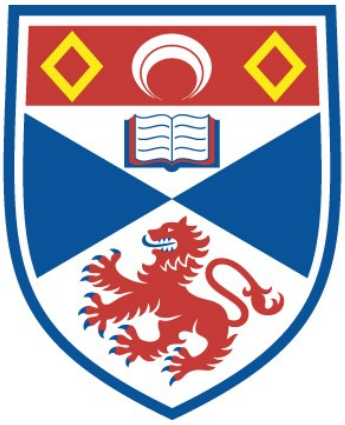
A THEORETICAL INVESTIGATION OF THE ELECTRONIC STRUCTURE AND PROPERTIES OF GLUTATHIONE, ASCORBIC ACID AND RELATED MOLECULES

Patricia R. Laurence

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1982

A THEORETICAL INVESTIGATION OF THE ELECTRONIC STRUCTURE AND PROPERTIES OF GLUTATHIONE, ASCORBIC ACID AND RELATED MOLECULES

A thesis

presented for the degree of

Doctor of Philoophy

in the Faculty of Science of the

University of St. Andrews

by

Patricia R. Laurence, B.Sc.

St. Leonard's College, 1982

TO MY PARENTS

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ABSTRACT

The methods of molecular quantum mechanics have been used to study the important biological molecules glutathione, ascorbic acid and related compounds. The preferred conformations of glutathione, its constituent amino acids and their residues have been investigated using the semi-empirical PCILO procedure. Comparisons with ab initio calculations are given for the smaller molecules. The electronic structure and properties of ascorbic acid and its metabolites have been studied using ab initio procedures. The calculations include geometry optimizations, electrostatic molecular potential maps, spin density calculations on the radical species and investigations of charge transfer interactions and of metal complexing.

Declaration

I declare that this thesis is my own composition, that the work of which it is a record has been carried out by me, and that it has not been submitted in any previous application for a higher degree.

The thesis describes the results of research carried out at the Chemistry Department, St. Leonard's College, University of St. Andrews, under the supervision of Dr. Colin Thomson since the 1st December 1978, the date of my admission as a research student.

I hereby certify that Patricia R. Laurence has twelve terms of research work under my supervision, has fulfilled the conditions of the Resolution of the University Court, 1967, no. 1, and is qualified to submit the following thesis in application for the degree of Doctor of Philosophy.

Dr. Colin Thomson

Director of Research

Ackowledgements

I should like to express my thanks to my supervisor, Dr. Colin Thomson, for his expert help and guidance throughout my time as a research student.

Also to Dr. John Ball, who for the last three years has endured a constant barrage of questions and complaints with characteristic patience, and has been the source of much helpful advice and encouragement.

I should also like to thank the National Foundation for Cancer Research for the award of a Research Studentship and for the generous provision of funds enabling me to attend the Summer Institute in Quantum Chemistry, held in Finland in September 1980, and the 8th International Symposium on Quantum Biology, held in Florida in March 1981. I should also like to thank Per-Olov Lowdin and the other members of staff for their expert tuition on the former occasion.

My thanks must also go to:

Prof. Alberte Pullman for providing a copy of the d-PCILO program.

Prof. Jacopi Tomasi for providing a copy of the PUMABAL program.

Prof. John Pople for providing a copy of the Gaussian 80 program.

The staff of the Aberdeen and Cambridge University Computing Centres and particularly the staff of the St. Andrews University Computing Laboratory for a very generous allocation of computing resources.

I should also like to thank my family and friends for their interest and encouragement, in particular George Kwun for some much-needed help with the figures and Ken Moran for (voluntarily!) reading this manuscript.

Lastly, I should like to thank my mother for her unselfish understanding and encouragement throughout my time at St. Andrews.

NOTES

1. For chapters 4 to 10 the figures and tables appear, in order, at the end of the appropriate chapter, with the figures preceding the tables.

2. Energies are quoted in kcal mol^{-1} or Hartrees (atomic units). The corresponding values in kJ mol^{-1} are:

$$1 \text{kcal mol}^{-1} = 4.2 \text{ kJ mol}^{-1}$$

$$1 \text{ Hartree} = 6.275 \times 10^{2} \text{ kJ mol}^{-1}$$

CONTENTS

GLUTATHIONE AND RELATED MOLECULES

ASCORBIC ACID AND RELATED MOLECULES

1. INTRODUCTION

The Schrodinger equation, which underlies much of modern science, may be written

$$H\Psi = E\Psi \qquad (1)$$

where Ψ is the wave function, H the Hamiltonian operator and E the energy of the system.

In principle the solution of this equation should provide answers to almost all problems in chemistry. However, the equation cannot be solved by analytical techniques for systems containing more than one electron and for many years calculations were restricted largely to two and three particle systems. The development of the independent-particle model by Hartree [1], Fock [2] and Slater [3], and of the self-consistent field procedure by Roothaan [4] and Hall [5] paved the way for calculations on polyatomic molecules. It was the advent of the electronic computer, however, which made such calculations feasible and it is the advances in computer technology together with the development of efficient computer programs, that has made the greatest impact on quantum chemistry in recent years.

There has been a rapid increase in both the accuracy of calculations and in the size of the molecular system which can be handled. For molecules containing six to ten atoms it is now possible to calculate properties such as the equilibrium geometry to a high degree of accuracy (references [6] and [7] contain many examples of calculations on small molecules including some that are highly

accurate). On the other hand calculations have been performed on large systems such as polycyclic aromatics [8,9], model enzymes [10] and sizeable fragments of the DNA helix [11].

Biology too has made many advances in recent years. The sciences of biochemistry and molecular biology have produced insight into many natural processes and most biologists now think and talk freely in terms of molecules and chemical reactions. However, there are those, with Prof. Albert Szent-Gyorgyi prominent among them, who believe that a true understanding of many biological processes can only be achieved by going one step further, to the submolecular or electronic level.

Szent-Gyorgyi´s ideas on electronic biology, with particular reference to the development of cancer, are presented in several publications [12-14] and a brief discussion is given in chapter 3 of this thesis. One argument as to why biologists should concern themselves with reactions at the electronic level is that the energy changes involved in these reactions are of comparable magnitude to those known to occur in biological systems. Another reason is that while molecular biology has been very successful in a number of areas it still does not provide an adequate explanation for the subtlety of many biological processes. In general, experimental biologists have been cynical about such ideas and it has been left largely to theoreticians to explore this area.

Quantum mechanics provides a detailed description of the electronic structure of molecules and of their properties and interactions. As such it is a useful tool for the investigation of electronic and molecular processes in biology. In recent years calculations have been performed on a multitude of different topics

from the permeation of biomembranes [15] to the action of neurotransmitters [16].

Theoretical calculations are most useful when used in conjunction with experimental results. They may have a predictive value, however, in cases where no results are available, suggesting further experiments to be carried out and sometimes providing details of structure and properties not readily obtainable from experimental investigation.

The studies reported in this thesis were largely inspired by the ideas and experimental observations of Albert Szent-Gyorgyi and co-workers [12-14]. The work was undertaken in an attempt to discover more about some of the molecules thought to be important in the control of cell proliferation, concentrating on the molecules glutathione and ascorbic acid, and species closely related to them.

The main part of the thesis falls into two sections. The first deals with the tripeptide glutathione and is largely concerned with the conformational flexibility of this molecule and of its constituent amino acids and their residues. Of necessity, since this is such a large molecule, the majority of calculations were performed using a semi- empirical procedure (PCILO) but some comparisons with ab initio results are also included. The second section deals with ascorbic acid and some of its metabolites. The results of a large number of ab initio calculations performed on these molecules, and on model compounds chosen to represent them, are reported. The calculations include geometry optimizations, electrostatic molecular potential maps, spin density calculations on the radical species and investigations of charge transfer interactions and of metal complexing.

An outline of the methods used in these calculations is given in the next section, together with a discussion of the properties investigated.

2. PROCEDURES AND PROPERTIES

A full account of the quantum theory and its applications to molecular calculations would be inappropriate here. This chapter outlines the various approximations that are normally made in the calculation of the molecular wave function, describes the computer programs used in this work and discusses the properties investigated.

2.1 Approximate Wave Functions

The Schrodinger equation may be solved exactly only for a two-particle system, such as the hydrogen atom or He^+ ion. For all molecules the Born-Oppenheimer approximation [17] is made. This separates nuclear and electronic motion and considers the latter with respect to a fixed nuclear system. (This approximation is justified by the large ratio of nuclear:electronic mass.) In addition, relativistic and magnetic effects are normally neglected and the resulting Hamiltonian, expressed in atomic units, is:

$$H = \sum_A \sum_{B>A} \frac{Z_A Z_B}{R_{AB}} + \sum_i \left\{ -1/2 \nabla_i^2 - \sum_A \frac{Z_A}{r_{iA}} \right\} + \sum_i \sum_{j>i} \frac{1}{r_{ij}} \qquad (2)$$

The first term in equation (2) describes the repulsions between all pairs of nuclei, with nuclear charges Z_A, Z_B and internuclear

separation R_{AB}. The second summation is over all electrons in the molecule and includes the one-electron operators for the kinetic energy and for the attraction between nuclei and electrons, where r_{iA} is the distance between electron i and nucleus A. The final term represents the repulsions between pairs of electrons, r_{ij} being the distance between electrons i and j. The energy of such a system, in any state described by the approximate (normalised) wave function ψ, is given by the expectation value of the Hamiltonian operator:

$$E = \int \psi^* H \psi \, d\tau \qquad (3)$$

Almost all methods for the calculation of the molecular wave function make use of the variation principle [18]. This is based on the fact that the energy calculated for any trial wave function is a rigorous upper bound to the true energy. In general ψ is expanded in terms of a number of parameters which are varied to give the lowest possible energy. The more flexible the wave function the closer this value approaches that of the exact solution.

The most general type of approximate wave function commonly used with the variation principle relies on expanding the wave function in terms of some finite set of known functions. The selection of this set of functions (usually referred to as the "basis set") is of prime importance in determining the quality of the wave function and this topic is dealt with in some detail below (section 2.2.2).

In the orbital approximation each electron is treated independently and the wave function ψ is expanded as an antisymmetrized product of one electron molecular orbitals:

$$\Psi = |\phi_1 \phi_2 \phi_3 \ldots\ldots \phi_m| \qquad (4)$$

where each ϕ_i is a function of the space-spin coordinates of one electron only, and the expression on the right represents the diagonal of a determinant. Minimization of the energy resulting from a single determinant wave function leads to a set of integrodifferential equations known as the Hartree-Fock equations.

In the Restricted Hartree-Fock model [19] constraints of equivalence and symmetry are applied to the wave function. The first corresponds to a pairing of the spin-orbitals, the orbitals occurring in pairs with a common spatial orbital but different spin factors. The symmetry constraint requires the orbitals to belong to one of the various irreducible representations of the symmetry point group of the system. Various methods are available for the solution of the RHF equations. The majority of calculations on closed-shell species reported in this work were performed using the method due to Roothaan [4], though some use was also made of the direct minimisation SCF procedure [20].

The restricted Hartree-Fock model has been extensively used to study a wide range of molecules and it has formed the basis for much of our understanding of the structure and properties of molecules. There are some properties, however, for which this model does not provide a good description. It is particularly poor in describing the energetics of dissociation processes [21]. Another property which it does not describe adequately is the spin density at the nucleus in open-shell species containing π electrons. This term makes an important contribution to the hyperfine structure of open-shell

molecules. The deficiency may be overcome by relaxing the constraint of double-occupancy, leading to the spin-unrestricted Hartree-Fock model (SUHF, or UHF). Most of the calculations on open-shell systems reported in this thesis were performed using the UHF model [22].

A problem that occurs with this model is that the wave function is not an eigenfunction of the spin operator S^2. This can lead to difficulties if one is trying to obtain an energy minimum, for example in geometry optimizations, since there may be contamination from states of higher multiplicity. For such calculations it may be preferable to use a restricted Hartree-Fock procedure to find the energy minimum and then perform an UHF calculation for the minimum energy configuration [23]. Alternatively annihilation [24,25] or projection [26,27] operators may be applied to the wave function to remove the contaminating spin states [28].

If the wave function is allowed to be completely flexible, the energy obtained will be the Hartree-Fock energy. This is not, however, the total energy of the system:

$$E_{total} = E_{Hartree-Fock} + E_{relativistic} + E_{correlation}$$

The relativistic term arises from the fact that the Schrodinger equation is not relativistically correct. Core electrons move at speeds approaching the speed of light and relativistic effects operate, affecting the mass and energy. Such effects are important only for electrons in the innermost orbitals and since chemical reactions involve only the valence orbitals the relativistic term will normally be approximately constant during chemical changes and may therefore be neglected.

The correlation energy presents a more serious problem [29,30]. It arises as a consequence of the approximations made in the Hartree-Fock model in which each electron is regarded as being independent of all other electrons. In reality, of course, the positions of the electrons will be correlated. In the calculation of properties, such as molecular conformation, which do not involve a change in the pairing of electrons the correlation energy will be approximately constant and is unlikely to have a significant effect on the final result. For other properties, for example ionization potentials and dissociation energies, the error due to the neglect of electron correlation may be important.

Improvements on the Hartree-Fock energy, including the problem of electron correlation, may be provided by such methods as configuration interaction [30,31] or perturbation theory [32-34]. The former allows the mixing in of some excited states to provide a more flexible wave function; the latter treats the difference between the Hartree-Fock model and the exact model as a perturbation on the Hartree-Fock model.

These methods of going beyond the Hartree-Fock wave function are now available as routine options in some computer packages and no doubt their use will become increasingly widespread. At the present time, however, they are too computationally expensive to be used on large biomolecules and their use in this thesis has been restricted to one or two calculations on very small molecules carried out for comparative purposes.

2.2 Computational Aspects

The methods currently available for the calculation of the molecular wave function and the computer programs derived from them may be broadly divided into the two categories of *ab initio* and semi-empirical. In general *ab initio* procedures employ only those approximations outlined in the preceding section. The wave function may still be a long way from the exact solution since, in addition to the approximations inherent in the method, it is usually necessary to choose a fairly small set of basis functions, thus placing a further restraint on the wave function. The great advantage of *ab initio* procedures is that the approximations are clearly defined and in principle can be eliminated, for example by using extended basis sets or including configuration interaction. They are, however, computationally expensive and some *ab initio* programs may also be severely limited as to the size of system they can handle.

Until recently the majority of calculations on biological molecules (to be understood here as molecules of less than 50 atoms occurring in or having an effect on biological systems) were performed using semi-empirical procedures. Historically these were the first molecular orbital methods to be exploited and today there is a wide selection to choose from.

Semi-empirical procedures can broadly be defined as those which employ, at some stage in the calculation, parameters derived either from experiment or from other calculations. In general they are computationally easier to handle and much less expensive than *ab initio* procedures. It has also been argued [35] that as an

investigative tool for chemists it is better to derive procedures that will reproduce experimental results and can be directly related to them. Such procedures may be useful provided the property being calculated is related to the source of parameterization. Procedures such as CNDO and INDO [36-38] have been designed to give agreement with ab initio results rather than with experiment and so avoid any bias towards a particular property. Such calculations cannot of course be argued to be "better" than the ab initio results on which they are based.

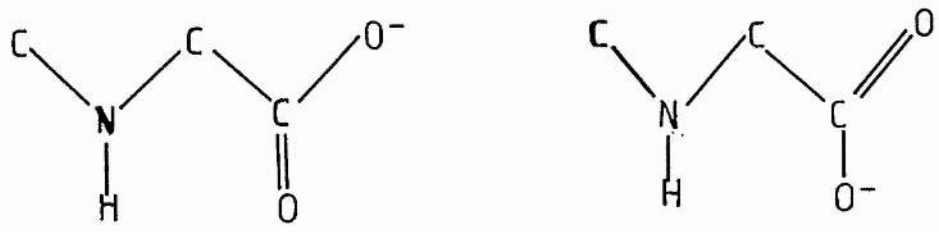
The majority of semi-empirical methods currently available, their applications and limitations, are discussed in reference [39] and a review, including comparison with ab initio methods, may be found in reference [40]. The only semi-empirical method used extensively in this work is the PCILO method and this is discussed below.

2.2.1 The PCILO Method

The acronym PCILO stands for Perturbative Configuration Interaction using Localised Orbitals. Full details of the method are given in reference [39] (vol.7, chap. 3) and in the original papers [41-44]. PCILO retains the CNDO approximation for the calculation of the electronic integrals but goes beyond the SCF approximation by performing limited configuration interaction. The starting point is a single determinant built from localized bonding orbitals. These are in turn constructed from hybrid atomic orbitals taken in pairs. At the same time a set of localized antibonding orbitals are formed. These are used to produce excited state configurations, which form the basis for the configuration

interaction. Rayleigh-Schrodinger perturbation theory is used up to third order to determine the energy and wave function.

The method is applicable to closed-shell molecules containing atoms which lie between H and Cl (inclusive, but excluding He and Ne) and for which a Lewis-type localised picture may be proposed. It may also be applied to a "closed-shell" ionic or zwitterionic species, the excess or lack of nuclear charge being distributed in the bonds and lone pairs of the atom concerned. Because it is necessary to supply the positions of the lone pairs as input (in the form of fictitious atoms), difficulties arise in describing species for which two or more canonical representations may be written. For example the two structures (a) and (b) in fig. 2.1 give rise to different energies.



(a) (b)

Fig. 2.1 Non-equivalent representations of ionized carboxylic acid group

Strategies that have beeen suggested [45] to overcome this difficulty are to take the structure with the best zeroth order energy and use it to calculate the third order energy, or to use the structure giving the best third order energy, or to take a mean of the third order energies. In dealing with ions of this type in this work the first two approaches were used and, in general, were found to give similar energy maps.

PCILO has been specifically designed for use in the calculation of molecular conformations. The program contains options to vary bond lengths and valence angles as well as dihedral angles. Its use to calculate molecular geometries, however, is not recommended. It is possible to vary two (or more) torsional angles simultaneously (corresponding to rotation about two different bonds in the molecule) and in this way a rotational potential energy surface may be constructed very readily. The method has found a wide application in the conformational study of biomolecules. Amino acids residues and dipeptides [46], nucleic acids [47], pharmacologically active compounds [48,49] and phospholipids [50] have all been studied with this method, and the problem of solvation has also been considered [51,52]. Possibly the largest molecules to be studied using PCILO are valinomycin [53] and glycine and alanine tripeptides [54]. These molecules contain 42 atoms, 132 electrons and 32 atoms, 92 electrons respectively. The tripeptide glutathione studied in this work contains 37 atoms and 162 electrons and is thus one of the largest molecules to be investigated by this method.

As compared with other methods, PCILO does considerably better than the empirical potentials which were previously used in

conformational analysis and at least as well as CNDO, usually better, with the added advantage that it is faster. Comparisons with ab initio calculations are reported in the literature for acetyl choline [55,56], serotonin [57,58], phenylethylamines [59-61] and neutral and zwitterionic glycine [45]. Chapter 5 of this thesis contains similar comparisons for glycine, cysteine and N-acetyl glycine. The general conclusion is that PCILO may be very useful as a rapid means of selecting a few "preferred" conformations from the multitude of possibilities available to a molecule.

The PCILO calculations reported in this work all use a modified version of the d-PCILO program [62]. The calculations were performed on the IBM 360/44 at St. Andrews University and the Honeywell 66 at Aberdeen University.

2.2.2 Ab initio procedures: the choice of a basis set

Although semi-empirical procedures may appeal because of their rapidity and close links with experiment the literature contains numerous examples of their shortcomings. Ab initio calculations, particularly with a minimal basis set, may be far from accurate but they do in general provide a better description of the molecular wave function than that available from semi-empirical calculations. A minimal basis set ab initio calculation is likely to provide a better starting point for the determination of most molecular properties than a semi-empirical calculation.

As mentioned earlier, the most important factor governing the quality of the wave function calculated by an SCF method is the choice of the set of functions used for the expansion of the molecular

orbitals. There is no way to overcome a poor choice of basis functions. A basis set may be "poor" not only if too few functions are used but also if there is a lack of balance between the basis sets describing the various centres in a molecule.

Slater type orbitals of the form $r^{n-1}e^{-\zeta r}Y_{lm}(\theta,\phi)$ were the first to be used in atomic and molecular calculations but the majority of calculations today use Gaussian type functions [63]: $x^l y^m z^n e^{-\alpha r^2}$. Gaussian lobe functions [64] may also be used but they have not been employed in this work and will not be discussed further.

Gaussian functions are computationally much easier to handle but are less appropriate for expanding the electronic wave function of molecules than Slater type functions. Consequently a larger set of functions must be used to obtain equivalent results.

The use of larger Gaussian basis sets increases computational requirements in two ways. Firstly the number of (two electron) integrals to be computed increases approximately as the fourth power of the number of basis functions. Secondly, the time required for the calculation of the electronic wave function is strongly dependent on the number of integrals. Although integrals over Gaussian functions can be rapidly computed the amount of disk space required to store the integrals may become prohibitive. A common way of reducing both the space and time requirements for a calculation is through the use of contracted functions. A contracted basis function may be defined [65] by

$$\chi_i = \sum_j \chi'_j \, c_{ji} \qquad (5)$$

where $\{\chi'_j\}$ denotes the original basis set, called the primitive set

and $\{\chi_i\}$ the contracted set. The contraction of a large set of Gaussian functions into a compact, but suitably flexible, set of expansion functions forms the basis for the modern, efficient use of Gaussian functions in molecular calculations.

The determination of optimized and balanced basis sets and the choice of suitable contraction schemes is a difficult and time-consuming process. Fortunately a large number of well-balanced basis sets are available in the literature, for example [66-76]. Reference [77] reviews some of the basis sets currently available and gives a useful bibliography.

The choice of basis set will depend on the problem to be investigated. Because of the large size of most biological molecules, calculations are frequently limited to the use of a minimal basis set. Minimal basis sets use the smallest possible number of functions in the expansion of the molecular orbitals. For Slater type orbitals this corresponds to one function per atomic orbital for each atom in the molecule, i.e. 1 for H, 2 for Li, 5 for C,O etc.. Such a basis set is often referred to as "single zeta". The widely used STO-NG basis sets developed by Pople and co-workers [67] retain the concept of a Slater type orbital but replace the STO by a linear expansion of Gaussian type functions determined by a least squares fit. Calculations have shown that N must be approximately 4 to produce energies equivalent to a minimal basis STO [67]. The STO-3G basis, however, has been very widely used for molecular calculations.

Details of the construction of these basis sets are given in the original paper [67]. As a means of reducing the time for the calculation of the integrals the orbital exponents of 2s and 2p are constrained to be equal. (Basis sets without such constraints have

also been constructed [78].) These basis sets for the atoms H to Ar are internal to the Gaussian 70 [79] and Gaussian 80 [80] program packages. The general form of, for example, an s-type function is

$$\phi_\mu(r) = \sum_{i=1}^{N} d_{i\mu} \exp(-\alpha_i f^2 r^2) \qquad (6)$$

where $d_{i\mu}$ are the contraction coefficients, $\{\alpha_i\}$ are the exponents and f is the scale factor.

The majority of ab initio calculations reported in this work were performed with the STO-3G basis set but the basis sets of Mezey and Csizmadia [71] and of Huzinaga et al. [72-75] were also investigated. Mezey's 6s3p basis with a minimal contraction scheme was found to be particularly useful. Although the authors recommend the use of split valence contraction schemes (see below) for the study of molecules, the minimal contraction scheme also gave a much better value for total energy and orbital energies than that obtained with STO-3G. In addition it is very economical in the amount of disk space required for integral storage. The basis set retains the constraint of equal orbital exponents and for this reason is particularly suitable as input to the Gaussian 80 program package.

The basis sets of Huzinaga et al. were investigated with a view to their use in the calculation of metal complexing. Basis sets for the atoms Li to Ne [74], Na to Ca [75], and Sc to Zn [73] are given. These basis sets do not attempt to fit Gaussian functions to a STO and the contraction procedure that is adopted [72,75] avoids the common problem with GTOs that too many functions are spent in the description of the region near the nucleus at the expense of the valence region. For this reason these basis sets are particularly useful for molecular calculations.

As an improvement on the minimal basis set, "split valence" sets

are often employed. These basis sets retain a minimal contraction for the innermost orbitals but allow more functions for the valence orbitals. The N-21G and N-31G basis sets of Pople's group [68,69] are of this type, as are the MIDI-i basis sets of Huzinaga et al. [74,81,82] and Mezey's contraction schemes B,C and D [71]. The use of this type of basis set is recommended for molecular calculations [71,74,77,81,82]. The values obtained for molecular properties are usually a considerable improvement on those obtained with a minimal basis set.

A further improvement of the basis set may be obtained by the addition of polarization functions. These are functions of p or d type (depending on the atoms present) which, because of their different symmetry, have a polarizing effect on the other orbitals. The inclusion of polarization functions is particularly important for ionic species and is essential for an accurate determination of molecular properties [77].

A split valence basis set with added polarization functions should provide a good description of a molecule and of its properties. Using such a basis set Huzinaga et al. have produced results of near Hartree-Fock accuracy for some properties [82].

Bearing all this in mind, the fact remains that for most biological molecules calculations with anything more than a minimal basis set are not feasible. However, if one is interested, for example, in the general shape of one molecule as compared to another, rather than a detailed and accurate prediction of geometry, then a minimal basis set may be adequate. Similarly, in a series of related molecules, trends in molecular properties predicted by minimal basis set calculations are likely to be reasonable, even if the numbers

themselves may not be accurate.

Because of the size of the systems investigated in this work, calculations have been largely restricted to the use of minimal basis sets. The limitations of such calculations have, however, always been born in mind, and wherever possible checks against more rigorous calculations have been made.

All the ab initio calculations reported in this work were performed using the Gaussian 70 and Gaussian 80 program packages and these are briefly discussed below.

2.2.3 Gaussian 70 and Gaussian 80

The Gaussian 70 program package used in these calculations is a modified version of that obtained from QCPE [79]. The modifications include facilities for performing damping and level shifting [83] in case of convergence difficulties in the SCF, and the energy decomposition analysis due to Morukuma [84]. The program can handle up to 35 atoms, 70 basis functions. STO-NG and 4-N1G basis sets for H to Ar and are internal to the program. For these basis sets optimization of the scale factors is possible. The atomic numbers of the component atoms and the geometry of the molecule must be provided as input. Facilities exist for limited geometry optimization and potential surface scanning.

The SCF procedure due to Roothaan [4] is used for closed-shell species and that due to Pople and Nesbet [22] for open-shells. The program calculates the total energy, molecular orbital coefficients and electric dipole moment, and performs a Mulliken population analysis [85] to give the net charge on the atoms and overlap

populations. For open-shell species atomic spin-densities and a value for S^2 are also given.

The Gaussian 80 program package [80] is an extension of Gaussian 70, with several additional features and facilities, and designed to run on a VAX computer. Additional basis sets that are available are N-21G, 6-31G, 6-311G [69,70], and LP-31G (local potential) [86]. Polarization functions may be added to most of the internal basis sets. A restricted open-shell procedure [87] is available as well as the usual RHF and UHF procedures. The UHF procedure is followed by annihilation of the largest spin contaminant. In case of convergence difficulty a direct minimisation SCF program is available [20] and our version of the program has recently been modified to allow damping. Mulliken population analysis is performed as in Gaussian 70. Spin densities and Fermi contact terms are calculated for open-shells.

Options are available to go beyond the Hartree-Fock level. Electron correlation energy may be calculated by Moller-Plesset perturbation [32] to second [33] or third [34] order and configuration interaction with all double substitions [88] from the Hartree-Fock reference determinant may also be performed.

One of the most useful features of the program is the availability of efficient geometry optimization procedures. Procedures based on the Fletcher-Powell technique [89,90] and on the Murtaugh-Sargent technique [91,92] are available as well as a procedure using analytic gradients [93,94].

Full details of options and limitations for both programs are contained in the program documentation.

2.3 Properties Investigated

Having obtained the electronic wave function for a stationary state of an N-electron system it is possible to calculate from it those properties which depend primarily on the electronic structure in this state. The calculation of molecular properties provides a link between theory and experiment, enabling quantitative tests of the relative merits of different models and approximations to be made and also providing a basis for the interpretation of experimental results.

In principle all primary properties may be computed as the expectation value of an appropriate operator:

$$O = \int \Psi^* \hat{O} \Psi \, d\tau \qquad (7)$$

It is possible to calculate a large number of molecular properties which may be directly or indirectly related to experimental observations, for example spectroscopic constants, dissociation energies and reaction pathways. The discussion here, however, will be restricted to properties investigated in this work.

2.3.1 Total energy and orbital energies

The total energy of a molecule in a given state may be computed from the wave function for that state using the Hamiltonian operator. In the orbital approximation the total energy is computed from the electronic wave function and the term for nuclear repulsion must then be added.

The electronic energy can be expressed in terms of integrals involving the molecular orbitals, ϕ_i:

$$E_{el} = \sum_i e_i^N + \sum_{i<j} J_{ij} - \sum_{i<j}' K_{ij} \qquad (8)$$

where $$e_i^N = \int \phi_i^*(1) H^N \phi_i(1)\, d\tau_1 \qquad (9)$$

$$J_{ij} = \iint \phi_i^*(1)\, \phi_i(1) \left(\frac{1}{r_{12}}\right) \phi_j^*(2)\, \phi_j(2) d\tau_1 d\tau_2 \qquad (10)$$

$$K_{ij} = \iint \phi_i^*(1)\, \phi_j(1) \left(\frac{1}{r_{12}}\right) \phi_i(2)\, \phi_j^*(2) d\tau_1 d\tau_2 \qquad (11)$$

In equation (8) the first term takes account of nuclear attraction and kinetic contributions while the second and third terms represent interelectronic interaction.

Orbital energies e_i are also related to these three terms and for closed-shell molecules

$$E_{el} = \sum_i (e_i^N + e_i) \qquad (12)$$

Many semi-empirical methods make the approximation $E_{el} = \sum e_i$. In addition these procedures employ parameters in the calculation of the total energy and it is thus not meaningful to compare values for the total energy obtained with different semi-empirical methods. Within one method, however, values may be compared and these energy differences may be compared between methods.

The orbital energies e_i are obtained from the solution of the secular determinant. Of particular interest are the highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbitals. The energy of

the HOMO gives an indication of how readily an electron may be removed from the molecule. For some anions the HOMO may have a slightly positive value. This implies that the molecule is unbound and in general the wave function will be a poor approximation to the Hartree-Fock wave function [95-97].

Koopmans' theorem [98] equates the ionization potential of an electron with the orbital energy e_i associated with that electron. While this is clearly a drastic approximation it does provide a useful means of comparing trends within a series of related molecules. Ionization potentials may be calculated more rigorously by computing total energies for the neutral and cationic species. If accurate values are to be obtained, however, it is necessary to perform geometry optimizations before and after ionization and to include correlation energy. Electron affinities may be calculated similarly [96,97], employing the anionic species. Again a rough indication of trends can be obtained from the energy of the LUMO. Care must be taken, however, as no definite physical meaning can be attached to the virtual orbitals and they are variable as a function of basis set size.

2.3.2 Dipole moment

The classical dipole moment of a system of N nuclei at positions R_A and m electrons at positions r_i is

$$\mu = e\sum_{A=1}^{N} Z_A R_A - e\sum_{i=1}^{m} r_i \qquad (13)$$

(independent of the positions of the origin of the coordinate system

if the molecule is electrically neutral). The value of the dipole moment in a state described by the normalized wave function Ψ is, for a single determinant wave function,

$$\mu = e\sum_{A=1}^{N} Z_A R_A - e\sum_{i=1}^{m} \mu_i \qquad (14)$$

where $-e_i$ is the contribution of an electron in the orbtial ϕ_i:

$$\mu_i = \int r\, \phi_i^* \phi_i \, dv \qquad (15)$$

The dipole moment provides a fairly sensitive measure of the quality of the wave function. The RHF model frequently gives better than qualitative agreement with experimental values. For some molecules with small dipole moments, however, an incorrect polarity may be predicted, for example CO [99].

2.3.3 Spin density and hyperfine coupling constants

The spin density is directly related to the one-electron density function. The density function $\rho(r,\sigma)$ of any N-electron wave function which has a well-defined value of total spin M_s can always be expressed as the sum of separate contributions from the two spin components:

$$\rho(r,\sigma) = P^{\alpha}(r)|\alpha(\sigma)|^2 + P^{\beta}(r)|\beta(\sigma)|^2 \qquad (16)$$

in which $P^{\alpha}(r) = |\psi^{\alpha}(r)|^2$ and $P^{\beta}(r) = |\psi^{\beta}(r)|^2$ are spinless densities associated with the spins +1/2 and -1/2 respectively. The

quantity $P^{\alpha}(r)dv$ is interpreted as the probability of finding an electron with spin +1/2 in the volume element dv at position r. The spin density may be regarded as the excess of density of spin +1/2 over spin -1/2, i.e.

$$P^{\alpha}(r) - P^{\beta}(r).$$

The spin density vanishes for a singlet state but for free radicals it is an important quantity, being related to the hyperfine coupling constants measured in electron spin resonance spectroscopy.

For radicals in solution the major contribution to the hyperfine coupling constant is the Fermi contact term, which depends on the spin density at the nucleus and is given by

$$A_{iso}(A) = (\frac{8\pi}{3})g_e \beta_e g_A \beta_A(P^{\alpha}(r) - P^{\beta}(r)) \qquad (17)$$

The Fermi contact term is a very good test of the behaviour of the wave function at the nucleus. There is, however, a serious deficiency in the RHF approximation in that for many open-shell species it predicts an identically zero Fermi contact term while the experimental spin density is known to be non-zero. Use of the UHF method ensures a prediction of non-zero spin densities. The values obtained, however, are seldom in good agreement with experiment. The semi-empirical procedure INDO gives reasonably good values for spin densities and hyperfine coupling constants and this has been used for their calculation in this work. A discussion of the theoretical calculation of hyperfine splittings is given in references [100,101].

2.3.4 Population analysis

The concept of population analysis, in which the total molecular charge density is partitioned into atomic and overlap (or bond) contributions, has proved to be a useful one in the interpretation of electronic structure calculations. The most widely used method is that due to Mulliken [85], which can be applied to any molecular wave function expressed in terms of atomic orbitals centred on the nuclei.

For a diatomic molecule with molecular orbital $\phi_i(r) = a_i A(r) + b_i B(r)$ the corresponding molecular-orbital density function is

$$P^i(r) = |\phi_i(r)|^2 = P_A^{(i)}(r) + P_B^{(i)}(r) + P_{AB}^{(i)}(r) \qquad (18)$$

where $$P_A^{(i)} = a_i^* a_i A(r)^* A(r)$$

$$P_B^{(i)} = b_i^* b_i B(r)^* B(r)$$

$$P_{AB}^{(i)} = a_i^* b_i A(r)^* B(r) + b_i^* a_i B(r)^* A(r)$$

The quantities $P_A^i(r)$ and $P_B^i(r)$ are atomic densities, while $P_{AB}^i(r)$ is an overlap density. For a polyatomic molecule in the orbital approximation the total molecular density function is the sum of the molecular orbital contributions:

$$P_{(r)} = \sum_i n_i P^{(i)}(r) \qquad (19)$$

where n_i is the occupation number of orbital ϕ_i ($n_i \leq 2$). This may

also be partitioned into contributions from the atoms and pairs of atoms:

$$P_{(r)} = \sum_{A} P_A(r) + \sum_{A>B} P_{AB}(r) \qquad (20)$$

where $P_A(r) = \sum_i n_i P_A^{(i)}(r)$ and $P_{AB}(r) = \sum_i n_i P_{AB}^{(i)}(r)$

are total atomic and total overlap density functions respectively. In this way the n_i electrons in the molecular orbital ϕ_i can be partitioned amongst the atoms and bonds to give net orbital atomic and overlap populations. These in turn may be combined to give a partitioning amongst the atoms only, which requires a separation of each overlap population into two parts to be associated with the relevant atoms. The resulting quantities are gross populations on each atom, N(A), and the quantity -eN(A) can then be interpreted as the electronic charge on atom A.

One defect of this method is the equal division of the overlap populations, which is clearly unrealistic unless the atoms are equivalent. A further drawback of the method is that the partitioning into atomic and overlap densities is highly sensitive to the numbers and types of atomic orbitals used as a basis and erroneous results may arise if the basis set is not balanced.

Various other schemes for the partitioning of the total molecular charge density are available [102-106]. The Mulliken population analysis, however, continues to be widely used and in spite of its deficiences has frequently been found to reproduce the essential features and trends of the electron distribution within a molecule or series of molecules.

2.3.5 Geometry optimization

The determination of molecular geometries by ab initio calculations may be used either to compare theoretical with experimental results, and hence assess the reliability of the former, or to predict structures for which insufficient experimental data is available. Even when geometries are known from experiment it is customary to perform geometry optimizations in order to obtain consistent theoretical results since calculations with a basis set of limited size at an experimental geometry will almost certainly produce values for the molecular energy higher than the minimum obtainable for that basis set.

The results of geometry optimizations are strongly dependent on the basis set used. Hartree-Fock theory with a minimal basis is moderately successful in reproducing geometries of neutral closed-shell molecules in the ground state. Bond lengths are accurately given to about 0.03Å and angles to within 4^{o}. The use of extended basis sets with polarization functions produces structures close to those of the Hartree-Fock limit [107].

Procedures for the prediction of equilibrium geometries usually involve the calculation of the total energy as a function of the geometric parameters and the use of some type of search procedure to obtain the structure of minimum energy. The geometry optimizations reported in this work were performed using the gradient technique.

The use of energy gradients for the determination of molecular geometries was first proposed and developed by Pulay [93,94,108]. In this method the energy is first differentiated analytically with

respect to the nuclear coordinates q_i, i.e. the force acting in the direction of q_i is calculated:

$$f_i = -\frac{\partial E}{\partial q_i} \qquad (21)$$

The force constant F_{ij}, the negative derivative of the force f_i as a function of q_j, is then determined numerically, carrying out calculations for different q_j's near the equilibrium configuration.

The equilibrium geometry is determined by allowing the nuclear coordinates to relax until the net forces on the atoms vanish. If $\underline{q}_0$ is the starting vector of internal coordinates and $\underline{F}_0$ a guess for the force constant matrix, a better approximation $\underline{q}_1$ may be obtained by calculating the forces $\underline{\phi}_0$ to form:

$$\underline{q}_1 = \underline{q}_0 + \Delta\underline{q} = \underline{q}_0 + \underline{F}_0^{-1}\underline{\phi}_0 \qquad (22)$$

If $\underline{F}_0$ and $\underline{q}_0$ are not too far from the correct values, a considerable improvement can be expected in $\underline{q}_1$ compared with $\underline{q}_0$. The iteration process:

$$\underline{q}_{i+1} = \underline{q}_i + \underline{F}_i^{-1}\underline{\phi}_i \qquad (23)$$

will then converge towards equilibrium.

The gradient technique has been incorporated into a number of ab initio programs [108-112,80] and has been widely used for the calculation of molecular geometries and force constants [94].

2.3.6 Conformational analysis

A knowledge of preferred conformation is of particular importance in the study of biological molecules. However, it is frequently of interest to know not only the optimum conformation of a molecule but also its conformational flexibility. The latter may be investigated by calculating the total energy of the molecule for different values of one or more torsional angles. If the molecule has only one degree of rotational freedom a curve of energy against torsional angle may be constructed. For two degrees of freedom results may be conveniently presented in the form of a conformational (or rotational potential) energy surface (see for example fig.4.5). Such a surface is produced by the simultaneous variation of the two torsional angles, ϕ,ψ, in, say, 30^{o} steps. The total energy of the molecule is calculated for each value of ϕ,ψ and "contours" linking points of equal energy are then drawn. The energies are usually given as values above the minimum energy conformation. In this work (in line with most of the literature) values are quoted in kcal mol^{-1} and contours are normally restricted to 5 or 6 kcal above the minimum.

For a molecule with more than two degrees of rotational freedom conformational analysis may be carried out by fixing all but two of the torsional angles and constructing an energy surface as described above. If the procedure is repeated for various combinations of the fixed torsional angles a series of "sub-maps" may be produced. These can then be combined to give a "global" map by taking for each value of ϕ, ψ the lowest energy presented by any sub-map (see for example [54,113,114] and chapter 4 of this thesis). Such a procedure is particularly useful when there are two angles of major importance, for example the backbone angles of a peptide fragment.

Ideally full geometry optimization should be performed for each point on the conformational energy surface. However, such a computationally expensive procedure is rarely followed , instead bond lengths and valence angles are normally fixed at values taken either from experiment, tables of standard values or previous calculations. While this is clearly an approximation it does appear that the overall features of a conformational energy surface are not influenced by small variations in geometry, and experimental values for barrier heights may be fairly well reproduced even without geometry optimization.

The information obtained from a conformational energy surface may be supplemented by the construction of a probability map [115]. In this work, however, such maps have not been used. A comprehensive review of the theoretical study of conformations may be found in reference [114].

2.3.7 Charge transfer interactions

In recent years quantum mechanics has been widely used to study intermolecular interactions. The interactions which are of interest in this work are weak interactions involving transfer of charge from donor to acceptor. This broad definition may be taken to include hydrogen-bonded complexes, which usually show a transfer of electronic charge from the hydrogen acceptor to the hydrogen donor.

There are a number of recent reviews on intermolecular interactions in general [116-119] and on hydrogen-bonding in particular [120-122]. These contain details of theory and methods as well as summaries of results.

The supermolecule approximation has been widely used in the study of intermolecule interactions. In this approach a molecular orbital calculation is performed for the molecular complex or cluster as if it were a single molecule, or "supermolecule". The interaction energy is given by the difference between the total energy of the complex and the sum of the energies of the isolated molecules:

$$\Delta E = E_{complex} - \sum E_{isolated\ monomers} \qquad (24)$$

The interaction energy is composed of five main terms. These are: (1) electrostatic energy, which is the classical interaction between the fixed nuclei and electron distribution of the individual molecules; (2) polarization energy, arising from the distortion of the charge cloud of one monomer by the presence of the other and vice versa; (3) exchange energy, which is a direct consequence of the Pauli principle and is due to repulsion between electrons of like spin; (4) charge transfer energy, a quantum mechanical attraction, which may be thought of as the interaction of occupied MO's of molecule A with vacant MO's of molecule B and vice versa, leading to electron dispersion and charge transfer; and (5) dispersion energy, the second order induced dipole-induced dipole attraction which is present even in rare-gas-rare-gas interactions.

Various decomposition schemes have been proposed [84,123-125] for the interaction energy. The most extensive is that due to Morokuma [84,123] which calculates electrostatic, polarization, charge transfer and exchange repulsion terms.

The majority of ab initio calculations on donor-acceptor complexes have been carried out within the Hartree-Fock (i.e. single

determinant SCF) approximation, which takes no account of the electron correlation. For weak intermolecular interactions the electron correlation may be divided into intramolecular correlation and intermolecular correlation. The first is unlikely to have an important effect on the energy since the formation of a weak complex will involve little change in electron pairing and the intramolecular correlation should thus remain approximately constant during complex formation. The intermolecular correlation, however, contains the dispersion energy as its leading term and the latter contributes significantly to the stabilization of weak complexes. This term cannot be calculated using a single-determinant wave function and it seems likely that extensive configuration interaction is needed for its accurate determination [117].

The ability of a single-determinant wave function to calculate the remaining terms will depend largely on the basis set used. Minimal basis sets often give reasonable values for the interaction energy. Detailed analyses reveal, however, that the charge transfer component of the interaction energy is unrealistically large [126]. This is due to the basis set superposition error which occurs when a molecule, described by a poor basis set, finds itself in the vicinity of another molecule and uses the additional basis functions to improve its own wave function. The error may be corrected by use of the counterpoise method [127,128].

The studies on donor-acceptor complexes in this thesis have been largely restricted to the use of minimal basis sets because of the size of the systems involved. Comparison with larger basis sets for some small model systems (see chapter 8) show that trends predicted by minimal basis sets are usually reliable and their widespread use in

the literature also supports this.

In the estimation of the amount of charge transferred from donor to acceptor the Mulliken population analysis has been used. The limitations of this method have already been discussed (section 2.3.4). However, a comparison with other procedures, with particular reference to the description of charge transfer processes, indicates that the Mulliken population analysis gives at least a useful qualitative picture of the direction and overall charge distribution when compared to more sophisticated procedures [106].

2.3.8 Electrostatic molecular potential

A recently developed tool for the theoretical investigation of molecular and structural properties is the electrostatic molecular potential [129-131]. The procedure characterizes a molecule by the global potential created in the surrounding space by the nuclear charge and the electron distribution. For a given wave function the value of the potential V(p) at a point p in space is:

$$V(p) = \sum_{\alpha} \frac{Z_\alpha}{r_{\alpha p}} - \int \frac{\rho(i)}{r_{pi}} d\tau_i \qquad (25)$$

where Z_α is the charge on nucleus α, $r_{\alpha p}$ and r_{pi} the distance of point p from nucleus α and electron i respectively, and $\rho(i)$ is the electron density distribution of the molecule. In the LCAO MO approximation $\rho(i)$ is given by

$$\rho(i) = \sum_r \sum_s P_{rs} \chi_r^*(1) \chi_s(1) \qquad (26)$$

where χ_r, χ_s are the atomic orbitals and P_{rs} is an element of the density matrix.

The equation (25) may be evaluated for an arbitrary wave function, semi-empirical or non-empirical, and is not restricted to an SCF framework. Ab initio SCF wave functions have been widely used in the calculation of electrostatic molecular potentials and since V(p) is the expectation value of a one-electron operator, Brillouin's theorem (see reference [98] p.137) implies that its SCF approximation is correct to one order higher than the SCF wave function employed.

V(p) also corresponds to the first order perturbation energy of a molecule under the influence of a unit positive charge. Molecular potentials are thus particularly useful in the study of protonation or other electrophilic reactions since a picture is provided of what the reagent "sees" when approaching the molecule.

Molecular potentials are most conveniently presented in the form of isopotential maps drawn in the space surrounding the molecule (see for example fig. 6.2). The values are usually expressed as energies (i.e. energy of interaction with a bare proton) in kcal or Hartrees. Since the maps give potential values in a particular plane, several maps may be necessary in order to characterize a molecule and if the molecule has no plane of symmetry, presentation of the results may be difficult. Having once calculated the wave function of a molecule any number of maps may be produced so that it is possible to study different areas or groups in a molecule.

Electrostatic molecular potentials have been calculated for a wide range of molecules including nucleic acid bases [132,133], the active sites of enzymes [134] and aromatic hydrocarbons [8,9,135].

Reviews giving the theoretical background with many examples may be found in references [136-138].

The electrostatic potential maps presented in this work were produced using the PUMABAL program [139] developed at Pisa university by Drs. Bonaccorsci, Petrongolo and Tomasi. The program is designed to use vectors from a wave function calculated with the Gaussian 70 program.

2.3.9 Solvent effects

The majority of molecular quantum mechanical calculations are performed on isolated molecules or molecular complexes. For biological molecules in particular this may be unrealistic since the environment is likely to have a strong influence on molecular properties and reactions. Comparisons between isolated-molecule calculations for a series of related molecules may be valid since the neglected environmental effects will be similar in all cases. For most studies, however, it is preferable to include some explicit consideration of the environment.

A large constituent of the biological environment is water and an increasing number of calculations on biological molecules include a consideration of the influence of water on the molecular properties under investigation.

There are basically two approaches to the study of solvent effects. The first attempts to account for the bulk effect of the solvent by use of a "continuum" model, essentially based on the classical electrostatic treatments of interacting systems by Born, Onsager and Kirkwood [140-142] inserted into a quantum mechanical

framework. This procedure has been developed in different ways by a number of authors [143-146].

The second method is the so-called "supermolecule" approach [147]. This concentrates on those water molecules which are likely to be tightly bound to the solute molecule. The procedure is to determine first the preferred position of hydration using a single water molecule as a probe. Further water molecules are then added to constitute a "first hydration shell" and the influence of the solvent on the properties of the solute molecule is then studied by performing a supermolecule calculation on the solute plus first hydration shell. The procedure may be repeated to give a second or third hydration shell. The two procedures may be combined to give a model in which the first hydration shell is determined as in the supermolecule approach and this is then inserted into a continuum representing the bulk effect of the solvent [148].

Another "combined" approach has been developed by Clementi and co-workers [149,150]. In this the interaction energy between solute and solvent molecules is computed for various orientations and fitted to a simple potential energy form which is then used in a Monte Carlo calculation. This method can provide a good description of the way in which solvent molecules will distribute themselves around a solute molecule. It is, however, computationally very expensive.

When dealing with biological molecules the inclusion of solvent effects, however desirable, is often prohibited by the size of the system. In this work it has not been possible to study the effects of solvent and, while it is acknowledged that such effects may play an important part in the systems investigated, calculations on the isolated molecules should at least provide a useful guide to the

properties of interest and serve as a starting point for more rigorous investigations.

3. SUB-MOLECULAR BIOLOGY AND CANCER

Sub-molecular biology was recently defined by R.J.P.Williams (in [14] p.1) as "...those chemical systems in biology at a lower level than whole molecules." The Shorter Oxford Dictionary defines cancer as "a malignant growth or tumour that tends to spread and reproduce itself". The man largely responsible for linking these two topics is Prof. Albert Szent-Gyorgyi. His "electronic theory of cancer" has been the subject of several publications [12-14] and a brief outline is given below.

The central proposition of the theory is that proteins in the normal living state behave as semi-conductors. It is this mobility of electrons that enables proteins to perform their diverse functions in nature. However, when proteins are isolated and purified they are found to be insulators, having filled valence bands. In semi-conductors the valence and conduction bands lie close together so that electrons may be excited into the conduction bands by heat agitation. In proteins, however, the valence and conduction bands are not sufficiently close together for this to occur and an alternative mechanism for semi-conduction must be found.

Szent-Gyorgyi has proposed that semi-conduction in proteins is brought about through "charge-transfer", that is, the removal of an electron by some suitable electron-acceptor, leading to desaturation of the valence band and allowing mobility of the electrons. (It is not necessary for the electron to be removed completely; sharing of electrons between donor and acceptor, such that there is a net transfer of electronic charge from the donor to the acceptor, is

sufficient to allow semi-conduction in a periodic molecule such as a protein.)

