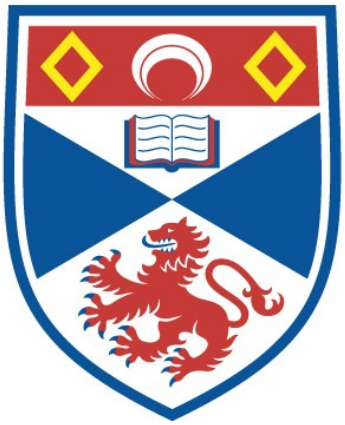
STRUCTURE ACTIVITY CORRELATION STUDIES OF ANTI-TUMOUR AGENTS BASED ON FLAVONE ACETIC ACID

Pierangela Sedda

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



1992

Structure Activity Correlation Studies of Anti-tumour Agents Based on Flavone Acetic Acid

A thesis presented for the degree of Doctor of Philosophy in the Faculty of Science of the University of St. Andrews

By

Pierangela Sedda

27-Jan-1992

DECLARATION

I Pierangela Sedda hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfilment of any other degree of professional qualification.

Signed Date 22.01.92

I was admitted to the faculty of Science of the University of St. Andrews under Ordinance General No 12 on 01 - Oct '88 and as a candidate for the degree of Ph. D. on Theoretical Chemistry

Signed Date 22.01.92

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate to the degree of Ph. D.

Signature of Supervisor Date 27.1.92

COPYRIGHT

ACKNOWLEGEMENTS

I would like to thank my supervisor Dr. C. Thomson for giving me the opportunity to work on this project and the National Foundation for Cancer Research for financial support. I am also grateful to the University of St. Andrews for the provision of research and library facilities and to my colleagues in the research group for their support and inspiration. Of the many people I am grateful to I would like to make a special mention of Derek Higgins and Michael Charlton for their help in unraveling so many computational problems. Alan Aitken, Ken Harris, Chris Glidewell and Joe Crayston, all provided an input to my work from which I have benefited. A special thanks goes to my parents, my family, my husband Peter, my friends Sergio & Anita, Paola Scano, and my teacher and mentor Pino Saba for their encouragement throughout the project. Mandy, Joy, Tricia, Colin, Graham, Andy, and all the other friends from University Hall, helped to make my stay in St Andrews memorable and I am grateful to them for being close to me, especially when I first arrived in St Andrews. Finally I would like to thank Audrey Lees for being such a wonderful companion, my father in law Dick Lukes, and all those who helped me in the final stages of this work.

To my father

ABSTRACT

In the first part of this thesis (chapters 1-8) the structure/activity (S/A) correlation studies on a class of anti-cancer drugs based on flavone acetic acid (FAA) by means of computer modelling techniques are reported. In particular, semiempirical and ab-initio quantum mechanical calculations have been performed on ten FAAs whose experimental anti-cancer activity was known. The results show that some calculated properties such as bond lengths, atomic charges, energies of the HOMO and the atomic orbitals involved in its formation, correlate with the anti-tumour activity. The correlations found were then used on another 38 molecules analogous to FAA whose anti-cancer activity had also been measured and of the 21 active molecules, 20 were predicted to be active by these SA correlations (95% success rate). From this study it also emerged that the pyrone ring may be directly involved in the anti-tumour mode of action of the FAA and it is suggested that vitamin-K may also play a role.

The second part (chapter 9) is a study of the dependence of the molecular electrostatic potential on the basis set. From this study it emerged that GEOSMALL and MINI-1 minimal basis sets produce MEPs that are more similar to those obtained with the 6-31G** basis set than the MEP obtained with the STO-NG basis set. GEOSMALL and MINI-1 also give better energies and better properties than STO-NG and their use is recommended when properties of large organic molecules are of interest. Also, from this study it emerged that the use of Mulliken charges for the calculation of the MEP with the point charge approximation is not advisable for it may lead to very different pictures of the electrostatic potential calculated directly from the wave function.

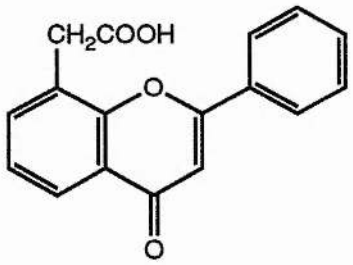
TABLE OF CONTENTS

CHAPTER 1:

Introduction

Introduction

The aim of the work reported in this thesis is to find structure-activity correlations for a class of anti-cancer drugs analogous to flavone acetic acid (FAAs) by means of computer modelling techniques.



Flavone Acetic Acid

Once correlations between the biological activity and the computed quantities have been found, new anti-cancer drugs with a mechanism of action similar to FAA but more powerful and more effective than FAA itself may be designed.

This project is financed by the NFCR (National Foundation for Cancer Research) and is part of a larger AICR (Association for International Cancer Research) project whose aim is to study the mechanism of action of FAA by *in vivo* (on the whole animal) and *in vitro* (in cell culture) studies of a large number of new synthetic analogues of FAA predicted to be active by computer modelling techniques. The synthesis of these compounds is done at the University of St. Andrews by Dr R. A. Aitken and Dr Sharma while the preclinical tests are conducted by Dr J. A. Double, Dr M. C. Bibby and Dr R. M. Phillips at the University of Bradford.

The interest in FAA and its derivatives arises from the fact that FAA has shown unique effects against experimental solid tumours, in particular the colon adenocarcinoma 38 (Co38) (a malignant tumour of the epithelial tissue on the colon) which is particularly resistant to conventional anti-cancer drugs. Furthermore, its toxicity is very limited and no side effects similar to those associated with conventional anti-cancer drugs (e.g. bone marrow damage) have been observed. However, FAA itself has not shown any

significant clinical activity and it is particularly important to develop new drugs which are able to treat solid tumours in humans. FAA represents a particularly interesting lead compound from which more effective drugs could be developed. To this end, a knowledge of the mode of action of FAA at the molecular level would be very useful. However, the mode of action of FAA still remains uncertain although FAA has been shown to stimulate the immune response of the host and to cause shut down of the blood vessels to the tumour (see chapter 2).

