.Phys.65,695.

Carroll,M.T.,Chang,C.,Bader,R.F.W.(1988).Mol.Phys.63,387.

Chandra,A.K. and Coulson,C.A.(1965).J.Chem.Soc.,2210.

Cheney,B.V.,Schukz,M.W.,Cheney,J.,Richards,W.G.(1988).J.Am.Chem.Soc.110, 4195.

Coulson,C.A.,March,N.H.,Altmann,S.(1952).Proc.Natl.Acad.Sci.U.S.A.38,372.

Cremer,D.,Kraka,E.,Slee,T.,Bader,R.F.W.,Lau,C.D.H.,Nguyen-Dang,T.T., MacDougall,P.J.(1983).J.Am.Chem.Soc.105,5069.

Dewar,M.J.S.(1949).The Electronic Theory of Organic Chemistry;Oxford University Press: London.

Ditchfield,R.,Hehre,W.C.,Pople,J.A.(1971).J.Chem.Phys.54,724.

Dyke,T.R.(1984).Top.Curr.Chem.120,85.

Ermler,W.C.,Mulliken,R.S.,Clementi,E.(1976).J.Am.Chem.Soc.98,388.

Exner,O.,(1978).In Correlation Analysis in Chemistry;Champman,N.B., Shorter,J.,Eds.;Plenum Press: New York. There is no reliably assigned σ_R^o value for NH_3^+.

Fischer,A.,Henderson,G.N.,Smyth,T.A.,Einstein,F.W.B.,Cobbledick,R.E.(1981). Can.J.Chem.59,584.

Frish,M.J.,Head-Gordon,M.,Schlegel,H.B.,Raghavachari,K.,Binkley,J.S., Gonzalez,C.,Defrees,D.J.,Fox,D.J.,Whiteside,R.A.,Seeger,R.J., Baker,R.Martin,Kahn,L.R.,Stewart,J.J.P.,Fluder,E.M.,Topiol,S., Pople,J.A.(1988)."GAUSSIAN 88",Gaussian Inc.,Pittsburgh,PA 15213.

Frisch,M.J.,Binkley,J.S.,Schlegel,H.B.,Raghavachari,K.,Melius,C.F., Martin,R.L.,Stewart,J.J.Py.,Bobrowicz,F.W.,Rohlfing,C.M.,Kahn,L.R., Defrees,D.J.,Seeger,R.,Whiteside,R.A.,Fox,D.J.,Fleuder,E.M., Pople,J.A.(1986). "GAUSSIAN 86",Gaussian Inc.,Pittsburgh,PA 15213.

Frisch,M.J.,Pople,J.A.,Del Bene,J.E.(1983).J.Chem.Phys.78,4063.

Gandour,R.(1981).Bioorg.Chem.10,169.

Gatti,C.,Barzaghi,M.,Simonetta,M.(1985).J.Am.Chem.Soc.107,878.

Gatti,C.,Fantucci,P.,Paccioni,G.(1987).Theoret.Chim.Acta,72,433.

Gleghorn,J.T. and McConkey,F.W.(1975).J.Chem.Soc.,Perkin Trans.2,1078.

Haney,M.A. and Franklin,J.L.(1969).J.Phys.Chem.73,4328.

Hammett,L.P.(1940).Physical Organic Chemistry;McGraw-Hill:New York.

Hammett,L.P.(1937).J.Am.Chem.Soc.59,96.

Hariharan,P.C. and Pople,J.A.(1973).Theo.Chim.Acta (Berlin),28,213.

Head-Gordon,M. and Pople,J.A.(1988).J.Chem.Phys.,89,5777.

Hehre,W.J.,Taft,R.W.,Topson,R.D.(1976).Prog.Phys.Org.Chem.12,159.

Hehre,W.C. and Pople,J.A.(1972).J.Am.Chem.Soc.94,6901.

Hehre,W.C.,Ditchfield,R.,Pople,J.A.(1972a).J.Chem.Phys.56,2257.

Hehre,W.J.,Radom,L.,Pople,J.A.(1972b).J.Am.Chem.Soc.94,1496.

Ingold,C.K.(1969).Structure and Mechanism in Organic Chemistry,2nd ed.; Cornell University Press: Ithaca,NY,1969.

Ingold,C.K.(1953).Structure and Mechanism in Organic Chemistry,Bell and Sons:London.

Joesten,M.D. and Schaad,L.J.(1974). "Hydrogen Bonding". Marcel Dekker:New York.

Lau,Y.K. and Kebarle,P.(1976).J.Am.Chem.Soc.98,7452.

Libit,L. and Hoffmann,R.(1974).J.Am.Chem.Soc.96,1370.

Lister,D.G.,Tyler,J.K.,Hog,J.H.,Larsen,N.W.(1974).J.Mol.Struct.23,253.

Lowry,T.H.,Richardson,K.S.(1981).Mechanism an Theory in Organic Chemistry,2rd ed.;Harper and Row: New York.

Luke,B.T.,Gupta,A.G.,Loew,G.H.,Lawless,J.G.,White,DH.(1984). Int.J.Quant. Chem.: Quant.Biol.Symp.11,117.