Methylglyoxal, $CH_3CO.HCO$, has been put forward as a possible electron-acceptor. The presence of two adjacent carbonyl groups suggest that the molecule may be able to accommodate additional electronic charge. Methylglyoxal is the smallest keto-aldehyde and occurs naturally in the human body. Previously no important biological function has been ascribed to the molecule in spite of the widespread occurrence of a reactive enzyme system which converts methylglyoxal to lactic acid. If methylglyoxal is indeed the universal electron acceptor then the existence of such an enzyme system would no longer be surprising.

It is suggested that methylglyoxal forms a Schiff base with the amino group of a lysine side-chain, which then folds in such a way that the glyoxal is brought close to the peptide backbone and can accept electrons from it. Theoretical calculations have shown [151] that the lysine chain could fold in such a way and there is experimental evidence that proteins "doped" with methylglyoxal do behave as semi-conductors [152]. Further experimental evidence is provided by e.s.r. studies [13] in which mixtures of methylamine and methylglyoxal are found to give a strong signal with a rich hyperfine structure. This signal is enhanced by the addition of ascorbic acid to the mixture. If casein is treated with methylglyoxal in the presence of ascorbate a new signal appears indicating that the ascorbate is also involved in the charge transfer complex. Szent-Gyorgyi has suggested that the ascorbic acid acts by giving up an electron to O_2, forming a cation radical [13]. The radical then complexes with methylglyoxal, making the latter into a strong acceptor

able to desaturate the protein and render it semi-conducting.

The link between semi-conduction in proteins and cancer is as follows. Szent-Gyorgyi views the history of life on the earth as divided into two periods. The first, the "α period" corresponds to a time when the globe was covered by a dense cloud of water vapour, the atmosphere was strongly reducing and there was no light or oxygen. Proteins had to be desaturated by weak acceptors such as methylglyoxal and only the simplest forms of life could develop. Such life forms had to survive by rapid proliferation, favoured by low cohesive forces and simplicity of structures. With the coming of light and oxygen the "β period" began in which the strong acceptor properties of oxygen made possible a much higher degree of unsaturation in proteins and a correspondingly high degree of development and differentiation.

The α <--> β transformation is reversible. During cell division structures are dismantled and there is a lowering of cohesive forces, corresponding to a partial return to the α state. After division the cell normally returns to the β state. If, however, the return to the β state is blocked in some way the cell will persist in the proliferative α state, continuing to divide and ultimately producing a tumour.

Szent-Gyorgyi regards control of cell proliferation as being brought about by antagonism between a "retarding" and a "promoting" substance. He identifies the former as methylglyoxal and suggests that the latter may be glutathione. The two are known to form a hemimercaptal [153] and glutathione acts as a coenzyme for the glyoxalase system which converts methylglyoxal to lactic acid. Cell division requires the presence of SH groups and methylglyoxal may inhibit cell division by binding the SH of glutathione, and possibly

other SH groups linked to proteins.

The theory outlined above clearly contains a good deal of speculation and is probably a gross simplification of the problem. However, it does represent a new approach, seeing cancer as a problem to be solved rather than a disease to be cured, and the lack of progress in conventional cancer research suggests that such a theory may at least be worth investigating. In recent years a good deal of research, both experimental and theoretical, has been directed towards an understanding of the cancer problem in terms of Szent-Gyorgyi's theory.(Much of the work has been carried out by members of the National Foundation for Cancer Research, an organization established in 1973 by Franklin C. Salisbury for this express purpose.)

E.s.r. spectroscopy has been used to study the reaction of methyl glyoxal with proteins [154] and with model amines [155,156]. The products which form in the latter system are not yet clearly identified but appear not to involve charge transfer, even though highly coloured solutions and precipitates are obtained and an e.s.r. signal is observed. A mechanism for the reaction between methylglyoxal and proteins, involving the lysine side-chain, has been proposed [154]. The presence of molecular oxygen and of ascorbic acid appear to be important, although the precise role of these two substances is not yet clear.

Experimental studies on doped proteins have been carried out [152,154], as well as theoretical band structure calculations [158,159]. The electron-acceptor properties of methylglyoxal [160-163] and of other ketoaldehydes [162,164] have been investigated by quantum mechanical methods.

This thesis reports the results of theoretical studies on two molecules which in different ways are important in the proposed charge transfer theory. The calculations on glutathione are the first to be reported for this molecule. It had been hoped to study the electronic structure of the adduct formed between methylglyoxal and glutathione but owing to the large size of the system this was not possible with the available computational resources. Although the study on glutathione largely takes the form of a semi-empirical conformational investigation it is hoped that this may provide the basis for a more detailed and accurate study of the molecule in the future. The studies on ascorbic acid and its metabolites include investigations of the electronic structure of neutral ascorbic acid, its anion, anion radical and dehydroascorbic acid. The donor/acceptor properties of these species are also reported and some suggestions are put forward as to the possible role of ascorbic acid in the charge transfer process.

GLUTATHIONE AND RELATED MOLECULES

4. CONFORMATIONAL STUDY ON GLUTATHIONE AND ITS CONSTITUENT AMINO ACID RESIDUES

4.1 Introduction

Glutathione, γ-glutamyl-cysteinyl-glycine, is widely distributed in animal and plant cells and in microorganisms. It is an important cellular constituent, occurring in relatively high concentrations (0.5-10mM). Glutathione has two characteristic structural features: a sulfhydryl (SH) group and a γ-glutamyl linkage. It is usually the most abundant intracellular thiol as well as one of the most abundant γ-glutamyl compounds. Several reviews of the biochemical properties of glutathione are available [165-167]. Among the functions ascribed to the molecule are protection of cell membranes and proteins through the maintainance of SH groups, destruction of peroxides and free radicals and detoxification of foreign compounds. Glutathione is also a coenzyme for the glyoxalase system and, as mentioned in the preceding section, it is this function and Szent-Gyorgi´s suggestion that glutathione is the "promotor" substance in cell regulation which are of particular interest in this work.

Glutathione was first isolated in 1888 [168,169]; later it was shown to be a tripeptide [170] and its structure established by synthesis [171]. The crystal structure of glutathione was determined by Wright [172] and more recently by Cole [173]. Two n.m.r. studies have been reported [174,175] but the two groups of authors reached different conclusions as to the preferred conformation of the molecule. No previous theoretical studies have been carried out on glutathione itself, although "residues" of its constituent amino acids

have been studied [176,113] and glycine, as the smallest amino acid, has been extensively studied [177-185,45]. A related theoretical study is that by Lavery and Pullman [186] on the mechanism of action of glyoxalase I but here the glutathione molecule is represented by a small model.

4.2 Procedure

The calculations were performed using the PCILO program (see above section 2.2.1). Conformational energy maps were produced as described in section 2.3.6 for pairs of torsional angles belonging to the backbone of the peptide chain (fig.4.1). (The torsional angle τ about the bond B-C in the sequence of bonded atoms A-B-C-D is the angle between the planes ABC and BCD. This angle is defined as positive for clockwise rotation around B-C when looking from B to C (fig.4.2).

In constructing the energy maps the convention is adopted that zero values of torsional angles correspond to fully eclipsed conformations [187], thus the fully extended peptide chain is represented by $\phi=\psi=180^{\circ}$. Stable conformations are described as C_5, C_7, M etc. following Pullman and Maigret [188], where C_n denotes an n-membered hydrogen-bonded ring and M conformations have adjacent peptide groups approximately perpendicular. C_5, C_7 and M conformations are illustrated for the glycyl model in fig. 4.3.

Initially optimization of bond lengths and angles was attempted using PCILO. The values obtained, however, were not in good agreement with experiment nor with ab initio results and for consistency

standard values of bond lengths and angles [189,190] were used throughout. For several of the molecules conformational energy surfaces produced with standard and calculated geometries were compared. In some cases slightly lower barriers to rotation were obtained for the calculated geometries but in general there were no significant differences, suggesting that small variations in geometrical parameters do not influence the conformational flexibility of a molecule (however, see comment at end of chapter 5).

The study of the preferred conformation of glutathione was carried out by considering first a series of model compounds representing fragments of the molecule. (These models are similar, but except in the case of cysteine not identical, to the "residues" used extensively by Pullman et al. in studies of di- and tri-peptides [188,54].) The division of the molecule into fragments is illustrated in fig.4.4.

The glycyl, cysteinyl and glutamyl fragments were first studied individually. Ionic as well as neutral species were considered for the glycyl and glutamyl models. The cysteinyl and glycyl framents were combined to give a cysteinyl-glycine model, which was again studied in neutral and ionic forms. In a similar way a glutamyl-cysteine model was formed by combining the glutamyl and cysteinyl fragments. Finally the complete molecule was studied in both neutral and ionic forms. The results of these calculations are presented in the sections which follow.

4.3 Glycyl Model

The neutral form of this molecule has three degrees of freedom (fig. 4.4). The barrier to ω rotation is ~4kcal, the preferred values being 0^{o} and 180^{o}. Two surfaces for ϕ, ψ rotation were produced for these values of ω (fig.4.5). Conformations giving rise to minima on the energy surfaces are listed in Table 4.1 and illustrated in fig. 4.3. The lowest energy is given by the M conformations for $\omega = 0^{o}$ but C_5 and C_7 conformations are also low in energy. ab initio calculations (see below, chapter 5) emphasize the importance of the C_5 and C_7 rings.

The ionized form of this molecule has only two degrees of rotational freedom. The problem of representing an ionized carboxylic acid group using PCILO and the approaches used to overcome it have been discussed above (section 2.2.1). THe energy maps produced indicate that the only important conformation is a planar C_5 arrangement in which there is hydrogen-bonding between NH and O^{-}. ab initio calculations confirm this.

4.4 Cysteinyl Model

There are four torsional angles to consider for this molecule. The side chain angles χ_1, χ_2 were restricted to staggered values i.e. 60^{o}, 180^{o}, 300^{o}. Nine sub-maps for ϕ, ψ rotation were constructed corresponding to the possible combinations of these angles. The global energy map is presented in fig. 4.6

This molecule is equivalent to the cysteinyl residue studied by Perahia, Pullman and Claverie [113] and the results follow closely those reported in their paper. The global minimum occurs for $\phi_1 = 90^o$, $\psi_1 = 30^o$ to 60^o which is the preferred conformation for all values of χ_1, χ_2 (illustrated in fig.4.7 for $\chi_1 = 60^o$, $\chi_2 = 180^o$). Perahia et al. report $\chi_1 = \chi_2 = 60^o$ to be the most stable combination of side-chain angles with $\chi_1 = 60^o$, $\chi_2 = 180^o$ 0.3kcal higher; the results of this study indicate $\chi_1 = 60^o$, $\chi_2 = 180^o$ to be the more stable with $\chi_1 = \chi_2 = 60^o$ 0.12 kcal higher. Within the error limits of the method (0.5 kcal) these two conformations must be regarded as equally stable. The arrangement with $\chi_2 = 60^o$ allows the formation of a hydrogen-bond between SH an O_2 but possibly this is too weak to be of any significance. Both conformations can form a weak hydrogen-bond between S and N_1 as well as the bond between N_2H and O_1. Other stable combinations of χ_1, χ_2 for these values of ϕ_1, ψ_1 are $\chi_1 = 300^o, \chi_2 = 180^o$ which offers the possibility of a hydrogen-bond between S and N_1H, and $\chi_1 = 180^o$, $\chi_2 = -60^o$ which could involve a bond between SH and O_2.

Other stable conformations are an M conformation at $\phi_1 = -30^o$, $\psi_1 = 120^o$ for $\chi_1 = -60^o$, $\chi_2 = 180^o$; and a number of conformations associated with $\phi_1 = \psi_1 = 180^o$ for various values of χ_1, χ_2, corresponding to the formation of a C_5 ring.

4.5 Glutamyl Model

In studying this molecule the aim was to find the preferred conformations of the glutamyl chain. Keeping the peptide link fixed

in the trans planar arrangement with the methyl group eclipsed there remain six torsional angles to consider. Each of these was varied in turn.

The preferred values of ϕ' and ω' are 60° and 0° respectively. ψ' showed a twofold minimum at 0° and 180° with values close to this ($\pm 30^{\circ}$) also being stable. χ'_1 variation shows the usual threefold minimum (60°,180°,300°) but the pattern for χ'_2 and χ'_3 is not so simple. Accordingly, for each preferred value of χ'_1 and ψ', χ'_2 and χ'_3 were varied simultaneously.

The most stable conformation occurs for $\psi' = 30^{\circ}$, $\chi'_1 = 60^{\circ}$, $\chi'_2 = 0^{\circ}$, $\chi'_3 = -120^{\circ}$ (illustrated in fig.4.8a). Some of the other stable conformations are listed in Table 2 and a typical example is illustrated in fig.4.8b.

Clearly there are a good many conformations that have not been studied and it may be that some important conformations have been overlooked. However, it seems possible to conclude that "closed" structures involving some interaction between the amino acid and peptide groups are favoured over "open" structures in which these groups are far apart.

The zwitterion of this molecule was studied in a similar way to the neutral species. $\phi' = \psi' = 0^{\circ}$ were found to be the preferred values for these angles. Combinations of χ'_1, χ'_2, χ'_3 giving rise to low energy conformations are as follows (fig.4.9):

(a) $\chi'_1 = 60^{\circ}$, $\chi'_2 = 270^{\circ}$, $\chi'_3 = 180^{\circ}$ to 270°

(b) $\chi'_1 = -60^{\circ}$, $\chi'_2 = 90^{\circ}$, $\chi'_3 = 90^{\circ}$ to 180°

(c) $\chi'_1 = 60^{\circ}$, $\chi'_2 = 120^{\circ}$, $\chi'_3 = -60^{\circ}$

(d) $\chi'_1 = 60^{\circ}$, $\chi'_2 = -60^{\circ}$, $\chi'_3 = 120^{\circ}$

The first two conformations represent arrangements in which

hydrogen-bonding can occur between the NH_3^+ group and the carbonyl of the peptide. The third conformation involves hydrogen-bonding between the NH of the peptide and the carboxylate oxygens, while the fourth has the two groups parallel so that both interactions may occur.

4.6 Cysteinyl-glycine Model

As well as ω and the two side-chain angles χ_1 and χ_2 there are two pairs of backbone angles ϕ, ψ to consider. Ideally a sub-map for ϕ_1, ψ_1 rotation should be produced for each set of values ϕ_2, ψ_2 and vice versa. Such a procedure would be very time-consuming and only a selected number of conformations have been studied. ϕ_1, ψ_1 rotations were performed for the preferred values of ϕ_2, ψ_2 determined in the study of the glycyl model; similarly for ϕ_2, ψ_2. The results were combined on "global" energy maps (fig.4.10). These maps cannot be regarded as complete descriptions of the conformational space, however they do illustrate the following points.

To a large extent the maps resemble those of the individual residues. The map for ϕ_2, ψ_2 variation with $\omega = 0^\circ$ is virtually identical to that obtained for the glycyl model. The map for $\omega = 180^\circ$ retains the main features of the corresponding glycyl model map but shows two new areas of low energy: $\phi_2 = 120^\circ$ to 150°, $\psi_2 = -60^\circ$ to 0° and, less noticeably, $\phi_2 = -90^\circ$, $\psi_2 = 0^\circ$ to 30°. These areas represent the formation of C_{10} rings (fig.4.12a,b). The corresponding values of ϕ_1, ψ_1 are: $\phi_1 = -30^\circ$ to 0°, $\psi_1 = 90^\circ$ to 120° and $\phi_1 = -60^\circ$, $\psi_1 = -30^\circ$ to 0° respectively. The global map for ϕ_1, ψ_1 variation does not indicate the areas represented by

these values to be of particularly low energy; however, an inspection of the component sub-maps (for example fig.4.11a) reveals that such low energy areas do exist for certain conformations of the cysteinyl residue.

Low energy areas corresponding to the formation of C_9 rings involving the S atoms (fig.4.12c) also appear on some sub-maps (e.g. fig.4.11b). The fact that such rings do not show up on the global energy maps possibly implies that they are of minor importance, however, it is not possible to assess accurately the relative importance of the various rings that may form without a thorough study of the entire conformational space.

The ion of the cysteinyl-glycine model was studied in a similar way to the neutral molecule. All the sub-maps for ϕ_2, ψ_2 variation are very similar to those obtained for the ion of the glycyl model, showing a marked preference for a planar arrangement. Any gain in stability from the formation of hydrogen-bonds between the carboxylate oxygens and -SH or -NH of cysteine is apparently insufficient to overcome the loss of stability that would result from a non-planar arrangement.

The map for ϕ_1, ψ_1 variation (fig.4.13) is rather different from that obtained for the cysteinyl residue or for the neutral cysteinyl-glycine molecule. The previously preferred C_7 conformation ($\phi_1 = -90^\circ$, $\psi_1 = 60^\circ$) is destabilised in favour of an arrangement with $\phi_1 = -150^\circ$, $\psi_1 = 30^\circ$. The C_5 conformation is still stable and represents the global minimum. There is a further low energy area around $\phi_1 = 120^\circ$ to 150°, $\psi_1 = -60^\circ$ to -30°. The reason for the higher energy of the C_7 arrangement could be a repulsion between the negatively charged carboxylate oxygens and the oxygen of the peptide

link. Similarly, the increased stability of the $\phi_1 = -150^\circ$, $\psi_1 = 30^\circ$ arrangement may be due to attraction between O^- and the -NH group of cysteine.

4.7 Glutamyl-cysteine Model

The object behind the study of this molecule was the investigation of the influence of the glutamyl chain on the preferred conformations of the cysteinyl residue. ϕ_1, ψ_1 rotation was performed for five different arrangements of the glutamyl chain. ϕ' was fixed at 60°, ω' at 0°, and χ'_1 at 60° for all five arrangements. The values of ψ', χ'_2 and χ'_3 were as follows:

(a) $\psi' = 0^\circ$ $\chi'_2 = 180^\circ$ $\chi'_3 = 180^\circ$

(b) $\psi' = 30^\circ$ $\chi'_2 = 0^\circ$ $\chi'_3 = -120^\circ$

(c) $\psi' = -30^\circ$ $\chi'_2 = 60^\circ$ $\chi'_3 = 120^\circ$

(d) $\psi' = 0^\circ$ $\chi'_2 = 60^\circ$ $\chi'_3 = -30^\circ$

(e) $\psi' = 180^\circ$ $\chi'_2 = 60^\circ$ $\chi'_3 = 90^\circ$

Conformation (a) represents the fully extended chain and the ϕ_1, ψ_1 map produced for this arrangement is similar to that of the cysteinyl model without the glutamyl chain. Conformation (b) represents the optimum conformation determined by the study on the glutamyl model. Here the C_7 ring is destabilized due to obstruction of the NH group by the free carboxylic acid group of glutamyl. A low energy area occurs around $\phi_1 = -90^\circ$, $\psi_1 = 90^\circ$ which may involve hydrogen-bonding between NH and the glutamyl carbonyl group. The area extends to enclose a second minimum, the C_5 arrangement, and includes some of the ϕ_1, ψ_1 values necessary for the formation of a C_{10} ring in

cysteinyl-glycine. This suggests that this arrangement of the glutamyl chain may favour C_{10} formation. Arrangement (c) is similar to (b) but in this case there is no obstruction of the NH group and the map produced is again very similar to that of the cysteinyl model. Conformations (d) and (e) have the carboxylic acid and peptide groups approximately parallel. For (d) the formation of the C_7 ring is hindered due to steric interaction. Similarly the C_5 ring is destabilized for conformation (e). A new minimum appears at $\varphi_1 = -60^o$, $\psi_1 = 150^o$ on the map for (d), probably involving a hydrogen-bond between the cysteinyl carbonyl and the OH group of glutamyl. The conformation at $\varphi_1 = 0^o$, $\psi_1 = 90^o$ is favoured on the map for (e), possibly due to hydrogen-bonding between cysteinyl NH and the carbonyl oxygen of glutamyl.

It is clear even from this brief study that the glutamyl chain can have a considerable influence on the conformation of the cysteinyl fragment both by altering the relative stability of the normally preferred conformations and by forming new stable structures.

4.8 Glutamyl-cysteinyl-glycine

This is the complete neutral glutathione molecule. It was studied by taking five arrangements of the glutamyl chain and for each selecting the preferred values of φ_1, ψ_1 and performing φ_2, ψ_2 variations. The five arrangements chosen are the four indicated by * in table 4.2 and the fully extended chain.

"Global" maps were constructed for φ_2, ψ_2 variation with $\omega = 0^o$ and 180^o and for φ_1, ψ_1 variation (fig.4.14). The map for

ϕ_2, ψ_2 variation with $\omega = 0^\circ$ shows the same principal features as that for the glycyl model, M and C_5 still being the preferred conformations. The map for $\omega = 180^\circ$ shows that C_7 is still of major importance while the relative energy of C_5 and M conformations has increased. A low energy area occurs around $\phi_2 = 120^\circ$, $\psi_2 = -90^\circ$ to -30° representing the formation of a C_{10} ring involving the glutamyl-cysteine peptide link; this is given added stability by interaction with the glutamyl amino and carboxylic acid groups. A further low energy area occurs at $\phi_2 = 90^\circ$ to 120°, $\psi_2 = 60^\circ$ to 90° also involving interactions with the glutamyl chain.

4.9 Ionized Glutathione

This was studied in a similar way to the neutral molecule. Five arrangements of the glutamyl chain were considered, these being the extended chain and the four stable conformations found in the study of the ionized glutamyl model (section 4.5). ϕ_2 and ψ_2 were fixed in the preferred planar arrangement as found in the study of the cysteinyl-glycine ion (section 4.6) and ϕ_1, ψ_1 rotations were performed. The results are presented in a global energy map (fig.4.15). Once again this map cannot be regarded as fully representative of all possible conformations, however, there does seem to be a marked preference for the conformation represented by $\phi_1 = -90^\circ$ to -60°, $\psi_1 = -30^\circ$ to 0°. This is illustrated in Fig.4.16 for the most stable arrangement of the glutamyl chain. This conformation does not appear to involve any "new" hydrogen bonds i.e. hydrogen-bonds other than those already present within the glutamyl

chain and the glycyl fragment. Its particular stability must thus be due to other factors.

The map also shows a local minimum for the C5 conformation ($\phi_1 = \psi_1 = 180^o$) and for a conformation with $\phi_1 = 60^o$, $\psi_1 = 150^o$ to 180^o. This latter conformation only occurs for a particular arrangement of the glutamyl chain and involves interaction between the NH_3^+ group and the carbonyl oxygen of the cysteinyl-glycine peptide link.

4.10 Discussion

Experimental data on the structure of glutathione is not extensive. In the crystal structure reported by Wright [172] the molecule is described as being in an S-shaped configuration with as angle between the two peptide links of 94.4^o. No intramolecular hydrogen-bonds are present.

The conformation of the central part of the molecule is in close agreement with the results for ionized glutathione. The arrangement of the glutamyl chain and the glycyl fragment is rather different. This is not surprising since in the crystal the terminal carboxylic acid and amino groups are involved in intermolecular hydrogen-bonds. The calculations are performed on an isolated molecule where no such intermolecular bonds can form; intramolecular bonds are thus of greater importance.

Of the two n.m.r. studies reported on the molecule Zenin et al. [174] conclude that the molecule must be in a "figure 8" conformation with interaction between the glutamyl NH_3^+ group and the glycyl COO^- as

well as between the glutamyl COO^- and the peptide NH groups. Attempts to find a stable conformation corresponding to such a structure were unsuccessful.

The second study, by Fujiwaro et al. [175], concludes that the preferred conformation of the fully dissociated molecule is one in which the carboxyl and amino groups of the glutamyl chain are far from the peptide backbone. This is in better agreement with the X-ray results but again different from the theoretical results which always indicate a preference for structures with the terminal groups of glutamyl interacting with the peptide link. The n.m.r. study, however, refers to a completely dissociated molecule i.e. an NH_2 group rather than NH_3^+ group is involved. The paper also concludes that changes in the functional groups of glutamic acid have an important influence on the preferred conformation of the molecule. This is supported by the theoretical results since the preferred conformation of neutral glutathione was found to be different from that of the ionized molecule.

It may be concluded that while the theoretical calculations do not reproduce exactly the available experimental results there is a substantial degree of agreement particularly with the X-ray results. The inclusion of solvent effects or some other means of allowing for the formation of intermolecular bonds would probably improve this further.

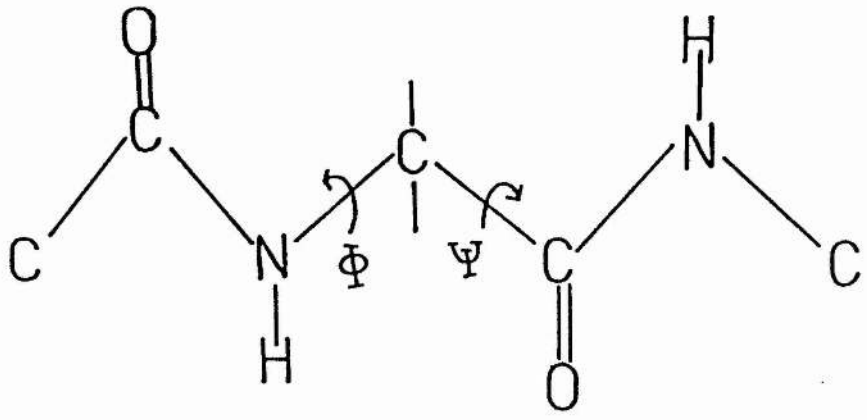


Fig. 4.1 An example of a typical fragment illustrating torsional angles

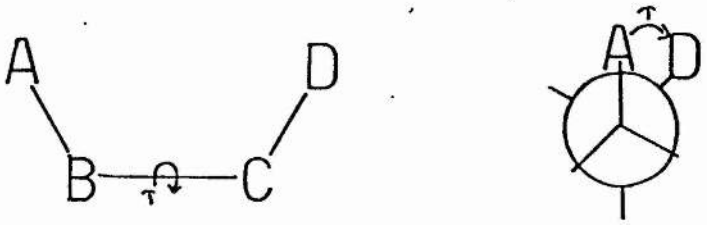


Fig. 4.2 Definition of torsional angle τ as seen
(a) perpendicular to the bond being rotated, and
(b) looking along the bond

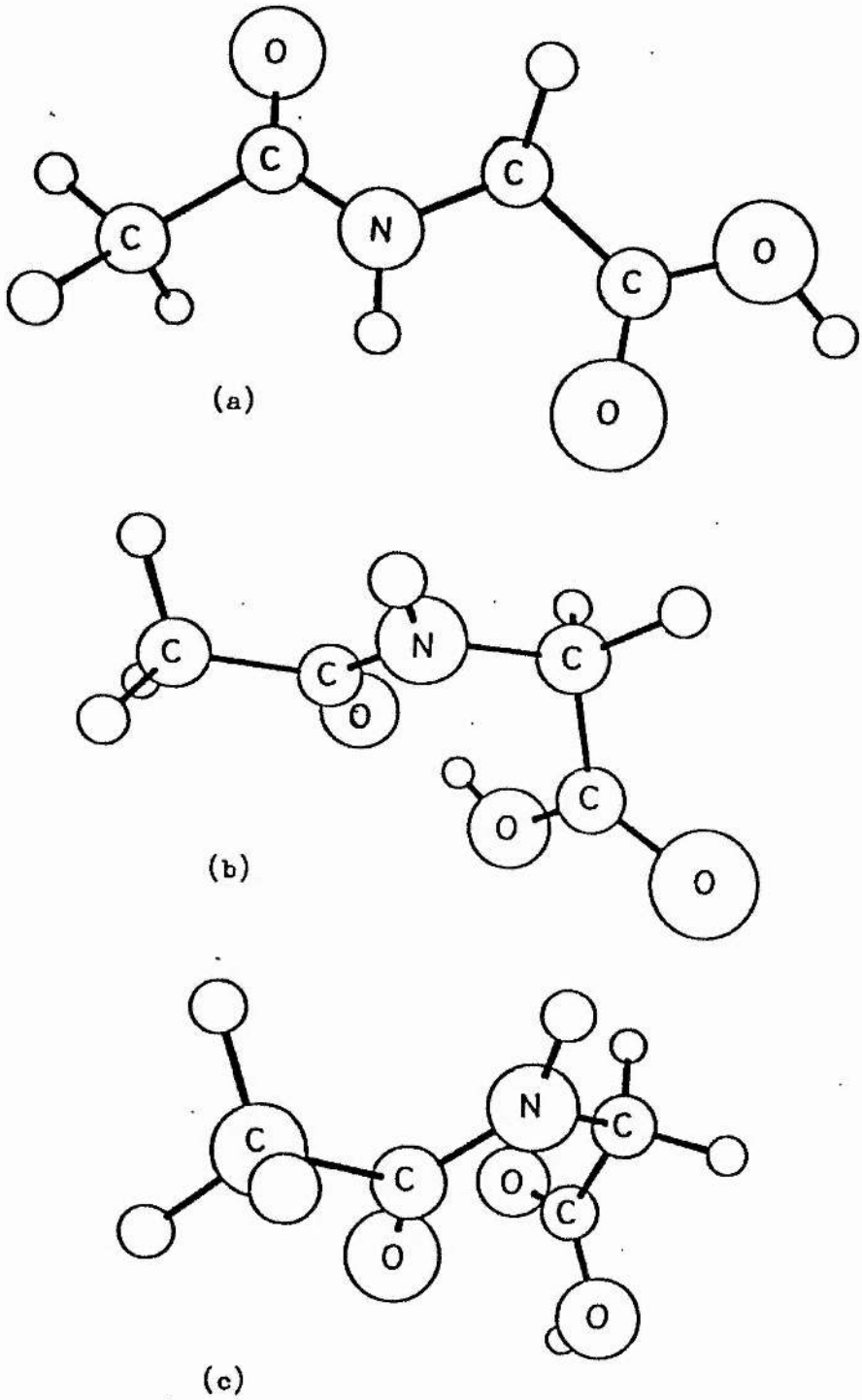


Fig. 4.3 Optimum conformations of the glycyl model:

(a) C_5 (b) C_7 (c) M

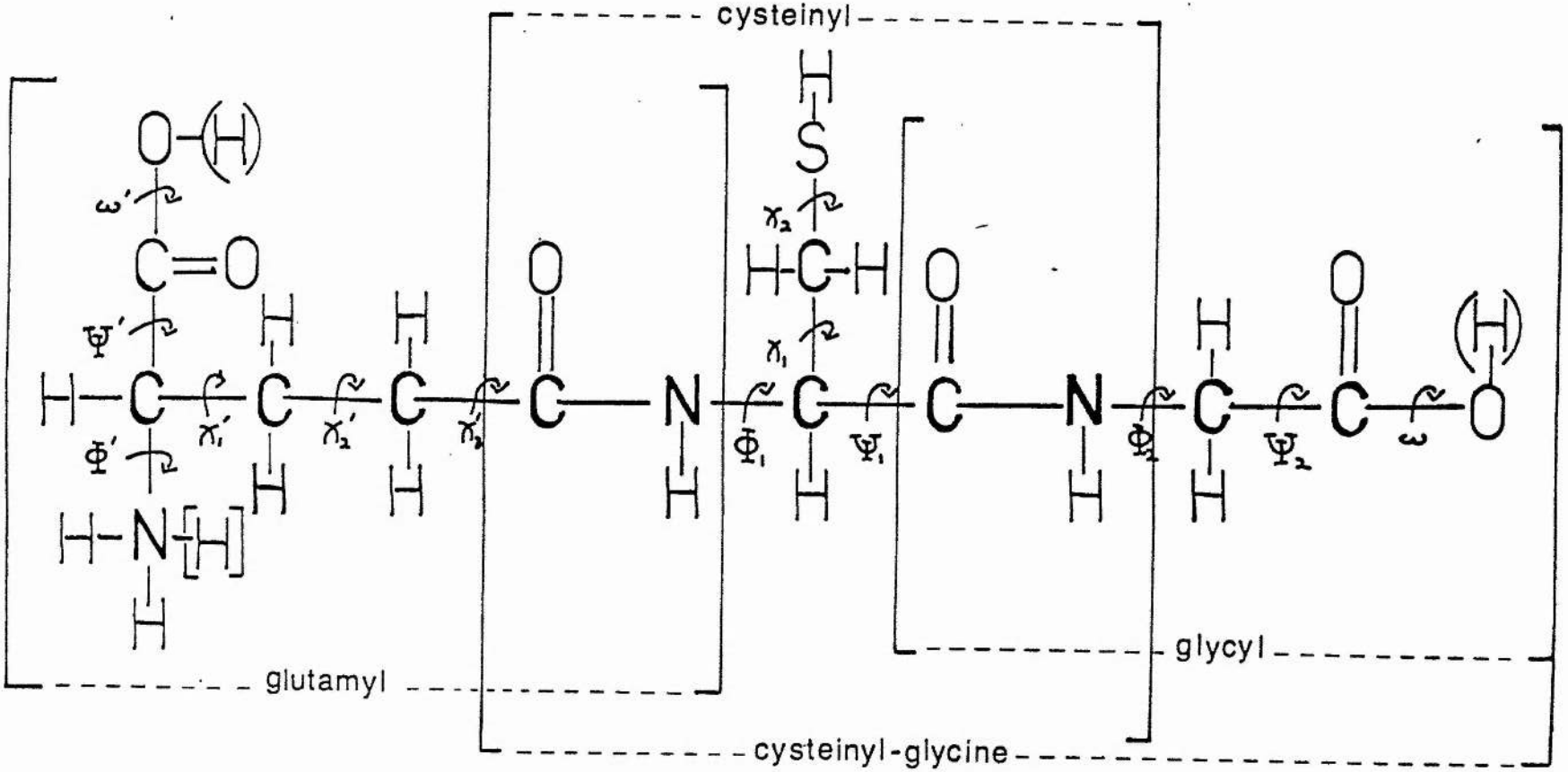


Fig. 4.4 Division of glutathione molecule into fragments. [H] present only in ionized form; (H) present only in neutral form.

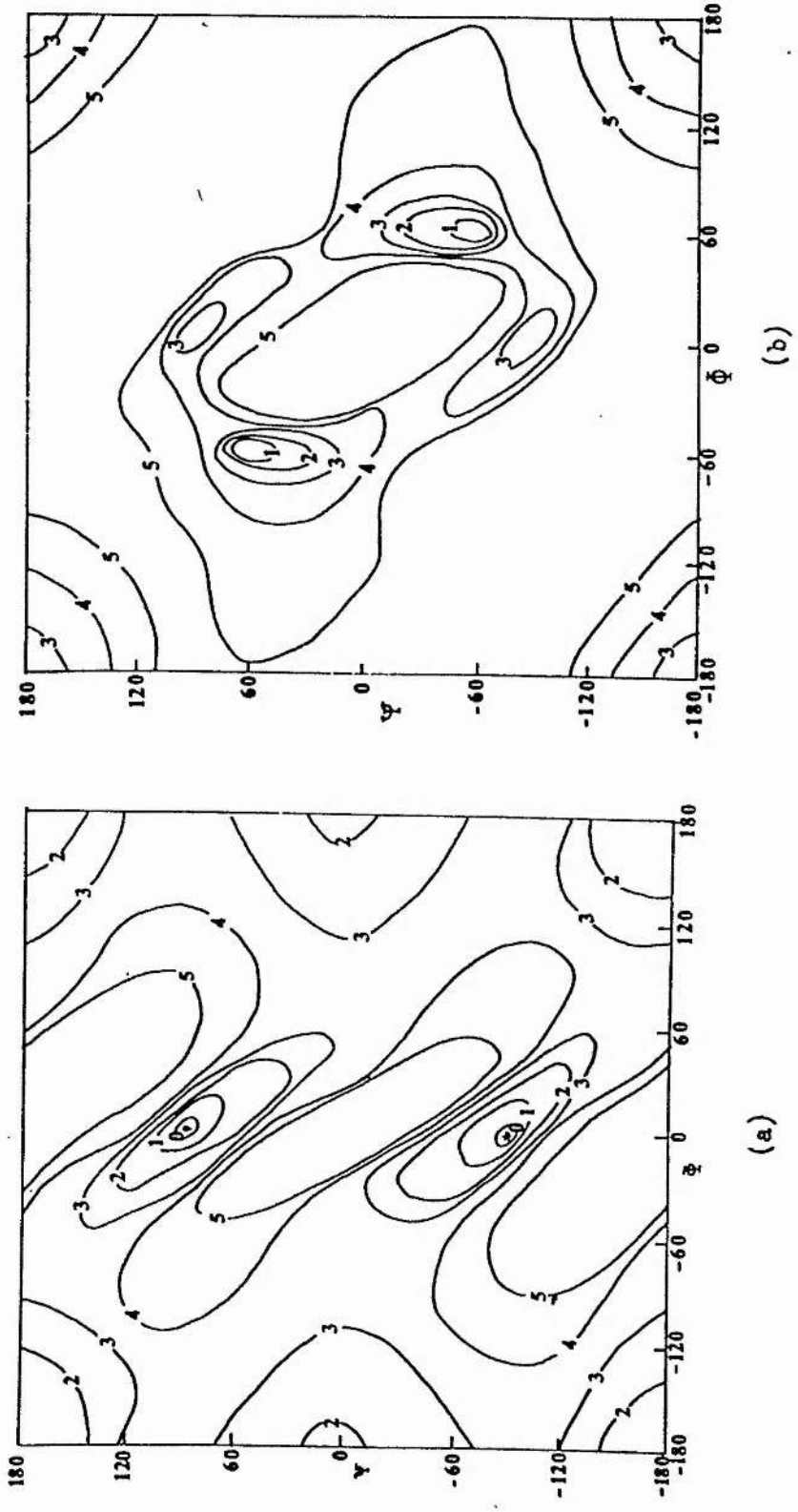


Fig. 4.5 Conformational energy maps for the glycyl model

(a) $\omega = 0^{\circ}$ (b) $\omega = 180^{\circ}$

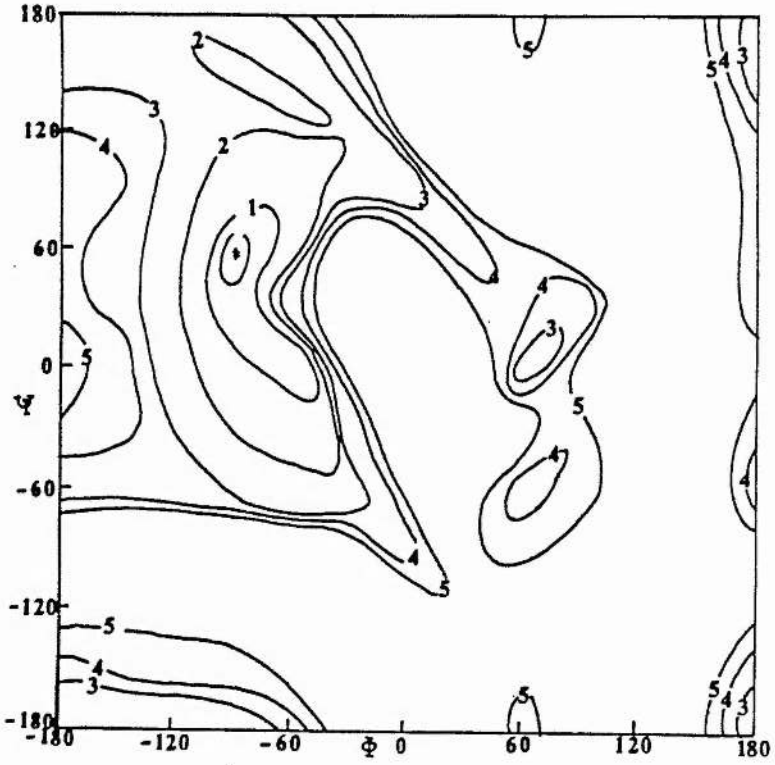


Fig. 4.6 Global conformational energy map for cysteinyl model

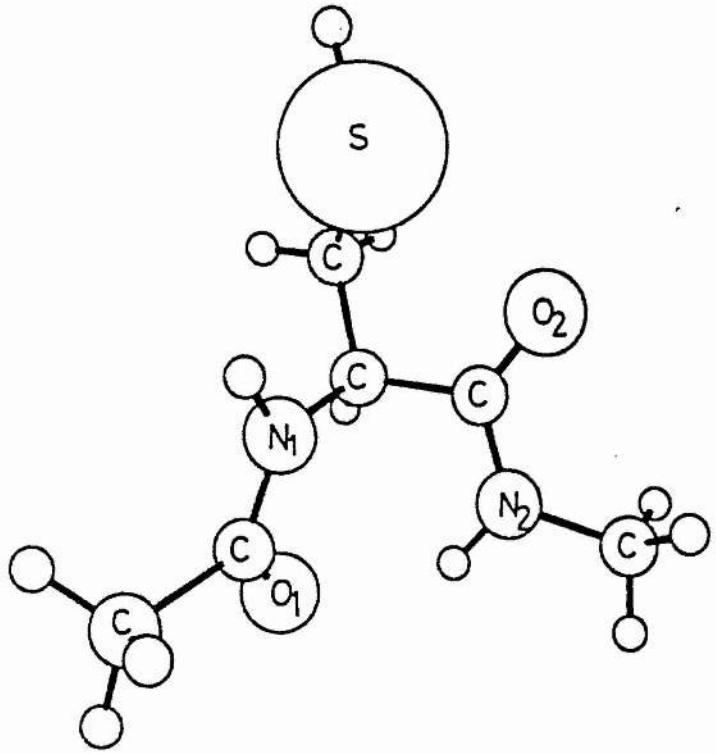
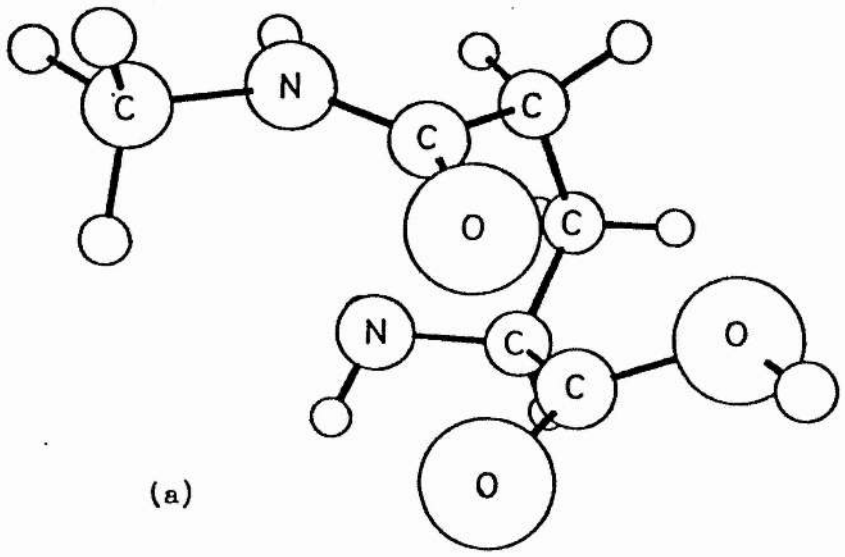
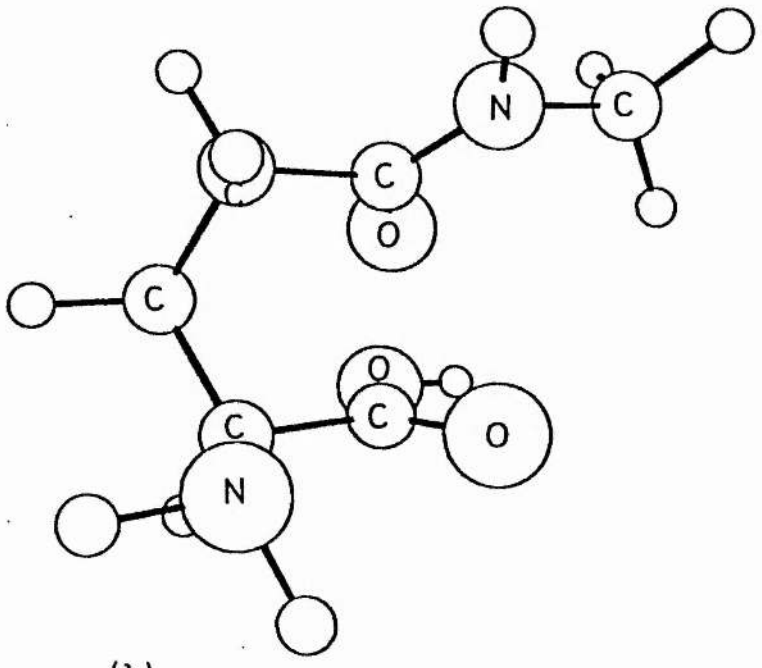


Fig. 4.7 Optimum conformation of cysteinyl residue



(a)



(b)

Fig. 4.8 Stable conformations of the glutamyl model

(a) optimum conformation

(b) another stable arrangement

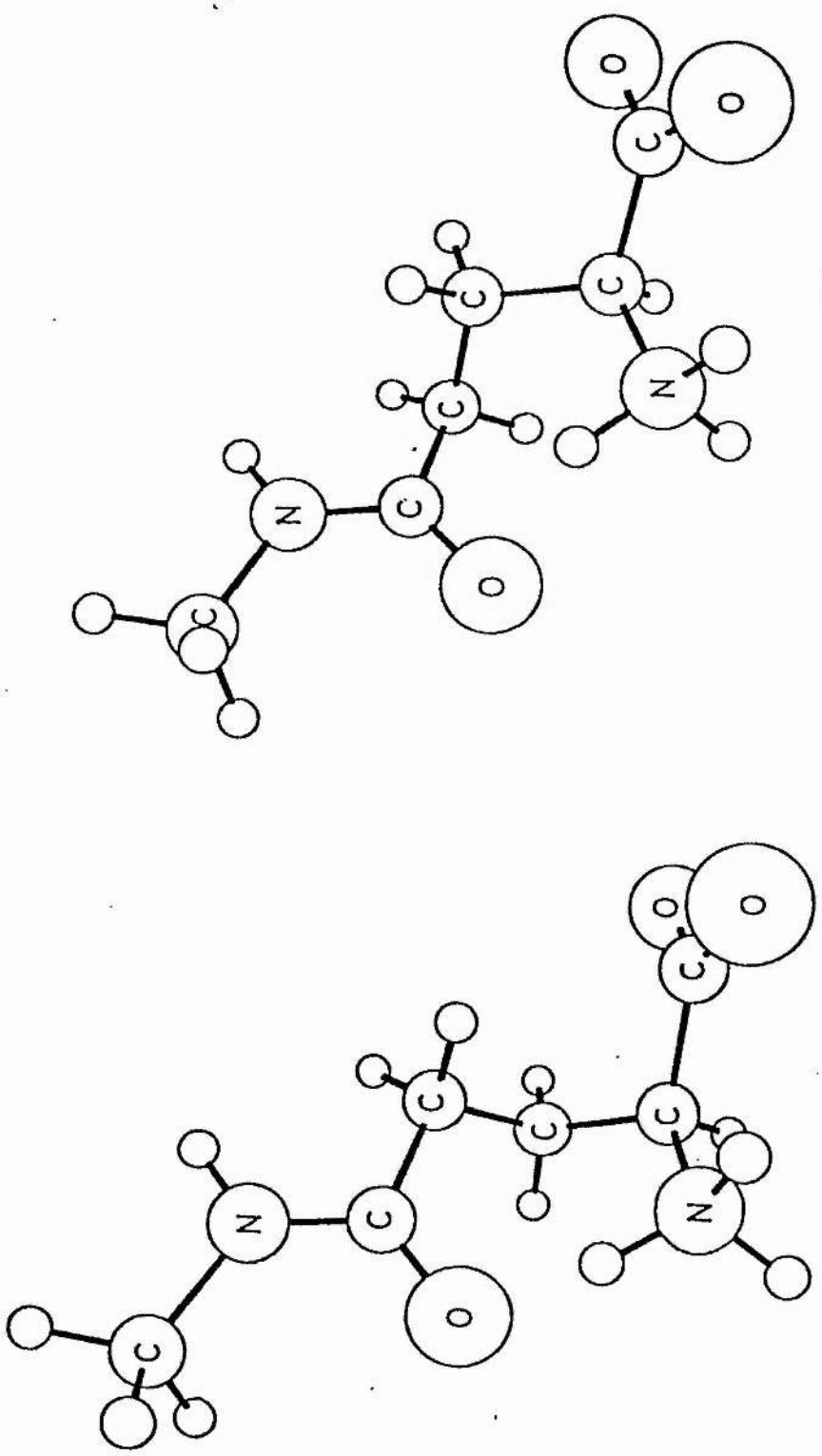


Fig. 4.9 Stable conformations of the glutamyl ion.

For descriptions see text.