When the mode of action is unknown, the only route to drug development is via structure-activity relationships (SAR) of related compounds. Some SAR's have already been reported in the literature for FAAs and also for XAAs (Xanthenone-4-Acetic Acid analogues). The later are a class of compounds that may act similarly to FAAs, while requiring lower doses in order to achieve the same level of activity (i.e. XAAs are more powerful than FAAs). The SAR studies (which are discussed in chapter 2) suggest that a narrow structure-activity relationships exists between the biological anti-cancer activity and electronic or conformational properties of this class of compounds. This relationship is, however, not very clear and further studies are necessary.

In this thesis, SAR among FAAs and XAAs are determined by means of computer modelling techniques. In particular, quantum mechanical calculations have been performed. These calculations will point to more active compounds than FAA and also will shed some light on their mechanism of action by suggesting which parts of the molecule are essential for the activity.

This introductory chapter gives a very brief introduction to cancer and cancer therapy, and an overview of how new drugs are generally discovered and developed, in particular how quantum mechanical methods can be used in drug development and their advantages and disadvantages. Chapter 2 describes FAAs and XAAs and describes hypotheses on the mechanism of action and the previous structure-activity studies which contain most of the activity data that have been used in this work. Chapter 3 deals with the basic theory of quantum mechanics relevant to this work the approximations used in

quantum chemistry and the calculation of molecular properties. Chapter 4 will describe the programs used and the computer facilities available in St. Andrews. Chapters 5,6, 7 and 8 will report the results obtained using semi empirical and ab-initio methods respectively and chapter 9 reports a study of the dependence of the molecular electrostatic potential (MEP) on the basis set used. The aim of this study was to see if a minimal basis set could give values of the MEP similar to those obtained with an extended basis set.

1.1 Cancer

Cancer is a disease that, after heart diseases, is the second major cause of death in modern societies. What causes it and how it develops is still under intense study and can not be summarised in a few words. A detailed knowledge of cancer is not necessary to follow the work in this thesis, and a more comprehensive description than that given here can be found in many books that have been written on the subject [1].

The human body contains millions of cells; some of them divide to produce new cells that, in turn, grow and divide again. Other cells stop dividing and develop specialized structures and functions, in a process called 'differentiation'. Cells which die are replaced by the division of other cells. In a healthy organism, there is an equilibrium between cell division and differentiation which is maintained by natural control mechanisms; cells grow or differentiate in response to the signals of hormones and other chemicals produced inside the body. If cells fail to differentiate properly and ignore the environmental signals and continue to divide, they form an abnormal mass of tissue called a tumour. Tumours can be benign or malignant and if the tumours are malignant they develope a disease called cancer. Cancer is classified according to the tissue from which it arises. For example a colon adenocarcinoma is a malignant tumour of the epithelial tissue of the colon.

A characteristic of cancer cells is their ability to invade surrounding tissues and thus disturb normal tissue structure and functions. They may also break away from the

point in which they originated and circulate around the body, through the blood or the lymph fluid, until they settle somewhere and grow into secondary tumours; this process is called 'metastasis'. Another important characteristic of cancer cells is that they ensure their own supply of oxygen and nutrients (necessary for their survival) by producing special substances which induce the growth of new capillary blood vessels into the tumour itself. The growth of new blood vessels is extremely rare in normal tissues. One of the main causes of death of the tumour cells treated with FAA seems to be the shut down of these new blood vessels (see chapter 2).

1.2 Cancer therapy

If the tumour is localized, it can be removed by surgery or a combination of surgery and radiation therapy. These treatments, however, cannot be used to cure advanced stages of cancer after metastasis, when cancer cells are found in different parts of the body. In this case, chemotherapy may be more successful for certain types of tumours. The chemotherapeutic approach is to use chemicals to selectively kill cancer cells while causing the minimal damage to normal cells. The target of most of the anti-cancer drugs is the metabolism because this is where cancer and normal cells differ the most. Unfortunately, anti-cancer drugs whose target is the metabolism, also tend to destroy fast-dividing normal cells such as those in the gastrointestinal tract, hair follicles, lymphocyte's, etc. causing the well known side-effects of chemotherapy such as nausea, hair loss, increased susceptibility to infections, etc. Drugs able to cure one type of cancer are not necessarily effective against others. Certain types of cancer such as leukaemias, Hodgkin's disease, choriocarcinoma etc, are chemosensitive and may be curable in a high percentage of cases by many different anti-cancer drugs. Most of the common solid tumours in contrast, are quite chemoresistent and are not curable with conventional anti-cancer drugs. Since its discovery, FAA was found to be particularly interesting because of its peculiar activity against solid tumours and the lack of the typical side effects shown

by the conventional anti-cancer drugs. A review of anti-cancer drugs is given by Pinedo [2].

1.3 Anti-cancer drug design.

The normal procedure for anti-cancer drug discovery which has been followed has been that of random screening. That is, a large range of compounds are tested against specific tumours (*in vitro* and *in vivo*) with the hope that some of them will manifest some activity. This is a very costly procedure and its effectiveness depends very much on the choice of the system used. For instance, the anti-cancer drugs known at the moment are not generally active against solid tumours and only recently the National Cancer Institute (NCI) screening program has been modified to include tests on solid tumours [3].

A rational approach to drug discovery may come from an in depth knowledge of metabolic pathways and biochemistry. It may be possible for instance, to design a compound able to act as an enzyme inhibitor of an unwanted specific reaction. To this end however, the structure of the active site of the enzyme and the reaction in question have to be known.