March,J.(1977).J.Advanced Organic Chem.2nd ed.;McGraw-Hill: New York.

Marsh,R.E.(1958).Acta Cryst.11,654.

Michel,P.,Koch,H.J,Germain,G.(1970).Acta Cryst.B26,410.

O'Brien,D.H.,Hart,A.J.,Russell,C.R.(1975).J.Am.Chem.Soc.97,4410.

Olah,G.A.,Staral,J.S.,Assencio,G.,Liang,G.,Forsyth,D.A.;Mateescu.G.D. (1978).J.Am.Chem.Soc.100,6299.

Olah,G.A. and Kuhn,S.J.(1958).J.Am.Chem.Soc.80,6535,6541.

Palla,P.,Petrongolo,C.,Tomasi,J.(1980).J.Phys.Chem.84,435.

Peters,D. and Peters,J.(1980).J.Mol.Struct.69,249.

Pimentel,G.C. and McClellan,A.D.(1960). "The Hydrogen Bond". Freeman: San Francisco.

Pople,J.A,Frisch,M.J.,Del Bene,J.E.(1982).Chem.Phys.Lett.91,185.

Quack,M. and Stockburger,M.(1972).J.Mol.Spectrosc.(1972).43,87.

Read,W.G. and Flygare,W.H.(1982).J.Chem.Phys.76,2238.

Rebek,Jr.,J.(1987).Science,235,1437.

Rebek,Jr.,J.,Duff,R.J.;Gordon,W.E.,Parris,K.(1986).J.Am.Chem.Soc.108,6068.

Sapse,A.M.,Fugler,L.M.,Cowburn,D.(1986).Int.J.Quant.Chem.29,1241.

Sapse.A.M. and Jain,D.C.(1984).J.Phys.Chem.88,4970.

Schafer,L.,Siam,K.,Klimkowski,V.J.,Ewbank,J.D.,Alsenoy, C.V.(1990). J.Mol.Struct.(THEOCHEM),204,361.

Schuster,P.(1976). In "The Hydrogen Bond" (eds.Schuster,P.,Zundel,G., Sandorfy,C.).North-Holland Publishing Company:Amsterdam,25.

Shea,J.A. and Flygare,W.H.(1982).J.Chem.Phys.76,4857.

Shipman,L.L. and Christoffersen,R.E.(1973).J.Am.Chem.Soc.95,4733. And references therein.

Siam,K.,Klimkowski,V.J.,Ewbank,J.D.,Van Alsenoy,C.,Schafer,L.(1984).J.Mol. Struct.(THEOCHEM),110,171.

Simpson,H.J. and Marsh,R.E.(1966).Acta Cryst.20,550.

Slee,T.S.,Larouche,A.,Bader,R.F.W.(1988).J.Phys.Chem.92,6219.

Slee,T. and MacDougall,J.(1988).Can.J.Chem.66,(2961).

Slee,T.(1986).Ph.D. Thesis, McMaster University.

Sordo,T. and Bertran,J.(1979).J.Chem.Soc.,Perkin Trans.2,1486.

Spiesecke,H. and Schneider,W.G.(1961).Tetrahedron Lett.468.

Srikrishnan,T.,Winiewicz,N.,Parthasarathy,R.(1982).Int.J.Pep.Protein Res. 19,103.

Stewart,R.F.(1979).Chem.Phys.Lett.65,335.

Streitwieser,A. and Heathcock,C.H.(1985).Introduction to Organic Chemistry, 3rd ed.;MacMillan: New York.

Swain,C.G. and Lupton,Jr.,E.C.(1968).J.Am.Chem.Soc.90,4328.

Taft,R.W.(1987).Ed.Progress in Physical Chemistry;Wiley: New York,Vol.16, pp 1287.

Taft,R.W.,Price,E.,Fox,I.L.,Lewis,I.C.,Anderson,K.K.,Davis,G.T.(1963).J.Am. Chem.Soc.85,709,3146.

Taft,R.W.(1960).J.Phys.Chem.64,1805.

Taft,R.W.(1956).In Steric Effects in Organic Chemistry;Newman,M.S.,Ed.; Wiley:New York,Chapter 13.

Wang,A.H.J. and Paul,I.C.(1979).Cryst.Struct.Commun.8,269.

Wiberg,K.B.,Bader,R.F.W.,Lau,C.D.H.(1987a).J.Am.Chem.Soc.109,1001.

Wiberg,K.B.,Bader,R.F.W.,Lau,C.D.H.(1987b).J.Am.Chem.Soc.109,985.

Wiberg,K.B. and Laidig,K.E.(1987).J.Am.Chem.Soc.109,5935.

Wiberg,K.B. and Wendolowski,J.J.(1981).Proc.Natl.Acad.Sci.U.S.A.78,6561.

Wright,L.R. and Borkman,R.F.(1982).J.Phys.Chem.86,3956. And references wherein.

Vinogradov,S.N. and Linnell,R.H.(1971). "Hydrogen Bonding". Van Nostrand Reinhold: New York.