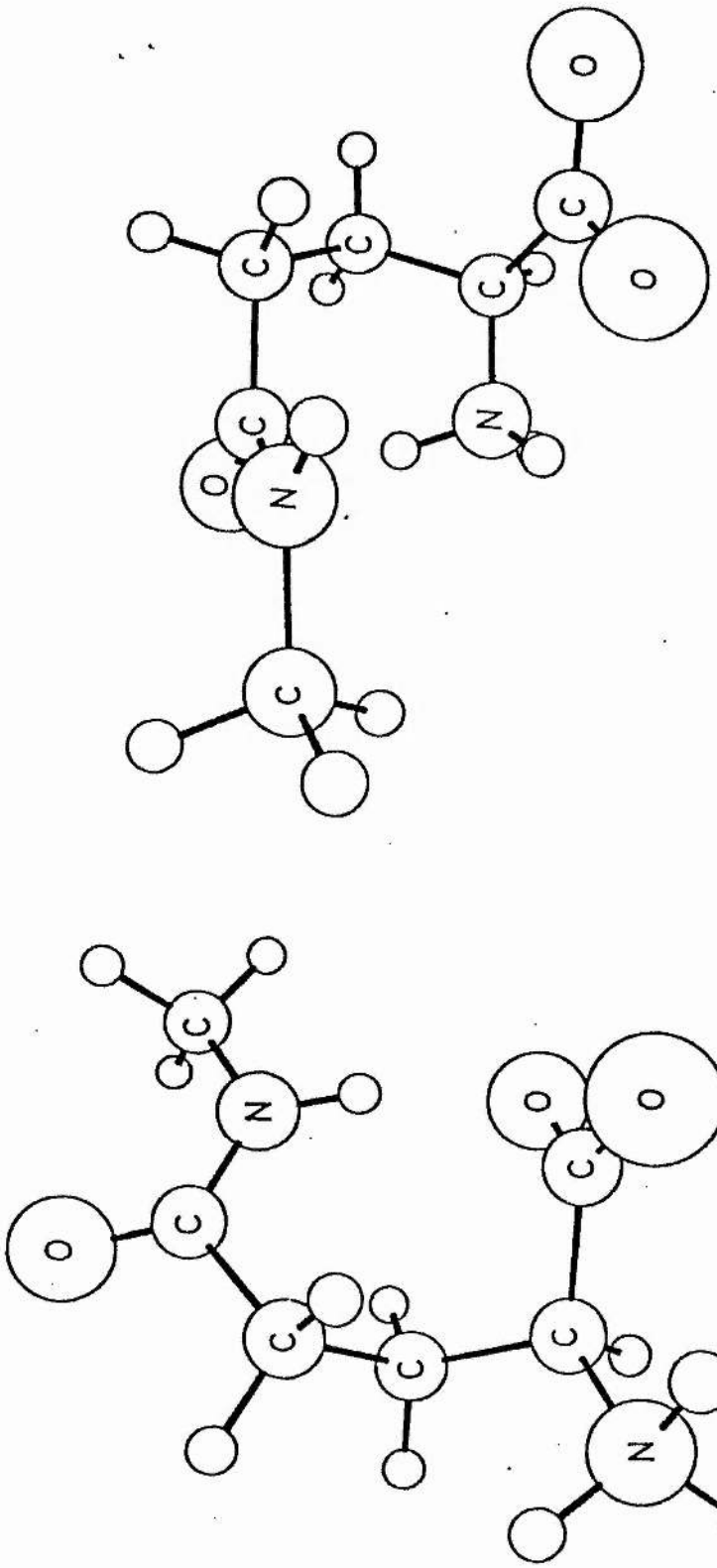


Fig. 4.9 continued

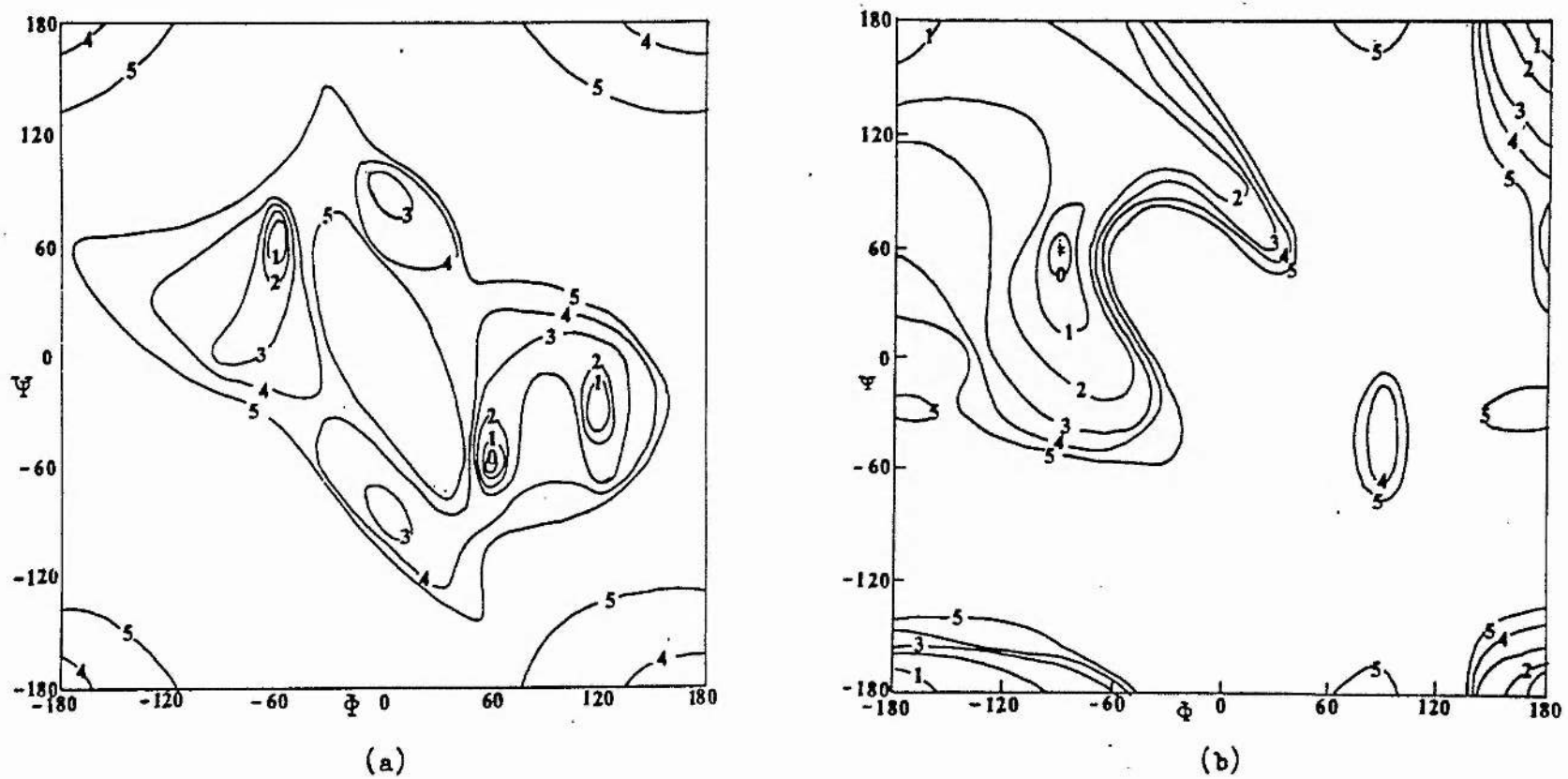


Fig. 4.10 Global energy maps for cysteinyl-glycyl model

(a) ϕ_2, ψ_2 rotation with $\omega = 180^\circ$

(b) ϕ_1, ψ_1 rotation

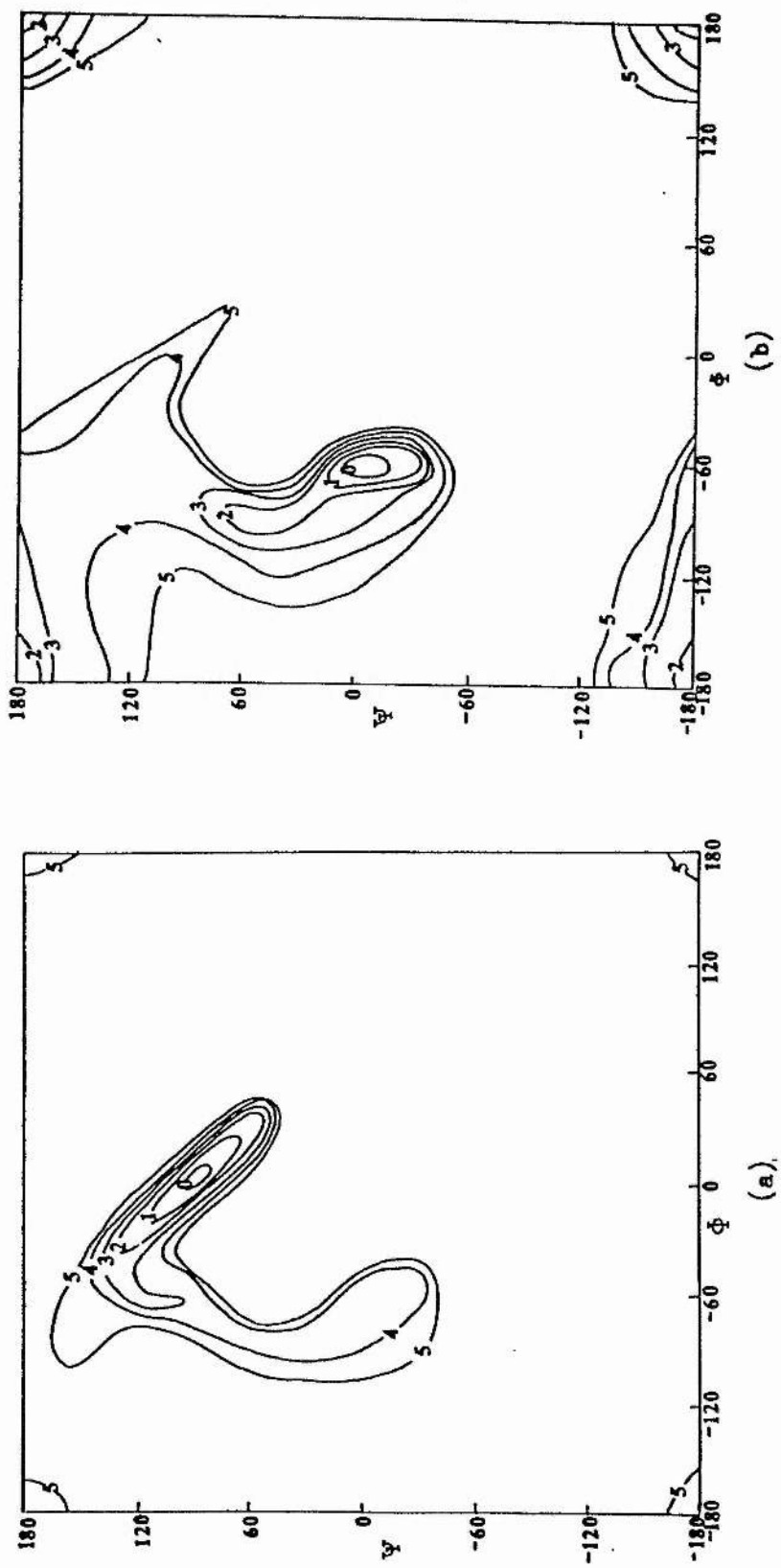


Fig. 4.11 Sub-maps for cysteinyl-glycyl model, Φ_1, Ψ_1 rotation

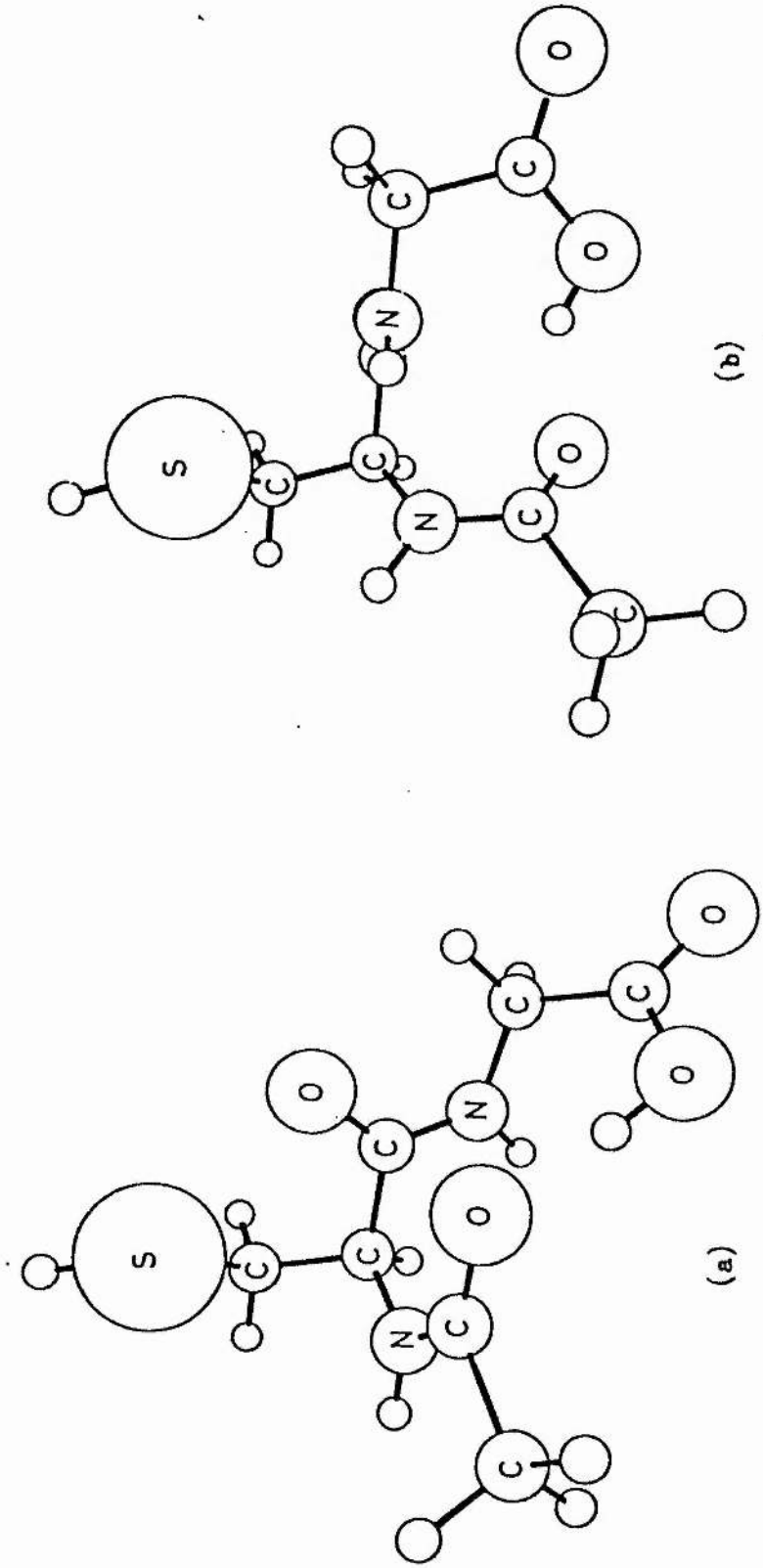


Fig.4.12 (a), (b) C_{10} rings

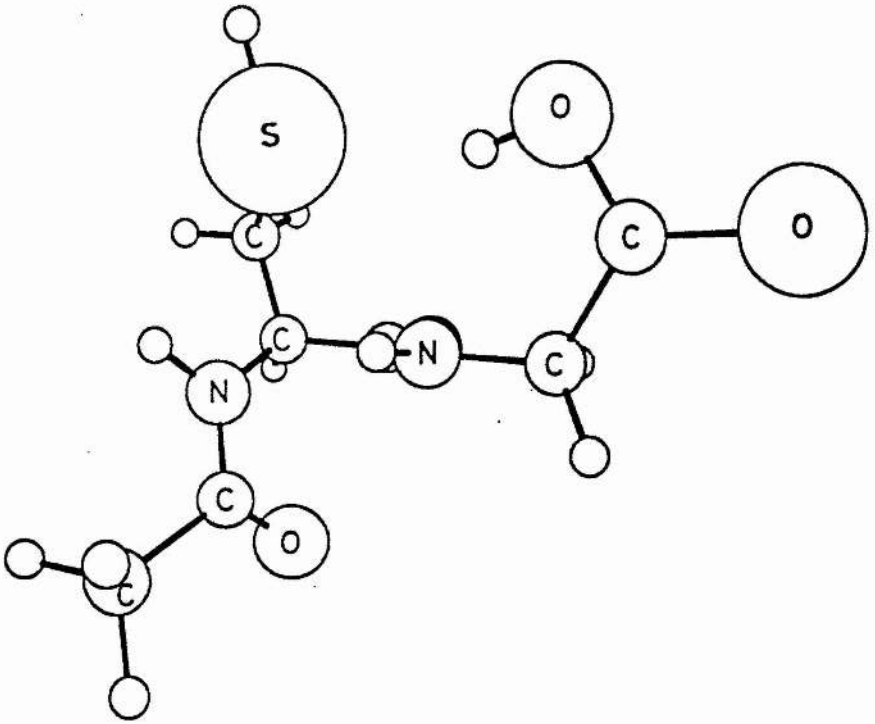


Fig. 4.12 (c) C_9 ring

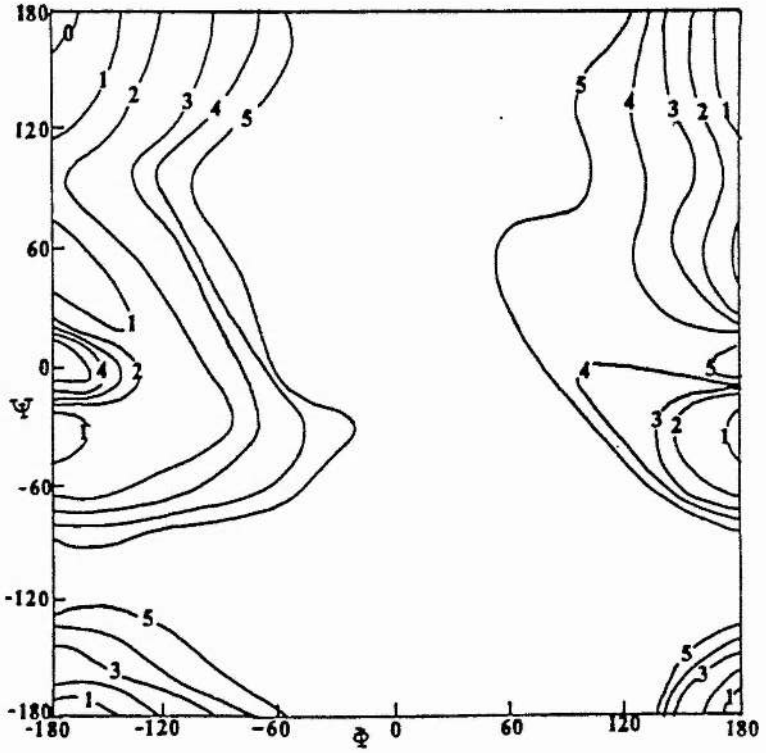


Fig. 4.13 Φ_1, Ψ_1 rotation for cysteinyl-glycyl ion

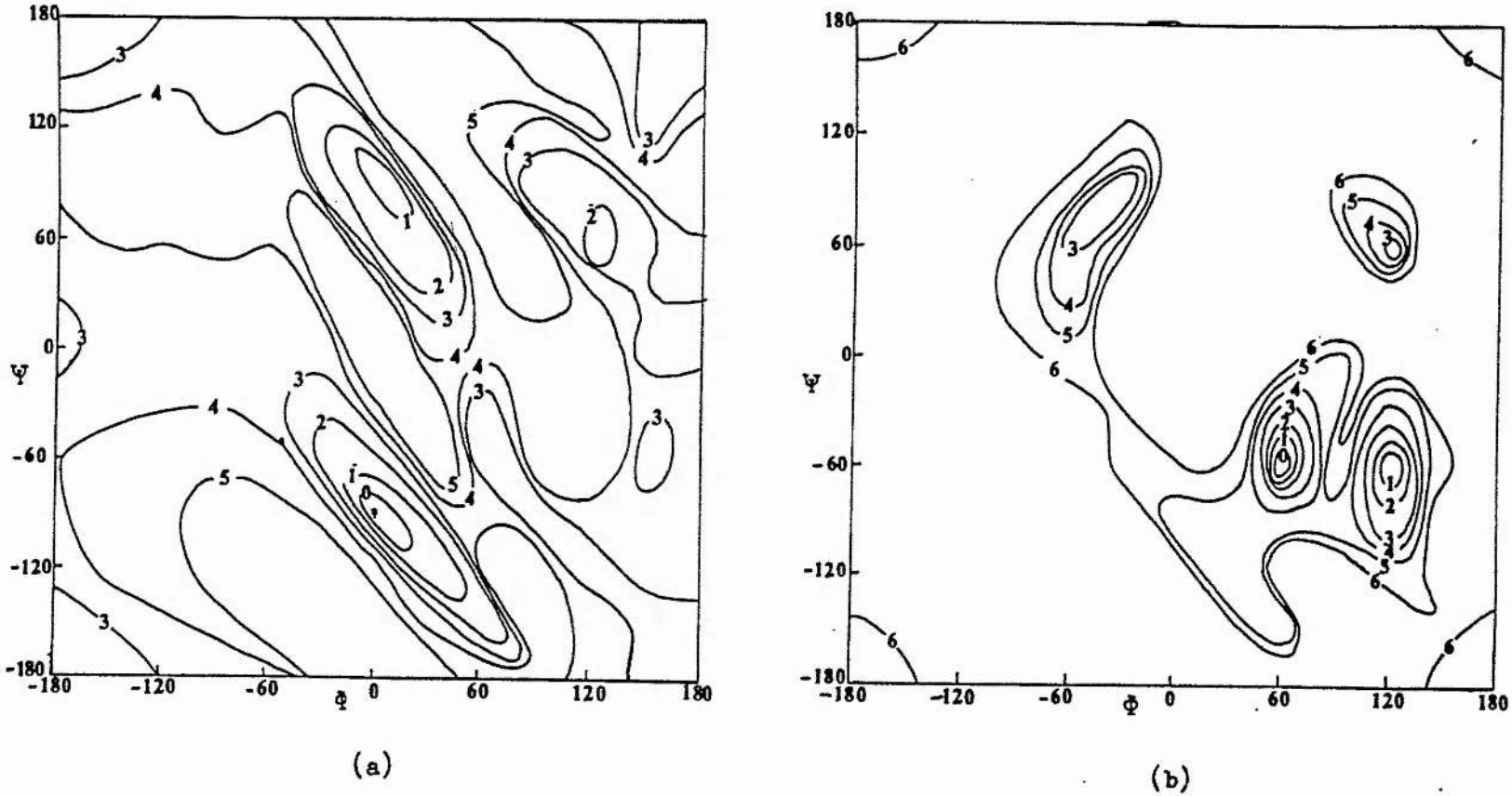


Fig. 4.14 Global maps for neutral glutathione

(a) ϕ_2, ψ_2 rotation with $\omega = 0^\circ$

(b) ϕ_2, ψ_2 rotation with $\omega = 180^\circ$

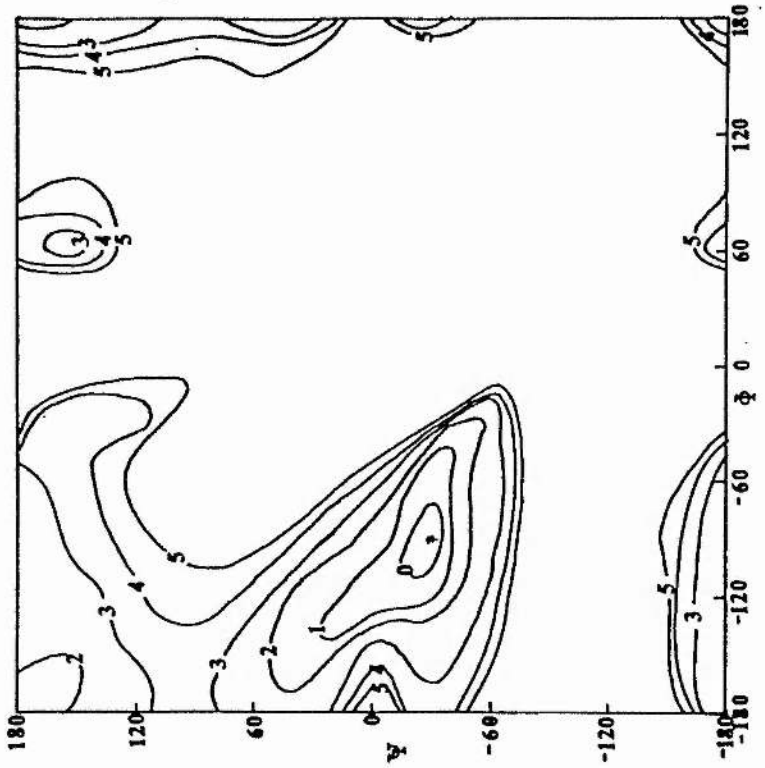


Fig. 4.15 Global map for ionized glutathione

Fig. 4.14 (c) map for Φ_1, Ψ_1 rotation

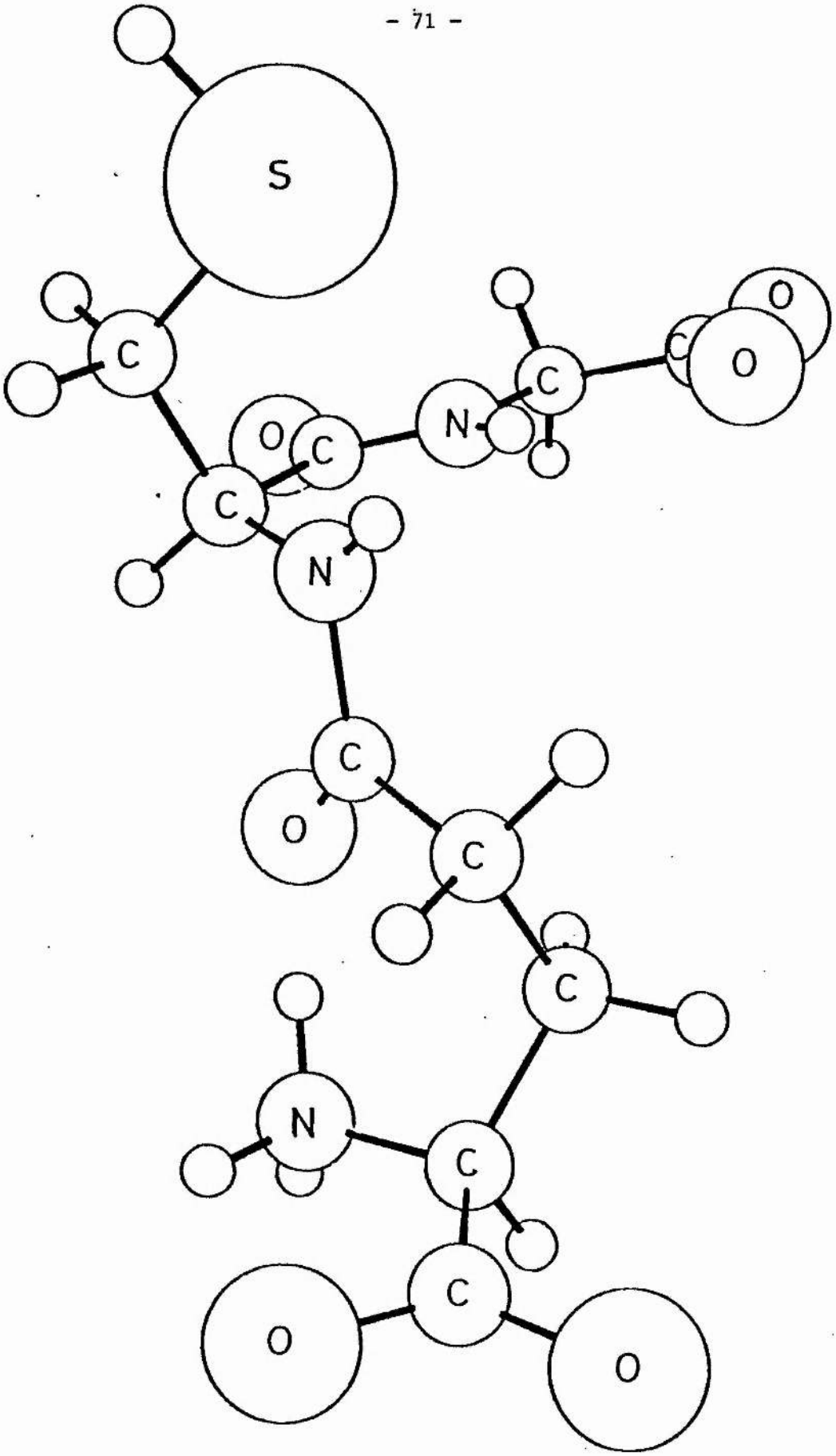


Fig. 4.16 Optimum conformation of ionized glutathione

Table 4.1

Stable conformations of the glycyl model, equivalent conformations indicated by }

VALUES OF TORSIONAL ANGLES (^{b})			DESCRIPTION OF CONFORMATION
ϕ	ψ	ω	
180	180	0	C_5
180	0	0	C_5
0	90	0 }	M
0	-90	0 }	M
180	180	180	C_5
0	-90	180 }	M
0	90	180 }	M
60	-60	180 }	C_7
-60	60	180 }	C_7

VALUES OF TORSIONAL ANGLES (^{o}) ($\chi=60^o$, $\chi'=0^o$ for all conformations)			
ψ'	χ_1'	χ_2'	χ_3'
30^o	60^o	0^o	-120^o *
-150^o	60^o	60^o	120 *
-150^o	60^o	0^o	-120^o
-150^o	60^o	30^o	120^o
-150^o	60^o	90^o	120^o
-30^o	60^o	60^o	180^o
30^o	60^o	30^o	120^o
0^o	60^o	60^o	-150^o *
180^o	60^o	60^o	90^o *
0^o	60^o	0^o	-120^o
0^o	60^o	60^o	-120^o

Table 4.2
Stable conformations of the glutamyl model

5 COMPARISON OF THE RESULTS OF PCILO AND AB INITIO SCF CALCULATIONS

5.1 Introduction

In addition to the study of the glutathione molecule described in the preceding chapter its constituent amino acids, glycine, cysteine and glutamic acid were investigated. Conformational energy maps were produced using the PCILO program and the preferred arrangements of the side-chain in cysteine and glutamic acid were investigated.

A limited number of ab initio calculations (using the Gaussian 70 and Gaussian 80 program packages) were also performed for these three amino acids and for the "glycyl model" (N-acetyl-glycine). For glycine, cysteine and N-acetyl-glycine conformational energy surfaces were produced using the STO-3G basis set. For N-acetyl-glycine the most interesting points on the energy surfaces were also investigated with 3-21G and 4-31G basis sets. For glutamic acid it was only possible to perform ab initio (STO-3G) calculations for a few of the preferred conformations determined by the PCILO calculations. The results of the two methods are presented and compared in the sections which follow. Standard geometries [189,190] have been used throughout.

5.2 Glycine

This molecule is important as it is the smallest amino acid and the prototype of all the others. The torsional angles are defined in

fig.5.1a. PCILO calculations give the preferred value of as 0° with a second minimum, 2.5 kcal higher, at 180°; the barrier to rotation is 3.9kcal. Conformational energy maps for ϕ , ψ rotation were produced for these two values of ω . The maps calculated with PCILO are presented in fig5.2 and those for the ab initio STO-3G calculations in fig.5.3.

The STO-3G maps indicate the global minimum to occur for $\phi = 60^{\circ}$, $\psi = 180^{\circ}$, $\omega = 0^{\circ}$. PCILO, on the other hand, gives two equivalent global minima at $\phi = -60^{\circ}$ and 180°, $\psi = 180^{\circ}$, $\omega = 0^{\circ}$; $\phi = 60^{\circ}$ represents a local minimum. Local minima also occur on the PCILO $\omega = 0^{\circ}$ map for $\psi = 0^{\circ}$, $\phi = 180^{\circ}, \pm 60^{\circ}$. These conformations represent local maxima on the corresponding STO-3G map. For $\omega = 180^{\circ}$ the STO-3G map indicates a minimum at $\phi = -120^{\circ}$, $\psi = 0^{\circ}$; the PCILO map shows two equivalent minima to either side of this point. PCILO also indicates local minima at $\psi = 180^{\circ}$, $\phi = 180^{\circ}$, $\pm 60^{\circ}$. STO-3G gives one other minimum at $\psi = 180^{\circ}$, $\phi = 60^{\circ}$.

The two methods thus agree in favouring planar arrangements of the molecule and in indicating certain conformations to be of low energy. However, they disagree as to the relative ordering of minima, and PCILO indicates a number of minima which do not appear at all on the STO-3G maps.

Glycine has been studied by several authors using both semi-empirical and ab initio procedures [177-185], and it appears that the "preferred" conformation of the molecule varies considerably depending on the method used. Caballol et al. [180] have compared the results of EHT calculations with those obtained by Scheraga's method, while CNDO, PCILO and STO-3G results have been compared by Palla et al. [45]. (The latter study agrees closely with the results

presented here.) Ab initio calculations using 4-21G [182,183], 4-31G [184] and 6-31G [185] basis sets have also been carried out for glycine. The extended basis sets predict the same energy minima as the STO-3G calculations. Calculations using a 4-21G basis set and performing complete geometry optimization for the two lowest energy conformers [182] confirm that the global minimum occurs at $\phi = 60^{o}$, $\psi = 180^{o}$, $\omega = 0^{o}$ with a second minimum, ~2kcal higher, at $\phi = -120^{o}$, $\psi = 0^{o}$, $\omega = 180^{o}$. Initially this appeared to contradict the results of microwave spectroscopy which found evidence only for the latter conformation [191,192]. The theoretical results, however, stimulated further experimental work [193] and the first conformation was also detected and found to be the lower in energy of the two. The theoretical results are thus in good agreement with experiment and for this molecule minimal basis set calculations, even without geometry optimization, appear to give an accurate indication of the preferred conformations.

Other properties such as ionization potential and proton affinity have been calculated with a 6-31G basis set and are discussed in reference [185], which includes a comparison with the zwitterion and four other amino acids. The solvation of glycine is discussed in reference [181].

5.3 Cysteine

The torsional angles for cysteine are defined in fig.5.1b. In the PCILO calculations two values of ω, 0^{o} and 180^{o}, were considered and the side-chain angles were restricted to staggered values,

i.e. $\pm 60^{\circ}, 180^{\circ}$. Altogether 18 submaps, corresponding to the various combinations of these values, were produced for ϕ, ψ rotation. These were combined into two "global" maps for $\omega = 0^{\circ}$ and 180° respectively. The maps resemble those obtained for glycine in overall shape, although the symmetry is lost due to the presence of the side-chain.

For the STO-3G calculations the side-chain angles were fixed at the optimum values of $\chi_1 = 60^{\circ}$, $\chi_2 = 180^{\circ}$ with ω at 0°. The map for ϕ, ψ rotation is presented in fig5.4 together with the corresponding PCILO submap. A comparison of these two maps shows that there is only minimum, that at $\phi = -60^{\circ}$, $\psi = 180^{\circ}$, which does not appear on both. This conformation represents the global minimum on the PCILO map whereas the STO-3G map has the global minimum at $\phi = 60^{\circ}$, $\psi = 180^{\circ}$. Both maps indicate local minima for $\psi = 0^{\circ}$ and for this value of ψ both show a threefold minimum for ϕ rotation. Although the PCILO map is considerably flatter than that for STO-3G it does appear that there is more agreement between the two methods in this case than was found for glycine.

No other conformational studies on cysteine have been found in the literature. However, ab initio calculations on the solvation of the molecule have been reported [181] and the results of a 6-31G basis set study, including a comparison with the zwitterion, are given in reference [185].

5.4 Glutamic acid

This molecule has seven degrees of freedom; torsional angles are

indicated in fig5.5a. Conformational energy maps for Φ, Ψ rotation were produced for staggered values of χ_1 (with the rest of the chain extended) and combined to give global energy maps. These are very similar in overall appearance to those obtained for glycine and cysteine.

In studying the preferred conformations of the side-chain an approach similar to that used for the glutamyl residue (section 4.5) was adopted. Thus ω, ω' and ϕ were fixed at their preferred values of 0^o, 0^o and 60^o respectively. Simultaneous rotation of χ_2 and χ_3 was performed for a series of values of χ_1 and Ψ. χ_2 and χ_3 were then fixed at their optimum values while χ_1 and Ψ rotations were performed. In this way it was hoped that most of the important conformations would be considered.

The PCILO calculations indicate a preference for conformations in which the two carboxylic acid groups are close together. Two of the most stable arrangements are illustrated in fig.5.6. Single ab initio calculations with an STO-3G basis set were performed for these conformations and for the extended chain arrangement. In contrast to the PCILO results the STO-3G basis set gives the lowest energy for the extended chain form. The folded conformations, with standard values for bond lengths and angles, are calculated to be highly unstable. The forces on the atoms were calculated for the extended chain form and one of the folded conformations; those for the folded conformation were found to be considerably higher than those for the extended chain form. A second force calculation was performed for the folded conformation with the three -C-C-C- bond angles altered to values between 114^o and 116^o, in line with the calculated forces on the atoms (the standard value for a -C-C-C- angle is 109.47^o). This calculation

showed a considerable decrease both in the forces and in the total energy. With these bond angles the folded conformation is only 7 kcal less stable than the extended chain form. Full geometry optimization would no doubt result in a further decrease in the total energy for all conformations and possibly a reversal in the relative stabilities. However, extended basis sets and/or configuration interaction might be needed for an accurate determination of preferred conformation. STO-3G basis sets have been found to favour planar arrangements when more accurate calculations and experimental evidence indicate non-planar forms to be preferred [194]. It could be that this tendency to favour planar arrangements results in the extended chain form (in which the atoms of the side-chain lie in a plane) being preferred over the folded forms. On the other hand the PCILO method has been criticized for its treatment of lone pairs [195] and it cannot be ruled out that these "stable" folded arrangements are an artifact of the PCILO method. At present it is not possible to perform extensive ab initio calculations on glutamic acid and the PCILO results must therefore be regarded with some reservation.

There do not appear to be any other studies of the conformational preferences of the glutamic acid side-chain. In the calculations on amino acid residues [176] the side-chain angles are fixed at staggered values and in the solvation study by Clementi [181] an experimental geometry is adopted. This latter is obtained from a neutron diffraction study by Sequira et al. [196] performed on L-glutamic acid hydrochloride. The molecule is found to be in an extended rather than folded conformation although the side-chain is twisted so that the atoms do not all lie in one plane and the α-carboxylic acid group is tilted out of the N-C-C plane. The most stable conformation

predicted by PCILO (fig.5.6a) also has the carboxylic acid group tilted. The experimental work has been done on a crystal where intermolecular forces are important in determining the structure. Extended conformations are more favourable for the formation of intermolecular hydrogen bonds, however for the isolated molecule intramolecular bonds are more important and folded arrangements may be preferred. At present there does not appear to be any suitable experimental data available for comparison with the theoretical results.

5.5 N-acetyl-glycine

The torsional angles for N-acetyl-glycine (the "glycyl model") are indicated in fig.5.5b. The PCILO results for this molecule were discussed in the preceding chapter (section 4.3). The corresponding STO-3G maps are presented in fig.5.7.

Both methods indicate the C_5 ($\phi = \psi = 180^\circ$, $\omega = 0^\circ$ or 180°) and C_7 ($\omega = 180^\circ$, $\phi = \pm 60^\circ$, $\psi = \mp 60^\circ$) arrangements to be stable. The main contrast is the emphasis PCILO gives to "M" conformations. For $\omega = 180^\circ$ they represent stable conformations equivalent in energy to the C_5 arrangement. For $\omega = 0^\circ$ they are ~1 kcal more stable than C_5. On the STO-3G map the "M" conformations occur at saddle points 5.7 and 11 kcal above the global minimum for $\omega = 0^\circ$ and 180° respectively. Apart from the "M" conformations both methods agree in predicting the order of stability to be C_5a ($\omega = 0^\circ$), C_7, C_5b ($\omega = 180^\circ$). These three points and the two "M" conformations, $\phi = -90^\circ$, $\psi = 0^\circ$, $\omega = 0^\circ$ (Ma) and $\phi = 90^\circ$, $\psi = 0^\circ$, $\omega = 180^\circ$ (Mb)

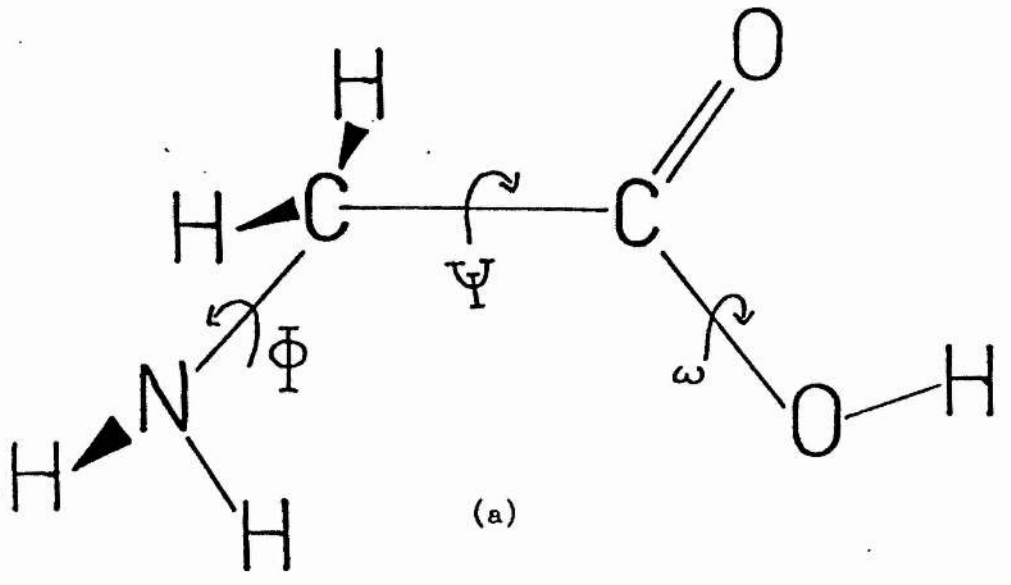
were investigated using 3-21G and 4-31G basis sets. Single calculations were performed at the proposed minimum energy conformations and at points close to them (in order to test whether the conformations represent local minima, maxima or saddle points). The results with these basis sets follow closely those obtained with the STO-3G basis set but with larger differences in energy between the conformations. For comparison the results are presented in Table 5.1 together with those for PCILO.

5.6 Discussion

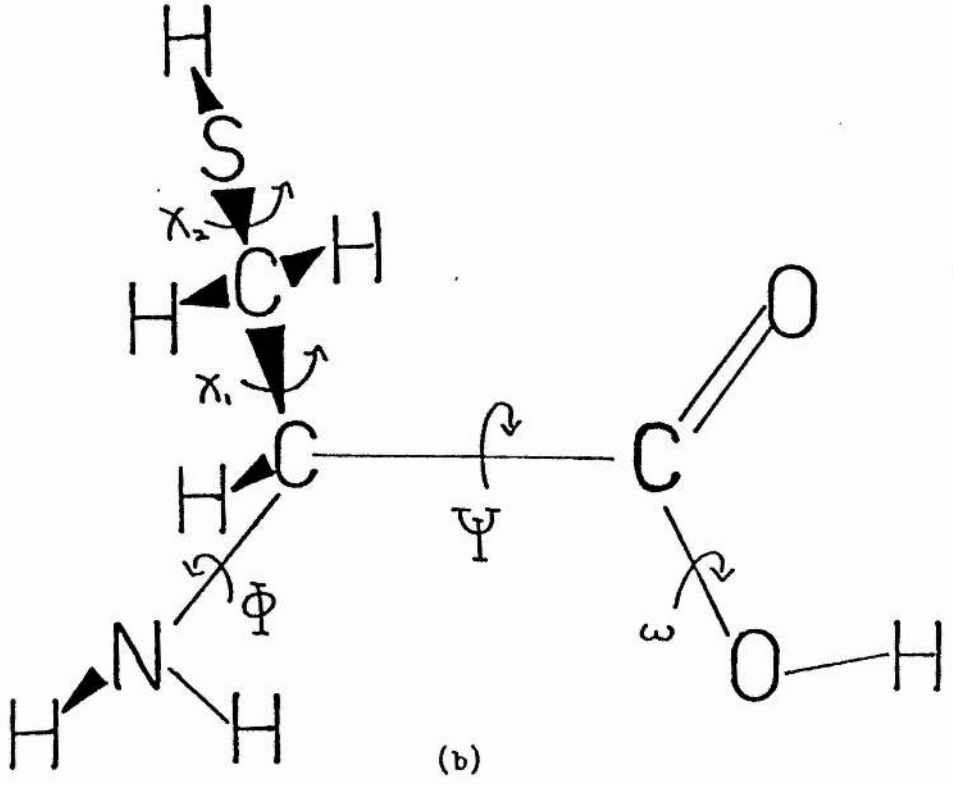
From the studies on glycine, cysteine and N-acetyl-glycine it appears that, in comparison to STO-3G *ab initio* calculations, PCILO tends to give low values for barrier heights, resulting in conformational energy maps that are rather flat. In some cases the two methods disagree as to the relative ordering of energy minima, it should be noted, however, that there is no minimum on any of the *ab initio* maps which does not have a corresponding minimum on the appropriate PCILO map. Thus while PCILO does not reproduce exactly the results of *ab initio* calculations it does provide a useful preliminary guide to the preferred conformations of a molecule. A two-dimensional conformational energy map can be produced very rapidly using PCILO and the optimum conformations can then be used as a starting point for more accurate calculations.

The results obtained for glutamic acid and for N-acetyl-glycine suggest that in some cases PCILO may indicate favoured conformations that are calculated to be of relatively high energy by

ab initio methods. This may be due to a shortcoming in the PCILO method which results in a underestimation of lone pair repulsion. In conclusion it seems that PCILO gives a reasonable overall picture of the conformational flexibility of a molecule but its predictions of absolute minima may not always be reliable and, if possible, should be checked against more accurate calculations. The ab initio results obtained for glutamic acid also suggest that using standard values for bond lengths and angles may not always be adequate.