Once a new drug has been discovered (lead compound), then there is another very important step to be taken in drug design, that is the 'drug development'. Drug development is necessary to establish the optimum structure of the drug and it generally leads to drugs which are more effective and less toxic than the lead compound itself. The strategy is to synthesize a large number of analogs of the lead compound, and study the substituent effects on the biological activity in order to find structure-activity relationships. The oldest approach to SAR studies is statistical, and establishes a correlation between the biological activity and measured properties associated with the chemical structure of the molecule. Measured properties used are, for instance, the Hammett index (of the acidity or basicity of substituted aromatic compounds) and the

Hansch parameter which has been extensively developed to take into account the lipophilicity of the substituents [4]. Many drugs have been optimised by using these and other indexes [4]. These methods are very expensive because they require the synthesis of a large number of compounds and the measurement of the properties. SAR studies of this kind have also been done for the XAAs (see chapter 2).

The advent of modern digital computers has made possible the calculation of properties of large molecules of biological interest by the methods of theoretical chemistry. Computer modelling in drug design is extensively used in SAR studies both by Universities and the pharmaceutical industry. Theoretical chemistry can be used to calculate the conformation and electronic properties of drugs, interaction energies between a drug and a receptor, and reaction mechanisms with enthalpies and entropies of reaction for suitable systems. Moreover, computer graphics provide a means to display the three dimensional structures of molecules such as enzymes, and many other biological compounds, as well as the visualisation of other properties such as frontier orbitals, electrostatic potentials, etc., that may be important for the biological activity.

The main techniques of theoretical chemistry which are used are empirical energy calculations and quantum mechanical methods. The first are very effective in conformational studies, particularly for systems with many degrees of rotational freedom, and they require small computer resources. In molecular mechanics, for example, a molecule is treated as being made of balls (atoms) and vibrating springs (bonds). The energy of the molecular system can be calculated given a set of parameters related to the strength of the different bonds in the molecule. Quantum mechanical methods, on the other hand, are based on the Schrodinger equation and can be divided into *ab-initio* and semi empirical methods (see chapter 3). Their use in drug design is particularly important because the calculation of electronic properties implicated in the reactions of the drug with the biological environment can only be achieved using quantum mechanical methods. Even if more expensive in term of computer resources, conformational studies may also be conducted by quantum mechanical methods with the advantage that changes in electronic properties associated with different conformations can also be calculated.

In this work, quantum mechanical methods have been used in preference to those of molecular mechanics. The reason for this is that it seems probable that the differences in activity of the FAAs are due mainly to differences in electronic properties rather than just in their conformation as can be seen from the table of the anti-cancer activities reported in chapter 2.

1.4 Quantum mechanical methods in drug design.

When applying quantum mechanical methods to pharmacological problems, it is important to understand that the model used contains many approximations. It is important to recognise the limitations of the method in order to analyze the results from the right perspective and obtain the most useful information that the method can offer. To ignore, or to underestimate these limitations may result in a misleading interpretation of the data and sometimes to erroneous conclusions. An introduction to quantum pharmacology is given by Richards [5].

The main point to be considered is that the results of a calculation are generally valid for a molecule in the gas phase. Although sometimes it is possible to include in the calculation the effect of a solvent (generally water), this will never give a model which is close enough to the real situation of a drug acting in the biological environment of a living organism. Furthermore, in a biological environment, enzymes can make possible some chemical reactions that would never occur in the gas phase. During the series of events that lead to the pharmacological effect it is possible, for instance, that the molecule responsible for the particular effect is no longer the molecule that originally entered the organism but some other metabolite.

Another point that should be taken into consideration when analysing the results of a calculation is the particular quantum mechanical method used to obtained those results. As is known, only approximate solutions of the Schrodinger equation can be obtained. In theory, molecular properties can be calculated to an arbitrary level of accuracy (sometimes

better that the experimental data) by increasing the basis set used (see chapter 3) and including in the calculation the effect of electronic correlation (see chapter 3). In practice, however, such calculations are very expensive in terms of computer resources and for molecules of biological interest, *ab-initio* calculations can be attempted only if powerful computers are available and the computer programs are effective in handling the enormous number of integrals necessary for the calculation. Even the best calculations on large molecules that can be performed at the moment are still very approximate and their predictions are not sufficiently reliable to be accepted without question. When possible, comparison of the calculated properties with experiment or with high-quality calculations should be made. Unfortunately, because of a lack of data it has not been possible to compare the results of the calculations with experimental properties of the FAAs. However, the calculated geometries of some flavones have been compared with their X-ray structures and also, the ^{13}C NMR chemical shifts of eight flavones have been related to the calculated atomic charges (see chapter 5). During the course of this study some experimental properties available for small molecules have been compared with the calculations but they are not reported here. A detailed study of the dependence of the molecular electrostatic potential on the basis set has been made with a smaller molecule (chapter 9).

This work required a large number of calculations for each molecule studied. In particular, the potential energy surface of these compounds would have been impossible to examine without the use of semi empirical methods. While the quality of an ab-initio method can be improved by enlarging the basis set, the quality of a semi empirical calculation is determined by the basic approximations and by the way the parametrization has been conducted. Among the different semi empirical methods, the AM1 and PM3 methods based on the NDDO approximation have been chosen because they have been parametrised to reproduce experimental heats of formation and they also succeed in predicting molecular properties on a large range of organic molecules (See chapter 3).

To summarize, the results of quantum mechanical calculations depend on the method used. The calculation of a specific molecular property using different methods

may give different values, therefore it is important that the comparison is done with properties calculated with the same method. For the same reason a calculation, or a series of calculations, on a single molecule would not be useful to explain the particular activity observed; this is also due to the fact that we do not know if the active species is the drug itself or a metabolite. Instead, more information can be obtained from the comparison of properties calculated with the same level of accuracy of a series of related molecules.

Perhaps the main advantage of using quantum mechanics in SAR studies is that, once a correlation has been found, then for any new compound the calculation can be performed before its synthesis and the pharmacological testing can be avoided if the molecule is predicted to be inactive from the calculation. This results in a saving of money and of animals that need to be used in preclinical studies.