(a)



(b)

Fig. 5.1 Definition of torsional angles for

(a) glycine (b) cysteine

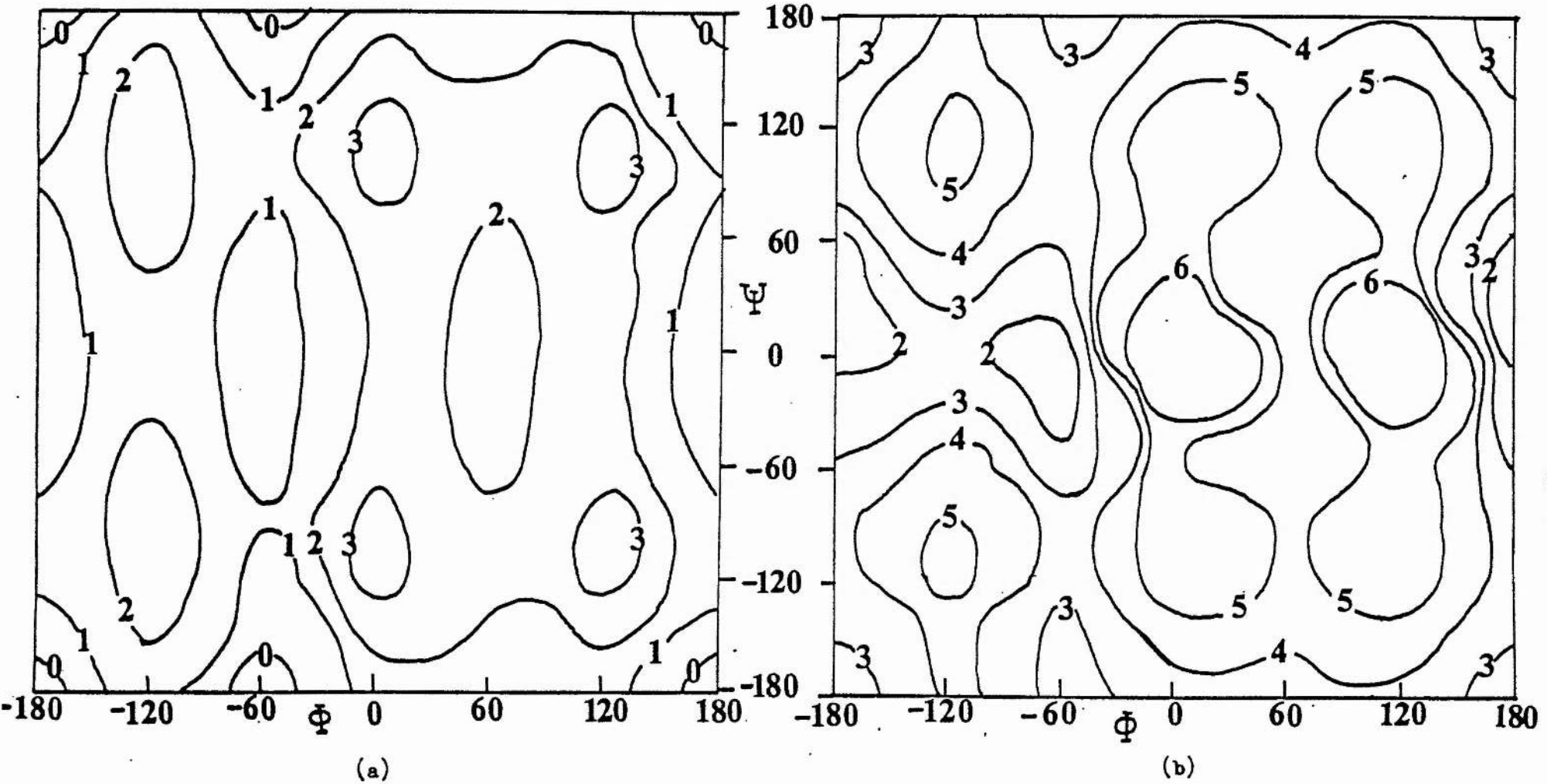


Fig. 5.2 PCILO conformational energy maps for glycine (a) $\omega = 0^{\circ}$ (b) $\omega = 180^{\circ}$

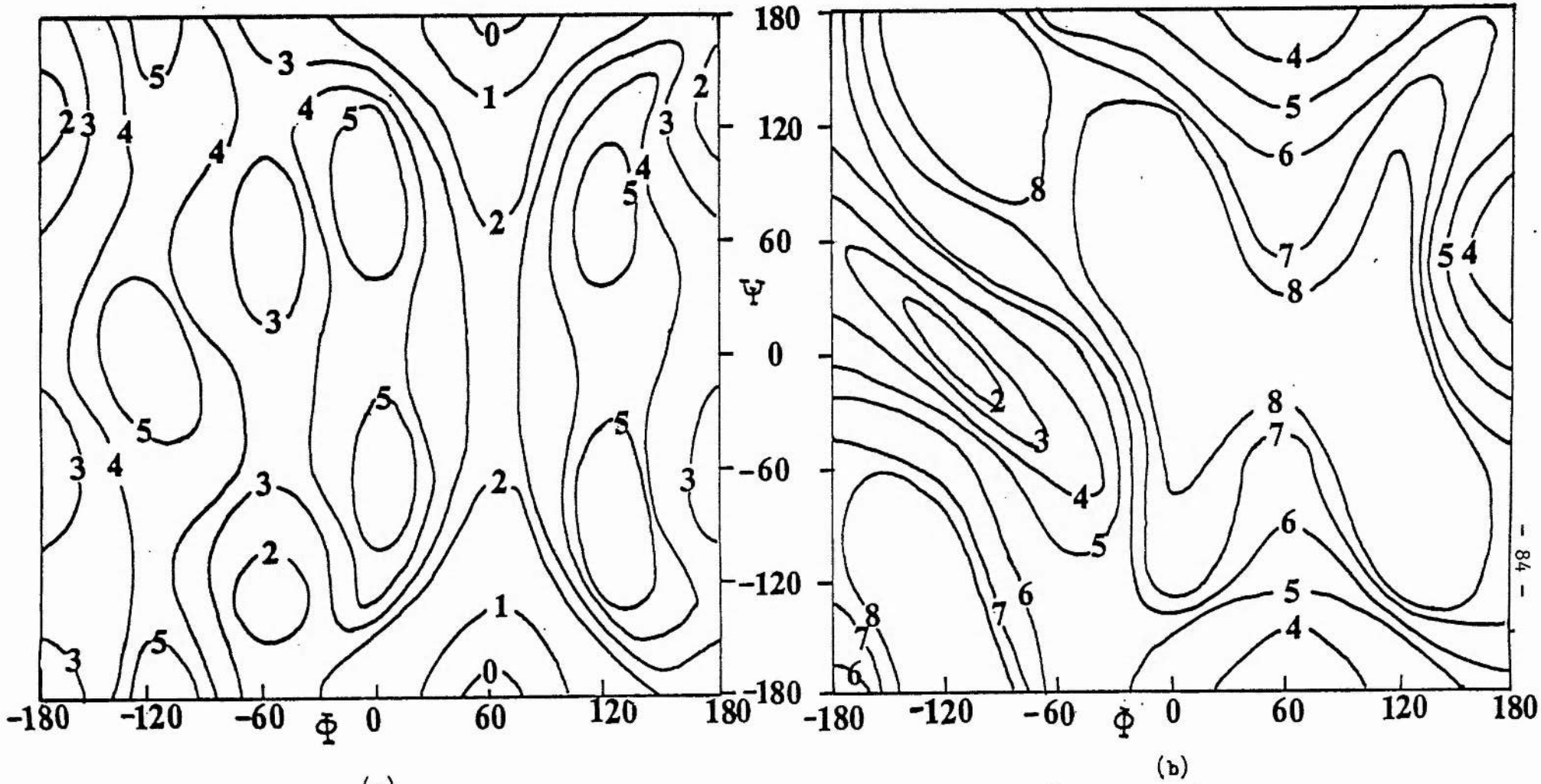


Fig. 5.3 STO-3G maps for glycine (a) $\omega = 0^\circ$ (b) $\omega = 180^\circ$

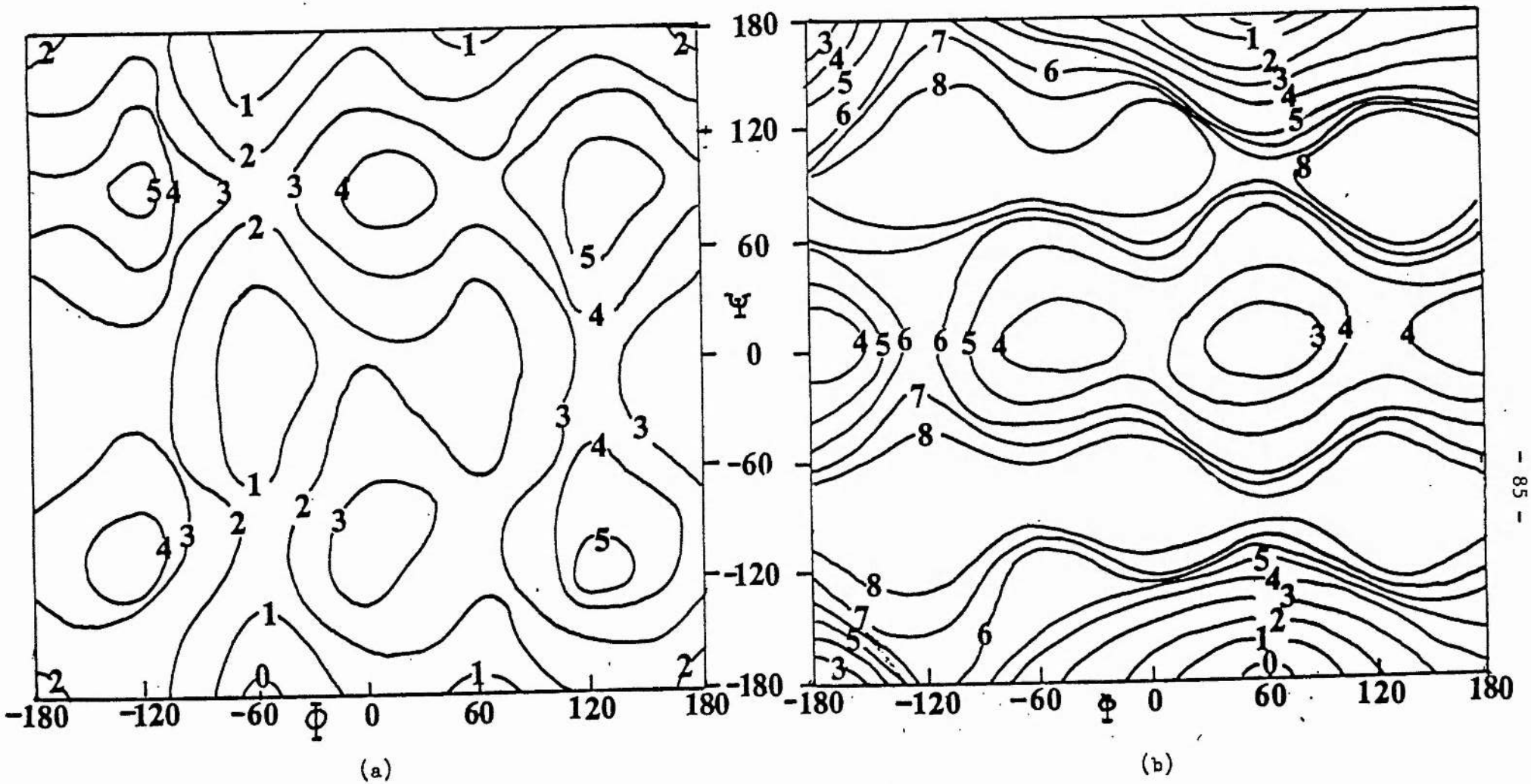
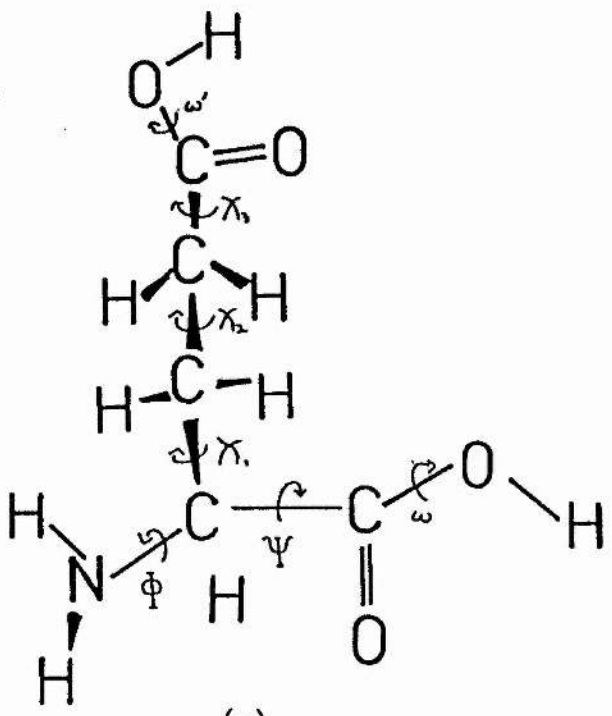
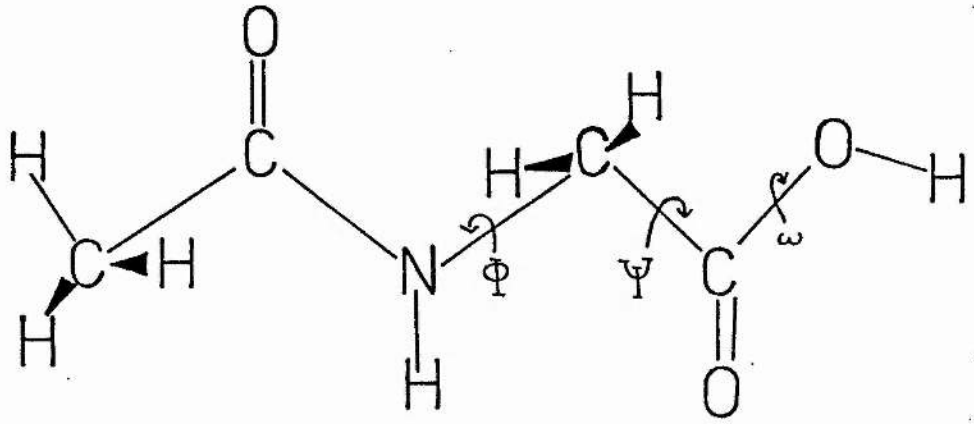


Fig. 5.4 Conformational energy maps for cysteine (a) PCILO (b) STO-3G



(a)



(b)

Fig. 5.5 Definition of torsional angles for

(a) glutamic acid (b) N-acetyl-glycine

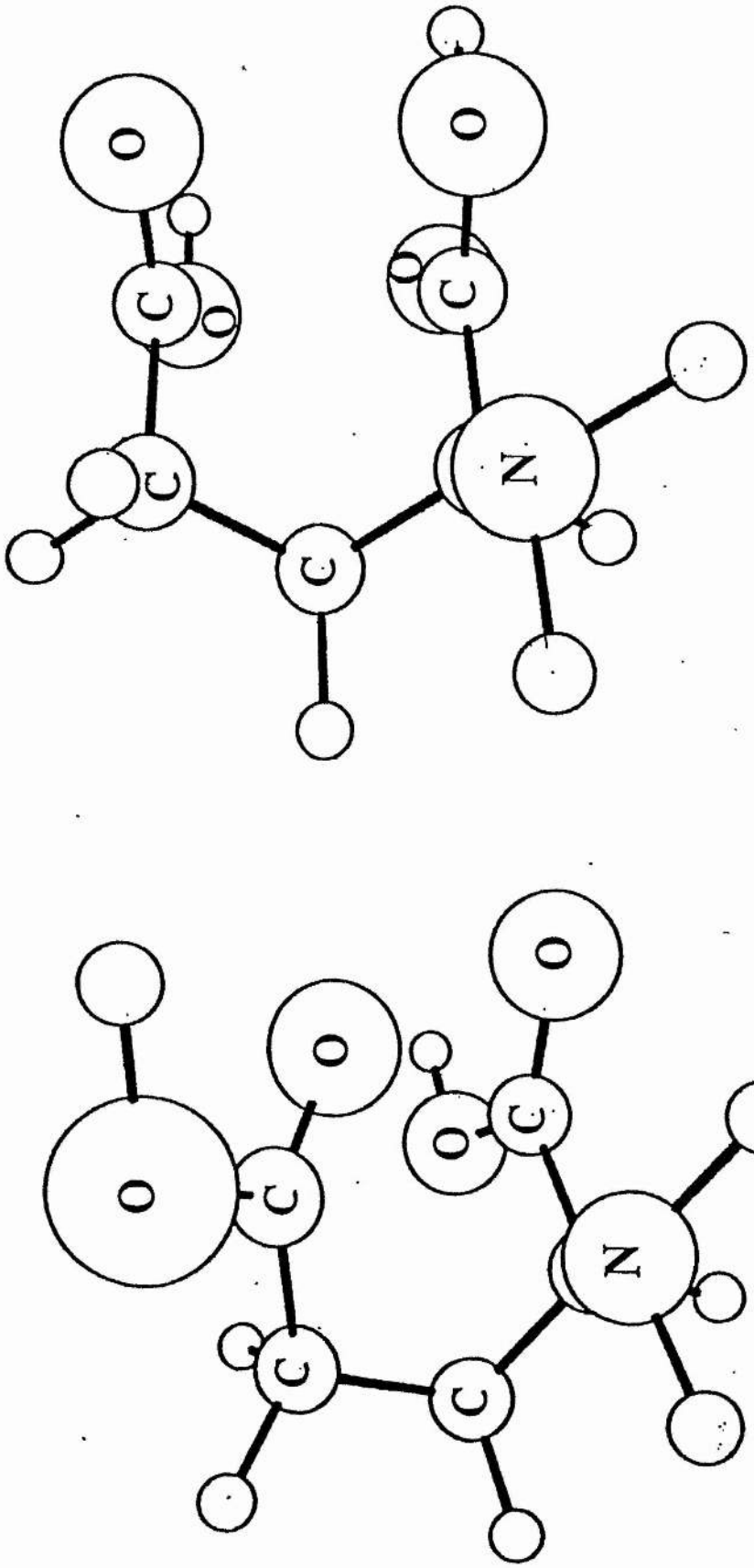


Fig. 5.6 Preferred conformations of glutamic acid

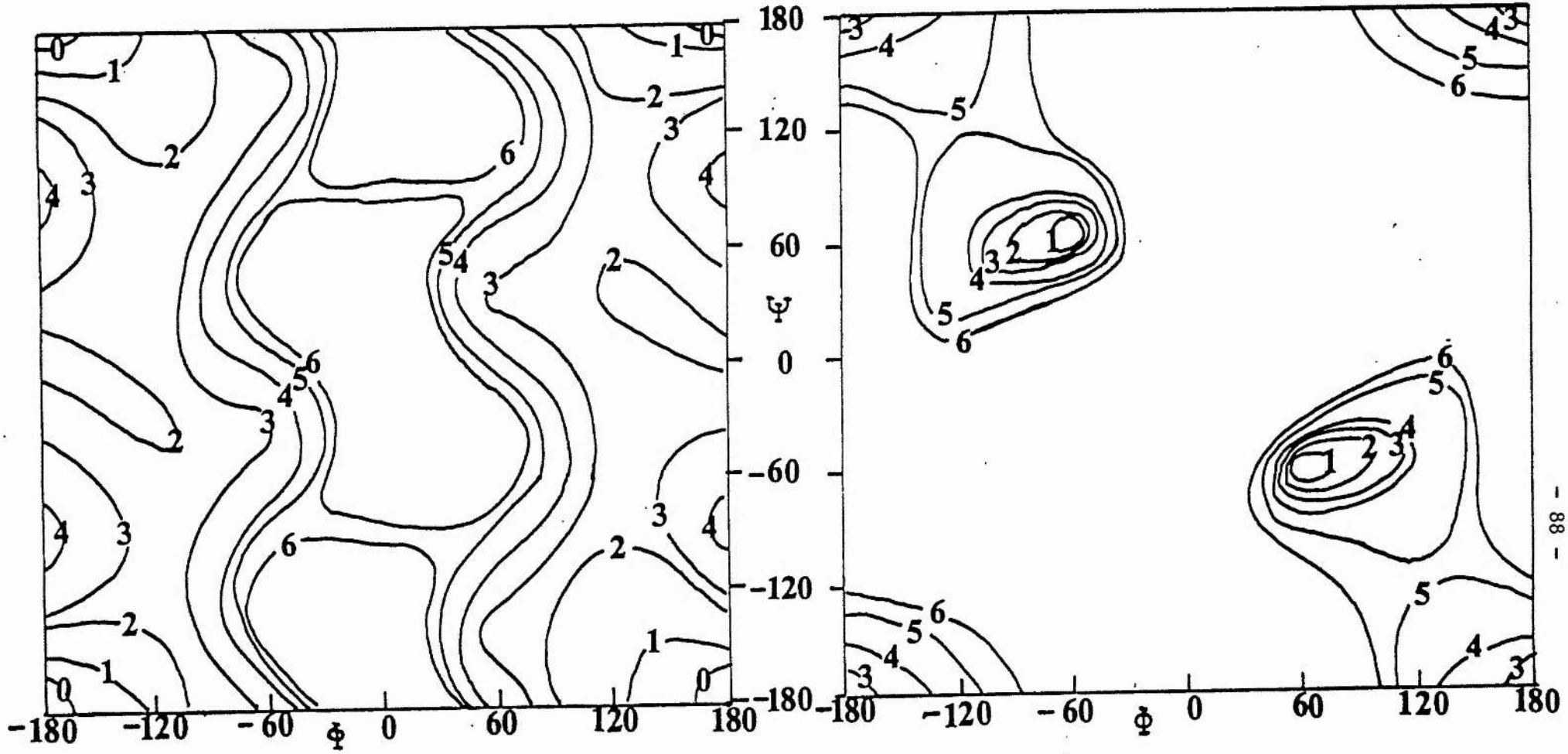


Fig. 5.7 STO-3G conformational energy maps for N-acetyl-glycine (a) $\omega = 0^{\circ}$ (b) $\omega = 180^{\circ}$

Table 5.1

Minimum energy conformations of N-acetyl-glycine

Method	PCILO	STO-3G	3-21G	4-31G
Conformation[a]	Energies in kcal mol^{-1} above global minimum			
C_5a	1.05(M)[b]	0.0(M)	0.0(M)	0.0(M)
C_5b	3.48(M)	2.92(M)	6.95(M)	7.68(M)
C_7	0.72(M)	0.50(M)	1.27(M)	3.68(M)
Ma	0.0(M)	5.71(S)[c]	12.89(S)	14.69(S)
Mb	3.33(M)	10.94(S)	23.32(S)	26.16(S)

a) for description see text

b) (M) indicates conformation occurs at minimum on potential energy surface

c) (S) indicates conformation occurs at saddle-point on potential energy surface

6. THE BORON ANALOGUE OF GLYCINE

6.1 Introduction

As mentioned in the preceding chapter, glycine is important as the "prototype" amino acid. Recently the synthesis and structure of a boron analogue of glycine, ammonia-carboxy-borane, $NH_3.BH_2.CO_2H$ (A), was reported in the literature [197]. Like its glycine counterpart the molecule possess the potential to form peptide linkages and hence to be incorporated into proteins. It is also reported to have biological activity, in particular it appears to have significant antitumour and antihyperlipidemic properties.

The results of a theoretical investigation into the structure, bonding and conformational flexibility of this molecule are reported here. Comparisons are made with similar results obtained for neutral and protonated glycine. The results of a brief conformational study on the boron analogue of N-acetyl-glycine ($[CH_3.CO.NH.BH_2CO_2H]^-$) are also presented.

The reported X-ray geometry [197] was used as a starting point for optimization of all geometric parameters using the Murtaugh-Sargent optimization technique [91] as implemented in the Gaussian 80 program package [80]. Since the X-ray analysis reports A to exist as a dimer in the crystal, geometry optimization was performed for both the isolated molecule and the dimer. For the isolated molecule both minimal (STO-3G) and extended (6-31G) basis sets were used; for the dimer, however, only minimal basis set calculations were possible. Optimization of the geometries of neutral

and protonated glycine was also carried out using the 6-31G basis set.

Electrostatic potential maps were produced as described in section 2.3.8 for the isolated molecules of glycine, protonated glycine and the boron analogue. Conformational energy maps were produced using the PCILO program for both X-ray and optimized geometries, the overall results being similar in both instances. For A an ab initio conformational energy surface was also produced using the Gaussian 80 program with an STO-3G basis.

6.2 Results

The results of the geometry optimizations are given in Table 6.1. For most of the parameters there is reasonable agreement between the X-ray and calculated values. The calculated value of the B-N bond length is slightly too large and that for the C-B-N angle too small, but it is possible that extended basis set calculations on the dimer would give better agreement with the X-ray results.

Comparing the computed STO-3G values for the monomer and dimer of A there appears to be little change in most of the parameters on dimerization (maximum change ~3%). Comparing calculated 6-31G values for the three isolated molecules, the B-N and B-C bonds are seen to be longer and the N-B-C angle smaller than the corresponding parameters in neutral and protonated glycine. The difference between the B-C-O(H) and O-C-O angles is also less than for either glycine or protonated glycine.

The dipole moments, energies of the highest occupied and lowest

unoccupied molecular orbitals and the electronic configuration of the highest occupied orbitals for the three isolated molecules and the dimer are given in Table 6.2. (These values were computed with the STO-3G basis set at the optimum geometries.)

The ionization potential computed from the orbital energy for A is slightly less than that for neutral glycine and the dipole moment is considerably higher (minimal basis sets do not always give reliable values for dipole moments but the 6-31G values for A and for glycine are 4.505 and 1.06 Debye respectively). The configuration of the highest occupied orbitals is the same for these two molecules and differs from that for protonated glycine.

The high dipole moment of A compared with that for glycine is presumably due to the presence of an NH_3 group in the molecule. The calculated charge distribution for A and for glycine is given in fig.6.1. Although the charge on the nitrogen is the same in both molecules, the hydrogens of A carry a higher positive charge than those of glycine and result in a net positive charge (+0.363) on the NH_3 group, in contrast to the almost neutral (-0.073) NH_2 group of glycine. The other notable difference is the positive charge on the boron atom and the polarity of the B-H bonds, compared to the non-polar CH bonds of glycine. The polarity of the B-N bond is also increased while that of the B-C bond is reduced when compared with the corresponding bonds of glycine.

Electrostatic potential maps were produced for A and for protonated and neutral glycine. The maps for neutral glycine and for A are given in fig.6.2 (the map for protonated glycine is not very interesting since it has a high positive potential at all points). The two maps show considerable similarity in the region of the

carboxylic group but differ at the nitrogen end of the molecule. Once again this difference is probably due to the presence of an additional hydrogen on the nitrogen atom, although the extended attractive region associated with the OH group on the map for A may be caused by the boron atom.

The molecule A has three degrees of rotational freedom. If the dimer is to form, however, the OH group must be in the cis position. The barrier to rotation of the OH group is higher than that for glycine (6.6 kcal as opposed to 3.93 kcal) and the cis position is more stable than the trans by 2.8 kcal. A conformational energy surface was produced for simultaneous rotation about the B-N and B-C bonds and compared with similar maps for neutral and protonated glycine. The barrier to rotation about B-N is very low (0.7 kcal compared with 1.5 kcal for glycine). The staggered arrangement is only 0.3 kcal more stable than the eclipsed and after re-optimization of the geometry this difference disappears. The barrier to rotation about B-C, on the other hand, is much higher than that for glycine (5.7 kcal compared with 1.04 kcal) and the two planar arrangements differ in energy by ~3 kcal (0.26 kcal for glycine). As for glycine the arrangement in which the carbonyl oxygen is cis to the nitrogen is preferred. The PCILO maps for protonated glycine are almost identical to those for A, having similar values for energy minima and barrier heights.

An ab initio conformational energy surface was produced for A using the STO-3G basis set. This is in very good agreement with the PCILO map, predicting identical minima and giving barriers to rotation of 0.3 and 4.7 kcal for rotation about B-N and B-C respectively.

Since the importance of A is its resemblance to the α-amino

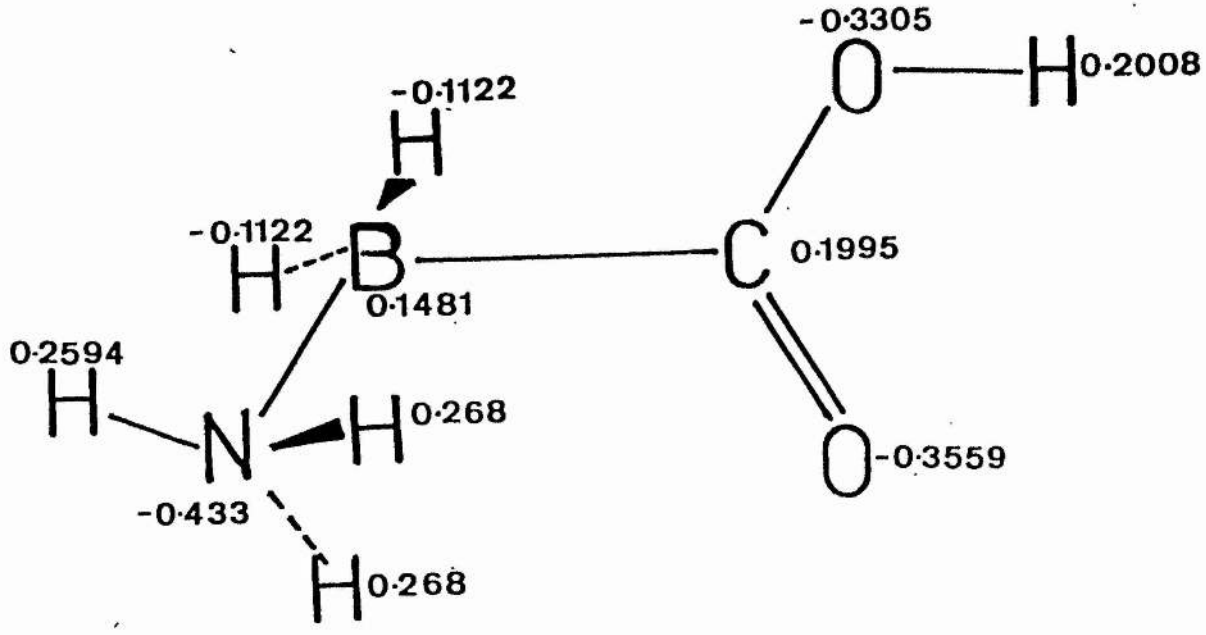
acids and its potential to form peptide linkages [197], it seemed of interest to investigate the conformational properties of a possible "peptide residue" of this molecule. As a comparison with earlier work (chapters 4 and 5) the boron analogue of N-acetyl-glycine was chosen. The geometry and torsional angles are given in fig.6.3. Conformational energy maps for ϕ, ψ rotation were produced for two values of ω : 0^{o} and 180^{o}. The maps are given in fig.6.4.

Comparison of the maps with those of N-acetyl-glycine (fig 4.5) reveals several important differences. Most obvious is the disappearance of the "M" conformations as important minima and the shifting of the C_7 minima on the $\omega = 180^{o}$ map. The $\omega = 0^{o}$ map is similar in overall appearance to that for N-acetyl-glycine but the C_5 arrangements are clearly preferred over the "M" conformations. On the $\omega = 180^{o}$ map the C_5 arrangements are again preferred, with further areas of low energy centred around $\phi = \mp 90^{o}$, $\psi = \pm 30^{o}$. These areas include the C_7 arrangement and the relative increase in energy of the latter may be due to the slightly different geometry which gives rise to high energy conformations at $\phi = \mp 60^{o}$, $\psi = \pm 30^{o}$. The geometries are not very different, however, (interatomic distances at the C_7 conformation differ by less than 2%), and it seems likely that factors other than simple steric ones are involved. It may be that the presence of the boron atom increases the preference for planar arrangements and that these low energy areas, which are more extensive than those for N-acetyl-glycine, represent a compromise between steric hindrance, H-bonding and the preference for planarity.

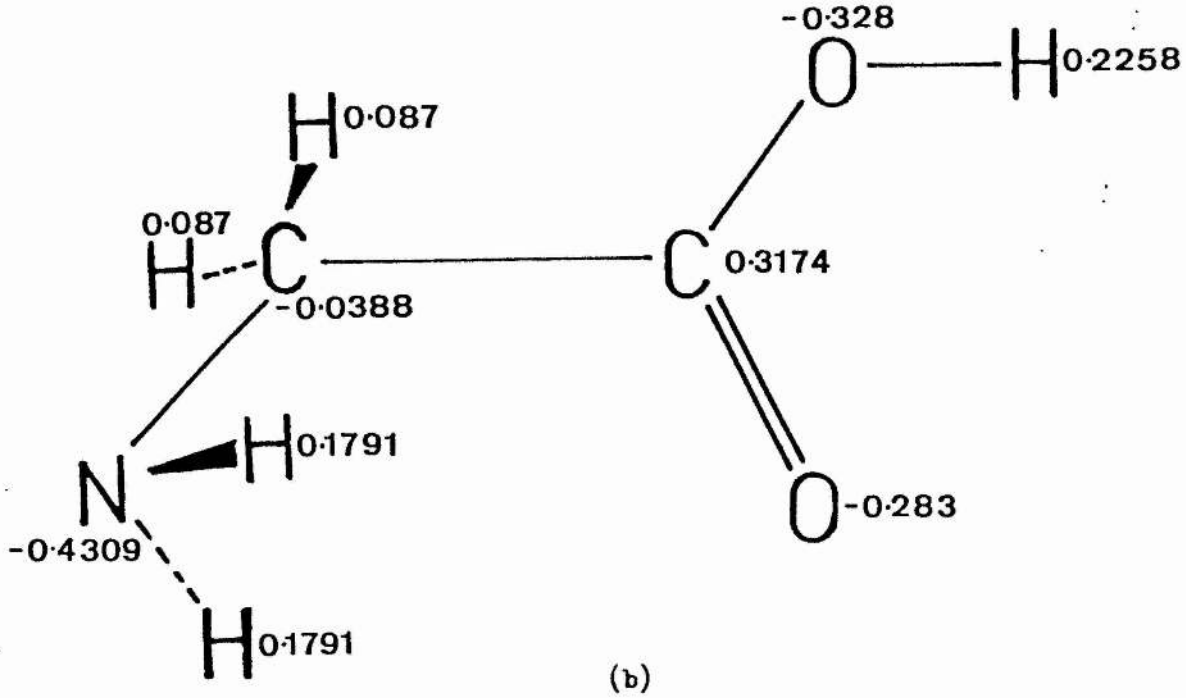
6.3 Conclusion

From this study it appears that the boron analogue of glycine differs in several respects from its glycine counterpart. The conformational properties of the molecule are almost identical to those of protonated glycine but in all other properties it is more like the neutral molecule. The differences which occur in the geometry, charge distribution, dipole moment and electrostatic potential appear to be due, at least in part, to the presence of the additional hydrogen on the nitrogen atom.

From the study of the boron analogue of N-acetyl-glycine it would appear that if glycine were to be replaced by its boron analogue in a protein, the resulting molecule would have conformational properties that were similar to those of its glycine counterpart, but with some important differences. If the increase in energy of the C_7 arrangement relative to C_5 is indeed a real effect, this could have serious consequences for protein structure since an extended rather than folded chain would then be favoured. Before reaching such a conclusion it is necessary to perform more accurate calculations, including full geometry optimizations, but these preliminary results do suggest that the problem may be worth investigating.



(a)



(b)

Fig. 6.1 Calculated charge distribution for

(a) boron analogue of glycine

(b) glycine

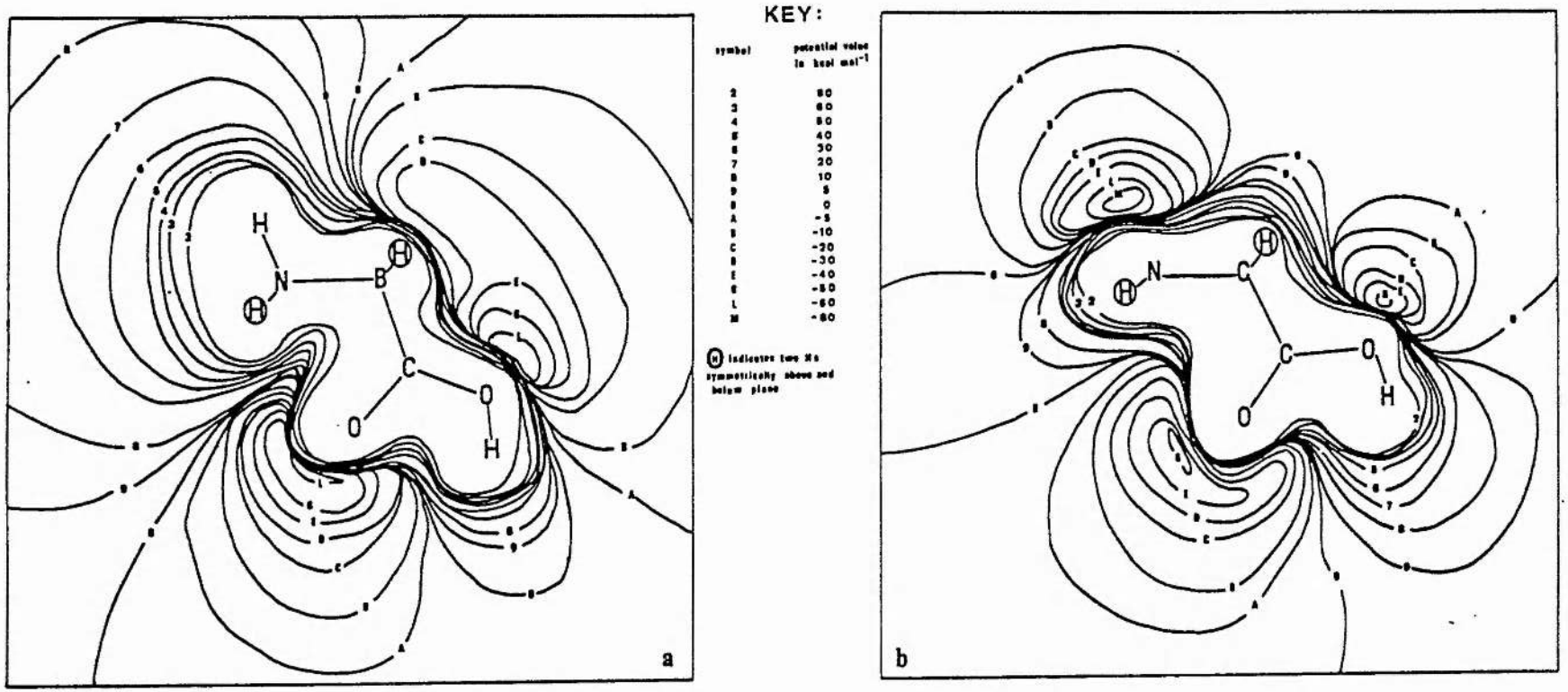


Fig. 6.2 Electrostatic molecular potential maps drawn in the plane of the molecule for (a) boron analogue of glycine (b) neutral glycine

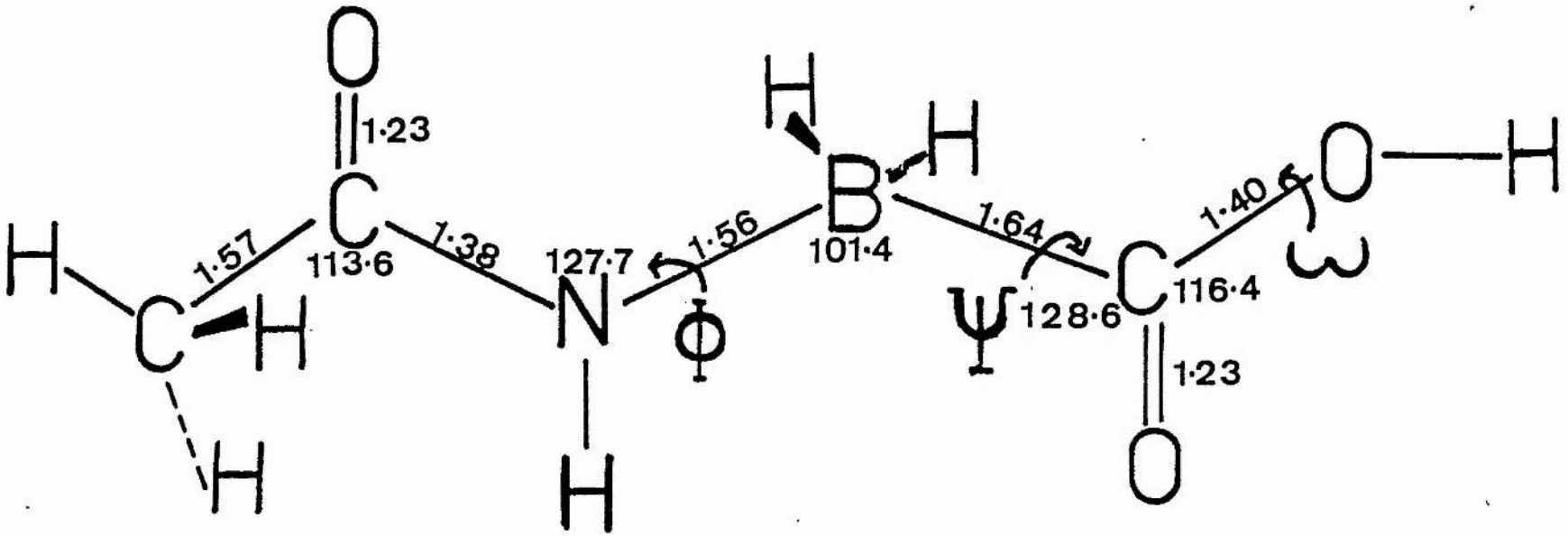


Fig. 6.3 Definition of torsional angles for boron analogue of N-acetyl-glycine

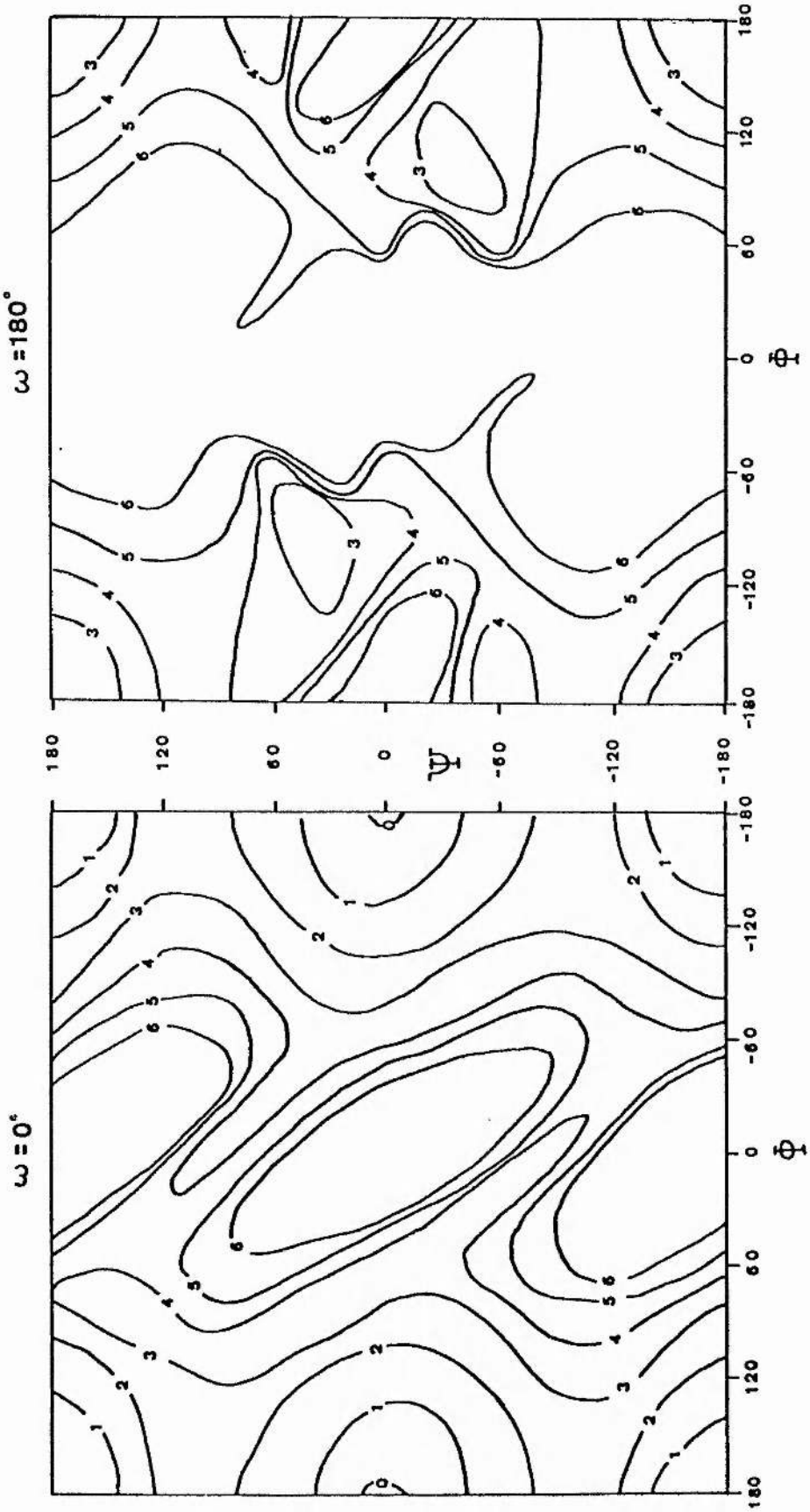


Fig. 6.4 Conformational energy maps for boron analogue of N-acetyl-glycine

Table 6.1

Bond lengths (Å) and valency angles (°) for $NH_3.BH_2.CO_2H$, its dimer, and glycine analogues

Molecule	X = B				X = C	
	$(NH_3.BH_2.CO_2H)_2$		$NH_3.BH_2.CO_2H$		$(NH_3.CH_2CO_2H)^+$	$NH_2.CH_2CO_2H$
Parameter	a	b	b	c	c	c
X-N	1.587	1.639	1.642	1.666	1.507	1.435
X-C	1.601	1.622	1.619	1.604	1.504	1.508
C=O	1.246	1.246	1.227	1.229	1.209	1.211
C-O(H)	1.321	1.356	1.396	1.354	1.320	1.353
C-X-N	108.8	102.7	102.3	102.0	107.3	114.9
X-C=O	124.2	122.4	126.4	124.4	121.6	126.3
X-C-O(H)	116.1	116.9	115.3	116.0	111.3	111.5
O=C-O(H)	119.1	120.7	118.4	119.6	127.1	122.1

a. Value from X-ray analysis [197]

b. Calculated optimum value with STO-3G basis set

c. Calculated optimum value with 6-31G basis set

Table 6.2

Orbital energies, dipole moments and electronic configuration for $NH_3.BH_2CO_2H$, its dimer, and glycine analogues

Molecule	E(HOMO)[a]	E(LUMO)[a]	dipole moment[b]	configuration of highest occupied orbitals
$NH_3.BH_2.CO_2H$	-0.2715	+0.3634	5.146	$(15\sigma)^2(4\pi)^2(16\sigma)^2$
$(NH_3.BH_2.CO_2H)_2$	-0.2328	+0.3588	0.002	$(7\pi)^2(32\sigma)^2(33\sigma)^2$
$[NH_3.CH_2.CO_2H]^+$	-0.5328	+0.1151	-	$(15\sigma)^2(4\pi)^2(16\sigma)^2$
$NH_2.CH_2.CO_2H$	-0.3056	+0.3139	0.712	$(15\sigma)^2(16\sigma)^2(4\pi)^2$

a. Energies in Hartrees

b. Dipole moment in Debyes

ASCORBIC ACID AND RELATED MOLECULES

7 MOLECULAR STRUCTURE

7.1 Introduction

Ascorbic acid (vitamin C) was first isolated in the 1930's and identified as the essential substance required to prevent scurvy [198]. The molecule appears to influence a wide range of biological activities including hormone production, enzyme biosynthesis and concentration levels of the cyclic nucleotides c-AMP and c-GMP. The molecule is also essential for the formation of collagen and possesses detoxificating and antihistamine activity; it also appears to have an enhancing effect on the immune system. (A discussion of the biochemical and physiological properties of ascorbic acid may be found in S. Lewin's book [199].)

Although vitamin C is essential for the normal functioning of the human body the molecule cannot be synthesized in significant quantities by the body and must therefore be ingested. The required daily intake of the vitamin has been the subject of some dispute. While 10 mg/day is sufficient to prevent scurvy, "official" recommendations for daily intake are usually between 30 and 70 mg/day [200] and some workers advocate much higher doses of 0.25-10g/day [201,202]. Some of the benefits that are reported to result from the ingestion of large quantities of vitamin C are discussed in reference [199]. Of particular interest in this work are the findings of Cameron and co-workers [203-205] who have administered large doses of vitamin C to terminal cancer patients with, in some cases, quite remarkable therapeutic results. While the significance of these

results has been disputed by others [206] the collection of case histories recently reported by Cameron [205] clearly shows that large doses of ascorbic acid have been of considerable benefit to a number of terminal cancer patients. In many cases relief of symptoms has been reported and some patients have experienced complete remission for periods of more than five years.

Further experimental evidence of the importance of vitamin C in controlling neoplastic disease has come from in vitro studies. The presence of ascorbic acid has been shown to suppress the growth of transformed cells [207] and to inhibit chemically induced transformation [208]. Reversion of the transformed phenotype was also observed in the latter study. Large doses of vitamin C have also been found to reduce the growth of chemically induced tumours in rats [209].

The mechanism by which ascorbic acid acts in these systems is not yet known and it may prove to be unrelated to the "charge transfer theory" put forward by Szent-Gyorgyi. However, the important role that he proposes for ascorbic acid in maintaining semi-conduction in proteins does offer one possible explanation for the effects observed by Cameron. As discussed in chapter 3 ascorbic acid is thought to act by enhancing the electron acceptor properties of methylglyoxal, thus increasing the charge transfer from proteins. The primary aim of the studies reported here was to investigate using quantum mechanical methods the possibility and extent of such charge transfer enhancement.

Neutral ascorbic acid is an enediol containing a double bond between the and carbon atoms and an acid ionizing group. It is reversibly oxidized to dehydroascorbic acid which can then

undergo further (irreversible) oxidation to oxalic acid. The relationship between ascorbic acid and dehydroascorbic acid is shown in fig. 7.1. The anion and anion radical are known to occur as intermediates in the oxidation/reduction process [210,211]. Other intermediates that have been proposed are neutral and cation radicals [212,213]; however, the e.s.r. evidence indicates that the anion is the stable radical form [214,215]. This has been substantiated by similar studies on α-hydroxytetronic acid [216]. The radical also occurs in biological systems [217,218].

The structure of ascorbic acid has been investigated experimentally by X-ray and neutron diffraction techniques [219,220] and by u.v. and n.m.r. spectroscopy [221-224]. Theoretical studies employing semi-empirical methods have been carried out to investigate the π-electron system of ascorbic acid and its anion [225] in order to correlate the structure of the orbitals with the observed u.v. spectra. The geometry and preferred tautomer of ascorbic acid have been calculated using the MINDO/3 and MNDO programs [226] and CNDO/2 has been used to investigate the electron affinity and possible binding to proteins [227]. The first ab initio study on ascorbic acid was that by Carlson et al. [228] in which ascorbic acid and its anion were investigated using an STO-3G basis set. The X-ray geometries [219,220] were employed and limited variation of the torsional angles of the side-chain was carried out. More recently ab initio and semiempirical (INDO) wave functions have been computed for the various possible radicals that may form from ascorbic acid and from α-hydroxytetronic acid [229]. The spin density and hyperfine coupling constants were calculated for the various radicals and compared with the experimental values. Only the values for the

anionic species were compatible with the experimental data. A number of investigations have also been carried out into the donor/acceptor properties of ascorbic acid and its metabolites [230,231]. These are discussed at the beginning of chapter 8.

7.2 Models for the ascorbic acid system

The neutral ascorbic acid molecule has 20 atoms and 92 electrons. ab initio calculations on such a molecule are computationally expensive, particularly if one is interested in its interactions with other molecules. For this reason the majority of calculations reported here have been performed using the model compound α-hydroxytetronic acid, its anion, anion radical and dehydro α-hydroxytetronic acid (furantrione) (figs. 7.2-7.6).

To investigate the suitability of these compounds as models for ascorbic acid and its metabolites single ab initio calculations were performed for the model compounds and for their respective ascorbic acid analogues. The total energies and energies of the HOMO and LUMO for each molecule are given in Tables 7.1 and 7.2. The charge distributions are illustrated in figs. 7.3-7.6. It can be seen that the orbital energies and charge distribution in the ring are very similar in the two sets of molecules. This suggests that for properties and reactions involving the ring system the α-hydroxytetronic acid compounds will provide suitable models for ascorbic acid and its metabolites.

An inspection of the wave function for ascorbic acid (see below) shows that the molecular orbitals of highest energy are associated

with the ring and not with the side-chain. That the anti-scorbutic properties of ascorbic acid are dependent on the presence of the side-chain is well known [232] and the absence of a chiral centre in ∝-hydroxytetronic acid will obviously affect its physiological distribution and metabolism. It seems likely, however, that electronic effects such as those occurring in charge transfer will involve only the ring system. This is substantiated by the recent observations of P. Gascoyne [233] that D-ascorbic acid has the same enhancing effect on the charge transfer between methylglyoxal and protein as the L isomer. Thus for the investigation of the charge transfer properties of ascorbic acid ∝-hydroxytetronic acid appears to offer a suitable model.

7.3 Geometry optimization

Geometry optimization was carried out for ∝-hydroxytetronic acid, its anion, anion radical and the dehydro compound using the reported X-ray results for ascorbic acid and its anions as a starting point. The calculations were performed with the STO-3G basis set using the Murtaugh-Sargent optimization procedure, as implemented in the Gaussian 80 program package. The results are presented in figs. 7.7 and 7.8. All the molecules are found to be approximately planar. The external oxygens of the anion are tilted slightly (4-5^{o}) out of the plane of the ring and the hydroxyl hydrogen lies at ~90^{o} to this plane. The OH groups of the neutral molecule lie in the plane.

The STO-3G basis set has been found to give reasonable values for bond lengths and angles in neutral closed shell species [107].It has

also been used in the study of anions but the values for bond lengths and angles are not so good [96]. For the radical there is the additional problem that the UHF wave function is not an eigenfunction of S^2. It was found that during geometry optimization the value of S^2 tended to rise, implying increased contamination by states of higher multiplicity. Attempts to determine the optimum geometry using a restricted open shell procedure were unsuccessful (the ROHF procedure available in the Gaussian 80 program package does not appear to be very efficient, particularly as it cannot be used with a gradient method of optimization). An inspection of the STO-3G wave function for the anion and anion radical reveals that the energies of the HOMOs are positive for both molecules. This implies that the unpaired electron in these molecules is unbound (see section 2.3.1). However, the 6s3p basis set of Mezey and Csizmadia with a minimal contraction scheme [71] was found to give negative values for the HOMOs of the anion and anion radical, as well as giving a much better value for the total energy than that obtained with the STO-3G basis set. Geometry optimization was repeated for these two molecules using this basis set. The results are shown in fig. 7.8. The 6s3p basis set gives larger values for bond lengths but the overall trends are the same. The C-O-C angle in the ring is slightly wider for both molecules but in general the values for bond angles and dihedral angles are in good agreement with the STO-3G results.

Reference [226] compares the results of calculations with MINDO/3 and MNDO to those obtained experimentally. A comparison of the STO-3G results with these values shows that the ab initio results for the bond angles are in general much closer to the experimental value than those obtained by the semi-empirical methods. The only angles for

which the STO-3G basis set gives poor values are the two H-O-C angles, where the calculated value is ~104° while the experimental value is ~113°. For the C_2-C_3, C_5-C_6 and C_7-O_8 bonds STO-3G gives values that are in good agreement with experiment. The other bond lengths, however, are calculated to be too long by up to 4%.