Furthermore, there may be some properties that are important for the activity but that cannot be investigated in any other way than by quantum mechanical calculations. For example, a particular non-equilibrium conformation or the charge on a particular atom etc.

In the calculations reported in this work, we seek correlations between the pharmacological activity and factors related to the electronic charge distribution of a series of (chemically) similar molecules with a range of *in-vivo* biological activity. Presuming that all of these share the same mode of action in their anti-tumour activity, we look for any trend in a particular calculated property that could explain the difference in the ultimate biological activity observed for the series of molecules taken into consideration. If the properties are calculated with the same level of accuracy then, quantum mechanics provides a powerful method to be used in molecular pharmacology. If a correlation is found then, based on this information, other active molecules can be designed and they should succeed in the goal of being more effective than the lead compound. All of the other interpretations that one may be tempted to make, given the amount of information produced by a calculation, should be regarded with caution, in particular if the calculation has been conducted with a semi empirical method or with a small basis set.

CHAPTER 2:

Flavone Acetic Acid and Related Molecules

Introduction

Flavone acetic acid is a synthetic flavonoid synthesized by LIPHA pharmaceutical (France) [6]. It was selected for clinical studies by the drug development program of the National Cancer Institute (NCI) because of its preclinical activity against a broad spectrum of solid tumours, especially the highly resistant colon adenocarcinoma 38 [7,8]. The goal of this new drug screening program of the NCI is to select new anti-cancer drugs active against solid tumours that, as mentioned previously, are very resistant to conventional chemotherapeutic agents.

Although FAA shows pronounced activity against advanced solid murine tumours (i.e. solid tumours in rats or mice) [8] as well as against human tumours transplanted in mice [10], clinical trials have proven unsuccessful [21,22,28]. The reason why FAA is inactive in man is still unclear [23,24,28].

2.1 A hypothesis on the mechanism of action of FAA

FAA is active against many different solid tumours (in mice) but not against leukemias [8]. *In-vitro* studies on human colon adenocarcinoma cell lines suggest that FAA does not act in the same way as conventional anti-cancer drugs, it does not act by causing substantial DNA breakage neither does it bind to DNA [9]. FAA is active both *in-vitro* and *in-vivo* although its activity is greater *in-vivo* than *in-vitro* [8,9]. This suggests that FAA acts with more than one mechanism of action.

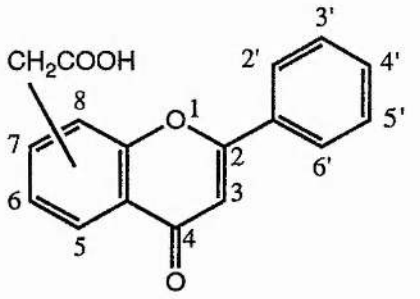
In-vivo studies proved that FAA does not act directly on the tumour cell but a host cellular component is necessary for the activity [10]. (see also Ref.[11] for a discussion). Different authors have suggested that FAA acts by stimulating the immune response [4] of the host and the latter would be responsible for the death of the tumour [10, 12,13,14,15]. Recent studies also suggest that one of the main components of the mode of action of FAA is its ability to cause the shut down of blood vessels inside the

tumour leaving it without nutrients and oxygen [16,17]. In particular, FAA activates the blood clotting cascade [18] after 15 minutes of drug administration giving rise to thrombus formation in the blood vessels [20].

2.2 Structure - Activity studies on FAA

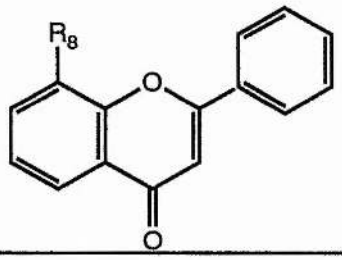
Previous structure-activity studies on FAAs were conducted by Atassi et al in 1984 when a range of derivatives of Flavone Acetic Acid (Fig. 2.1) where synthesized and their activity against colon adenocarcinama 38 (Co38) measured [19]. The anti-tumour activity was expressed as Tumour Growth Inhibition % (TGI %) and the minimum percentage requested by NCI to demonstrate a significant anti-tumour effect is 58 %. The 6- and 8- flavone acetic acids were prepared but only the 8-flavone acetic acid was active (98 %TGI), the other was completely devoid of Co38 anti-cancer activity. Throughout this work, the 8-Flavone Acetic Acid will be simply called FAA and the derivatives will be referred generally as FAAs.

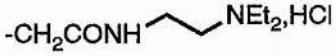
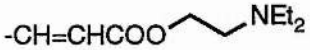
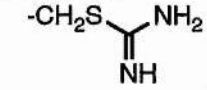
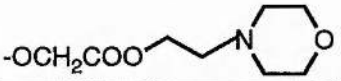
FIG. 2.1: Structure and numbering scheme of Flavone Acetic Acids.



Flavones substituted in the 8- position with groups other than the acetic acid were also prepared and tested; the results are reported in Table 2.1 and show that the compound **23**, which has a carbamidothioate function in position 8, had a significant activity. This may suggest that the acetic acid group is not strictly necessary for the activity.