Comparing the STO-3G geometries for the four molecules it can be seen that the removal of a proton causes a lengthening of the C_2-C_3 bond and a shortening of the C_2-C_7 bond. In addition the C_3-O_4 bond length is decreased to a value comparable to that of the C_7-O_8 bond. The C_5-C_3-C_2 angle is decreased and the C_3-C_2-C_7 angle is increased in the anion. There is also a lengthening of the C_2-O_1 bond and the hydroxyl hydrogen lies approximately at right angles to the plane of the ring.

In the anion radical the three external C-O bond differ from one another by ~0.04 Å and increase in the order C_7-O_8 < C_3-O_4 < C_2-O_1. The C_2-C_3 and C_2-C_7 bonds are slightly longer than in the anion but both are still considerably shorter than the C_3-C_5 bond.

In the dehydro model compound the C_3-C_5 and C_2-C_7 bonds are almost equal in length and the C_2-C_3 bond is only slightly shorter. The three C=O bonds are approximately the same length and the C-C-O angles are all close in value. No resonance is possible in this molecule.

7.4 Electronic Configuration and Bonding

The complete molecules of ascorbic acid and its metabolites were studied by taking the optimized ring structure of the appropriate

model compounds and replacing one of the hydrogens at the C_5 position by the side-chain $CHOH.CH_2OH$. For ascorbic acid and dehydroascorbic acid the geometry of the side-chain was taken from the X-ray study of neutral ascorbic acid [219], while for the anion and anion radical the results of a similar study on the anion were used [220]. A single calculation was performed for each molecule using both the STO-3G and the 6s3p basis sets.

The total energy, the energy of the four highest occupied orbitals and the energy of the lowest unoccupied orbital for the four molecules are listed in Table 7.2. It can be seen that 6s3p basis set gives lower values for the total energy and orbital energies than the STO-3G basis set. An inspection of the wave-function, however, shows that the composition of the molecular orbitals is generally similar for both basis sets, as are the trends in orbital energies for the four molecules.

The charge distribution calculated with the 6s3p basis set is shown in figs. 7.9 and 7.10. Compared with the values calculated with the STO-3G basis set (figs. 7.3-7.6) the 6s3p basis set gives greater polarity for most bonds. The total dipole moment for the neutral molecules is also higher. Since the Mulliken population analysis is so dependent on the basis set it is to be expected that there will be differences in the total atomic charge and overlap populations. However, the differences are not very large and the trends within the four molecules are similar for both basis sets.

The absence of a plane of symmetry in the molecule means that orbitals cannot be strictly classified as σ or π. However, molecular orbitals which consist largely of contributions from the p_y atomic orbitals of carbon and oxygen have been designated π in Table

7.2. Similarly orbitals consisting largely of contributions from s, p_x and p_z orbitals have been labelled σ. Those which contain substantial (>0.3) contributions from both types of atomic orbitals are indicated by σ/π. The letters r and s are used to indicate whether the orbital is associated largely with the ring or with side-chain. The label ro means "ring only" and indicates that a molecular orbital is localized on the atoms of the ring, while rm and sm are used to indicate orbitals that are associated mostly with the ring or mostly with the side-chain. The label s+r indicates an orbital that has substantial contributions from both ring and side-chain atomic orbitals.

An inspection of Table 7.2 shows that the HOMO and LUMO for all four molecules are associated only with the atoms of the ring. For all but dehydroascorbic acid the HOMO is a π type orbital; for this molecule the HOMO is a σ orbital with a π orbital lying immediately below it.

The four highest occupied orbitals in the anion and anion radical and the two highest in the neutral molecules are associated predominantly with the ring. The HOMO of neutral ascorbic acid contains approximately equal contributions from the p_y orbitals of atoms 1,2,3 and 4, and slightly less from atom 8. In the anion the contributions from atoms 2 and 8 have increased while the contribution from 1 has decreased and atom 3 does not contribute at all. In the radical the contribution from 2 and 8 have decreased compared with the anion and there is a larger contribution from atom 1, as well as a contribution from atom 3. The values for atoms 1,4 and 8 are similar to those of neutral ascorbic acid, while the contribution from 3 is smaller and that from 2 is larger. The HOMO of dehydroascorbic acid

contains contributions from all the atoms of the ring system, the largest contributions coming from atoms 1,4 and 8.

The energies of the HOMOs of the four molecules increase in the order DHA < ascorbic < radical < anion, while the LUMOs increase in the order DHA < ascorbic < radical $\simeq$ anion. With the 6s3p basis set the LUMO of DHA is calculated to be slightly negative and that of neutral ascorbic acid only slightly positive. This suggests that these two molecules should be good electron acceptors. The high energy of the HOMOs of the anion and anion radical imply that these molecules will be good electron donors.

A rough estimate of ionization potential and electron affinities may be obtained from the energies of the HOMOs and LUMOs respectively (see section 2.3.1). Calculated in this way the anionic species, as might be expected, have low ionization potentials and high (positive) electron affinities, while the reverse is true for the neutral molecules. The 6s3p basis set predicts a negative value for the electron affinity of dehydroascorbic acid.

An inspection of the overlap populations as calculated with the Mulliken population analysis shows that there is no significant overlap between the atoms of the side-chain and those of the ring (except, of course, between C_5 and the first carbon of the side-chain). The overlap populations in the side-chain are similar for all four molecules. The values calculated with the STO-3G basis set for the ring are shown in figs. 7.11 and 7.12. The 6s3p basis set in general gives higher values but the trends are similar.

As might be expected the greatest changes in overlap populations for the four molecules are between atoms 1,2,3,4,7 and 8. In neutral ascorbic acid the highest overlap is between C_2 and C_3, the values for

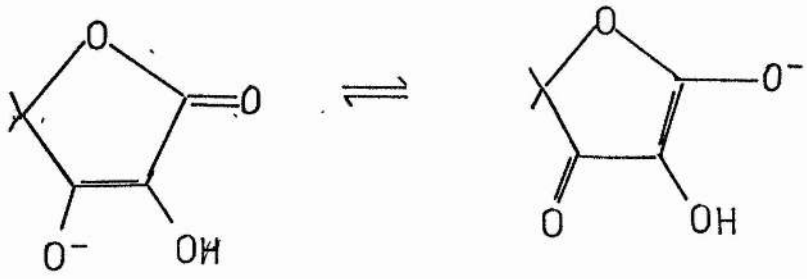
the other C-C bonds being considerably less. The overlap between C_7 and O_8 is also higher than any of the other C-O bonds. In the anion the overlap between C_2 and C_3 has decreased while that between C_2 and C_7 has increased. The overlap between C_3 and O_4 is similar to that between C_7 and O_8 while the overlap between C_2 and O_1 remains close to the value for the neutral molecule; the overlap between O_1 and H_{10} has decreased slightly. In the radical the highest overlap is still between C_2 and C_3 with a value close to that for the anion. The value for C_7-O_8 is also similar while that for C_3-O_4 has decreased and that for C_2-O_1 has increased. In dehydroascorbic acid the highest overlaps occur for the three C=O bonds, which are approximately equivalent. The three C-C bonds also have roughly equal values which are less than those for the C-C bonds in ascorbic acid.

The charge distribution in the side-chain is similar for all four molecules. The values for the ring atoms calculated with the STO-3G basis set are shown in figs. 7.3-7.6. The main difference between ascorbic acid and its anion are an increased negative charge on all the oxygens, particularly O_4 and O_8 which have almost equal charges. There is also an increase in the positive charge at C_3 and a decrease at C_7, while C_2 has acquired a negative charge. There is a decrease in the positive charge on H_{10}. In the radical the charges on C_2 and C_3 are small and positive, as for the neutral molecule. The charge on O_6 is similar to that for the anion but the charges on the other oxygens have decreased. Between the radical and dehydroascorbic acid there is a further decrease of charge on the oxygens and an increase in positive charge for carbons 2,3 and 7.

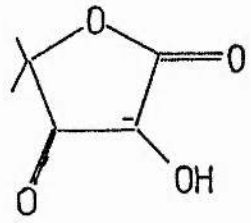
It is interesting that H_{10} and H_{11} have almost equal charges in the neutral molecule. The overlap populations for the two O-H bonds

are also very similar, as are the bond lengths and angles. Experimental evidence [220] indicates that it is H_{11} that is ionized to form the anion, although Lewin has suggested [199] that the two anions may exist in equilibrium. To investigate this further the geometry of a model of the alternative anion was optimized using the 6s3p basis set and the Murtaugh-Sargent optimization technique. The results are shown in fig.7.8.

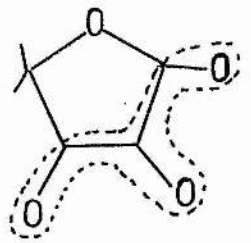
Calculations with the 6s3p basis set show this second anion to be ~18 kcal mol^{-1} less stable than the first, it thus seems unlikely that the two would exist in equilibrium under normal conditions. In the second anion there is a considerable difference in bond length between C_2-C_3 and C_2-C_7, and C_2-O_1 is longer than C_3-O_8. Since the charge distribution in the neutral molecule suggests that both OH groups are equivalent, the difference in energy between the two anions is probably due to resonance stabilization of the anion which forms by removal of H_{11}:



The calculated charge distribution, overlap populations and geometry for this molecule support this resonance picture. The negative charge on C_2 suggests that there may also be a contribution from the structure:



While similar resonance is possible for the radical the charge distribution, overlap populations and geometry do not show the same pattern as for the anion. In this case the unpaired electron appears to be delocalized over the tricarbonyl system:



7.5 Spin Density and Hyperfine Coupling Constants

Spin densities and hyperfine coupling constants for the ascorbyl anion radical and for the corresponding model compound were computed using the INDO program. Ab initio SCF calculations using both the STO-3G and 6s3p basis sets were also carried out to determine spin densities.

Previous calculations on these two molecules [229] gave reasonable agreement with experiment for the proton hyperfine coupling

constants but the values for carbon suggested that the experimental assignments for C_2 and C_7 should be reversed. The calculated value for C_5 was also much lower than the experimental value. It was suggested that this discrepancy might be due to the use of an unoptimized geometry. To investigate this suggestion the hyperfine coupling constants for the model anion radical were calculated using the optimum geometry computed with the STO-3G and 6s3p basis sets. Calculations were also performed using the X-ray geometries for neutral ascorbic acid and its anion. The results are presented in Table 7.3.

It is apparent from the calculations that changes in geometry do have a marked effect on the hyperfine coupling constants, however, the results for the optimized geometries cannot be said to be in better agreement with experiment than those for the X-ray geometry. The value for C_5 has increased but is still less than the experimental value. The value for C_3, which was previously in good agreement with experiment, has also increased and is calculated to be much too high, particularly for the 6s3p optimum geometry. The values for C_2 and C_7 are much closer to one another, especially for the STO-3G geometry, but neither is in good agreement with the experimental value. The marked difference in the values for the two protons at the anion X-ray geometry is presumably due to the fact that the ring is not quite planar. The dihedral angle $C_6-C_5-O_4-C_3$ is $\sim 9^{\circ}$ in the experimental geometry. For the optimum 6s3p geometry this angle is $\sim 4^{\circ}$ and for STO-3G $\sim 3^{\circ}$. The values for the protons at the 6s3p geometry are much too high. The values for the side-chain protons at this geometry (Table 7.4), however, are too low. The hyperfine coupling constants for the oxygen atoms do not vary so much with the changes in geometry

and the ordering is the same for all geometries: $a_1 > a_4 > a_8 > a_6$. Unfortunately no experimental data is available at present with which to compare these values.

The calculated results for the complete molecule (Table 7.4) are again in poor agreement with experiment. This suggests that the "optimum" geometry calculated using the UHF procedure may not be the true geometry of the ground state. The high value for S^2 (0.81) suggests that there is contamination from states of higher multiplicity. The only way to overcome this problem is to perform geometry optimization using a RHF procedure. Unfortunately, although this task was begun, the calculations proved to be too computationally expensive and time-consuming to pursue systematically.

Ab initio atomic spin densities for the model ascorbyl radical have been calculated with the STO-3G basis set and are presented in Table 7.5 for the optimum STO-3G geometry and for the anion and neutral X-ray geometries. S-orbital spin densities calculated with the INDO program are also presented.

For all the geometries the largest spin density is calculated to be at O_1 with a substantial value also at O_4 and rather less at O_8. The values at C_2, C_3 and C_7 are also quite high, suggesting that the unpaired electron is to some extent delocalized over the tricarbonyl system. The s-orbital spin densities follow roughly the same pattern as the hyperfine coupling constants.

In conclusion, these calculations using the UHF optimum geometries of the model ascorbyl radical do not show the improvement expected when compared to calculations using unoptimized geometries. The poor agreement of the calculated hyperfine coupling constants with those determined experimentally suggests that the geometry of the

model radical should be reoptimized using a restricted open shell procedure. In this present work this has not been possible and the UHF geometry has been used in the calculations discussed in the following chapters. Previous calculations have shown that variations in the geometries of the constituent molecules do not have any marked effect on supermolecule calculations of the type described here [163]. However, it would be desirable to repeat the calculations using a more accurate geometry when one becomes available.

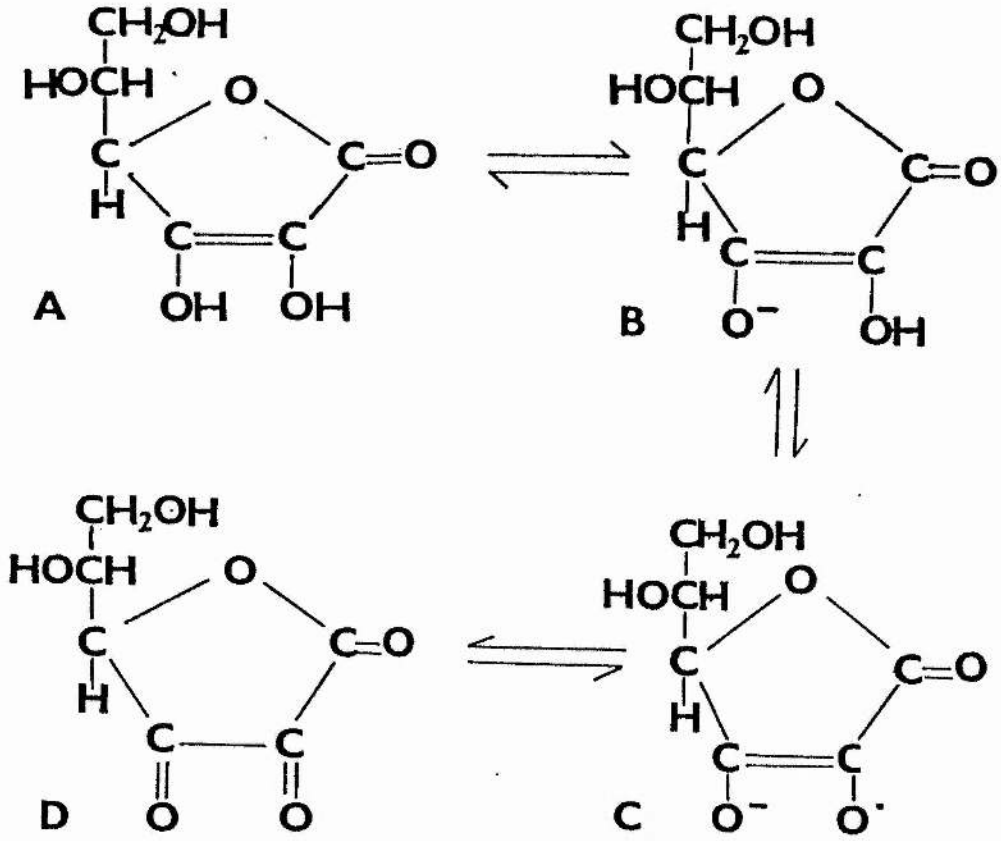


Fig. 7.1 Relationship between ascorbic acid and its metabolites (a) ascorbic acid (b) ascorbate anion (c) ascorbyl radical (d) dehydroascorbic acid

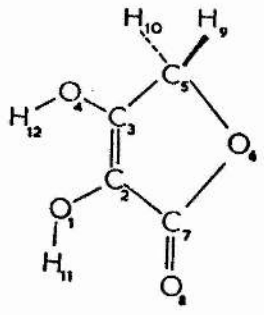


Fig. 7.2 ∝-hydroxytetronic acid, showing numbering scheme used for model compounds

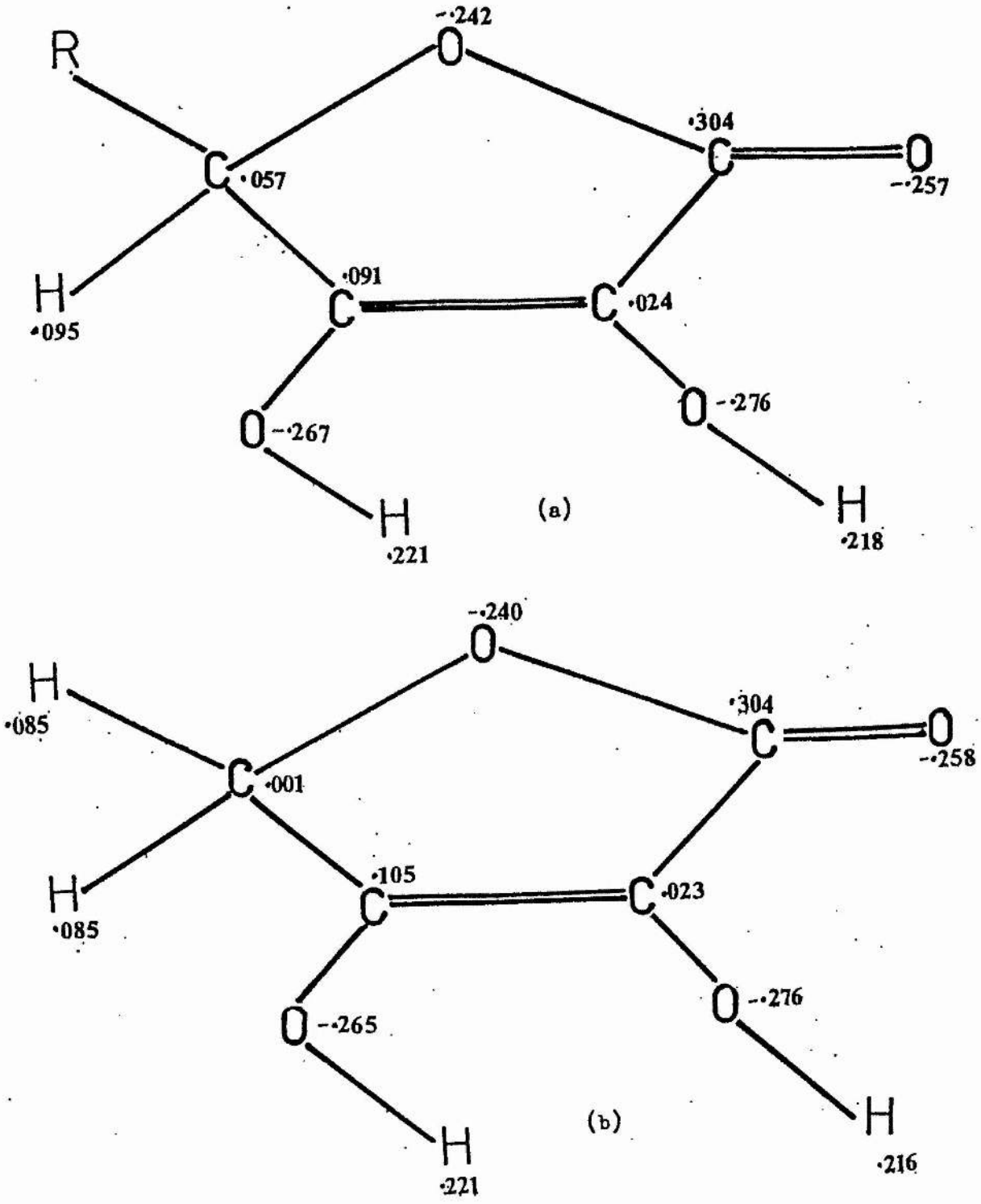
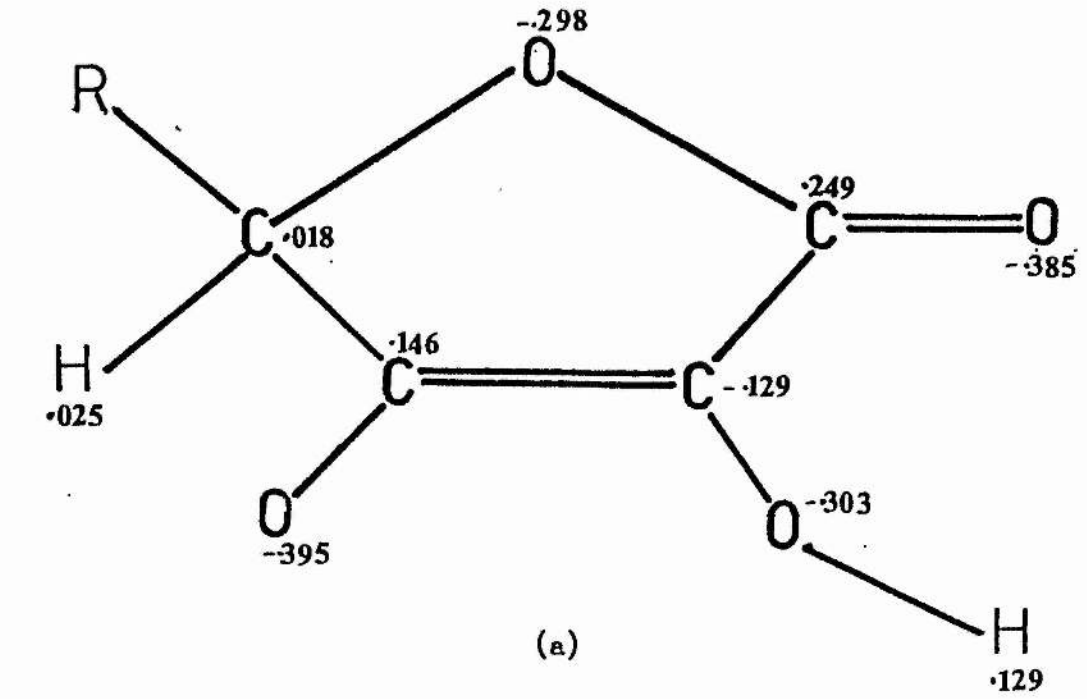
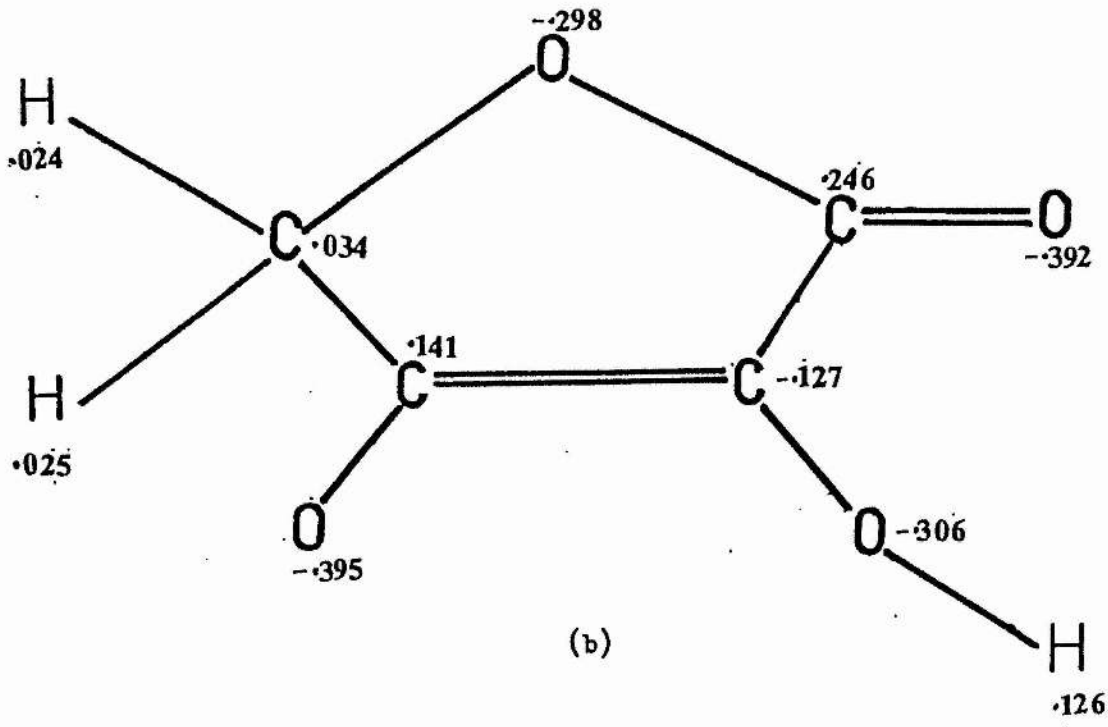


Fig. 7.3 Charge distribution calculated with STO-3G basis set for (a) ascorbic acid (b) α-hydroxytetronic acid

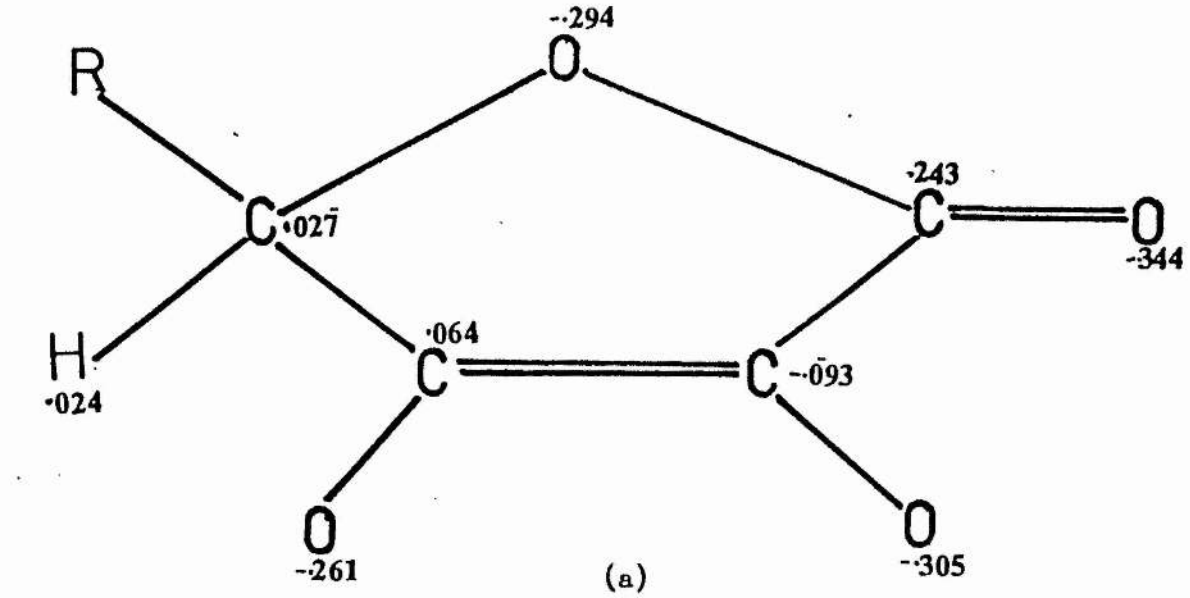


(a)

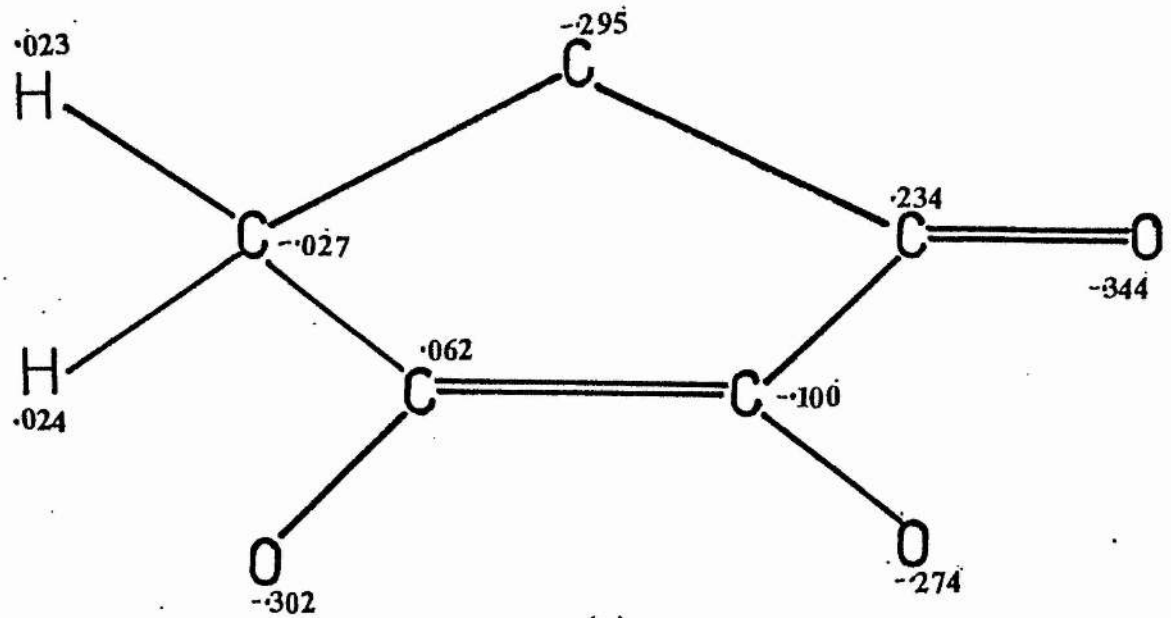


(b)

Fig. 7.4 Charge distribution calculated with STO-3G basis set for (a) ascorbate anion (b) α-hydroxytetronate anion

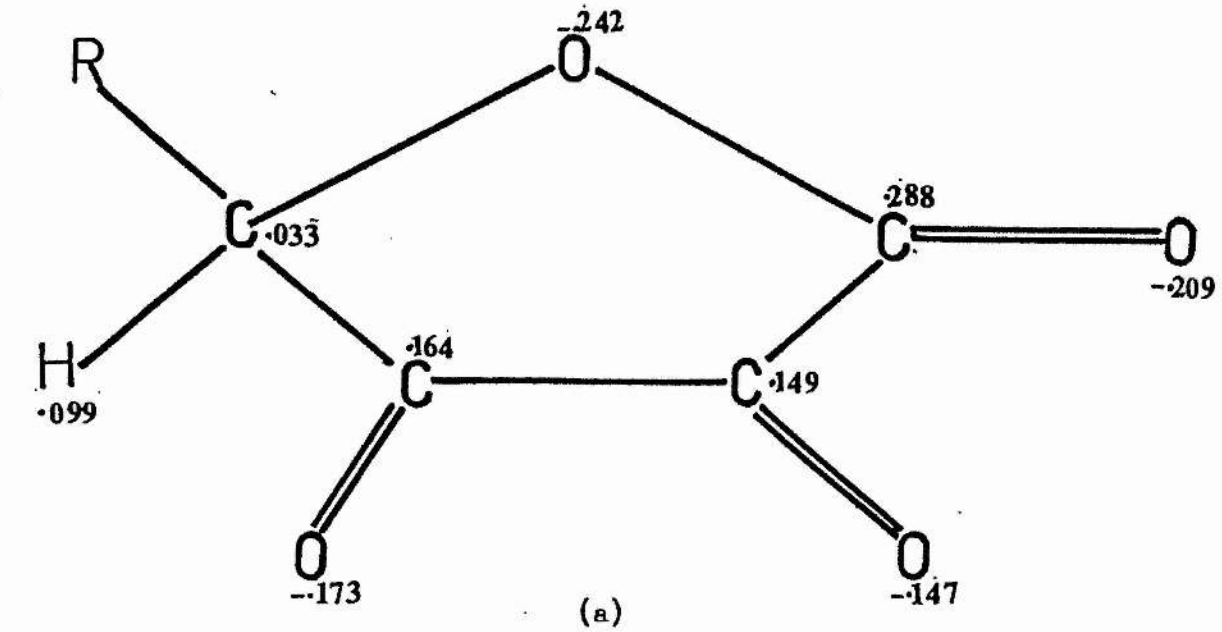


(a)

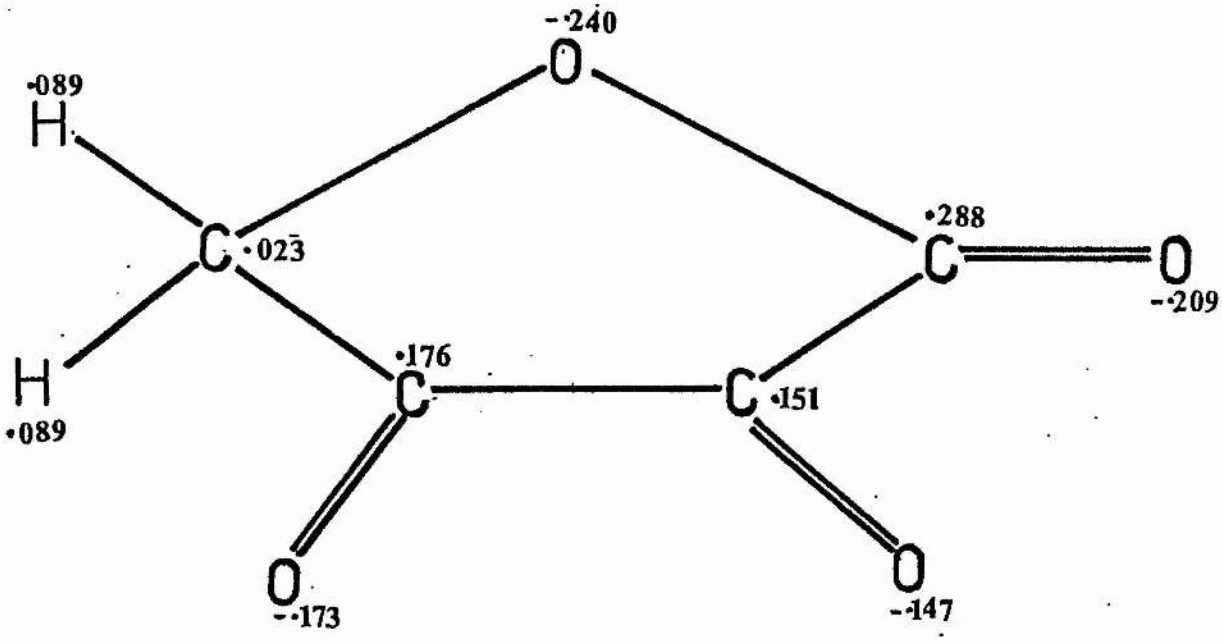


(b)

Fig. 7.5 Charge distribution calculated with STO-3G basis set for (a) ascorbyl radical (b) ∝-hydroxytetronate radical



(a)



(b)

Fig. 7.6 Charge distribution calculated with STO-3G basis set for (a) dehydroascorbic acid (b) furantrione (dehydro-α-hydroxytetronic acid)

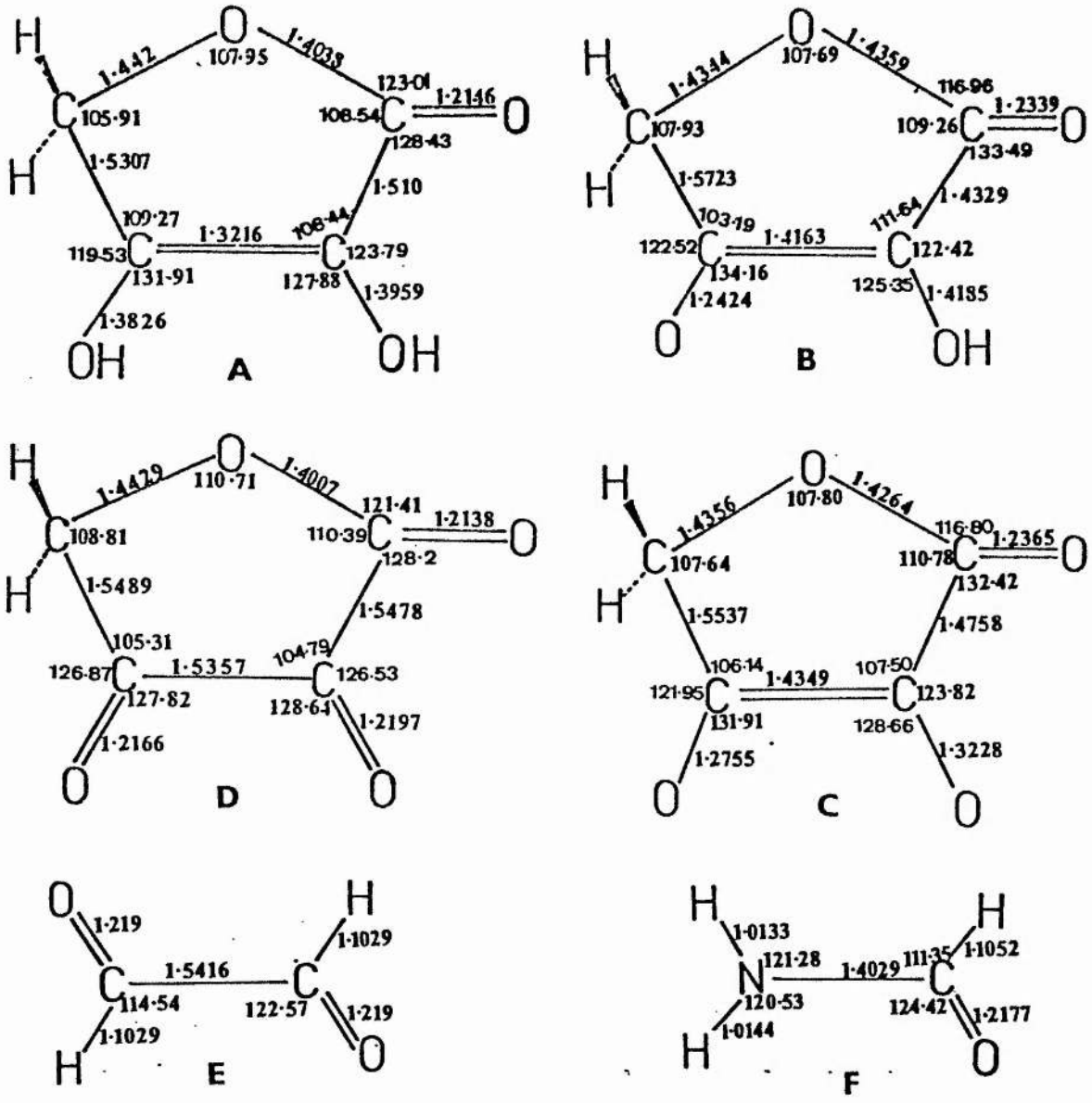
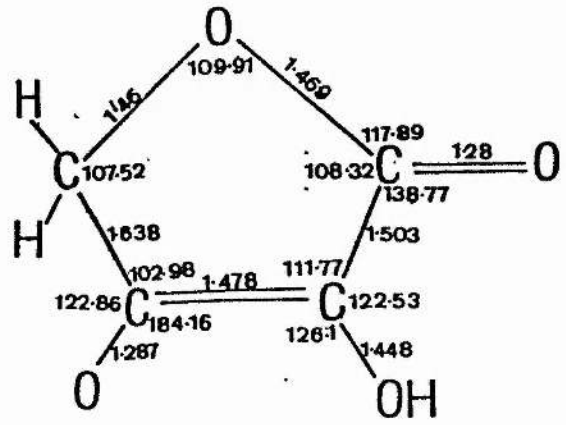
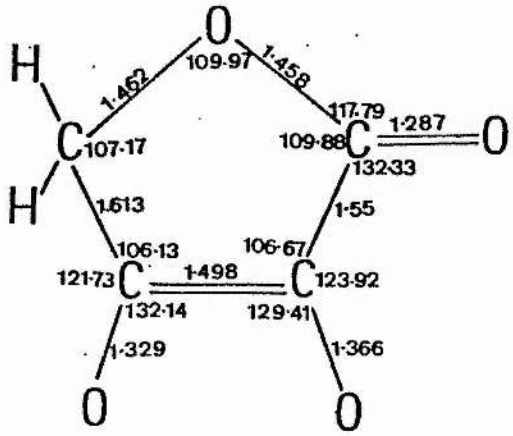


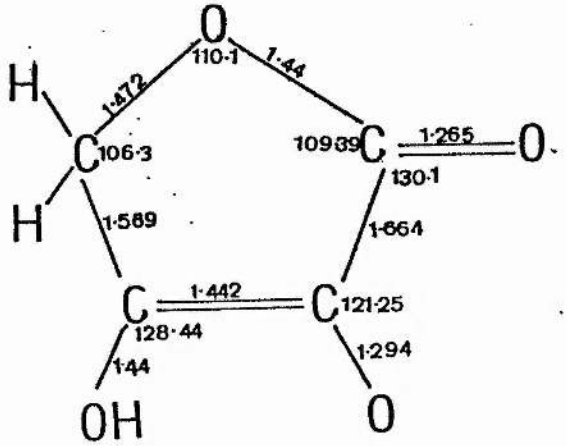
Fig. 7.7 Optimum geometries for the model compounds calculated with the STO-3G basis set; model compounds represent: (a) ascorbic acid (b) ascorbate anion (c)ascorbyl radical (d) dehydroascorbic acid; (e) is glyoxal (f) formamide



a

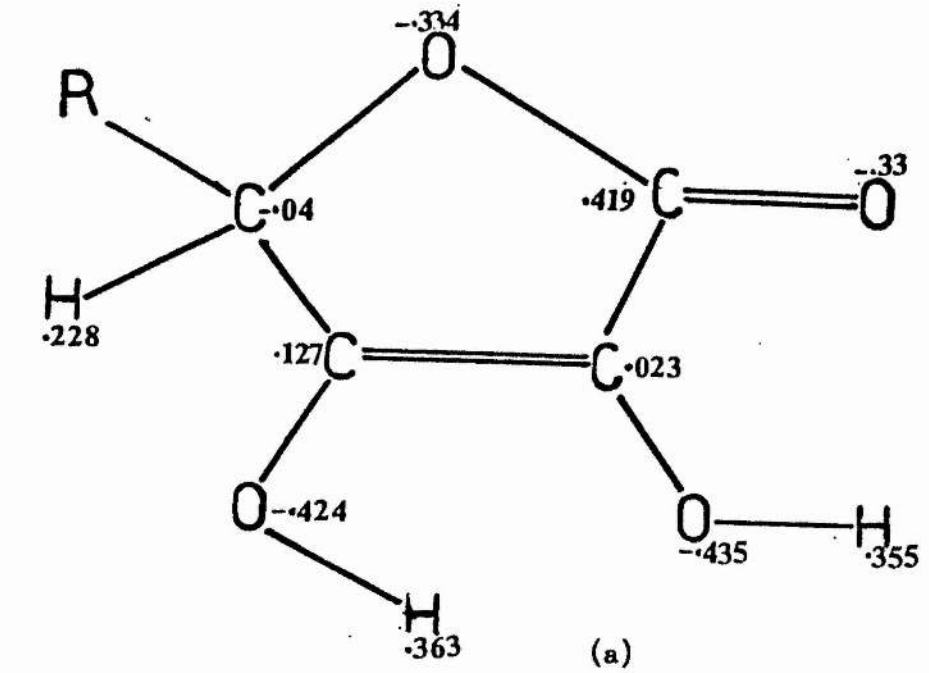


b

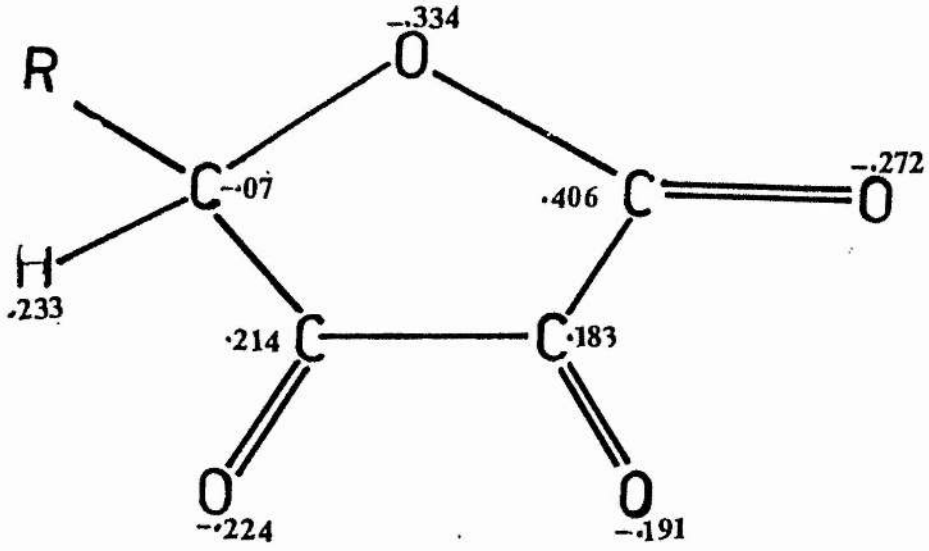


c

Fig. 7.8 Optimum geometries calculated with 6s3p basis set for model compounds representing (a) ascorbate anion (b) ascorbyl radical (c)alternative ascorbate anion



(a)



(b)

Fig. 7.9 Charge distribution calculated with 6s3p basis set for (a) ascorbic acid (b) dehydroascorbic acid

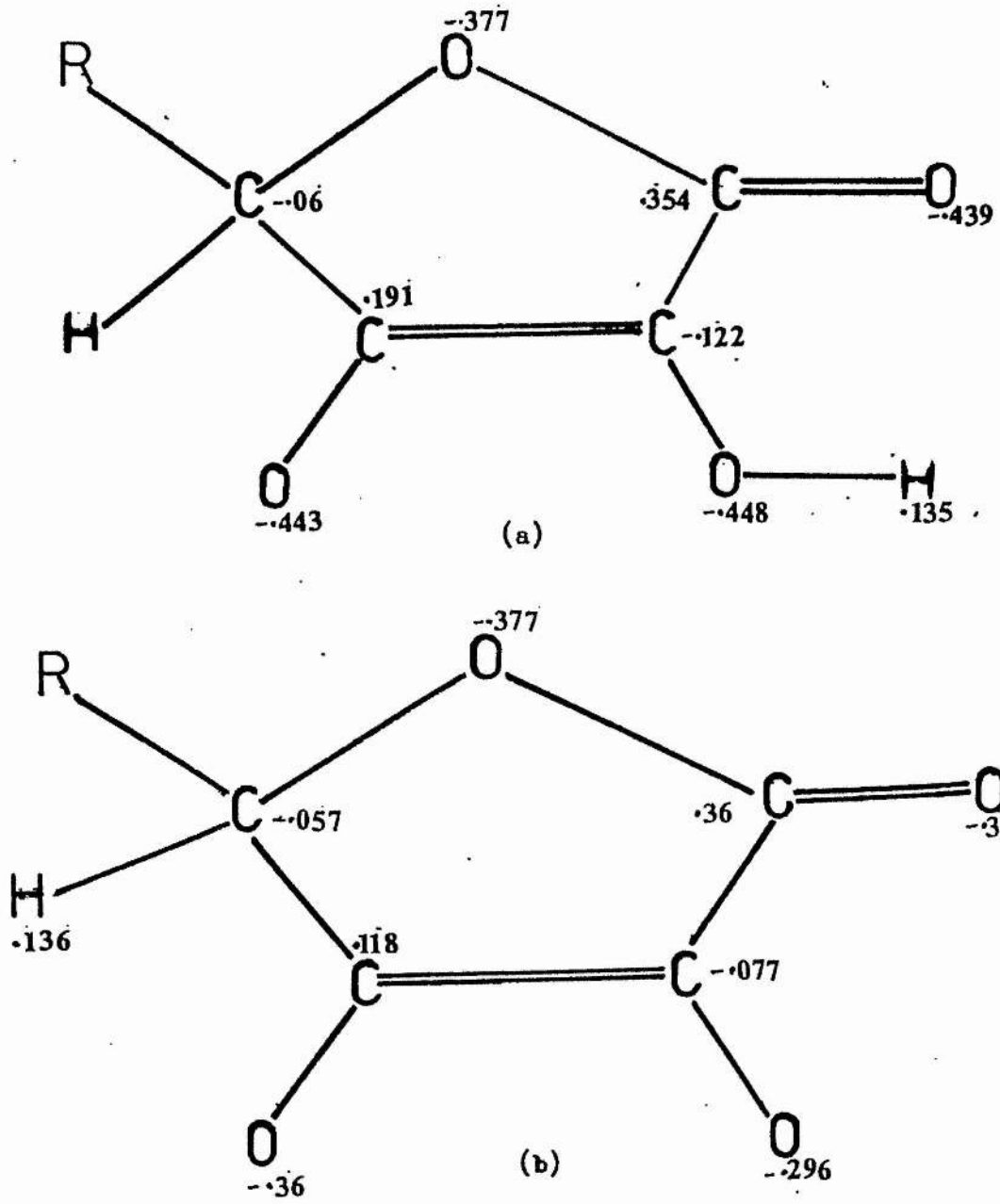
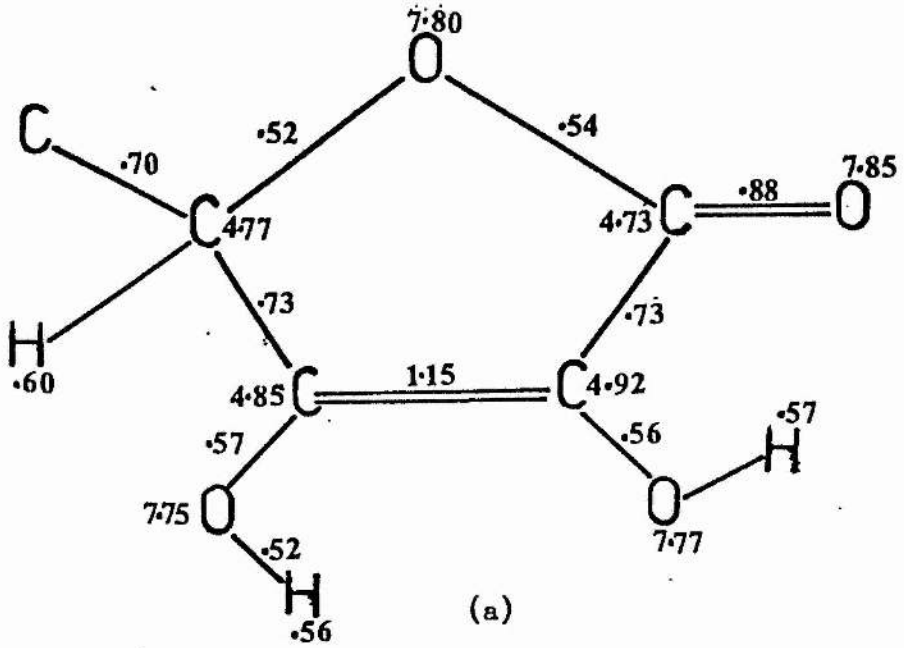