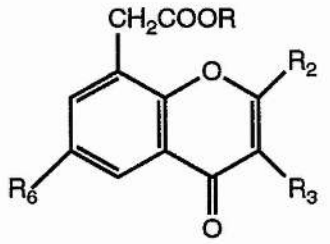
TABLE 2.1: Anti-tumour activity of some flavones substituted in 8- position by functions other than acetic acid. (Data from ref.[19])

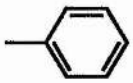
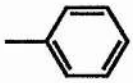
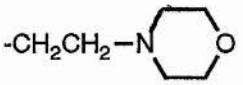
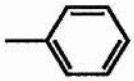
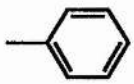
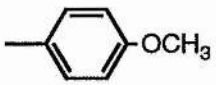
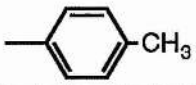


No	R_8	% TGI
19	$-CH_2CH_2COOH$	40
20	$-O-CH_2COOH$	43
21	$-CH_2CONH$ 	32
22	$-CH{=}CHCOO$ 	48
23	$-CH_2S$ 	77
24	$-OCH_2COO$ 	57

The esters of FAAs were also tested and the activity data are reported in Table 2.2. Atassi suggested that the activity could be correlated with the length and the nature of the ester function [19]. From this study, the compound **12** (Table 2.2) was selected for clinical evaluation but it was found to be toxic at the doses required for anti-tumour activity in human plasma. However it was also shown that the ester was hydrolysed to FAA in mouse plasma and the actual anti-cancer drug was FAA while the ester function was responsible for the toxicity [25].

TABLE 2.2: Anti-tumour activity of the FAAs esters. (data from ref. [19])



No	R_2	R_3	R_6	R	TGI %
12		H	H	$CH_2CH_2NEt_2$	100
13		H	H		100
14		H	H	$-(CH_2)_3$-NMe2	19
15		H	H	$CH_2CH_2NEt_2$	18
16		H	OMe	$CH_2CH_2NEt_2$	63
17		H	H	$CH_2CH_2NEt_2$	-*
18		H	H	$CH_2CH_2NEt_2$	-*

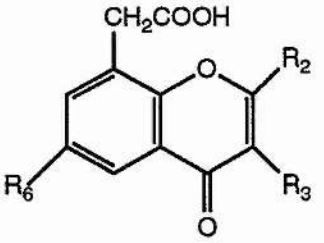
* Mentioned to be inactive in the text but no data is reported in the table on Co38 activity [19].

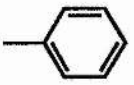
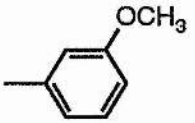
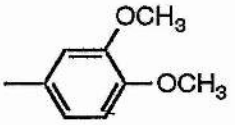
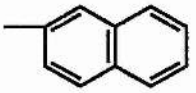
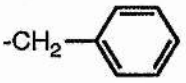
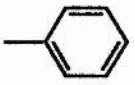
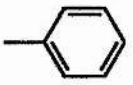
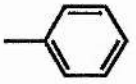
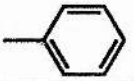
Finally, Table 2.3 reports activity data on FAAs (published in the same paper [19]) showing how the activity depends on the substituents in the R_2 position. Of the 10 compounds listed in the Table 2.3, six are active (**1-3**, **5**, **6** and, **8**) and, four are inactive (**4**, **7**, **9** and, **10**). This demonstrates that variations in the nature of the substituents in the R_2 position give rise to dramatic variation in the biological activity. However, on the basis of this data alone it is not possible to distinguish whether electronic or conformational effects are important in determining the anti-tumour activity of FAAs. For instance the introduction of an *m*-methoxyphenyl group into the R_2 position of FAA

(compound 2, Table 2.3) results in a compound of 100% activity, while the activity of a naphthyl substituted analog (compound 4, Table 2.3) is 41%. This effect could be ascribed to the steric bulk of the naphthyl derivative, inhibiting access to an active site, or to the mesomeric effect of the methoxy group in increasing the electron density in the ortho position. Quantum mechanical methods are particularly appropriate in this case because they allow simultaneous calculation of conformation, and electronic properties.

Previous calculations on the compounds listed in Table 2.3 were carried out in this laboratory using AM1 semi empirical methods [26, 27]. The results were encouraging enough to suggest that calculations should be extend to a larger number of compounds and the properties calculated with more sophisticated ab-initio methods.

TABLE 2.3: Anti-tumour activity of FAAs (Data from Ref. [19])

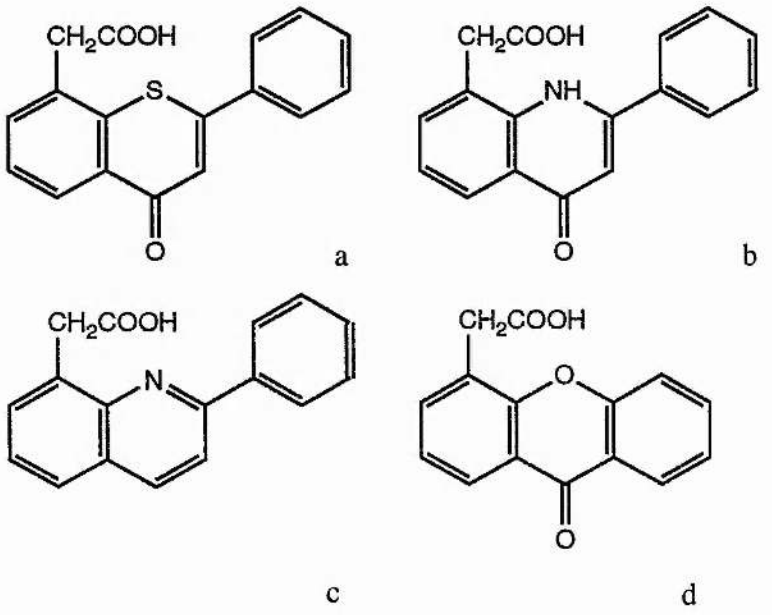


No	R_2	R_3	R_6	TGI %
1		-H	-H	96
2		-H	-H	100
3		-H	-H	70
4		-H	-H	41
5		-H	-H	100
6		-H	-H	97
7	$-CH_3$	-H	-H	0
8			-H	100
9		-H	-OH	0
10	-H		-H	24

Further structure activity studies on FAAs have been conducted by Atwell et al. [29]. Their results point out the importance of electronic effects in the anti-tumour activity

of FAAs suggesting that a narrow structure-activity relationship exists. Of the compounds reported in fig. 2.2, only xanthenone-4-acetic acid, **2d**, was found to be active; the compounds **2a-2c** which retain the same overall topology of FAA but differ in the electronic character were found to be inactive [29].

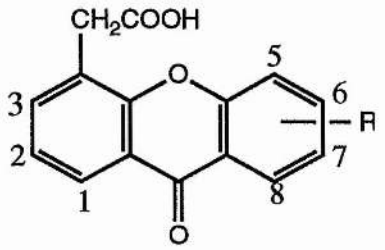
FIG. 2.2: Some of the analogues of FAA studied by Atwell et al [29].



2.3 Structure-Activity Studies on Xanthenone Acetic Acids (XAAs)

Xanthenone acetic acid and analogous were found to have anti-tumour activity against colon adenocarcinoma 38 in a similar fashion to flavone acetic acid [34,35]. They cause haemorrhagic necrosis of the tumour and stimulate the immune system as does FAA and it has been suggested that FAAs and XAAs share the same mode of action in killing cancer cells [34,35]. Extensive structure-activity studies on XAAs have been conducted by Atwell et al.[30-33] and some of their results are summarised in Table 2.4-2.6. In all the tables reported in this section, the symbols + and ++ mean that the treated tumour showed 50-90% or > 90% respectively of hemorrhagic necrosis; **OD** is the optimum drug dose expressed in mg/Kg (Details can be found in reference [30]). Particularly interesting is the observation of Atwell et al. that the potency of the compounds is related much more

TABLE 2.4: Co38 anti-tumour activity of some substituted XAAs (Data from Ref.[30]).

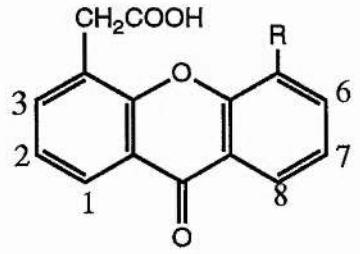


No	R	OD	Activity	No	R	OD	Activity
1	FAA	330	++	**24**	5-OH	750	-
11	XAA	220	++	**37**	6-Me	220	+
13	1-Me	150	+	**47**	6-OMe	150	+
15	1-OMe	500	-	**36**	6-Cl	150	+
12	1-Cl	330	++	**35**	6-OH	750	±
14	1-OH	750	-	**40**	7-Me	500	+
17	2-Me	500	+	**38**	7-OMe	500	+
19	2-OMe	750	+	**39**	7-Cl	500	+
16	2-Cl	330	+	**42**	7-OH	500	-
18	2-OH	750	-	**41**	7-NO_2	500	-
20	3-Me	150	+	**45**	8-Me	330	-
22	3-OMe	220	-	**43**	8-OMe	750	-
46	3-Cl	500	++	**44**	8-Cl	750	+
21	3-OH	500	-				
28	5-Me	45	++				
25	5-OMe	150	+				
23	5-Cl	150	++				
34	5-NO_2	220	+				

to the position than to the nature of the substituents, with 5-substituted compounds being the most dose potent (see table 2.4); for example, only 45 mg/kg of **28** or 150 mg/kg of **23** are required to produce the same level of activity of FAA for which the optimum dose is 330 mg/kg.

Extended structure-activity studies, using Hammet and Hansch parameters, have been conducted for a range of 5-substituted XAAs [31]. According to this study the activity broadly correlates with the lipophilicity and, in order to retain high activity and increase the potency it is necessary to have in the 5-position small, lipophilic substituents. The activity data is reported in Table 2.5 where the substituents are listed in order of increasing hydrophilicity. Compound **31,** for instance, is inactive because it contains the phenyl group which is very lipophilic but also very large. The ethyl group, or better, the methyl group are both lipophilic and small, and when inserted in the 5-position give rise to compounds (**30** and **28**) which are more potent than XAA and FAA while retaining the same activity.

TABLE 2.5: Co38 anti-tumour activity of some 5-substituted XAAs (Data from Ref.[31]).

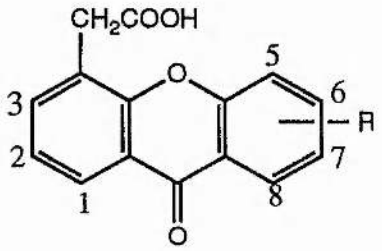


No	R	OD	Activity	No	R	OD	Activity
11	H	220	++	**27**	OPr	150	+
29	CF_3	220	+	**25**	OMe	150	+
48	Br	100	-	**49**	OBu	220	+
23	Cl	150	++	**34**	NO_2	330	+
31	Ph	750	-	**33**	NHCOMe	330	-
30	Et	100	++	**50**	CH_2COOH	330	-
28	Me	45	++	**24**	OH	330	-
26	OEt	330	++	**32**	NH_2	1125	+

Some disubstituted XAAs were found to be very active and, some also very potent [31] as reported in Table 2.6. In particular the 5,6 dimethyl derivative **68** is about ten

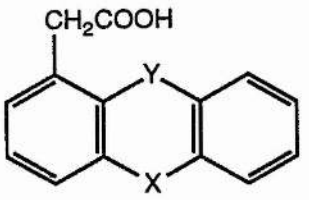
times more potent than FAA. This data confirms that the potency of the compounds can be increased by increasing the global lipophilicity of the molecules with small lipophilic groups in key positions. However, structure-activity relationships are still not clear. For instance, of the 30 compounds in Table 2.4 only five of them are very active (++) (**11, 12, 46, 28** and, **23**), fourteen of them moderately active (+) and ten completely inactive (-). Why does the methyl group give a compound which is very active when it is in 5-position (**28**) but a completely inactive one if it is in 8-position (**45**)?.

TABLE 2.6: Co38 anti-tumour activity of some Disubstituted XAAs (Data from Ref.[32]).