Fig. 7.10 Charge distribution calculated with 6s3p basis set for (a) ascorbate anion (b) ascorbyl radical



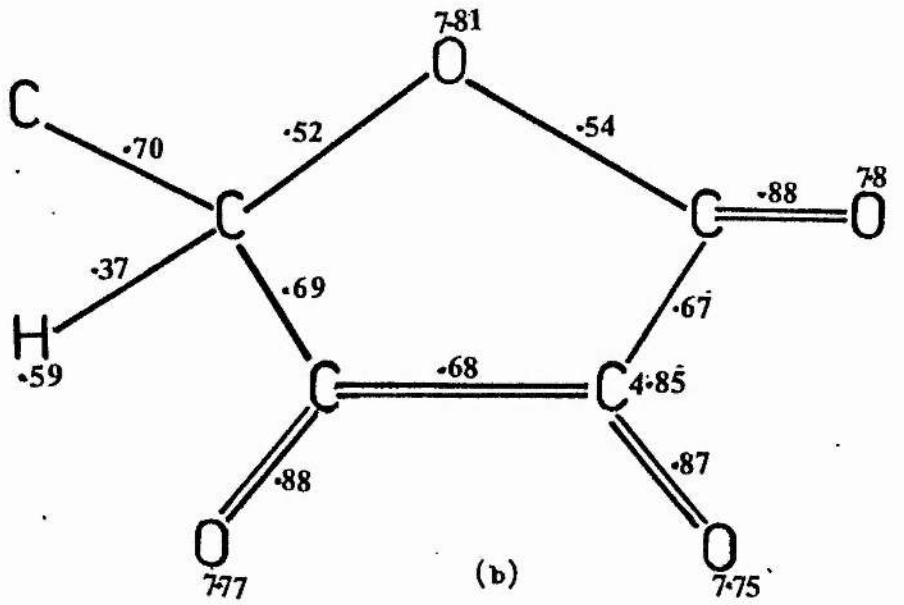


Fig. 7.11 Overlap populations calculated with STO-3G basis set for

(a) ascorbic acid (b) dehydroascorbic acid

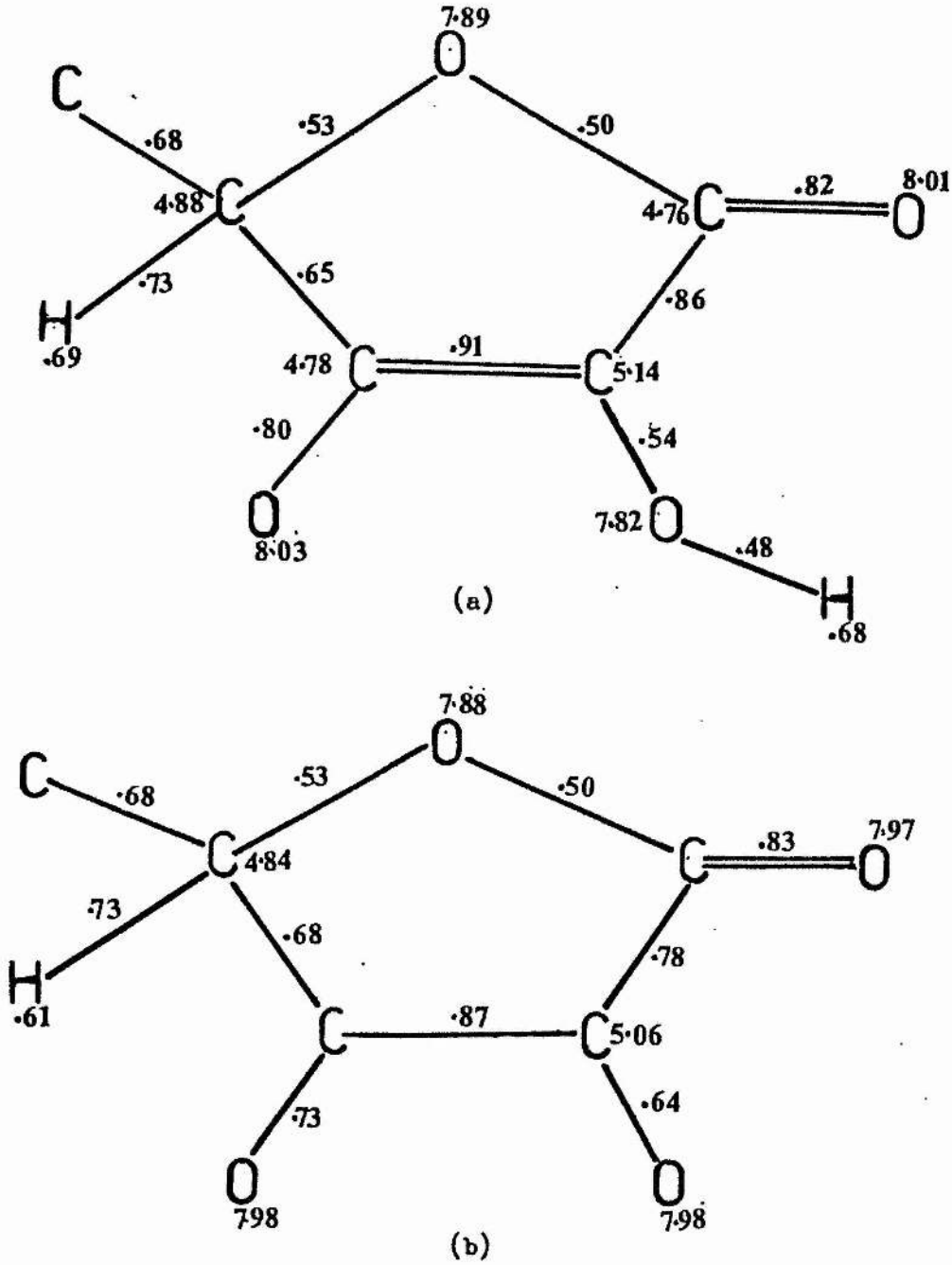


Fig. 7.12 Overlap populations calculated with STO-3G basis set for (a) ascorbate anion (b) ascorbyl radical

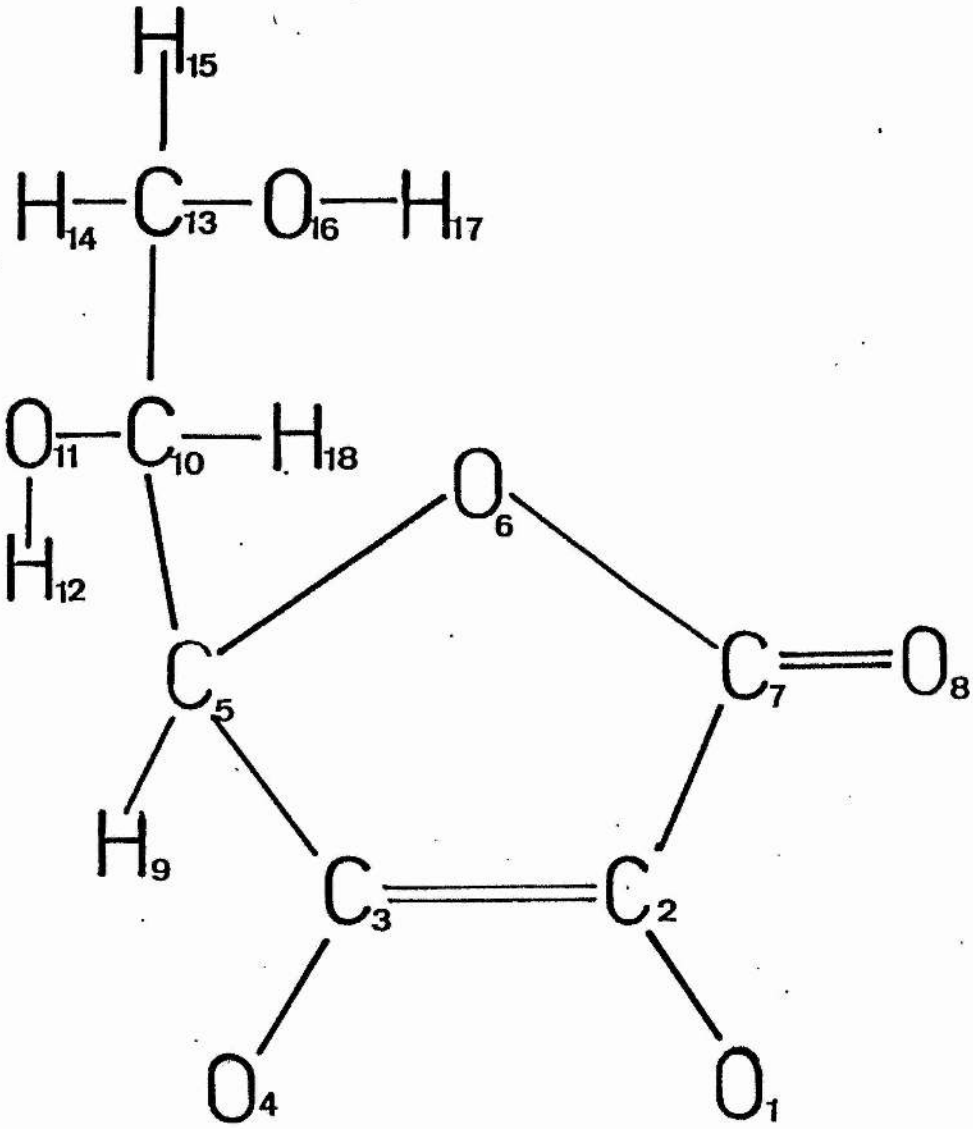


Fig. 7.13 Numbering scheme used for ascorbyl radical

Table 7.1

Energies of the isolated model molecules

Molecule	total energy[a]	E(HOMO)[a]	E(LUMO)[a]
Glyoxal	-223.58239	-0.3277	+0.1971
Formamide	-166.68822	-0.3082	+0.3273
Model ascorbic acid	-447.2879	-0.2633	+0.2352
Model ascorbate anion	-446.5847	+0.0378	+0.4960
Model ascorbyl radical	-446.03233	α +0.0297 β +0.0011	α +0.4738 β +0.5144
Model dehydro-ascorbic acid	-446.1066	-0.3036	+0.1494

a) energies in Hartrees

Table 7.2

Energies and orbitals for complete molecules

A with STO-3G, B with 6s3p basis set

Molecule	total energy[a]	E(HOMO-3)	E(HOMO-2)	E(HOMO-1)	E(HOMO)	E(LUMO)
A						
ascorbic	-672.09848	-0.3586	-0.3422	-0.3373	-0.2650	+0.2325
		σ/π sm[b]	σ/π rm	σ/π rm	π ro	π ro
anion	-671.38002	-0.133	-0.1174	-0.0603	+0.026	+0.4787
		π rm	σ ro	σ ro	π ro	π ro
radical	-670.83191 α	-0.1587	-0.1504	-0.0732	+0.02	+0.4556
		σ/π rm	σ/π ro	σ ro	π ro	π ro
	β	-0.1336	-0.1229	-0.0507	-0.0151	+0.4992
		σ/π ro	π rm	σ ro	π ro	π ro
dehydro-ascorbic	-670.91559	-0.3864	-0.3662	-0.346	-0.3032	+0.1469
		σ/π s+r	σ/π sm	π rm	σ ro	π ro
B						
ascorbic	-677.56372	-0.4721	-0.461	-0.4554	-0.3928	+0.0797
		σ/π rm	σ/π sm	σ/π rm	π ro	π ro
anion	-676.98808	-0.2669	-0.2602	-0.2014	-0.1231	+0.3178
		σ/π rm	σ/π rm	σ ro	π ro	π ro
radical	-676.44009 α	-0.2965	-0.2838	-0.2148	-0.1347	+0.2865
		σ ro	π rm	σ ro	π ro	π ro
	β	-0.2787	-0.2625	-0.1938	-0.1784	+0.3332
		σ ro	π ro	σ ro	π ro	π ro
dehydro-ascorbic	-676.35642	-0.4998	-0.4736	-0.4644	-0.4401	-0.0127
		σ/π sm	σ/π s+r	π rm	σ ro	π ro

(a) All energies are quoted in Hartrees

(b) For description of orbitals see text

Table 7.3

Calculated INDO hyperfine coupling constants for model radical at various geometries

hfcc[a,b]	neutral X-ray	anion X-ray	optimum STO-3G	optimum 6s3p	experimental[c] hfcc
a_2	-5.29	-3.52	-2.37	-2.52	1.03
a_3	-3.51	-2.50	-5.06	-8.21	3.65
a_5	-0.06	-0.24	0.38	1.35	2.85
a_7	-0.43	-1.26	-2.27	-3.28	5.72
a_1	11.56	11.83	11.52	10.58	-
a_4	5.80	3.39	5.73	6.65	-
a_6	-0.09	-0.30	-0.22	-0.27	-
a_8	0.29	1.12	1.65	1.75	-
a_9	-2.62	-1.57	-3.81	-6.37	2.32
a_{10}	-2.59	-0.72	-2.52	-4.66	2.32

a) a_i isotropic hyperfine coupling constant for nucleus i in Gauss

b) for numbering scheme see fig. 7.13

c) values from reference [214]

Table 7.4

Calculated INDO hyperfine coupling constants for ascorbyl radical at various geometries

hfcc[a,b]	neutral[d] X-ray	optimum STO-3G	optimum 6s3p	experimental[c] hfcc
a_2	-5.39	-2.65	-2.92	0.96
a_3	-2.85	-4.68	-7.54	3.62
a_5	-0.37	0.07	0.78	2.78
a_7	-0.46	-2.23	-3.24	5.74
a_{10}	-0.63	-0.52	-0.76	-
a_{13}	-0.05	0.06	0.08	-
a_1		11.63	10.69	-
a_4		5.57	6.44	-
a_6		-0.24	-0.26	-
a_7		1.65	1.76	-
a_{11}		-0.08	-0.08	-
a_{16}		0.0	-0.01	-
a_9	-2.08	-3.81	-6.90	1.76
a_{12}		-0.07	-0.07	-
a_{14}	0.01	-0.01	-0.02	0.19
a_{15}	-0.04	-0.04	-0.06	0.19
a_{17}		-0.01	-0.01	-
a_{18}	0.05	0.12	0.24	0.07

a),b),c) see footnotes to Table 7.3

d) values from reference [229]

Table 7.5

Spin densities calculated for the model radical at various geometries

atom	neutral X-ray		anion X-ray		optimum STO-3G	
	a	s	a	s	a	s
C_2	-0.24	-0.007	-0.18	-0.004	-0.12	-0.003
C_3	-0.37	-0.004	-0.32	-0.003	-0.41	-0.006
C_5	0.04	0.0	0.03	0.0	0.04	0.001
C_7	0.12	0.001	-0.03	-0.002	-0.22	-0.003
O_1	0.87	0.013	0.89	0.013	0.81	0.013
O_4	0.72	0.007	0.48	0.004	0.63	0.006
O_6	0.01	0.0	-0.01	0.0	-0.02	0.0
O_8	-0.13	0.0	-0.12	0.001	0.31	0.002
H_9	-0.01	-0.001	-0.01	-0.03	-0.01	-0.007
H_{10}	-0.01	-0.001	-0.01	-0.001	-0.01	-0.005

a atomic spin densities from ab initio UHF calculations with STO-3G basis set

s s-orbital spin densities from INDO calculations

8 CHARGE TRANSFER INTERACTIONS

8.1 Previous calculations

As mentioned in the introduction the property of ascorbic acid that is of particular interest in this work is its ability to enhance charge transfer from proteins to methylglyoxal. Preliminary theoretical studies to investigate this property have been carried by J.R.Ball and C.Thomson [161]. These authors used ∝-hydroxytetronic acid as a model for neutral ascorbic acid and investigated its donor/acceptor properties with respect to formamide and glyoxal by performing ab initio calculations with an STO-3G basis set. (Glyoxal is used as a model for methylglyoxal since the two have been found to possess similar charge transfer properties [234,235].) Four different arrangements were considered and it was found that the molecule can accept charge from formamide and behave as either a donor or acceptor towards glyoxal depending on the relative positions of the two molecules. More recently the anion radical has been found to behave as a donor towards glyoxal [231].

Otto, Ladik and Szent-Gyorgyi [230] have investigated the acceptor properties of ascorbic acid, ∝-hydroxytetronic acid and the cations derived from them. The results of ab initio calculations with a minimal STO-3G basis set are reported for these molecules and for various adducts formed between methylglyoxal, ∝-hydroxytetronic acid and methylamine. The charge distribution and orbital energies are presented and discussed and the electron acceptor properties of the molecules are inferred from the latter.

In this work an extensive series of calculations have been performed to investigate the interactions of model compounds representing ascorbic acid, its anion, anion radical and dehydroascorbic acid with glyoxal and formamide. The supermolecule approach has been used (see section 2.3.7) and the majority of calculations were performed using the STO-3G basis set.

For each ascorbic model compound a series of stacked and planar arrangements with glyoxal and with formamide were considered. In the stacked arrangements the formamide (or glyoxal) was placed in a plane parallel to the ring of the model compound; typical arrangements are shown in fig. 8.1a,b. Typical in-plane arrangements are depicted in fig. 8.1c,d. The relative orientation of the molecules was varied although no systematic optimization of the supermolecule was carried out, the previously determined optimum STO-3G geometries being used throughout.

For each arrangement studied the total energy and interaction energy, ΔE, were calculated (see equation 24, section 2.3.7). The extent of charge transfer from donor to acceptor was determined by comparison of the total charge on the constituent fragments in the supermolecule with those on the non-interacting isolated molecules. The charge densities were computed using the Mulliken population analysis (see section 2.3.4).

In order to investigate the influence of the basis set on the results a series of calculations was performed for some small model systems using a variety of basis sets and including some configuration interaction. A number of the supermolecule calculations were repeated using the 6s3p basis set of Mezey and Csizmadia [71], with a minimal contraction scheme. Finally a few calculations were performed on

"trimers" consisting of a model compound with both glyoxal and formamide. The results are presented in the sections which follow.

8.2 Supermolecule calculations with the STO-3G basis set

8.2.1 Stacked arrangements with glyoxal

Approximately 160 different stacked arrangements of the model compounds with glyoxal were investigated. In a typical arrangement (fig. 8.1a) the C-C bond of glyoxal lies below the C_2-C_7 bond of the ring, the distance between the two planes being R. Some arrangements were also studied in which the C-C bond of glyoxal was placed below and parallel to the C_3-C_2 bond in the ring and in which the glyoxal molecule was reversed. Various non-parallel arrangements of the C-C bonds were also considered, and in all cases the interplane distance R was varied. A representative selection of the results is given in Table 8.1.

The model compounds for ascorbic acid and dehydroascorbic acid form no stable stacked arrangements with glyoxal. ΔE increases with decreasing R and at low values of R (~2.6 Å) both molecules act as donors towards glyoxal. At larger values of R the magnitude and direction of charge transfer is dependent on the relative orientation of the molecules but in most cases it is so small as to be neglible.

In contrast the model ascorbate anion forms a large number of stable arrangements with values of ΔE between -0.002 and -0.0048 a.u. The anion acts as a donor to glyoxal with the transfer of up to 0.051e. It was not possible to perform an extensive series of calculations on the model anion radical because of convergence

difficulties. However, a number of calculations were performed successfully and the results indicate that the anion radical can also form stable arrangements with glyoxal and act as a donor to it, transferring up to 0.016e.

8.2.2 Planar arrangements with glyoxal

Two planar hydrogen-bonded arrangements are possible for the ascorbic acid model. One is depicted in fig. 8.1c, the second is similar but with bonding to H_{11} instead of H_{12}. Both arrangements were studied for various values of R and α. The results for the optimum values of these parameters are shown in Table 8.2. It can be seen that the ascorbic acid model acts as an acceptor to glyoxal.

At the optimum geometry of the ascorbate anion model the remaining hydroxyl hydrogen lies out of the plane and planar hydrogen-bonded arrangements are thus not possible. Positive interaction energies were found for all coplanar arrangements of the two molecules and attempts to locate a stable non-planar hydrogen-bonded complex were unsuccessful.

The model compounds for the anion radical and dehydroascorbic acid do not possess any hydroxyl groups and are therefore unable to form hydrogen-bonded arrangements with glyoxal. Coplanar arrangements of these molecules with glyoxal were found to be unstable.

8.2.3 Stacked arrangements with formamide

A series of stacked arrangements of the model compounds with formamide were investigated, the arrangements studied being similar to those described above for glyoxal. Stable arrangements were found for all four model compounds and a selection of typical results is shown

in Table 8.3.

It can be seen that in the arrangements having the lowest interaction energies, all the models act as acceptors towards formamide. In general the acceptor properties of the models increase at smaller interplane distances, but the ascorbate anion and anion radical models can act as donors in some arrangements. Although the amounts of charge transferred are small they are in general at least an order of magnitude greater than those calculated for similar methylglyoxal/formamide supermolecules [161,162].

8.2.4 Planar arrangements with formamide

No planar arrangements with formamide were investigated for the model anion radical. The other three model compounds were found to form a number of stable hydrogen-bonded in-plane arrangements with formamide. A typical example is shown in fig.8.1d. R and $\propto$ were both varied and bifurcated arrangements in which both hydrogens of the NH_2 group participate in bonding were also considered. A selection of results is presented in Table 8.4. It can be seen that all three molecules act as donors to formamide, donating up to 0.03e.

Formamide is the model for a peptide in these calculations and it is doubtful whether, in a protein molecule, ascorbic acid would be able to approach sufficiently closely to form a bond with the hydrogen of the peptide group. However, the results do suggest that ascorbic acid, or one of its metabolites might complex with a side-chain NH_2 group in the protein, should one be available.

In the optimum arrangements the planar complexes with formamide are calculated to be between 2 and 10 times more stable than the stacked arrangements. The complexes for the anion are particularly

stable. It may be that this hydrogen-bonding between ascorbate and -NH is one means by which ascorbic acid is incorporated into proteins. A recent study by Molloy and Wilson [236] demonstrates that ascorbic acid has two or more different binding sites in the protein bovine serum albumin (BSA). In one group of sites the ascorbic acid is weakly bound while in the other the binding is relatively strong. Although the energy differences between the planar and stacked arrangements reported here are not as great as the experimental values for the differences in binding strengths, it could be that the strongly bound sites involve complexing with NH_2 groups while the weakly bound sites correspond to the stacked arrangements. Such suggestions are, of course, speculative and further investigation is required before reaching any conclusions.

8.3 The effect of basis set size

The shortcomings of minimal basis sets in the study of intermolecular interactions have been discussed above (section 2.3.7). To check that the use of such basis sets can provide a reliable qualitative indication of the direction and extent of charge transfer some small model systems were studied with a variety of basis sets. The supermolecules studied were as follows:

1. planar NH_3 with CO
2. HCN with CO
3. H_2CO (formaldehyde) with planar NH_3
4. $H_2CO^{\mp}$ with planar NH_3
5. cis-glyoxal with formamide

In the first supermolecule CO was placed below NH_3 with the C-O bond parallel to one of the NH bonds, the C atom being below N at a distance of 3Å. An arrangement with the C-O bond at right angles to one of the NH bonds was also investigated. A selection of the basis sets internal to the Gaussian 80 program package were used to calculate the total energy, interaction energy and net charge on CO in the supermolecule. For one arrangement calculations were also performed using Moller-Plesset perturbation theory to second and third order [32] and configuration interaction with all double substitutions [34]. The results are shown in Table 8.5.

For the HCN/CO supermolecule the intermolecule distance was fixed at 3 Å with the C-O bond below the N-C bond. Again a series of internal basis sets were used. The results are shown in Table 8.6.

For supermolecules 3 and 4 formaldehyde (or its anion radical) was placed in a plane 3 Å below NH_3 with C below N and the C-O bond either parallel ($\theta = 0^o$) or at right angles ($\theta = 90^o$) to one of the N-H bonds. The results are shown in Tables 8.7 and 8.8.

For supermolecule 5 cis-glyoxal was given the geometry of the $O_1-C_2-C_3-O_4$ part of the dehydroascorbic acid model. The N-C bond of formamide was placed below the C-C bond of glyoxal with N below C_2, the interplane distance being 3.4 Å. For this supermolecule some of the basis sets of Huzinaga et al. [72-75] were used. The results are given in Table 8.9.

An inspection of the results shows that increasing the basis set size usually increases the binding interaction energy and the amount of charge transferred. The maximum interaction energy and charge transfer is usually obtained for split-valence basis sets such as 3-21G and 6-21G. In some cases the addition of polarization functions

decreases the binding interaction energy but this does not appear to be a general trend. The application of perturbation theory decreases the binding interaction energy for the STO-3G basis set but increases it for the 3-21G and 6-31G* basis sets. The application of configuration interaction has a similar but less marked effect. In all cases the trends shown by the STO-3G basis sets appear to be reliable. It also appears that in comparison with larger basis sets, STO-3G tends to underestimate the amount of charge transferred and give values for ΔE that are too small.

8.4 Calculations with the 6s3p basis set

The 6s3p basis of Mezey and Csizmadia [71] with a minimal contraction scheme was found to give better values for the total energy and orbital energies of the ascorbic acid model compounds. A selection of the supermolecule calculations reported in section 8.2 were repeated using this basis set. The results are presented in Tables 8.10 and 8.11. In every case the results obtained with this basis set parallel those obtained with the STO-3G basis. The total energy is lower while the binding interaction energy and amount of charge transferred are, in general, higher.

8.5 Supermolecule trimers

The results of the supermolecule calculations reported in sections 8.2 and 8.4 show that models of the ascorbate anion and anion

radical can act as electron donors to glyoxal. The neutral molecule, on the other hand, can act as an electron acceptor towards glyoxal through the formation of a hydrogen-bonded dimer. All four model compounds can accept electronic charge from formamide.

Szent-Gyorgyi has proposed that ascorbic acid acts to enhance the charge transfer between protein and methylglyoxal through the formation of a cation radical. Such a species would of course be a good electron acceptor, however, all the evidence from e.s.r. studies on ascorbic acid implies that it is the anion radical which is the stable species. This radical is so stable that some have argued that it is the neutral acid or its anion that is the physiologically active form. The results presented in the above sections suggest two ways in which ascorbic acid might act to enhance charge transfer without forming a cation radical.

The formation of a hydrogen-bonded complex between neutral ascorbic acid and methylglyoxal should enhance the acceptor properties of the latter with respect to formamide. Alternatively, if it is the anion or anion radical which is active, a triple "stacked" arrangement might form with the ascorbate anion in the centre, glyoxal above and formamide below. Since both the anion and anion radical can donate charge to glyoxal and accept charge from formamide this would enhance the overall charge transfer from formamide to glyoxal (i.e. from protein to methylglyoxal).

The 6s3p basis set was found to be sufficiently economical in its integral file space requirements to allow a limited investigation of these "trimer" systems to be carried out. Unfortunately the Gaussian 80 program cannot handle open shell systems of this size so the arrangement with the anion radical could not be studied. The

arrangements with the neutral molecule and with the anion were studied as follows. In the first, the optimum planar hydrogen-bonded arrangement of ascorbic acid with glyoxal was used and formamide was added below the glyoxal in the position calculated to be the most stable for methylglyoxal/formamide [162]. For the anion, formamide was placed below the plane of the ring and glyoxal above in positions found to be favourable in the dimer calculations.

The results are given in Table 8.12 and show that the presence of ascorbic acid in either its neutral or anionic form (and by extrapolation its anion radical form) can increase the net charge on formamide by a factor of 6 compared to that found in the glyoxal/formamide dimer. More extensive calculations are needed before reaching any conclusions as to which arrangement might be preferred. In the triple stacked arrangement the position of the side-chain may be important. Also no consideration of environmental effects has been included and extrapolation from these simple calculations to semi-conduction in proteins is highly speculative. The calculations do show, however, that enhancement of charge transfer in the glyoxal/ formamide system by ascorbic acid is theoretically possible and can take place without the formation of a cation radical.

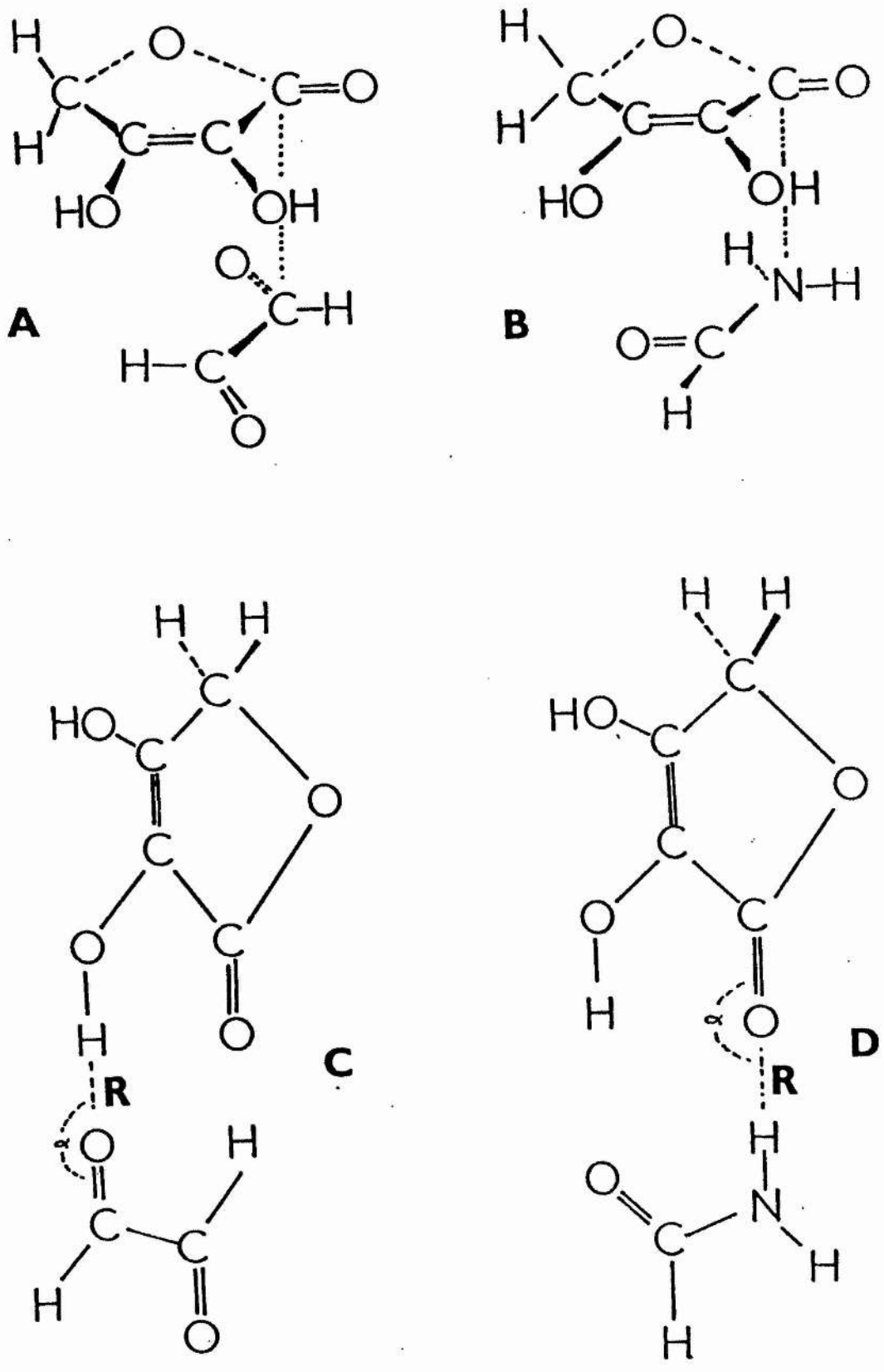


Fig. 8.1 Examples of typical supermolecules studied in the charge transfer calculations (see text for further details)

Table 8.1

Results of "supermolecule" calculations for stacked arrangements with glyoxal

Model compound representing:	Position of C-C bond of glyoxal	R(Å)	total energy[a]	ΔE[a]	net charge on glyoxal[b]
Ascorbic acid	below C_2-C_7	2.6	-670.84677	+0.02350	-0.0163
		3.0	-670.86476	+0.00551	-0.0034
	below C_2-C_3*	3.6	-670.86950	+0.00077	+0.0001
Dehydro-ascorbic acid	below C_2-C_7	2.6	-669.66199	+0.02704	-0.0009
	below C_2-C_3	2.8	-669.67280	+0.01623	+0.0018
		3.2	-669.68369	+0.00534	+0.0006
Ascorbate anion	below C_2-C_7	2.8	-670.17029	-0.00320	-0.0512
		3.0	-670.17186	-0.00477	-0.0244
	below C_2-C_3*	3.4	-670.17137	-0.00428	-0.0044
Ascorbyl radical	below C_2-C_3	3.2	-669.61739	-0.00267	-0.0089
	below C_2-C_3*	3.4	-669.61786	-0.00314	-0.0037
	below C_2-C_7	3.2	-669.61730	-0.00258	-0.0074

* Indicates glyoxal reversed

[a] Energies in Hartrees

[b] Units of electronic charge

Table 8.2

Results for optimum planar arrangements of ascorbic acid with glyoxal

Bonding to:	R(Å)	α	total energy[a]	ΔE^a	net charge on glyoxal[b]
H_{11}	4.2	108°	-670.87545	-0.00518	+0.0156
H_{12}	4.0	108°	-670.87578	-0.00551	+0.0303

a,b see footnotes to Table 8.1

Table 8.3

Results of "supermolecule" calculations for stacked arrangements with formamide

Model compound representing:	Position of N–C bond of formamide	R(Å)	total energy[a]	ΔE[a]	net charge on formamide[b]
Ascorbic acid	below C_2–C_7 (N below C_2)	3.0	–613.97386	+0.00223	+0.0013
		3.6	–613.97697	–0.00089	+0.0002
	below C_2–C_3 (N below C_2)	3.6	–613.97646	–0.00037	+0.0006
Dehydro-ascorbic acid	below C_2–C_7 (N below C_2)	3.0	–612.79456	+0.00030	+0.0034
	below C_2–C_3 (N below C_2)	3.2	–612.79572	–0.00086	+0.0031
	below ring (M below C_2)	3.4	–612.79555	–0.00069	+0.0022
Ascorbate anion	below C_2–C_7 (N below C_2)	3.2	–613.27485	–0.00193	+0.0022
	*	3.2	–613.27377	–0.00085	+0.0013
	below C_2–C_3 (N below C_3)	3.0	–613.27528	–0.00236	+0.0005
Ascorbyl radical	below C_2–C_3 (N below C_2)	3.2	–612.71951	+0.00104	+0.0010
		3.4	–612.72102	–0.00047	+0.0007
	below C_2–C_7*	3.6	–612.72107	–0.00052	–0.0001

* Indicates formamide reversed

a,b see footnotes to Table 8.1

Table 8.4

Results for planar arrangements with formamide

Model representing:	bonding to:	R(Å)	α	total energy[a]	ΔE[a]	net charge on formamide
Ascorbic acid	O_8	4.2	125°	-613.98213	-0.00603	-0.0174
	O_8 B	3.2		-613.97808	-0.00200	-0.0010
Ascorbate anion	O_8	3.8	133°	-613.28915	-0.01723	-0.0770
	O_4	3.8	133°	-613.29157	-0.01865	-0.0802
	O_8 B	3.0		-613.28292	-0.01000	-0.0030
	O_4 B	3.0		-613.28304	-0.01012	-0.0031
Dehydro-ascorbic acid	O_8	4.2	120°	-612.79870	-0.00384	-0.0157
	O_1	4.2	125°	-612.79725	-0.00239	-0.0155
	O_4	4.2	133°	-612.79837	-0.00351	-0.0157
	O_8 B	3.2		-612.79674	-0.00188	-0.0010
	O_1 B	3.2		-612.79619	-0.00133	-0.0009
	O_4 B	3.2		-612.79607	-0.00122	-0.0010

B indicates bifurcated bonding involving both hydrogens of formamide

a,b see footnotes to Table 8.1

Table 8.5

Results for NH_3/CO with various basis sets

Position of C–O bond	type of calculation	basis set	total energy[a]	ΔE[a]	net charge on CO[b]
parallel to N–H	HF	STO-3G	-166.65991	-0.00080	-0.0034
	"	3-21G	-167.96564	-0.00331	-0.0179
	"	4-31G	-168.65892	-0.00108	-0.0111
	"	6-21G	-168.67573	-0.00319	-0.0179
	"	6-31G*	-168.91181	-0.00006	-0.0102
	MP2	STO-3G	-166.82606	-0.00045	-0.0034
	"	3-21G	-168.28045	-0.00374	-0.0179
	"	6-31G*	-169.36245	-0.00109	-0.0102
	MP3	STO-3G	-166.83376	-0.00049	-0.0034
	"	3-21G	-168.27437	-0.00390	-0.0179
	"	6-31G*	-169.37310	-0.00105	-0.0102
	CID	STO-3G	-166.84139	-0.00050	-0.0034
	"	3-21G	-168.28299	-0.00340	-0.0179
	"	6-31G*	-169.37973	-0.00007	-0.0102
at 90° to N–H	HF	STO-3G	-166.65994	-0.00083	-0.0034
		3-21G	-167.96570	-0.00337	-0.0178
		4-31G	-168.65895	-0.00111	-0.0111
		6-21G	-168.67579	-0.00325	-0.0178
		6-31G	-168.83190	-0.00048	-0.0107
		6-31G*	-168.91181	-0.00005	-0.0101

a,b see footnotes to Table 8.1

Table 8.6

Results for HCN/CO "supermolecule"

Basis set	total energy[a]	ΔE[a]	net charge on CO[b]
STO-3G	-202.89457	+0.00163	-0.0011
3-21G	-204.44602	+0.00065	-0.0021
6-21G	-205.33107	+0.00090	-0.0021
4-31G	-205.28199	+0.00161	-0.0019
6-31G	-205.49217	+0.00224	-0.0017
6-31G*	-205.61033	+0.00224	-0.0011

a,b see footnotes to Table 8.1

Table 8.7

Results for NH_3/H_2CO "supermolecule"

θ	basis set	total energy[a]	ΔE[a]	net charge on NH_3[b]
0^o	STO-3G	-167.79013	-0.00003	+0.0028
	3-21G	-169.09429	-0.00295	+0.0173
	6-21G	-169.80205	-0.00280	+0.0174
	4-31G	-169.79961	-0.00105	+0.0161
	6-31G	-169.97376	-0.00048	+0.0114
	6-31G*	-170.03960	-0.00054	+0.0111
90^o	STO-3G	-167.79013	-0.00004	+0.0028
	3-21G	-169.09422	-0.00288	+0.0177
	6-21G	-169.80198	-0.00273	+0.0177
	4-31G	-169.79957	-0.00101	+0.0121
	6-31G	-169.97373	-0.00045	+0.0116
	6-31G*	-170.03959	-0.00053	+0.0116

a,b see footnotes to Table 8.1

Table 8.8

Results for NH_3/H_2CO "supermolecule"

θ	basis set	total energy[a]	ΔE[a]	net charge on NH_3[b]
0°	STO-3G	-167.54842	+0.00978	-0.0005
	3-21G	-168.98908	+0.01276	-0.0047
	6-21G	-169.69570	+0.01288	-0.0049
	4-31G	-169.71149	+0.01282	-0.0109
	6-31G	-169.89077	+0.01371	-0.0115
	6-31G*	-169.93881	+0.01306	-0.0124
90°	STO-3G	-167.54835	+0.00985	-0.0003
	3-21G	-168.98869	+0.01315	-0.0033
	6-21G	-169.69531	+0.01327	-0.0034
	4-31G	-169.71115	+0.01316	-0.0093
	6-31G	-169.89046	+0.01403	-0.0099

a,b see footnotes to Table 8.1

Table 8.9

Results for cis-glyoxal/formamide "supermolecule"

Basis set	total energy[a]	ΔE[a]	net charge on formamide[b]
STO-3G	-390.26669	-0.00070	+0.0004
MINI-3	-394.01918	-0.00117	+0.0014
MIDI-3	-394.57610	-0.00516	+0.0072
SMIDI-3	-394.34501	-0.0018	+0.0033
SMIDI-3*	-394.62643	-0.00259	+0.0038

a,b see footnotes to Table 8.1

Table 8.10

Stacked arrangements with glyoxal, results for 6s3p basis set

Model compound representing:	position of C-C bond of glyoxal	R(Å)	total energy[a]	ΔE[a]	net charge on glyoxal[b]
Ascorbic acid	below C_2-C_3	3.0	-676.33683	+0.00661	-0.0008
	" *	3.0	-676.33576	+0.00768	-0.0038
	below C_2-C_7	3.0	-676.33404	+0.00940	-0.0018
	" *	3.0	-676.33630	+0.00714	-0.0047
Dehydro-ascorbic acid	below C_2-C_3	3.0	-675.12848	+0.01019	+0.0055
	" *	3.0	-675.12746	+0.01121	+0.0015
	below C_7-C_2	3.0	-675.12902	+0.00960	+0.0011
	"	3.0	-675.12865	+0.01003	+0.0003
Ascorbate anion	below C_2-C_7	2.8	-675.78834	-0.00566	-0.0668
	"	3.0	-675.78927	-0.00660	-0.0358
	"	3.2	-675.78888	-0.00620	-0.0186
	below C_2-C_3	3.0	-675.77849	+0.00419	-0.0234
	" *	3.0	-675.78836	-0.00569	-0.0167
Ascorbyl radical	below C_2-C_3	3.2	-675.23307	-0.00232	-0.0119
	"	3.4	-675.23397	-0.00322	-0.0062
	"	3.0	-675.23044	+0.00031	-0.0229
	below C_2-C_7*	3.0	-675.23042	+0.00033	-0.0198

a,b see footnotes to Table 8.1

* indicates glyoxal reversed

Table 8.11

Stacked arrangements with formamide, results with 6s3p basis set

Model compound representing:	position of N-C bond of formamide	R(Å)	total energy[a]	ΔE[a]	net charge on formamide[b]
Ascorbic acid	below C_2–C_7 (N below C_7)	3.6	-618.98624	-0.00089	+0.0003
	" *	3.6	-618.98581	-0.00046	0.00
	below C_3–C_2 (N below C_3)	3.6	-618.98477	-0.00013	+0.0003
Dehydro-ascorbic acid	below C_7–C_2 (N below C_7)	3.4	-617.82336	-0.00176	+0.0016
	below C_2–C_3 (N below C_2)	3.4	-617.78273	-0.00215	+0.0043
	below C_2–C_7* (N below C_2)	3.4	-617.78070	-0.00013	+0.0016
Ascorbate anion	below C_7–C_2* (N below C_7)	3.0	-618.42972	-0.00513	+0.0008
	"	3.2	-618.42957	-0.00498	+0.0019
	below C_2–C_3 (N below C_3)	3.0	-618.42872	-0.00413	+0.0008
	"	3.2	-618.42888	-0.00429	+0.0019
Ascorbyl radical	below C_2–C_3 (N below C_2)	3.4	-617.87373	-0.00107	+0.0013
	"	3.2	-617.87219	+0.00050	+0.0022

* indicates formamide reversed

a,b see footnotes to Table 8.1

Table 8.12

Results of calculations on supermolecule "trimers"

Complex[c]	total energy[a]	ΔE[a]	net charge on glyoxal[b]	net charge on formamide[b]
Model ascorbic + glyoxal (planar) +formamide (stacked)	-844.38843	-0.01215	+0.0410	+0.0006
Model anion + glyoxal + formamide (stacked)	-843.82028	-0.00480	-0.0223	+0.0029
glyoxal + formamide (stacked)	-393.42390	-0.00011	-0.0004	+0.0004

a,b see footnotes to Table 8.1

c for description of conformation see text

9. ELECTROSTATIC MOLECULAR POTENTIALS

Electrostatic molecular potential maps have been produced using the PUMABAL program (see section 2.3.8) for ascorbic acid, its anion, anion radical and dehydroascorbic acid, and for the corresponding model compounds. Maps drawn in the plane of the molecule and in several planes below it have been produced and a selection are presented here. In addition electrostatic potential maps have been produced for formamide and for glyoxal. A brief investigation of the change in electrostatic potential brought about by the formation of charge transfer complexes has also been made.

The maps for the model compounds and the corresponding complete molecules are shown in figs. 9.1-9.4. As in previous calculations the geometry of the side-chain for the two neutral molecules is taken from X-ray data on ascorbic acid and that for the two anionic species from similar data on the anion. It is difficult to represent the side-chain on the maps since the atoms do not lie in the plane of the ring. For the anionic species most of the atoms lie between 1 and 2 Å above the plane of the ring, while for the neutral molecules the second oxygen atom of the side-chain lies almost in the plane and gives rise to an additional attractive region. In the anion radical there is a small attractive region due to this same oxygen but here the atom lies ~2 Å above the plane and the effect is less noticeable. The disappearance of the attractive potential near O_6 in the anion radical (compared to the model) is presumably due to the CH_2 group of the side-chain which lies above the plane and fairly close to O_6. The decrease in potential at O_8 may also be due to this group. Apart from

these effects the presence of the side-chain does not appear to influence the electrostatic potential of the ring system in any of the four molecules.

In the anion the strong attractive potential obscures all other features and the electrostatic potential map is not very informative. The map for the anion radical shows a high negative potential at O_1 and O_4, with a less attractive region around O_8 and a relatively low attractive potential at O_6 which disappears when the side-chain is present. In dehydroascorbic acid attractive potentials are also found at these four oxygen atoms. However, the potentials are not so high and the values at O_1 and O_4 are slightly less than those at O_6 and O_8. In neutral ascorbic acid the highest negative potentials are again associated with O_6 and O_8. The attractive regions at O_1 and O_4 are smaller and there are areas of positive potential associated with the hydrogens of the OH groups.

Maps drawn in planes below the plane of the ring were produced for all the molecules. Those for the ascorbic acid model (fig.9.5) illustrate the typical trends. An inspection of the maps shows that at 2 Å below the ring there is little attractive potential except in the case of the anion. In the charge transfer complexes discussed in the preceding chapter the glyoxal or formamide is usually at a distance of at least 3 Å below the plane of the ring. From these maps it seems unlikely that an attractive potential from the ring would be experienced at this distance. However, an investigation of the change in electrostatic potential in and below the plane of the ring when glyoxal or formamide is present was carried out as follows.

Seven complexes, described in Table 9.1, were chosen for study. Maps were produced in the planes of the two molecules (figs.

9.9-9.12) and in planes between them. The maps for isolated glyoxal and formamide are shown in fig. 9.13..

For complexes 1 and 2 a slight decrease in the attractive potential of the model radical is observed, particularly at O_8, both in the plane of the ring and 1.1 Å below it. For complex 1 there is decrease of 20 kcal mol^{-1} in the attractive potential at the oxygen atom of formamide compared to the isolated molecule and the extent of the attractive potential is also reduced. In complex 2 an increased negative potential is observed at 2.1 Å below the plane of the ring and in the plane of glyoxal. The increase occurs below O_1 of the ring and has a marked effect on the potential around the oxygen of glyoxal which lies below this atom. An increase of 30 kcal mol^{-1} in the attractive potential of this oxygen is observed.

In complex 3 there is a slight decrease in the negative potential 2.1 Å below the plane but the most obvious effect is the large increase in negative potential around formamide in the plane of this molecule. A similar large increase in negative potential is observed for glyoxal in complex 4.

In complex 5 there is a slight decrease in negative potential in the plane of the ring and at 1.1 Å below it. In the plane of formamide there is a marked decrease in the potential at the oxygen atom (-60 to -40 kcal mol^{-1}) and the attractive area is less extensive. In complex 6 little effect is observed in the plane of formamide. There is a slight increase in the attractive potential around O_1 in the plane of the ring and at 1.1 Å below it. In complex 7 the negative potential at O_4 and O_8 has increased. The potential at O_1 is slightly distorted due to the presence of the glyoxal. The potential at the hydrogen-bonded oxygen of glyoxal is decreased but

the rest of the molecule appears to be unaffected.

These results indicate that changes in electrostatic potential do occur on formation of charge transfer complexes and, although the changes are small, they may be important. On the whole the potentials of formamide and glyoxal appear to be more affected than those of the model compounds. The changes are usually as one would expect: an increase in attractive potential when the molecule is behaving as an electron acceptor and a decrease when it is behaving as an electron donor.

Although the electrostatic molecular potential is strictly a measure of the interaction of a proton with the molecule it does also have some bearing on solvation [136]. Changes of 20 to 30 kcal mol^{-1}, such as those observed here, may well alter the pattern of hydration around a molecule. A study of hydration in the ascorbic acid system has not been possible in this thesis. The electrostatic potentials presented here suggest that there may be important differences in hydration patterns between the various metabolites of ascorbic acid. The results also suggest that complex formation may alter the hydration pattern around formamide and glyoxal.

NOTE

Figs. 9.1 to 9.13 are electrostatic molecular potential maps, drawn in the plane of the molecule unless otherwise stated.