No	R	OD	Activity	No	R	OD	Activity
65	1,5-Me_2	65	++	**72**	5-Me,6OMe	30	++
66	2,5-Me_2	220	++	**73**	5-Me,6-NMe_2	45	-
67	3,5-Me_2	45	++	**74**	5,6-Cl_2	100	+
68	5,6-Me_2	30	++	**75**	5-Cl,6-Me	100	++
69	5-Me,6-F	45	++	**76**	5,6-$(CH_2)_3$	45	++
70	5-Me,6-Cl	30	++	**77**	5,6-benzo	65	++
71	5-Me,6-Br	45	++	**78**	6,7-benzo	330	+

Quantum mechanical calculations have been performed for most of the molecules in Tables 2.4-2.6 in an attempt to understand these correlations better and compare them with the results obtained for the FAAs. Very recently, a paper by Atwell et al. has been published reporting AM1 semi empirical calculations for some tricyclic acetic acid compounds of the following general formula:



X=CO,$COCH_2$,CH_2CO,O,CH_2,S,SO,SO_2,-CH=,-N=; Y=O,S,SO,SO_2,CO,-CH= and, -N=.

From this study it emerged that the direction of the dipole moment is important for the activity of these molecules [33]; this will be discussed later in chapter 8.

CHAPTER 3:

Theoretical Methods

Introduction

The calculations reported in this work use both *ab initio* and semi empirical methods. This chapter describes the approximations used to calculate the wave function of the molecules and their properties but is not intended to be exhaustive; a more comprehensive description of the theory involved in quantum chemical calculations can be found in one of the standard references [36]. This chapter is divided into three sections and section 3.1 contains the basic approximations used in order to solve the time-independent Schrodinger equation [37]

$$\mathcal{H}\Phi=E\Phi \tag{3.1}$$

where Φ is the wave function of the stationary state of energy E; the Hamiltonian operator of the system, $\mathcal{H}$, will be discussed later in this chapter. Section 3.2 describes further approximations used in semi empirical methods. Finally, section 3.3 describes the calculation of molecular properties that can be obtained from the wave function.

Unless otherwise specified, atomic units are used throughout this work. The atomic unit of length is generally referred to as the Bohr; it is defined as:

$$a_0 = \frac{h^2}{4\pi^2 me^2} = 0.529167 \times 10^{-8}\text{cm} = 0.529167\ \text{Å} \tag{3.2}$$

were h is Planck's constant. The atomic unit of the electric charge is the protonic charge: e= 4.80298 X 10^{-10} esu. The atomic unit of energy is the *Hartree*, it corresponds to the interaction energy of two unit charges separated by one Bohr radius:

$$\varepsilon_0 = \frac{e^2}{a_0} = 4.35942 \times 10^{-11}\text{erg} = 27.21\ \text{eV} = 627.5\ \text{Kcal/Mol} = 2.1948 \times 10^5\text{cm}^{-1} \tag{3.3}$$

The atomic unit of mass is the electron mass: m = 9.09191 x 10^{-28} g.

3.1 Approximate solution of the Schrodinger equation

3.1.1 The non-relativistic Hamiltonian operator

The operator corresponding to the total energy of the system is called the Hamiltonian operator and is denoted by $\mathcal{H}$. The non relativistic Hamiltonian operator for a molecule of N electrons and M nuclei (in atomic units) is:

$$\mathcal{H} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \qquad (3.4)$$

Were M_A is the ratio of the mass of nucleus A to the mass of an electron, and Z_A is the atomic number of nucleus A. ∇_i^2 and ∇_A^2 are called Laplacian operators and contain second derivative operators with respect to the coordinates of the *ith* electron and the *Ath* nucleus. The first term in Eq. (3.4) is the operator for the kinetic energy of the electrons; the second term is the operator for the kinetic energy of the nuclei; the third term represents the Coulomb attraction between electrons and nuclei; the fourth and fifth terms represent the repulsion between electrons and nuclei, respectively. A first approximation to the Hamiltonian operator is achhieved by keeping the last term constant, i.e the repulsion between nuclei, and in neglecting the kinetic energy of the nuclei. This is called the Born-Oppenheimer approximation [39] and is described in 3.1.2 below.

3.1.2 The Born-Oppenheimer approximation [39]

The justification for the Born-Oppenheimer approximation is that the mass of protons is about 1860 times bigger than the mass of electrons and therefore the nuclei can be thought to move much more slowly compared to the electrons. Within the Born-Oppenheimer approximation, the kinetic energy of the nuclei can be neglected and the

repulsion between nuclei can be considered to be constant for every molecular configuration (see eq. (3.4)). What is left from (3.4) is called the electronic Hamiltonian and can be written as:

$$\mathcal{H}_{elec} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} \tag{3.5}$$

The electronic wave function $\phi(r)$ is an eigenfunction of the electronic Hamiltonian

$$\mathcal{H}_{elec}\phi(r) = \varepsilon_{elec}\,\phi(r) \tag{3.6}$$

which is a function of the electronic coordinates and depends parametrically upon the nuclear coordinates. It describes the state of the molecule with fixed nuclear positions. The total energy is obtained by adding the nuclear repulsion energy to the electronic energy ε_{elec}. It also depends parametrically upon the nuclear coordinates and when it is calculated for different nuclear positions, defines the potential energy surface of the molecule.

$$\varepsilon_{tot} = \varepsilon_{elec} + \sum_{A=1B>A}^{M}\sum^{M}\frac{Z_A Z_B}{R_{AB}} \tag{3.7}$$

From now on (3.5) will be referred to as the Hamiltonian operator and (3.6) as the Schrodinger equation. Also, the electronic wave function will be simply called the wave function.