The key to the maps is as follows:

symbol	potential value in kcal mol^{-1}
2	80
3	60
4	50
5	40
6	30
7	20
8	10
9	5
0	0
A	-5
B	-10
C	-20
D	-30
E	-40
G	-50
L	-60
M	-80

H indicates two hydrogens symmetrically above and below the plane

R represents the side-chain $CH(OH)CH_2OH$ (see text for details)

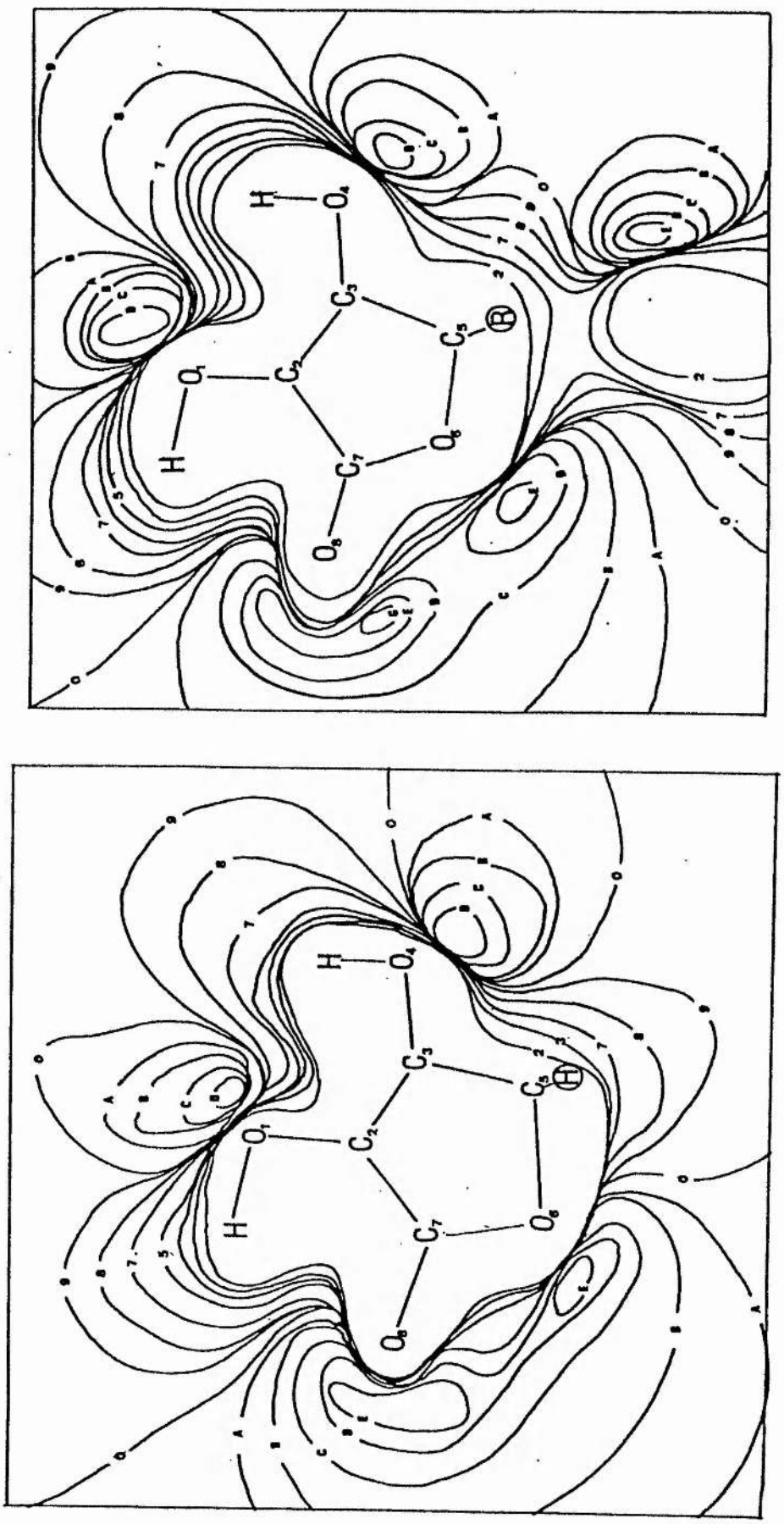
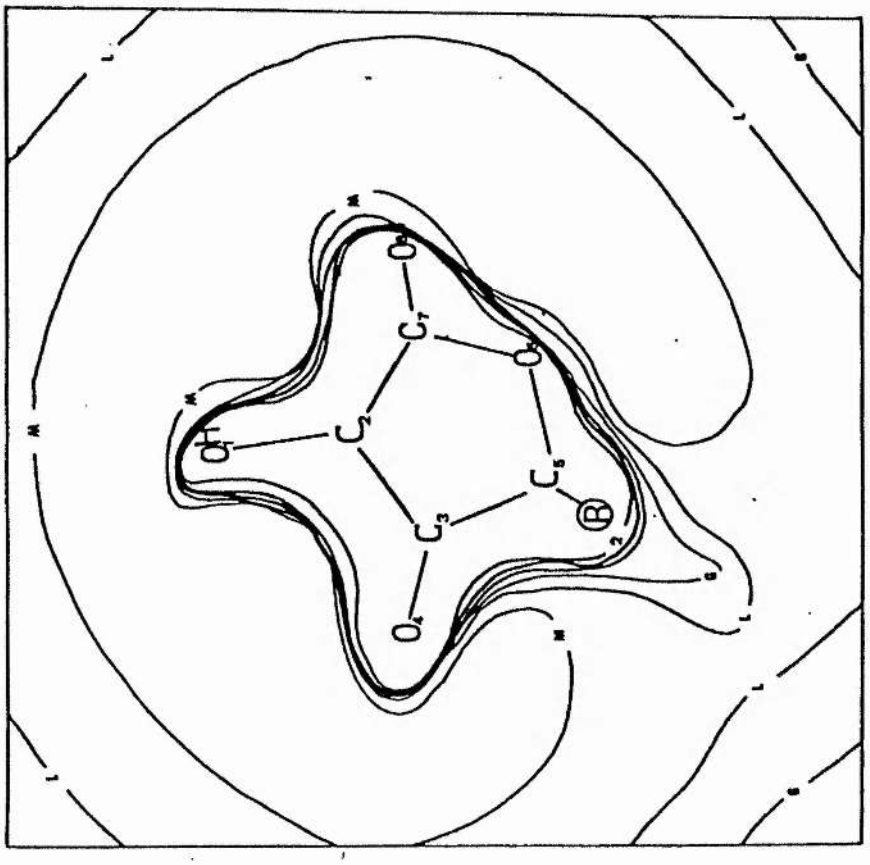


Fig. 9.1 Maps for ascorbic acid (a) model compound (b) complete molecule



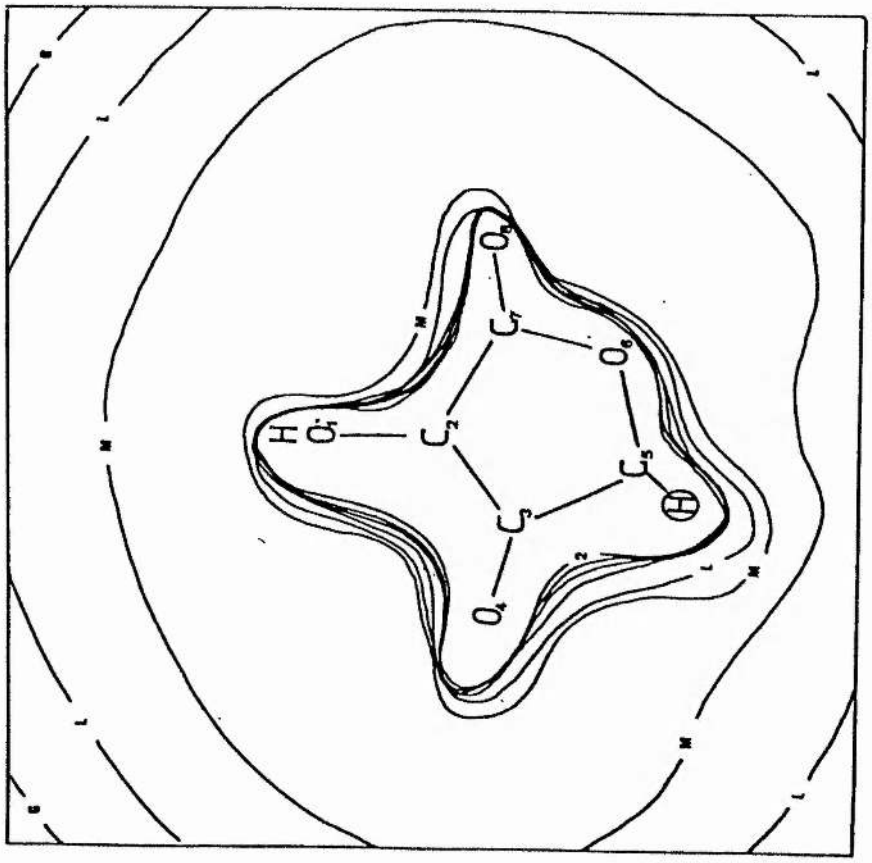


Fig. 9.2 Maps for ascorbate anion (a) model compound (b) complete molecule

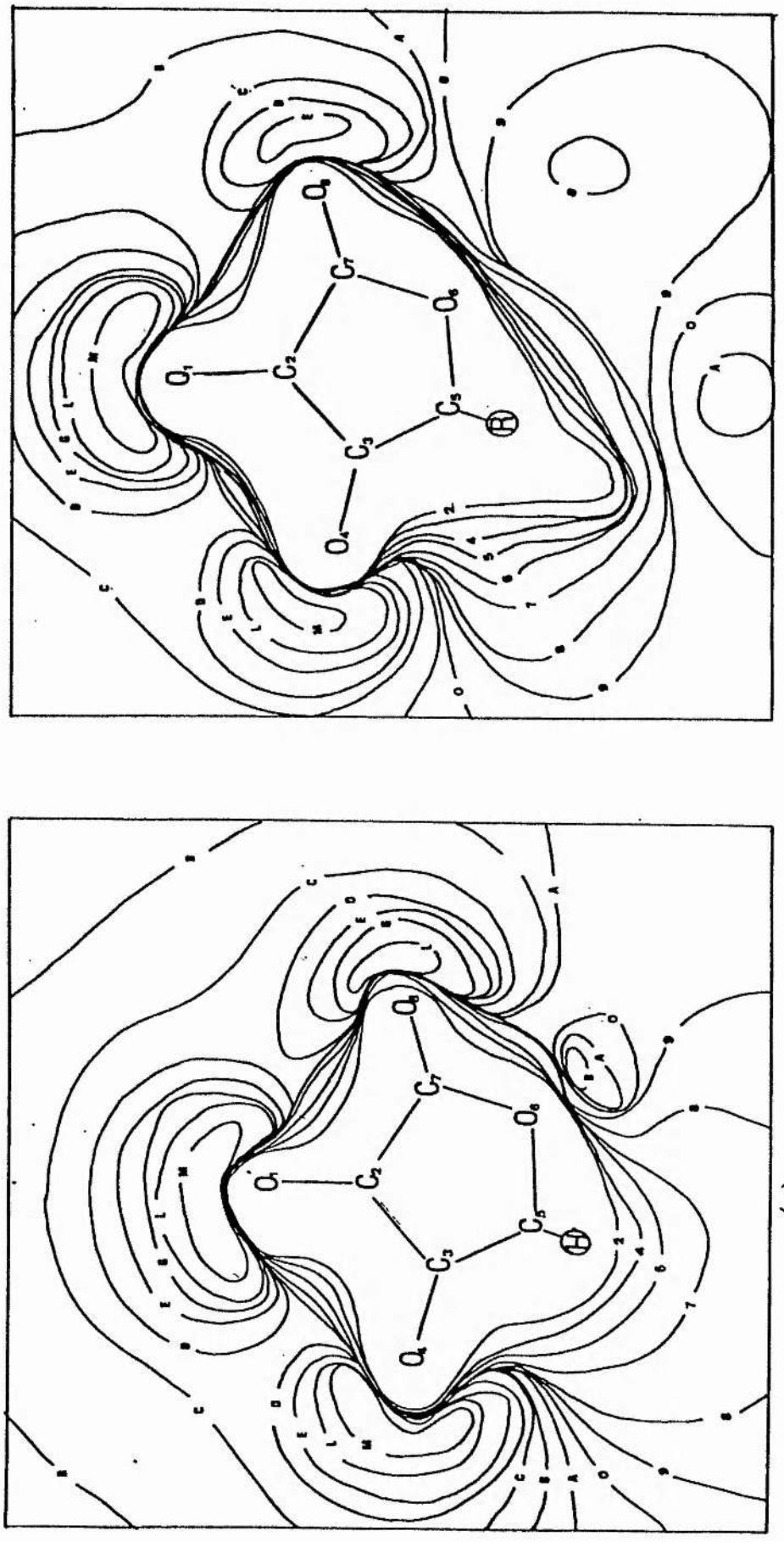
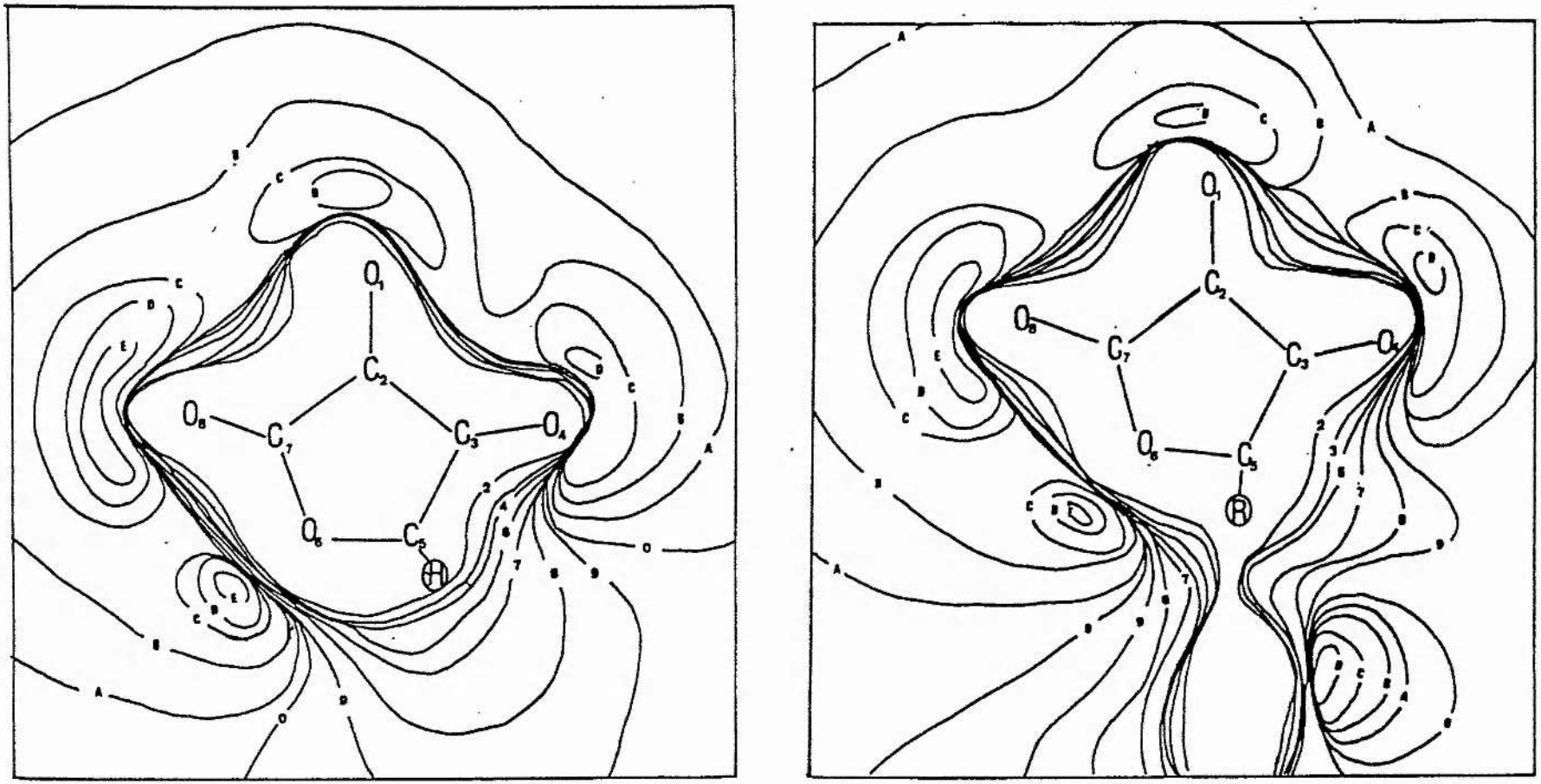


Fig. 9.3 Maps for ascorbyl radical (a) model compound (b) complete molecule



(a) (b)

Fig. 9.4 Maps for dehydroascorbic acid (a) model compound (b) complete molecule

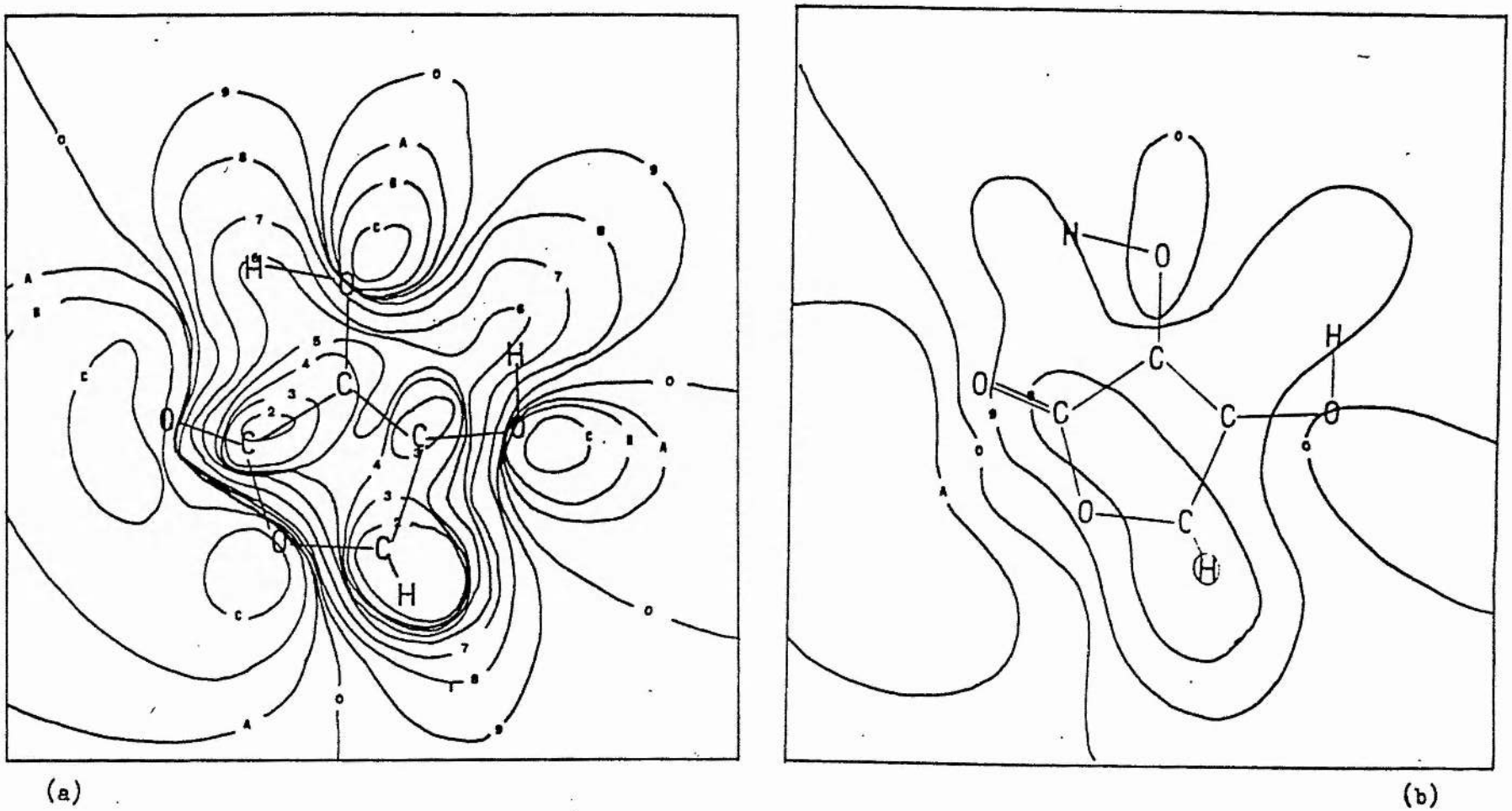


Fig. 9.5 Maps for ascorbic acid (a) 1.1 Å below plane of molecule

(b) 2.1 Å below plane of molecule

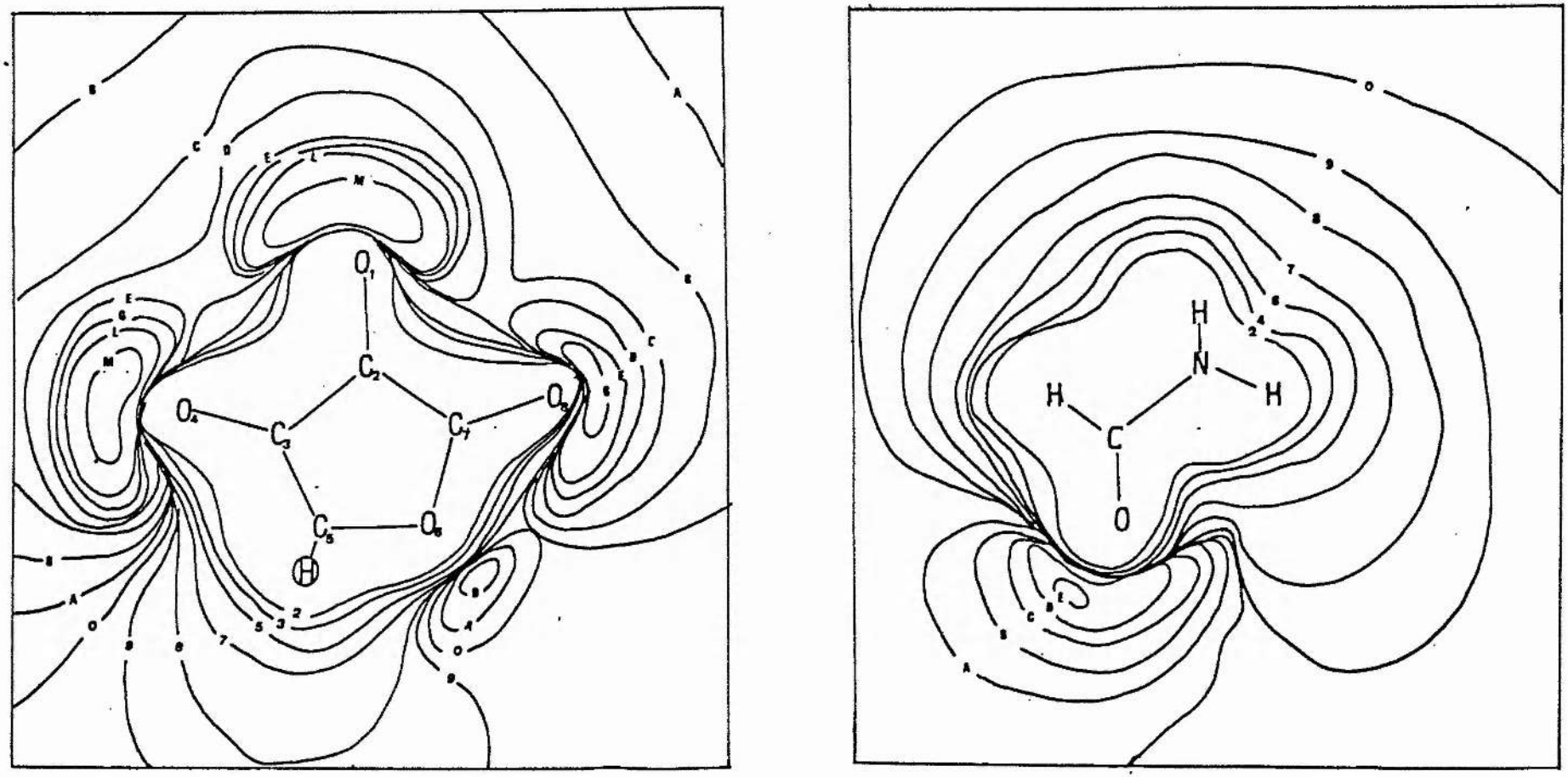


Fig. 9.6 Maps for complex 1 (a) in plane of radical (b) in plane of formamide

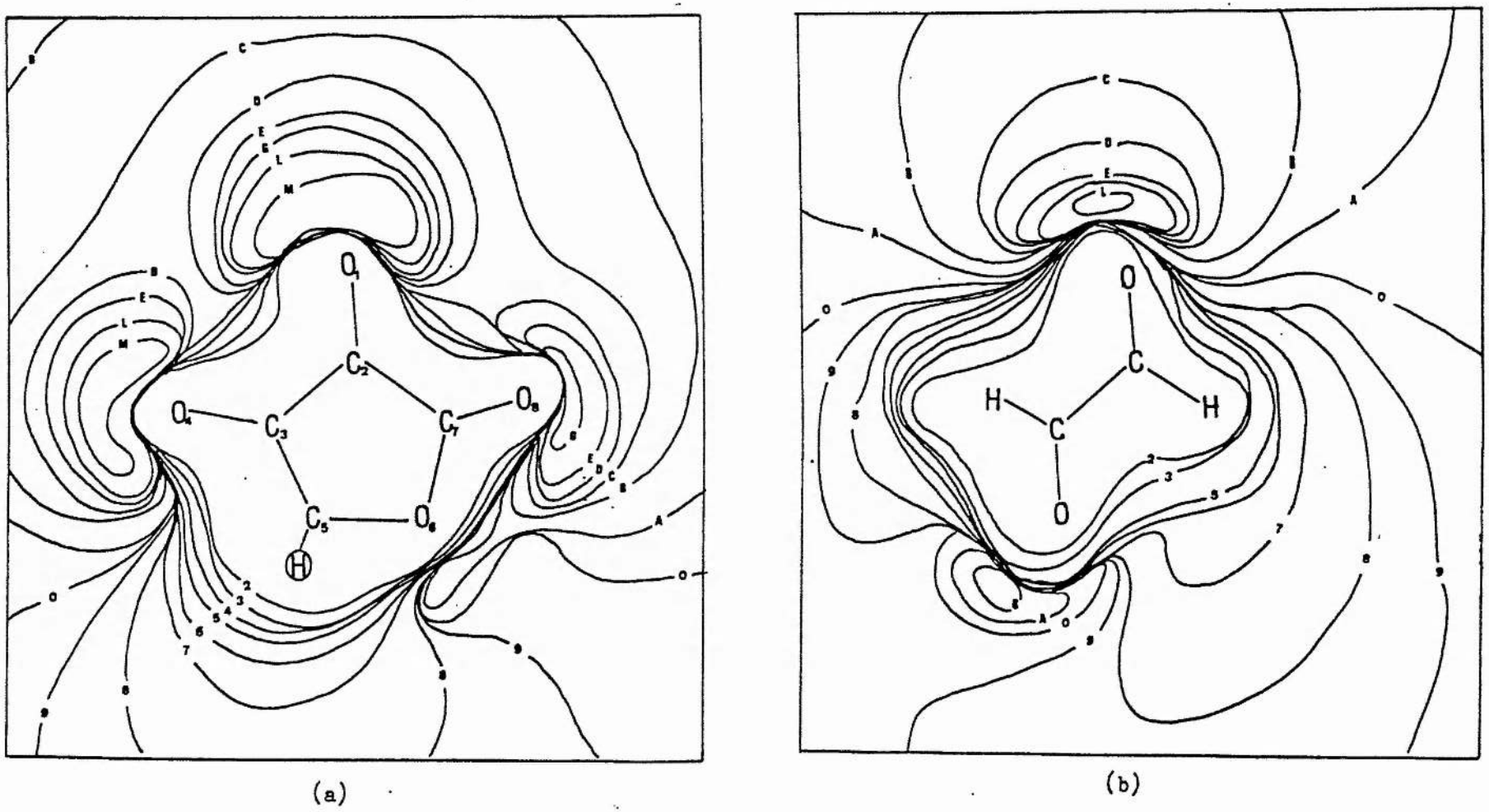
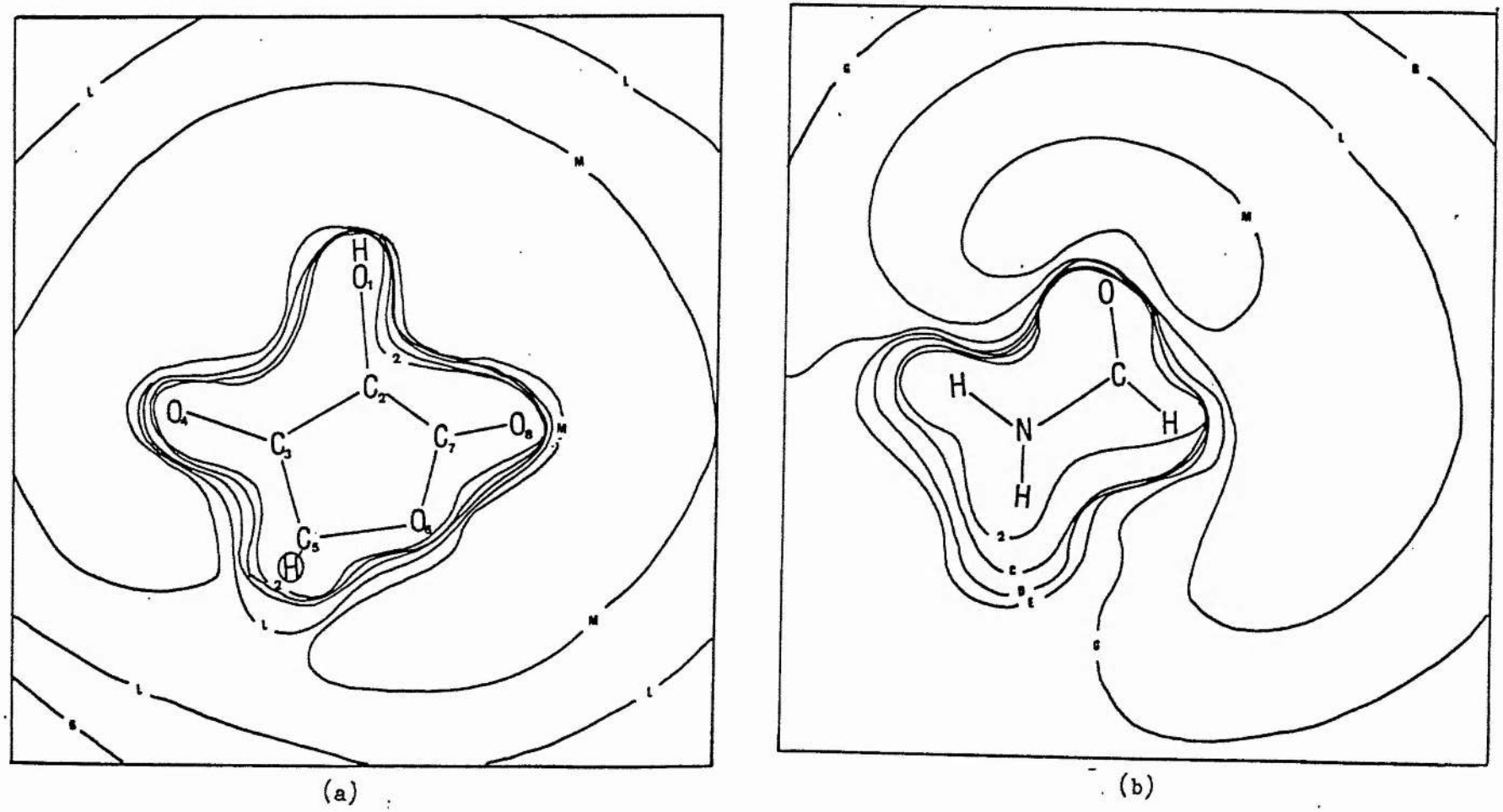


Fig. 9.7 Maps for complex 2 (a) in plane of radical (b) in plane of glyoxal



(a) (b)

Fig. 9.8 Maps for complex 3 (a) in plane of anion (b) in plane of formamide

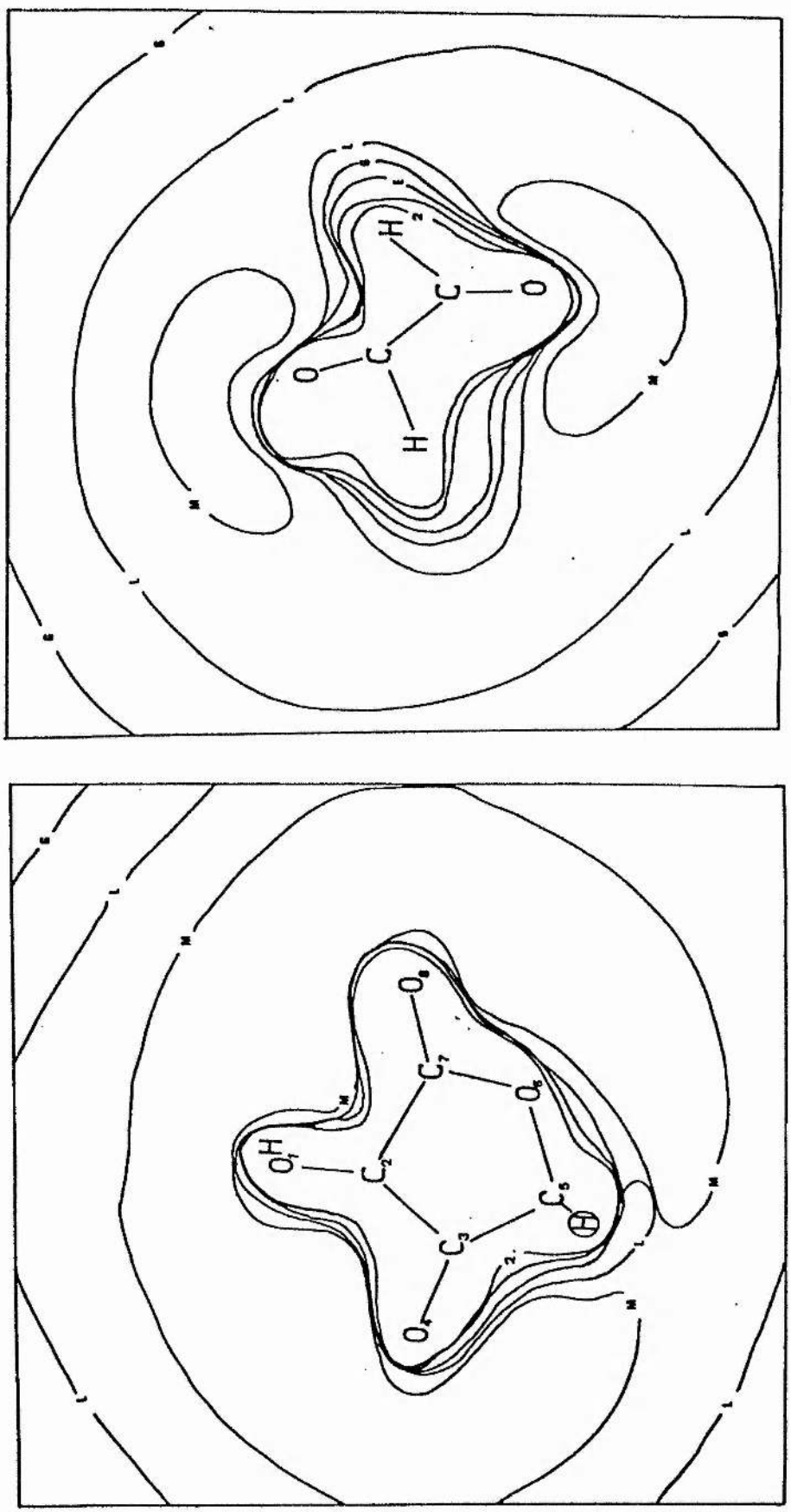


Fig. 9.9 Maps for complex 4 (a) in plane of anion (b) in plane of glyoxal

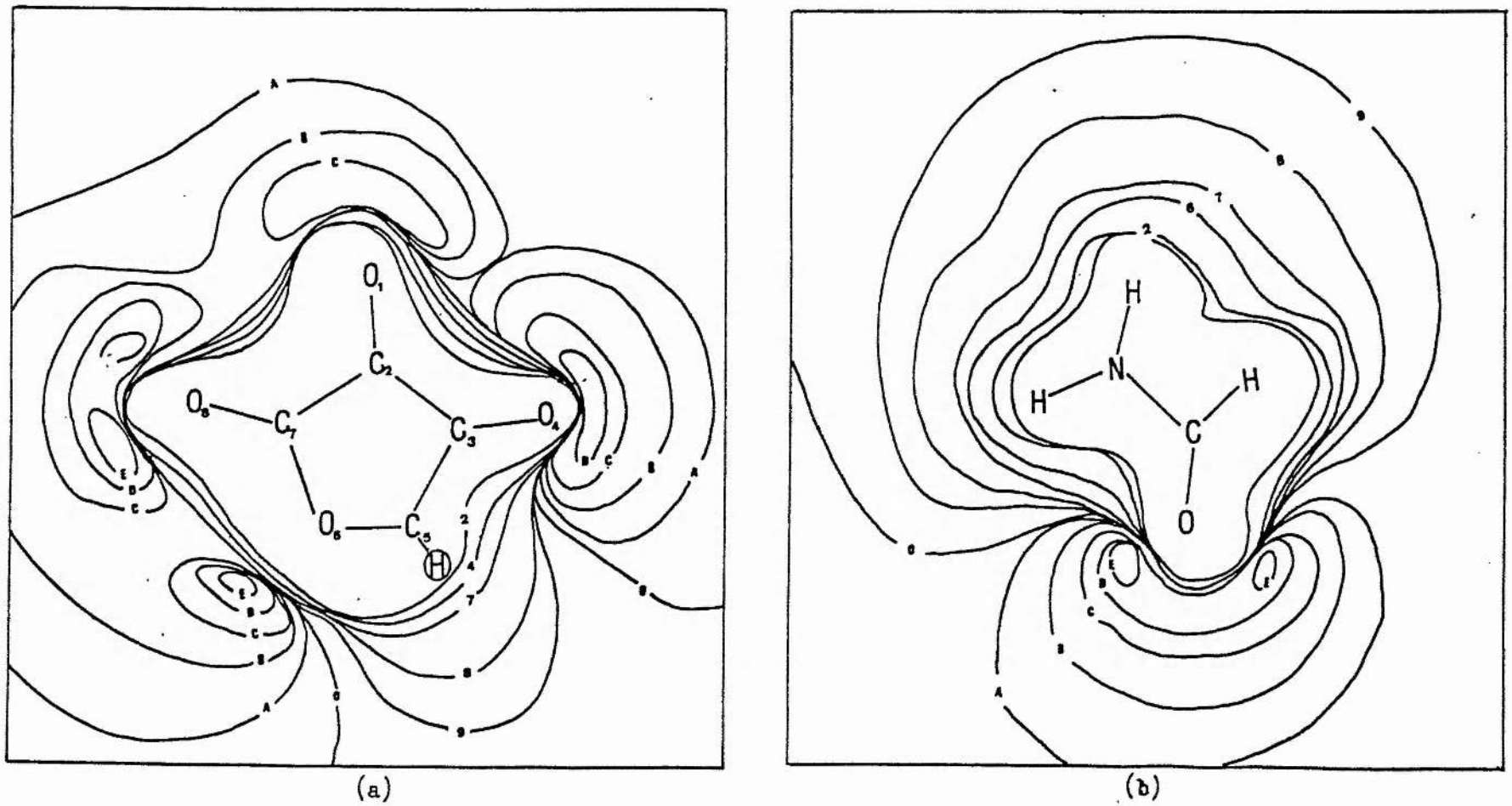
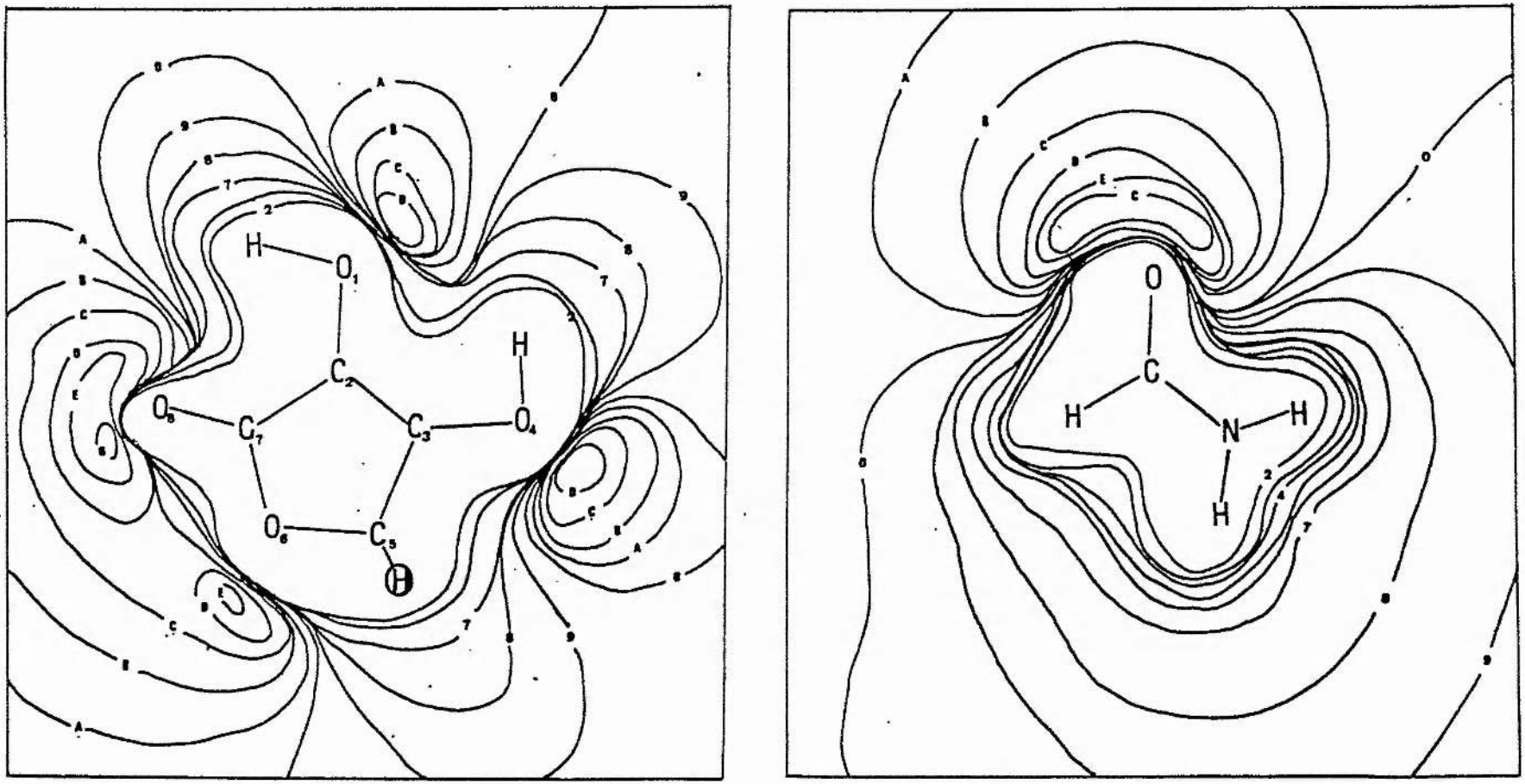


Fig. 9.10 Maps for complex 5 (a) in plane of dehydroascorbic acid model

(b) in plane of formamide



(a) (b)

Fig. 9.11 Maps for complex 6 (a) in plane of ascorbic acid model

(b) in plane of formamide

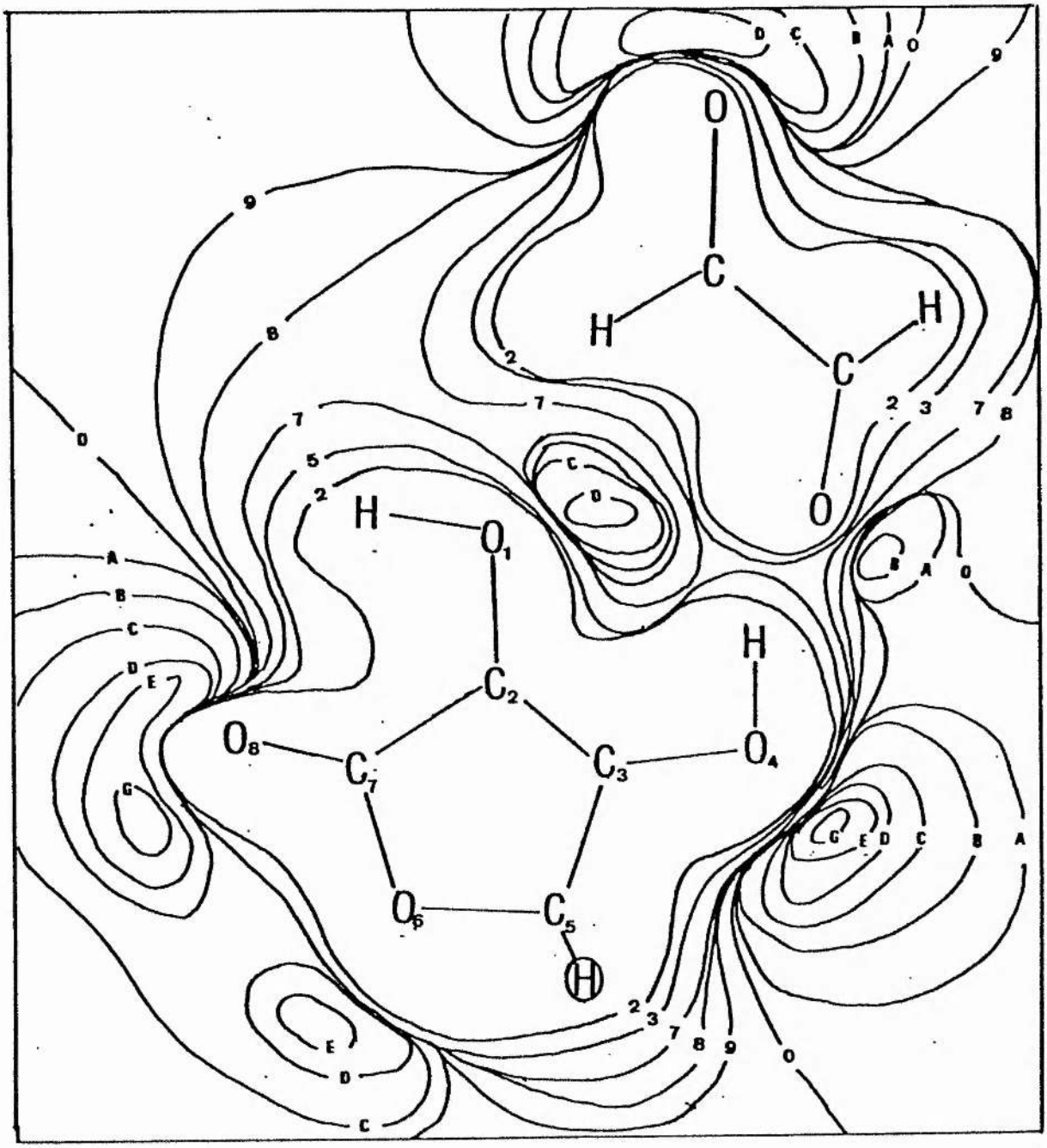


Fig. 9.12 Map for complex 7 in plane of both molecules

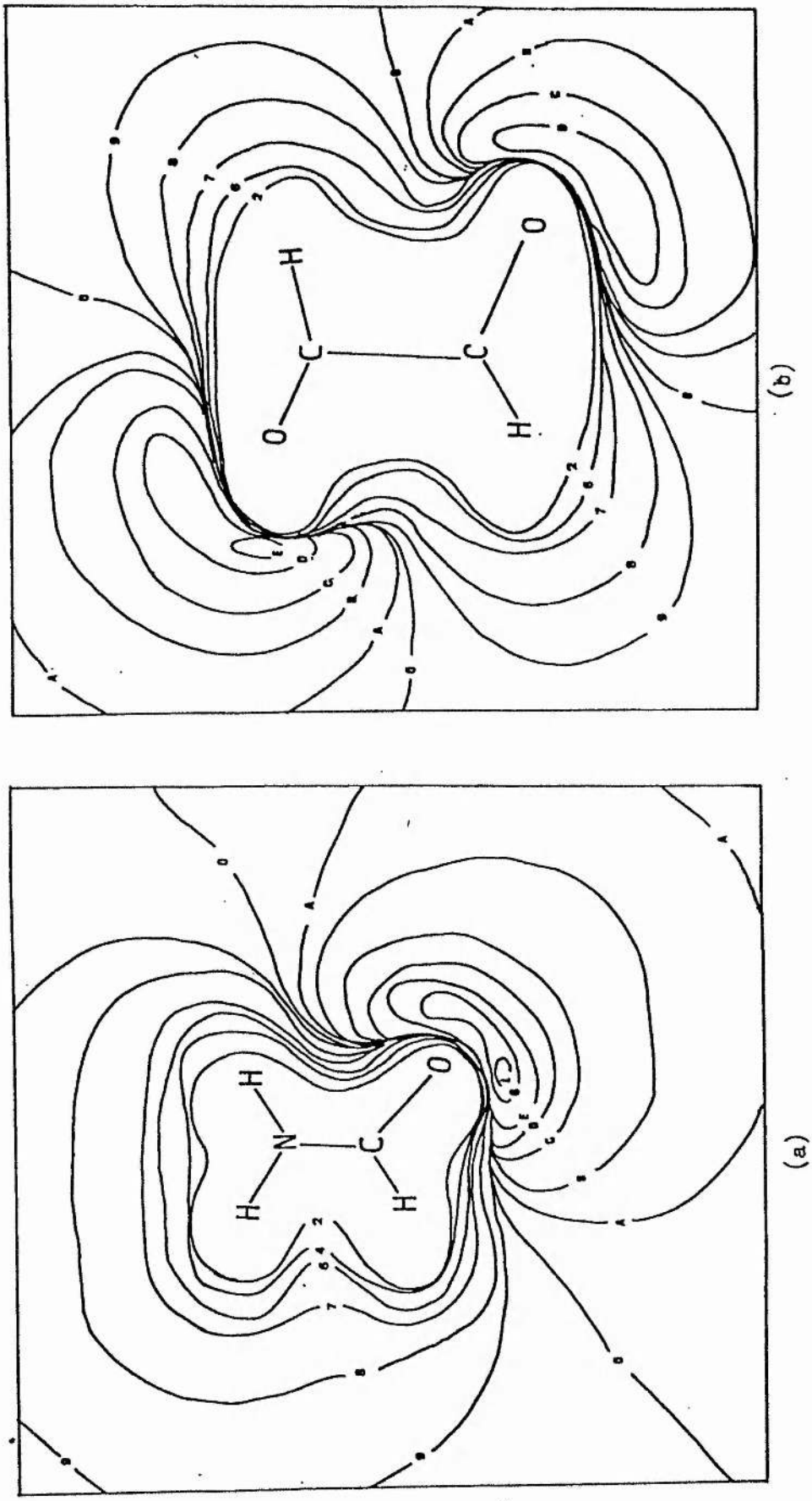


Fig. 9.13 Maps for (a) formamide (b) glyoxal

Table 9.1

Description of charge transfer complexes for which electrostatic molecular potentials have been calculated

Complex	Description
1.	model radical + formamide (stacked) N below C_2, C below C_3 R=3.4
2.	model radical + glyoxal (stacked) C–C below C_2–C_3 R=3.4
3.	model anion + formamide (stacked) N below C_3, C below C_2 R=3.2
4.	model anion + glyoxal (stacked) C–C below C_7–C_2 R=3.0
5.	model dehydroascorbic + formamide (stacked) N below C_2 , C below C_3 R=3.4
6.	model ascorbic + formamide (stacked) N below C_3, C below C_2 R=3.6
7.	model ascorbic + glyoxal (planar) bonding to H_{12}, R=4.0, α=108°

10. METAL COMPLEXING

10.1 Models and basis sets

Ascorbic acid is known to form complexes with a number of metal ions. Metals such as iron and copper are thought to play an important part in the oxidation/reduction of the ascorbic acid system [237,238] and it has also been suggested that complexes with divalent ions such as Ca^{2+} and Mg^{2+} may have important physiological properties [199]. Originally it had been planned to study the possible complexes formed between a variety of metal ions and ascorbic acid and its metabolites. However, the available time and resources have permitted only a limited investigation of complexes involving Mg^{2+} and Ca^{2+}. It is hoped that these calculations may provide a starting point for a wider study in the future.

The compounds used in the previous investigations, even without the side-chain, are too large for studies involving metal ions. Smaller model compounds were constructed by taking a 2 or 3 carbon fragment of the ring and replacing the adjacent carbon atoms by hydrogens (fig. 10.1). The bond lengths and angles of the appropriate model compounds were used in the fragments. To investigate the suitability of these "fragments" as models for ascorbic acid and its metabolites the charge distribution for each molecule was calculated (fig. 10.1) and compared with that of the original molecule (figs. 7.3-7.6).

It can be seen that while the charge distribution of the complete molecules is not exactly reproduced in most cases the fragments do

provide a reasonable qualitative picture. The negative charges on the oxygens, particularly for the anionic species, are generally too high. In the 2-carbon dehydroascorbic acid fragment the oxygens have approximately the same charge, whereas in the complete molecule a difference of ~0.03e is observed. In spite of such discrepancies a study using these molecules should at least provide an initial guide to the preferred positions for metal binding and an indication of the relative binding energies.

There have been a number of studies of the binding of metal ions to organic molecules (for example references [239-242]) and in general the authors have found it necessary to use fairly large basis sets and to include polarization functions in order to obtain a correct description of the behaviour of the metal in the complex. Some workers have reoptimized the atomic basis sets for excited states in order to obtain a basis set for the cation [243]. In general, however, the atomic basis sets have been used.

The basis sets used here are those of Huzinaga et al. [72-75]. Several basis sets for the first and second row atoms and the first transition series are given with a variety of contraction schemes. Full details of the basis sets may be found in the original papers [72-75,81,82]. The terms MINI, MIDI and SMIDI are used to denote the different contraction schemes. MINI indicates a minimal contraction scheme while MIDI retains the minimal contraction scheme for the inner shells but splits the outer shells into contractions of N-1 and 1. The SMIDI contraction scheme is similar but splits only the outermost shell. The same exponents are used for each contraction scheme and no scale factors are employed. The numbers 1,2,3,4 refer to the particular basis set used. Thus MINI-1 indicates basis set 1 with a

minimal contraction scheme, and so on. Four basis sets are given for the atoms of the first two rows but only two for the transition metals. Since it was hoped to include transition metals in the study the majority of calculations were performed with basis set 1. Some calculations were also performed with basis 3 since this gives a better value for the total energy. Basis sets which include polarization functions on all the atoms are indicated by *; a basis set with polarization functions only on the metal is indicated by (*). The exponents of the polarization functions for the first row atoms and magnesium are given in the original papers. Since an exponent for calcium is not given the value calculated to be optimum for beryllium and magnesium was used.

10.2 Binding of Mg^{2+} and Ca^{2+}

The binding of Mg^{2+} to fragment B (fig. 10.1) was investigated for a variety of basis sets. Two positions of the metal atom were investigated (fig.10.2). In the first the metal atom was placed symmetrically between the two oxygens (fig. 10.2a) and distance R was varied. The angles α and ϕ were also varied ($\phi = 180^{o}$ corresponds to the arrangement with the metal ion in the plane of the molecule). In the second arrangement (fig. 10.2b) the Mg^{2+} was placed close to one oxygen and R and α were again varied. The results using the MIDI-3* basis set are shown in Table 10.1

It can be seen that the symmetrical bonding to both oxygens is preferred. Variation of α in the second arrangement shows that the metal ion tends to move into a position where bonding to both oxygens

is possible. In the first arrangement movement from a symmetrical position is not favoured and the in-plane position of the metal is preferred.

The same system was investigated using the SMIDI-1, SMIDI-1* and SMIDI-1(*) basis sets. The values of α and ϕ were fixed at 90^{o} and 180^{o} respectively. R was varied and the two arrangements were compared. The pseudopotential basis set available in the Gaussian 80 program package, LP-31G, was also used and the results obtained are included in Table 10.2 along with those of the other basis sets. No binding energies for this basis set could be obtained, however, since calculations on the isolated magnesium cation failed.

The calculations show that while all the basis sets agree in predicting the optimum values of R for the two arrangements only the basis sets which include polarization functions calculate the symmetrical position to be more stable. For MIDI-3* the difference in binding energy between the two arrangements is 0.0248a.u. while for SMIDI-1* and SMIDI-1(*) it is 0.0241 and 0.0045a.u. respectively. Thus, the SMIDI-1* basis set gives a reliable description of the binding of the metal ion in the complex when compared to the larger MIDI-3* basis set. Unfortunately, because of limitations on the number of basis functions in the Gaussian 80 program it is not possible to use even this basis set for calculations involving either calcium or the radical species and for these complexes the less satisfactory SMIDI-1(*) basis set had to be used. Results obtained with this basis set for the binding of Ca^{2+} to fragment B and the binding of Mg^{2+} and Ca^{2+} to fragment D are given in Table 10.3.

As one would expect, the metal ion is more strongly bound to the anion radical than to the neutral fragment and the complexes with the

radical have smaller optimum values for R. For both fragments the Mg^{2+} ions are more tightly bound than the Ca^{2+} ions, having larger binding energies and smaller optimum values for R.

10.3 Influence of metal ions on charge transfer

The presence of the metal ions considerably reduces the energies of the outermost orbitals of the fragment molecules, making the LUMOs negative. This suggests that if metal complexing to ascorbic acid or its metabolites does occur it will enhance the electron-acceptor properties of these molecules. Also, if the ascorbyl radical complexes with a divalent metal ion the resulting complex will be a cation radical and could fulfil the role which Szent-Gyorgyi proposes for such a species [13]. Since it is clearly impossible to study the charge transfer complexes discussed in chapter 8 with a large basis set that includes polarization functions, the possibility of using fragments B and D as models for the rings in such complexes was investigated. The results of calculations on fragment B and formamide with a variety of basis sets have been presented in Table 8.7 (section 8.3). Similar calculations were attempted with fragment D. This molecule, however, was found to behave as as electron donor towards formamide. The 3-carbon fragment (C) was also investigated but it too was found to behave as a donor. Even with the small fragments it is not possible to perform calculations with basis sets that include polarization functions when the metal ions are present. Because of this only a very limited study using basis set MINI-1 has been carried out.

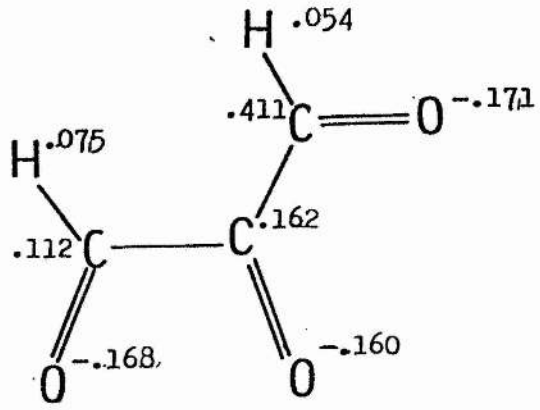
The complete dehydroascorbic acid ring was used with Mg^{2+} or Ca^{2+} placed in the optimum positions as determined by the calculations on the fragment molecule. Formamide was placed below the ring with N below C_2 and C below C_3. The total energy, interaction energy and net charge on formamide were calculated in the usual way (section 8.2). Similar complexes using the anion radical ring were also studied. The results are shown in Table 10.4.

The presence of the metal ions appears to enhance the stability of the complexes and increase the amount of charge transfer from formamide. More detailed calculations are necessary before reaching any final conclusions but this preliminary study does support the suggestion that metal ions may play an important part in the charge transfer process.

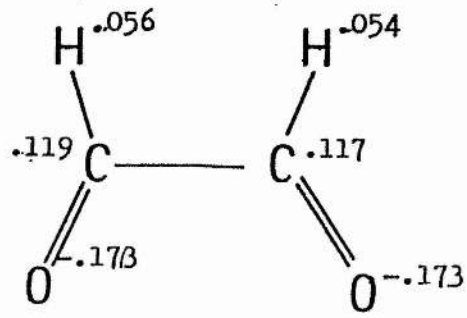
It has been suggested [226] that ascorbic acid and the anion may form complexes in which the metal ion lies above the ring with the side-chain also involved. Complexing with the anion in which the hydroxyl hydrogen shifts to O_4 and the metal is positioned between O_1 and O_8 has also been proposed [199]. This anion was found to be unstable relative to that ionizing at O_1, however, the presence of a metal ion might alter this. If possible these suggestions should also be investigated in a future study of complexing in the ascorbic system.

Of the basis sets used here the SMIDI-1* basis set seems particularly useful since it is relatively compact and economical in computational requirements but at the same time gives a good description of the molecule or complex. Pseudopotentials are being used increasingly in studies of large molecules, particularly those which include metals. For molecules which have mainly first row

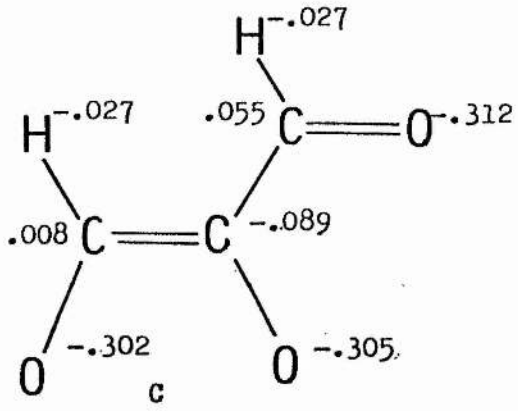
atoms, however, use of pseudopotentials does not result in much saving of time. The LP-31G basis set available in the Gaussian 80 package does not appear to be very efficient and in this study there was not sufficient time to investigate other pseudopotentials that may now be found in the literature [244]. Some of these may prove suitable for calculations on these systems.



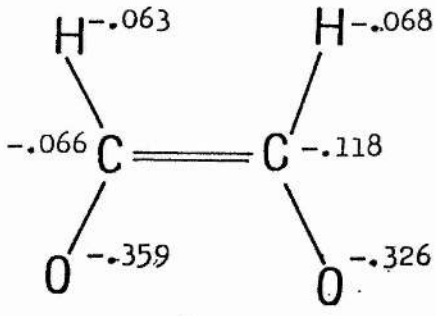
A



B

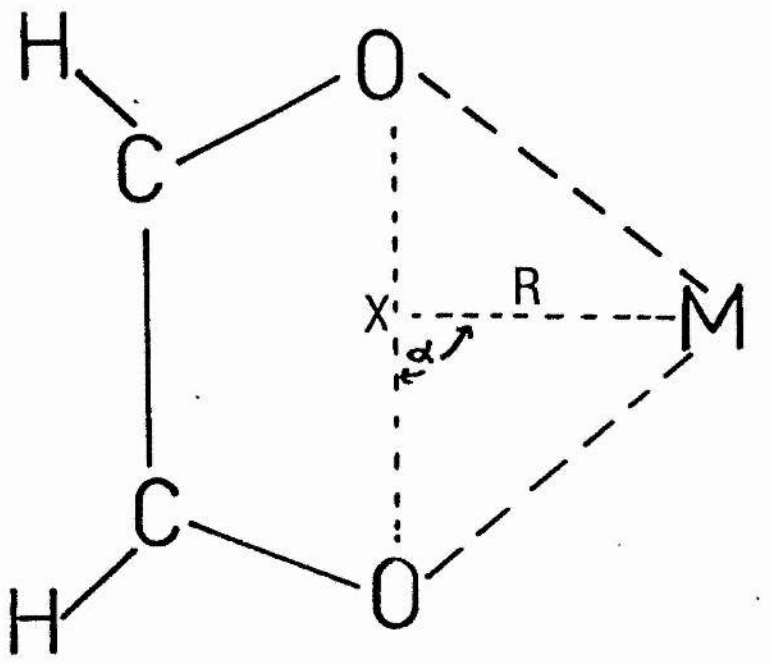


C

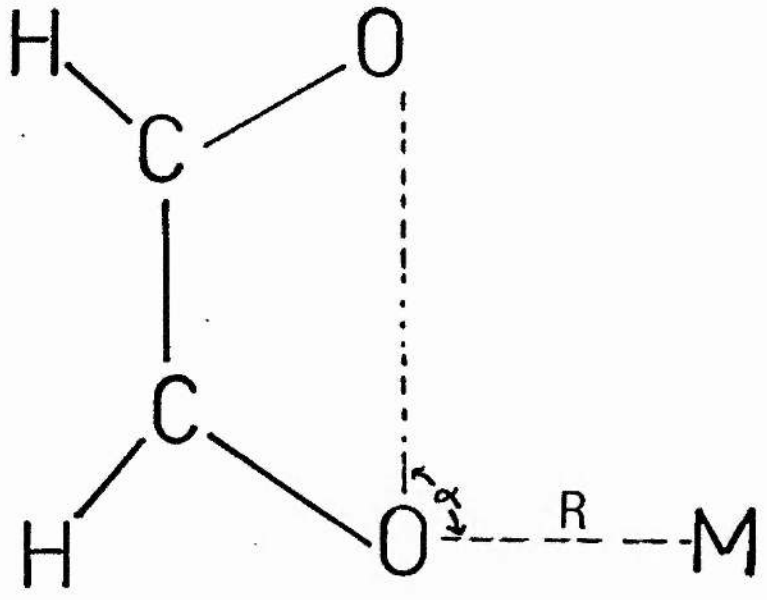


D

Fig. 10.1 Molecules representing fragments of dehydroascorbic acid (A,B) and ascorbyl radical (C,D); molecules C and D carry a net charge of -1.



(a)



(b)

Fig. 10.2 Arrangements for the binding of divalent cation M to fragments B and D. ϕ is the dihedral angle M–X–O–C in (a) and M–O–O–C in (b). $\phi = 180^{\circ}$ corresponds to the in-plane position of M.

Table 10.1

Binding of Mg^{2+} to fragment B, results using MIDI-3* basis set

Arrangement[a]	R(Å)	(α^o)	(ϕ^o)	total energy[b]	binding energy[b]
1.	1.2	90	180	-424.67673	-0.1575
	1.6	90	180	-424.69604	-0.1768
	2.0	90	180	-424.67480	-0.1555
	1.6	90	175	-424.69521	-0.1759
	1.6	90	160	-424.68891	-0.1696
	1.6	80	180	-424.69301	-0.1737
	1.6	110	180	-424.67734	-0.1581
2.	1.6	90	180	-424.63909	-0.1198
	1.8	90	180	-424.67123	-0.1520
	2.0	90	180	-424.66972	-0.1505
	2.6	90	180	-424.62291	-0.1036
	2.0	80	180	-424.67322	-0.1540
	2.0	110	180	-424.66483	-0.1456
	2.0	70	180	-424.67867	-0.1594
	2.0	60	180	-424.68657	-0.1673
	2.0	50	180	-424.69364	-0.1744

a see fig. 10.2

b energies in Hartrees

Table 10.2

Binding of Mg^{2+} to fragment B, results for a variety of basis sets

Arrangement[a]	basis set	R(Å)	total energy[b]	binding energy[b]
1.	SMIDI-1*	1.2	-423.11553	-0.1470
		1.6	-423.13208	-0.1635
		2.0	-423.10984	-0.1413
2.	SMIDI-1*	1.6	-423.07605	-0.1075
		1.8	-423.10795	-0.1394
		2.0	-423.10675	-0.1382
1.	SMIDI-1(*)	1.4	-422.90722	-0.1382
		1.6	-422.91415	-0.1451
		1.8	-422.91001	-0.1410
2.	SMIDI-1(*)	1.6	-422.87655	-0.1075
		1.8	-422.90922	-0.1402
		2.0	-422.90786	-0.1388
1.	SMIDI-1	1.2	-422.83346	-0.0892
		1.6	-422.85911	-0.1149
		2.0	-422.84587	-0.1016
2.	SMIDI-1	1.6	-422.84174	-0.0975
		1.8	-422.87298	-0.1287
		2.0	-422.86961	-0.1254
1.	LP-31G	1.4	-43.89014	
		1.6	-43.89396	
		1.8	-43.88785	
2.	LP-31G	1.8	-43.89564	

a,b see footnotes to Table 10.1

Table 10.3

Binding of Mg^{2+} to fragment D and Ca^{2+} to fragments B and D, results using SMIDI-1(*) basis set

Fragment	ion	arrangement	R(Å)	total energy[a]	binding energy[a]
B	Ca^{2+}	1.	1.8	-398.29274	-0.0857
			2.0	-398.30041	-0.0934
			2.2	-898.29894	-0.0919
		2.	2.2	-898.29246	-0.0854
			2.4	-898.29260	-0.0855
			2.6	-898.28576	-0.0787
D	Mg^{2+}	1.	1.2	-423.38178	-0.5801
			1.4	-423.38249	-0.5809
			1.6	-423.36792	-0.5663
		2.	1.6	-423.33102	-0.5294
			1.8	-423.34275	-0.5411
			2.0	-423.32179	-0.5202
	Ca^{2+}	1.	1.6	-898.70580	-0.4661
			1.8	-898.70792	-0.4682
			2.0	-898.69643	-0.4567
		2.	1.8	-898.62495	-0.3853
			2.0	-898.66286	-0.4232
			2.2	-898.66275	-0.4231

[a] energies in Hartrees

Table 10.4

Influence of metal ions on charge transfer, results using MINI-1 basis set

Complex	total energy[a]	ΔE[a]	net charge on formamide[b]
model dehydroascorbic + formamide	-616.30754	-0.0035	+0.0040
model dehydroascorbic + formamide + Mg^{2+}	-814.07538	-0.0887	+0.0119
model dehydroascorbic + formamide + Ca^{2+}	-1289.47016	-0.0495	+0.0087
model radical + formamide	-616.42563	-0.0014	+0.0010
model radical + formamide + Mg^{2+}	-814.56205	-0.4556	+0.0045
model radical + formamide + Ca^{2+}	-1289.92412	-0.3837	+0.0029

[a] energies in Hartrees

[b] in units of electronic charge

11. CONCLUSION

Recent developments in computer hardware and software mean that it is now possible to perform quantum mechanical calculations on medium-sized molecules of biological interest. Such calculations can provide information which may supplement experimental data and serve as a stimulus to further experimental work.

In this thesis two specific problems have been investigated. The first being the conformational flexibility of the tripeptide glutathione and the second the electronic structure and charge transfer properties of ascorbic acid and its metabolites. These two problems were chosen because of their relevance to the "electronic theory" of cancer proposed by Albert Szent-Gyorgyi [13].

In the study of glutathione the theoretical results are in partial agreement with the available experimental data. As well as predicting the "optimum" conformation, the calculations indicate other conformations that may be accessible to the molecule and provide information about the changes in conformational flexibility that occur on ionization.

The studies on ascorbic acid and its metabolites have demonstrated that it is theoretically possible for these molecules to enhance the transfer of charge from formamide (the simplest model of a protein) to glyoxal (a model for methylglyoxal). The results also suggest that the formation of such charge transfer complexes may be enhanced by the presence of metal ions and may influence, and be influenced by, the hydration of the constituent molecules.

In the course of these investigations comparisons have been made

between results obtained with the semi-empirical PCILO method and more rigorous ab initio calculations. These comparisons show that PCILO gives a reliable qualitative picture of the conformational flexibility of a molecule, although it does not always predict the minimum energy conformation correctly. Similarly, the STO-3G basis set has been shown to provide a reasonable account of trends in molecular properties within a series of molecules when compared with calculations performed using larger basis sets. Other minimal basis sets, such as those put forward by Huzinaga et al. [72-75] and Mezey and Csizmadia [71], have been shown to be at least as computationally efficient as the STO-3G basis set and, in general, to provide a better description of the electronic structure of a molecule.

As with most theoretical investigations on biological molecules the work reported here must be regarded as a preliminary rather than a definitive study. The results obtained from these calculations suggest several further experimental and theoretical studies that might usefully be carried out.

For glutathione it would be helpful to have additional experimental information, possible obtained from n.m.r. work, on the preferred conformations of the molecule. It would also be useful to check the PCILO results against more accurate calculations, particularly for the glutamyl residue. Further information might also be gained by studying the changes in the conformational flexibility of the molecule that occur when glutathione forms an adduct with methylglyoxal.

The calculations on ascorbic acid have already stimulated further experimental work. The recent investigations by P.Gascoyne [233] comparing the effects of the D and L isomers of ascorbic acid on charge

transfer in methylglyoxal/protein systems were carried out as a result of the calculations showing that the charge transfer properties of the molecule are likely to depend only on the ring system. Similar studies to investigate the influence of metal ions are currently underway, and e.s.r. studies using ^{17}O labelled ascorbic acid are planned for the near future [245].

As well as a more extensive study of the complexing of metal ions in the ascorbic acid system, the most useful theoretical study that could be carried out would be an investigation of solvent effects. One of the most conspicuous differences between normal and cancerous cells is the disorder in the intracellular water content of the latter [246,247]. It could be that the highly structured arrangement of water molecules in the normal cell favours and is favoured by the formation of charge transfer complexes. A theoretical study of the preferred positions of hydration in the isolated molecules and in the complexes might provide a first step in the investigation of such an hypothesis.

In conclusion, while these calculations have neither proved nor disproved the "electronic theory" of cancer, they have provided information on the electronic structure and properties of two important biological molecules which may be of use to both theoreticians and experimentalists in future investigations.

REFERENCES

1. D.R.Hartree, Proc. Camb. Phil. Soc. 24 89, 111 (1928)

2. V.Fock, Zeits. f. Physik 61 126 (1930)

3. J.C.Slater, Phys. Rev. 35 210 (1930)

4. C.C.J.Roothaan, Rev. Mod. Phys. 23 69 (1951)

5. G.G.Hall, Proc. Roy. Soc. (London) Ser. A 205 541 (1951)

6. R.A.Whiteside, J.S.Binkley, R.Krishnan, D.J.DeFrees, H.B.Schlegal, and J.A.Pople, "Carnegie-Mellon Quantum Chemistry Archive" (Pittsburgh, 1980)

7. W.G.Richards, T.H.Walker, R.Hinkley, "Bibliography of ab initio molecular wave functions" (O.U.P., London 1971, 1977)

8. P.C.Harihan, J.J.Kaufman and C.Petrongolo, Int. J.Quantum Chem. Quantum Biol. Symp. 6 223 (1979)

9. P.C.Harihan, C.Petrongolo and J.J.Kaufman, Int. J.Quantum Chem. Quantum Biol. Symp. 8 in press (1981)

10. M-M.Rohmer and G.H.Loew, Int. J.Quantum Chem. Quantum Biol. Symp. 6 93 (1979)

11. A.Pullman, Int. J.Quantum Chem. Quantum Biol. Symp. 8 in press (1981)

12. A.Szent-Gyorgyi, "Electronic Biology and Cancer" (M.Decker, New York, 1974)

13. A.Szent-Gyorgyi, "The Living State and Some Observations on Cancer" (M.Decker, New York, 1978)

14. "Submolecular Biology and Cancer", CIBA Foundation Symposium 67 New Series, (Excerpta, Medica, Amsterdam, 1979)

15. S.Scheiner, in "Quantum Chemistry in Biomedical Sciences", p.493 (H.Weinstein and J.P.Green Eds, New York Academy of Sciences, New York, 1981)

16. H.Weinstein, R.Osman, S.Topiol and J.P.Green, ibid. p.434

17. M.Born and J.R.Oppenheimer, Ann. Physik. 84 457 (1927)

18. L.Pauling and E.B.Wilson, "Introduction to Quantum Mechanics", (McGraw-Hill, New York, 1935)

19. R.K.Nesbet, Proc. Roy. Soc. (London) A230 312 (1935)

20. R.Seeger and J.A.Pople, J.Chem. Phys., 65 265 (1976)

21. E.Steiner, "The Determination and Interpretation of molecular wave functions" (Cambridge University Press, 1976)

22. J.A.Pople and R.K.Nesbet, J.Chem. Phys., 22 571 (1954)

23. A.Hinchcliffe and D.G.Bounds, J.Mol. Struct., 54 231 (1979)

24. A.T.Amos and G.G.Hall, Proc. Roy. Soc. A 263 483 (1961)

25. A.T.Amos and L.C.Snyder, J.Chem. Phys., 41 1773 (1964)

26. J.E.Harriman, J.Chem. Phys., 40 2827 (1964)

27. D.H.Phillips and J.C.Schug, J.Chem. Phys., 61 1031 (1974)

28. D.M.Chipman, J.Chem. Phys., 71 761 (1979)

29. P.O.Lowdin, Advan. Chem. Phys., 2 207 (1959)

30. H.F.Schaefer, "The Electronic Structure of Atoms and Molecules" (Addison-Westley, Reading, Mass. , 1972)

31. I.Shavitt, in "Electronic Structure Theory", Mod. Theo. Chem., vol. 6, chap. 6 (ed. H.F.Schaefer, Plenum Press, New York, 1979)

32. C.Moller and M.S.Plesset, Phys. Rev. 46 618 (1934)

33. J.S.Binkley and J.A.Pople, Int. J. Quantum Chem. 9 229 (1975)

34. J.A.Pople, J.S.Binkley and R.Seeger, Int. J. Quantum Chem. S10 1 (1976)

35. M.J.S.Dewar, Science, 187 1037 (1975)

36. J.A.Pople ,D.A.Santry and G.A.Segal, J.Chem. Phys., 43 5129 (1965)

37. J.A.Pople and G.A.Segal, J.Chem. Phys., 43 5136 (1965)

38. J.A.Pople and D.L.Beveridge, "Approximate Molecular Orbital Theory" (McGraw-Hill, New York, 1970)

39. "Semi-empirical Methods of Electronic Structure Calculations", Mod. Theor. Chem., Vols. 7 and 8 (ed. G.A.Segal, Plenum Press, New York, 1974)

40. K.Jug, Theoret. Chim. Acta , 54 263 (1980)

41. S.Diner, J.P.Malrieu and P.Claverie, Theoret. Chim. Acta , 13 1 (1969)

42. J.P.Malrieu, P.Claverie and S.Diner, Theoret. Chim. Acta , 13 18 (1969)

43. S.Diner, J.P.Malrieu, F.Jordan and M.Gilbert, Theoret. Chim. Acta 15 100 (1969)

44. F.Jordan, M.Gilbert, J.P.Malrieu, and V.Pincelli, Theoret. Chim. Acta 15 211 (1969)

45. P.Palla, K.Petrongolo and J.Tomasi, J.Phys. Chem. 84 435 (1980)

46. B.Pullman and A.Pullman, Adv. Prot. Chem. 28 347 (1974)

47. B.Pullman and A.Saran, Prog. Nucleic Acid Res. Mol. Biol. 18 215 (1978)

48. B.Pullman, Adv. Quantum Chem. 10 251 (1977)

49. B.Pullman, in "Molecular and Quantum Pharmacology", p.9 (eds. E.Bergmann and B.Pullman, Reidel, Dordrecht, 1974)

50. B.Pullman and A.Saran, Int. J.Quantum Chem. Quantum Biol. Symp. 2 71 (1975)

51. A.Pullman and B.Pullman, Quart. Rev. Biophys. 7 505 (1975)

52. B.Pullman and H.Berthod, Theoret. Chim. Acta 36 317 (1973)

53. B.Maigret and B.Pullman, Theoret. Chim. Acta 37 17 (1975)

54. B.Maigret and B.Pullman, Theoret. Chim. Acta 35 113 (1974)

55. B.Pullman, Ph.Courriere and J.L.Coubeils, Mol. Pharmacol. 7 397 (1971)

56. A.Pullman and G.N.J.Port, Theoret. Chim. Acta 32 77 (1973)

57. Ph.Courriere, J.L.Coubeils and B.Pullman, C.R.Acad. Sci. Paris 272 1697 (1971)

58. G.N.J.Port and B.Pullman, Theoret. Chim. Acta 33 275 (1974)

59. B.Pullman, J.L.Coubeils, Ph.Courriere and J.P.Gervois, J.Med. Chem. 15 17 (1972)

60. B.Pullman, H.Berthod and Ph.Courriere, Int. J.Quantum Chem. Quantum Biol. Symp. 1 93 (1974)

61. M.Martin, R.Carbo, C.Petrongolo and J.Tomasi, J.Am.Chem.Soc. 97 1338 (1975)

62. A.Pullman, private communication

63. S.F.Boys, Proc.Roy.Soc. (London) Ser. A 200 542 (1950)

64. H.Preuss, Z.Naturforsch. 11 823 (1956)

65. T.H.Dunning, J.Chem Phys. 53 2823 (1970), 55 716 (1971)

66. F.B.Van Duijneveldt, IBM Technical Research Report No. RJ 945 (1971)

67. W.J.Hehre, R.F.Stewart and J.A.Pople, J.Chem. Phys. 51 2657 (1969), 52 2769 (1970)

68. R.Ditchfieid, W.J.Hehre and J.A.Pople, J.Chem. Phys. 54 724 (1971), 56 4233 (1972)

69. J.S.Binkley, J.A.Pople and W.J.Hehre, J.Am.Chem.Soc. 102 939 (1980)

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

71. P.G.Mezey and I.G.Csizmadia, Can. J.Chem. 55 1181-99 (1977)

72. A.N.Tavouksoglou and S.Huzinaga, J.Chem. Phys. 72 1385 (1980)

73. H.Tatewaki and S.Huzinaga, J.Chem. Phys. 71 4339 (1979)

74. H.Tatewaki and S.Huzinaga, J.Comp. Chem. 1 205 (1980)

75. Y.Sakai, H.Tatewaki and S.Huzinaga, J.Comp. Chem 2 100-107 (1981)

76. A.D.M^{c}lean and G.S.Chandler, J.Chem. Phys. 72 5639 (1980)

77. T.H.Dunning and P.J.Hay, in "Methods of Electronic Structure Theory", Mod.Theor.Chem. vol.3 (ed. H.F.Schaefer, Plenum Press, New York, 1977)

78. R.Ditchfield, W.J.Hehre and J.A. Pople, J.Chem. Phys. 52 5001 (1970)

79. W.J.Hehre, W.A.Lathan, R.Ditchfield, M.D.Newton and J.A.Pople, "Gaussian 70", QCPE 236, Indiana University

80. J.S.Binkley, R.A.Whiteside, R.Krishnan, R.Seeger, D.L.DeFrees, H.B.Schegel, S.Topiol, L.R.Kahn and J.A.Pople, "Gaussian 80", Dept. Chem. Carnegie-Mellon University, Pittsburgh, PA 15213

81. H.Tatewaki and S.Huzinaga, J.Chem.Phys.72 399 (1980)

82. Y.Sakai, H.Tatewaki and S.Huzinaga, J.Comp.Chem. 2 108 (1981)

83. M.F.Guest and V.R.Saunders, Mol.Phys. 28 819 (1974)

84. K.Kitaura and K.Morokuma, Int. J.Quantum Chem.10 325 (1976)

85. R.S.Mulliken, J.Chem.Phys. 22 1833, 1841, 2338, 2343 (1955)

86. S.Topiol, unpublished results

87. J.S.Binkley, J.A.Pople and P.A.Dobosh, Mol.Phys. 28 1423 (1974)

88. J.A.Pople, R.Seeger and R.Krishnan Int. J.Quatum.Chem. Symp. 11 149 (1977)

89. R.Fletcher and M.J.D.Powell, Compt.J. 6 163 (1963)

90. J.B.Collins, P.V.R.Schleyer, J.S.Binkley and J.A.Pople, J.Chem.Phys. 64 5142 (1976)

91. B.A.Murtaugh and R.W.H.Sargent, Compt.J. 13 185 (1980)

92. N.S.Ertz, QCPE 283, Indiana University

93. P.Pulay, Mol.Phys. 17 197 (1969)

94. P.Pulay, in "Electronic Structure Theory", Mod. Theo. Chem. vol. 4, p.153 (ed. H.F.Schaefer, Plenum Press, New York, 1977)

95. W.B.England, J.Chem.Phys. 72 2108 (1980)

96. L.Radom, in "Electronic Structure Theory", Mod. Theo. Chem. vol. 4, p.333 (ed. H.F.Schaefer, Plenun Press, New York, 1977)

97. J.Chandresekhar, J.G.Andrade and P.v.R.Schleyer, J.Am. Chem. Soc. 103 5609 (1981)

98. R.M^{c}Weeny and B.T.Sutcliffe, "Methods of Molecular Quantum Mechanics" (Academic Press, London and New York, 1969)

99. M.Yoshimine and A.D.McLean, Int. J. Quantum Chem. S1 313 (1967)

100. C.Thomsom in "Electron Spin Resonance" vol. 1, p.1 (Specialist Periodical Report, Chemical Society, London, 1973)

101. C.Thomson in "Electron Spin Resonance" vol.3, p.1 (Specilaist Periodical Report, Chemical Society, London, 1976)

102. P.O.Lowdin, J.Chem. Phys. 18 365 (1950), 21 374 (1953)

103. P.Politzer and R.R.Harris, J. Am. Chem. Soc. 92 6451 (1970)

104. R.W.F.Bader, S.G.Anderson and A.J.Duke, J. Am. Chem. Soc. 101 1389 (1979)

105. R.Carbo and C.Arnau, Gazz. Chim. Ital. 108 171 (1978)

106. P.Otto and J.Ladik, Int. J. Quantum Chem. 18 1143 (1980)

107. J.A.Pople in "Electronic Structure Theory", Mod. Theo. Chem. vol.4, p.1 (ed. H.F.Schaefer, Plenum Press, New York, 1977)

108. P.Pulay, Theoret. Chim. Acta 50 299 (1979)

109. W.Meyer and P.Pulay, MOLPRO description (Munchen and Stuttgart, Germany, 1969)

110. H.B.Schlegel, S.Wolfe and F.Bernadi, J.Chem. Phys. 63 3632 (1975)

111. A.Kormonicki, I.Kazahiro, K.Morokuma and R.Ditchfield, Chem. Phys. Lett. 45 595 (1977)

112. H.Huber, P.Carsky and R.Zahradnik, Theoret. Chim. Acta 41 217 (1976)

113. D.Perahia, B.Pullman and P.Claverie, Int. J. Quantum Chem., 6 337 (1972)

114. B.Pullman "Quantum Mechanics of Molecular Conformations" (Wiley-Interscience, 1976)

115. W.G.Richards "Quantum Pharmacology" p.121 (Butterworths, London, 1977)

116. K.Morokuma and K.Kitaura in "Molecular Interactions", vol.1, p.21 (eds. H.Ratajczak and W.J.Orville-Thomas, Wiley, 1980)

117. P.A.Kollman in "Methods of Electronic Structure Theory", Mod. Theo. Chem. vol.4, p.109 (ed. H.F.Schaefer, Plenum, New York, 1977)

118. P.A.Kollman, J.Am. Chem. Soc. 99 4875 (1977)

119. K.Morokuma, Acc. Chem. Res. 10 294 (1977)

120. P.Schuster in "Intermolecular Interactions: from Diatomics to Biopolymers", p.363 (ed. B.Pullman, Wiley, New York, 1978)

121. L.J.Schaad in "Hydrogen Bonding", chap.2 (eds. M.D.Joesten and L.J.Schaad, Dekker, New York, 1974)

122. P.A.Kollman and L.C.Allen, Chem. Rev. 72 283 (1972)

123. K.Morokuma, J.Chem. Phys. 55 5085 (1971)

124. P.A.Kollman and L.C.Allen, J.Chem. Phys. 52 5085 (1970)

125. M.Dreyfus, B.Maigret and A.Pullman, Theoret. Chim. Acta 17 109 (1970)

126. A.Johansson, P.A.Kollman and S.Rothenberg, Theoret. Chim. Acta 29 167 (1973)

127. M.Bulski and G.Chalasinski, Theoret. Chim. Acta 44 399 (1977)

128. W.Kolos, Theoret. Chim. Acta 51 219 (1979)

129. R.Bonaccorsci, E.Scrocco and J.Tomasi, J.Chem. Phys. 52 5270 (1970)

130. R.Bonaccorsci, C.Petrongolo, E.Scrocco and J.Tomasi in "Quantum Aspects of Heterocyclic Compounds in Chemistry and Biochemistry" (Jerusalem Symp. on Quantum Chem. Biochem. II, eds. B.Pullman and E.D.Bergmann, Israel Academy of Sciences and Humanities, Jerusalem, 1970)

131. R.Bonaccorsci, A.Pullman, E.Scrocco and J.Tomasi, Theoret. Chim. Acta 24 51 (1972)

132. R.Bonaccorsci, E.Scrocco, J.Tomasi and A.Pullman, Theoret. Chim. Acta 36 339 (1975)

133. D.Perahia and A.Pullman, Theoret. Chim. Acta 48 263 (1978)

134. L.C.Allen in "Quantum Chemistry in Biomedical Sciences", p.383 (eds. H.Weinstein and J.P.Green, New York Academy of Sciences, New York, 1981)

135. P.Politzer and K.C.Daiker, Int. J. Quantum Chem. Quantum Biol. Symp. 4 317 (1977)

136. P.Politzer and D.G.Truhlar "Chemical Applications of Atomic and Molecular Electrostatic Potentials" (Plenum, New York, 1981)

137. E.Scrocco and J.Tomasi, Adv. Quantum Chem. 2 115 (1978)

138. E.Scrocco and J.Tomasi, Top. Curr. Chem. 42 95 (1978)

139. R.Bonaccorsci, C.Petrongolo and J.Tomasi, personal communication

140. M.Born, Z. Phys. 1 45 (1920)

141. J.G.Kirkwood, J.Chem. Phys. 2 351 (1934)

142. L.Onsager, J.Am. Chem. Soc. 58 1486 (1936)

143. J.Hylton, R.E.Christoffersen and G.Hall, Chem. Phys. Lett. 24 51 (1974)

144. A.J.Hopfinger, "Conformational Properties of Macromolecules" (Acad. Press, New York, 1973)

145. D.L.Beveridge, R.J.Radna, G.W.S.Schnuelle and M.M.Kelly in "Molecular and Quantum Pharmacology", p.153 (7^{th} Jersalem Symp., Reidel, Dordrecht, 1975)

146. O.Tapia and O.Goscinski, Mol. Phys. 291653 (1975)

147. A.Pullman and B.Pullman, Quart. Rev. Biophys. 7 505 (1975)

148. J.L.Burch, K.S.Raghuveer and R.E.Christoffersen in "Environmental Effects in Molecular Structure and Properties" p.17 (8^{th} Jerusalem Symp., eds. B.Pullman and E.D.Bergmann, Reidel, Dordrecht, 1976)

149. L.Carozzo, G.Corongiu, C.Petrongolo and E.Clementi, J.Chem.Phys. 68 787 (1978)

150. E.Clementi "Lecture Notes in Chemistry" (Springer-Verlag, New York 1980)

151. P.Otto, J.Ladik, K.Laki and A.Szent-Gyorgyi, Proc. Natl. Acad. Sci. U.S.A. 75 3548 (1978)

152. S.Bone and R.Pethig in [14] p.83

153. M.P.Schubert, J.Biol.Chem. 114 341 (1936)

154. P.R.C.Gascoyne, Int. J. Quantum Chem. Quantum Biol. Symp. 7 93 (1980)

155. P.R.C.Gascoyne, Int. J. Quantum Chem. Quantum Biol. Symp. 8 to be published (1981)

156. A.Freestone and C.Thomsom, unpublished work

157. J.Eden, P.R.C.Gascoyne and R.Pethig, J.Chem. Soc. Faraday 1 76 426 (1980)

158. S.Suhai, J.Kaspar and J.Ladik, Int. J. Quantum Chem. 17 995 (1980)

159. J.Ladik in [14] p.51

160. S.F.Abdulner in [14] p.195

161. J.R.Ball and C.Thomson in [14] p.143

162. J.R.Ball and C.Thomson, J. Comp. Chem. 1 275 (1980)

163. J.R.Ball calculations on excited states of methylglyoxal (to be published)

164. J.R.Ball and C.Thomson, J. Comp. Chem. 1 400 (1980)

165. A.Meister in "Metabolism of Sulphur Compounds" (ed. D.M.Greenberg, Academic Press, New York, 1975)

166. A.Meister and S.S.Tate, Ann. Rev. Biochem. 1976 559

167. S.Colowick, A.Lazaro, E.Racker, D.R.Schwarz, E.Stadtman and H.Waelsch "Glutathione" (Academic Press, New York, 1954)

168. J.de Rey-Pailhade, C.R.Acad. Sci. 106 1683 (1888)

169. J.de Rey-Pailhade, C.R.Acad. Sci. 107 43 (1888)

170. F.G.Hopkins, J.Biol. Chem. 84 269 (1929)

171. C.R.Harington and T.H.Mead, Biochem. J. 29 1602 (1935)

172. W.B.Wright, Acta Cryst. 11 632 (1958)

173. F.E.Cole, Am. Cryst. Assoc. Summer Meeting 1970 p.34

174. S.V.Zenin, G.I.Chuprina and A.Y.Krylova, Zh. Obshchci Khimii 45 1337 (1975)

175. S.Fujiwara, G.Formicita-Kozlowsita and H.Kozlowski, Bull. Chem. Soc. Japan 50 3131 (1977)

176. B.Pullman and A.Pullman, Adv. Protein Chem. 28 347 (1974)

177. K.Chung, R.M.Hedges and R.D.MacFarlane, J.Am. Chem. Soc. 98 7523 (1976)

178. A.Immamura, H.Fujita and C.Nagata, Bull. Chem. Soc. Japan 42 3118 (1969)

179. W.Oergle and J.Sabin, J.Mol. Struct. 15 131 (1973)

180. R.Caballol, J.Riera and R.Carbo, J. Chim. Phys. Physicochim. Biol. 70 150 (1973)

181. E.Clementi, F.Cavallone and R.Scordamaglia, J.Am. Chem. Soc 102 5531 (1977)

182. M.Sellars and L.Schafer, J.Am. Chem. Soc. 100 7728 (1978)

183. L.Schafer, H.L.Sellars, F.J.Lovas and R.D.Suenram, J. Am. Chem. Soc. 102 6566 (1980)

184. S.Vishveshwara and J.Pople, J.Am. Chem. Soc. 99 2422 (1977)

185. L.R.Wright and R.R.Borkman, J.Am. Chem. Soc 102 6207 (1980)

186. R.Lavery and B.Pullman, Int. J. Quantum Chem. Quantum Biol. Symp. 6 467 (1979)

187. IUPAC-IB Commission of Biochemical Nomenclature, J.Mol. Biol. 52 1 (1970)

188. B.Pullman and B.Maigret, Theoret. Chim. Acta 35 113 (1974)

189. H.A.Scheraga, Adv. Phys. Org. Chem. 6 103 (1968)

190. J.A.Pople and M.Gordon, J.Am. Chem. Soc. 89 4253 (1967)

191. R.D.Brown, P.D.Godfrey, J.W.V.Storey and M.P.V.Bassez, J. Chem. Soc. Chem. Comm. 1978 547

192. R.D.Suenram and F.J.Lovas, J.Mol. Spect. 72 372 (1978)

193. R.D.Suenram and F.J.Lovas, J.Am. Chem. Soc. 102 7180 (1980)

194. C.Thomson calculations on fluorocarbons (to be published)

195. J.J.Kaufman, Int. J. Quantum Chem. Quantum Biol. Symp. 1 197 (1974)

196. A.Sequira, H.Rajagopal and R.Chidambaram, Acta Cryst. B28 2514 (1972)

197. B.F.Speilvogel, M.K.Das, A.T.M^{c}Phail and K.D.Onan, J.Am. Chem. Soc. 102 6343 (1980)

198. A.Szent-Gyorgyi, Ann. Rev. Biochem. 32 1 (1963)

199. S.Lewin "Vitamin C: its Medical Biology and Medical Potential" (academic Press, London, 1976)

200. D.E.Metzler "Biochemistry" p.621 (Academic Press, New York, 1977)

201. L.Pauling "Vitamin C and the Common Cold" (Freeman, San Francisco, 1970)

202. I.Stone "The healing Factor: Vitamin C Against Disease" (Grosset, New York, 1972)

203. E.Cameron and A.Campbell, Chem. Biol. Interact. 9 285 (1974)

204. E.Cameron, A.Campbell and T.Jack, Chem. Biol. Interact. 11 387 (1975)

205. E.Cameron and L.Pauling "Cancer and Vitamin C" (Linus Pauling Institute of Science and Medicine, California, 1979)

206. E.T.Creagan, C.G.Moertel, J.R.O´Fallan, A.J.Schutt, M.J.O´Connell, J.Rubin and S.Frank, New England J. Med., 27th Sept. 1979

207. C.H.Park, M.Amare, M.A.Savin and B.Hoogstraten, Cancer Res. 40 1062 (1980)

208. W.F.Benedict, W.H.Wheatley and P.A.Jones, Cancer Res. 40 (1980)

209. G.Kallistratos and E.Fasske, J.Cancer Res. Clin. Oncol. 96 91 (1980)

210. K.M.Bansal, M.Schoneshofel and M.Gratzel, Z.Naturforsch. 28B 528 (1973)

211. A.Aldaz and A.M.Alquie, J.Electroanal. Chem. Interfac. Electrochem. 47 532 (1973)

212. H.J.Bielski, D.A.Comstock and R.A.Bowen, J.Am. Chem. Soc. 93 5624 (1971)

213. Y.Kirino and T.Kwan, Chem. Pharm. Bull. 19 718 (1971)

214. G.P.Laroff, R.W.Fessenden and R.H.Schuler, J.Am. Chem.Soc. 94 9062 (1972)

215. B.H.Beilski, A.O.Allen and M.Schwarz, J.Am. Chem. Soc. 103 3516 (1981)

216. M.A.Schuler, K.Bhatia and R.H.Schuler, J.Phys. Chem 78 1063 (1974)

217. P.S.Duke, Exp. Mol. Path. 8 112 (1968)

218. N.J.F.Dodd, Br. J. Cancer 28 257 (1973)

219. J.Hvoslef, Acta Cryst. B24 23, 1431 (1968)

220. J.Hvoslef, Acta Cryst. B25 2214 (1969)

221. Y.Ogata and Y.Kosugi, Tetrahedron 26 4711 (1970)

222. J.H.Billman, S.A.Sojka and P.R.Taylor, J.Chem.Soc. Perkin Trans 2 2034 (1972)

223. S.Berger, Tetrahedron 33 1587 (1977)

224. T.Radford, J.G.Sweeny and G.A.Iocobucci, J.Org. Chem. 44 658 (1979)

225. E.Flood and P.Skancke, Acta Chem. Scand. 27 3069 (1973)

226. P.Bischof, M.Eckert-Maksic and Z.B.Maksic, Z.Natursforsch. 36A 502 (1981)

227. M.Estrada and G.DelConde, Int. J. Quantum Chem. Quantum Biol. Symp.

6 89 (1979)

228. G.L.Carson, H.Cable and L.G.Pedersen, Chem. Phys. Lett. 38 75 (1976)

229. C.Thomson, J.Mol. Spect. 67 133 (1980)

230. P.Otto, J.Ladik and A.Szent-Gyorgyi, Proc. Nat. Acad. Sci. U.S.A. 76 3849 (1979)

231. C.Thomson, J.Biol. Phys. 7 39 (1979)

232. W.N.Haworth and E.L.Hirst, Ergenbnisse der Vitamnin und Hormonefors. 2 160 (1938)

233. P.R.C.Gascoyne personal communication

234. P.Otto, S.Suhai and J.Ladik, Int. J. Quantum Chem. Quantum Biol. Symp. 4 451 (1977)

235. C.Thomson unpublished work

236. T.P.Molloy and C.W.Wilson, Int. J. Vit. Nutr. Res. 50 380 (1980)

237. E.V.Shtamm, A.P.Purmal and Yu.I.Skurlatov, Z.Fizich. Khim. 48 2233 (1980)

238. P.Martinez and D.Uribe, J.Chim. Phys. 78 47 (1981)

239. I.Fischer-Hjalmars and A. Henriksson-Enflo, Int. J. Quantum Chem. 18 409 (1980)

240. D.Perahia, A.Pullman and B.Pullman, Theoret. Chim. Acta 42 23 (1976)

241. O.Demoulin, I.Fischer-Hjalmars and A.Henriksson-Enflo, Int. J. Quantum Chem. 12 (supplement 1) 351 (1977)

242. A.Dedieu, M-M.Rohmer, A.Veillard in "Metal-Ligand Interactions in Organic Chemistry and Biochemistry" (eds. B.Pullman and N.Goldblum, Reidel, Dordrecht, 1977)

243. B.Roos, A.Veillard and G.Vinot, Theoret. Chim. Acta 20 1 (1970)

244. S.Topiol, R.Osman and H.Weinstein in "Quantum Chemistry in

Biomedical Sciences" p.17 (eds. H.Weinstein and J.P.Green, New York Academy of Sciences, New York, 1981)

245. P.R.C.Gascoyne work in progress

246. R.Damadian, Hosp. Practice, July 1977 p.63

247. R.Damadian, K.Zaner, D.Hor, T.D.Maro, L.Minkoff and M.Goldsmith, Ann. New York Acad. Sci. 222 1048 (1973)