3.1.3 The antisymmetry principle

The antisymmetry principle states that the wave function for an N electron system must be antisymmetric with respect to the interchange of the coordinates (space and spin) of any two electrons i.e.

$$\phi(q_1,q_2,..,q_i,..q_j,..,q_N)=-\phi(q_1,q_2,..,q_j,..q_i,..,q_N)$$

(3.8)

where q_i denotes the four coordinates of space and spin collectively. The logic behind this principle is that the electrons are indistinguishable particles and therefore, no physical property of the system can be affected if the electrons are simply renamed or renumbered. The Pauli exclusion principle [40] derives naturally from the antisymmetry principle. The wave function of a N-electron system has to satisfy both equations (3.6) and (3.8).

3.1.4 The variational principle

Given a normalized N-electron wave function $\Psi(r_1r_2...r_N)$, the energy E is the expectation value of the Hamiltonian operator; i.e.

$$E = \int\int.....\int \Psi^*(r_1r_2....r_N)\ \mathcal{H}\Psi(r_1r_2...r_N)\ dr_1dr_2...dr_N = <\Psi|\mathcal{H}|\Psi>$$

(3.9)

where on the right hand side, the integral is expressed in the matrix notation of Dirac [38]. In this notation, which will be used throughout the rest of this work, the complex conjugation is implied for the left-hand elements enclosed in brackets.

The variational theorem states that the energy ε, calculated with an approximate wave function Φ is an upper bound to the energy calculated with the true wave function Ψ of the system; i.e.

$$< \Psi \mid \mathcal{H} \mid \Psi > = E \leq \varepsilon = < \Phi \mid \mathcal{H} \mid \Phi >$$

(3.10)

where $\varepsilon = E$ only in case $\Phi=\Psi$. This principle is used to obtain approximate wave functions with the Hartree-Fock method described later; in fact, if the wave function Φ includes parameters that can be optimized, then the 'best ' wave function (in a variational sense) is that for which all parameters have been optimized to yield the lowest energy.

3.1.5 The molecular orbital method and the Slater determinant

The Schrodinger equation can be solved exactly only for one electron systems, for other systems approximate solutions of eq. (3.6) are required. What makes the solution difficult is the electron repulsion term in the Hamiltonian operator (3.5) which depends on the instantaneous relative coordinates of the two electrons i and j. The idea behind the molecular orbital and the Hartree-Fock approximation (see later) is to write the wave function as a product of one-electron functions or a linear combination of them such that the Hamiltonian can be written as a sum of one-electron operators and it is possible to obtain solutions for the Schrodinger equation by a straightforward separation of variables.

The Molecular Orbital approximation [41] was introduced by Mulliken in 1928 in an attempt to construct a satisfactory many electron wave function from a combination of molecular spin orbitals (MSOs) dependent upon the coordinates of only one electron:

$$\psi_k(q_i)$$

The subscript k labels the different MSOs and q represents collectively spatial and spin coordinates. The Hamiltonian operator (3.5) does not contain spin coordinates therefore the MSO can be written as a product of a molecular orbital (MO) and a spin function. The MO being a function of only the spatial coordinates r_i of the electron i:

$$\varphi_u(r_i)$$

The spin function can be represented by two mutually orthogonal spin wave functions $\alpha(s_i)$ and $\beta(s_i)$. For closed shell systems containing 2n electrons a further approximation

consists of assigning two electrons with different spin functions to the same molecular orbital, that is, from each MO φ_u can be obtained two spin orbitals $\psi_u=\varphi_u\alpha$ and $\psi_{u+1}=\varphi_u\beta$. In all the calculations reported in this work, only closed systems have been considered and these approximations have always been used.

As mentioned earlier, the molecular wave function must be antisymmetric with respect to the interchange of the coordinates of any two electrons. A simple product of MSO does not satisfy this condition. For a 2n-electron system, with two electrons per spatial orbital, a suitable wave function can be written as a Slater determinant of the 2n spin orbitals involved [44]:

$$\phi(r_1 r_2 \ldots r_N) = (2n!)^{-\frac{1}{2}} \begin{vmatrix} \varphi_1(r_1)\alpha(s_1) & \varphi_1(r_1)\beta(s_1) & \ldots\ldots & \varphi_n(r_1)\beta(s_1) \\ \varphi_1(r_2)\alpha(s_2) & \ldots\ldots & & \\ \ldots\ldots & & & \\ \varphi_1(r_{2n})\alpha(s_{2n}) & \ldots\ldots & & \varphi_n(r_{2n})\beta(s_{2n}) \end{vmatrix}$$

(3.11)

where $(2n!)^{-1/2}$ is the normalization factor. Interchanging the coordinates of two electrons corresponds to interchanging two rows of the Slater determinant, which changes the sign of the determinant. Therefore a Slater determinant satisfies the antisymmetry principle. In a Slater determinant both the MSOs and the MOs are linearly independent, that is they can always be transformed to form an orthonormal set i.e.

$$\left\langle \psi_u(r_i) \mid \psi_k(r_i) \right\rangle = \left\langle \varphi_u(r_i) \mid \varphi_k(r_i) \right\rangle = \delta_{uk}$$

(3.12)

A short hand notation for eq (3.11) is

$$\phi = \left| \varphi_1(r_1)\alpha(s_1)\ \varphi_1(r_2)\beta(s_2)\ \varphi_2(r_3)\alpha(s_3)\ \ldots\ldots\ \varphi_n(r_{2n-1})\alpha(s_{2n-1})\ \varphi_n(r_{2n})\beta(s_{2n}) \right|$$

(3.13)

which only shows the diagonal elements of the Slater determinant and includes the normalization constant.

3.1.6 The Hartree-Fock approximation.

The Hartree-Fock (HF) approximation was introduced in 1930 by Fock [42] and it is based on a earlier work by Hartree [43]. In this approximation the variational principle is applied to a wave function written as a Slater determinant; by minimising its energy, a set of "effective" one-electron operators (the Fock operators [42]) is obtained and the many-electron problem is therefore simplified to a set of one-electron eigenvalue equations known as the Hartree-Fock equations. A brief outline of the HF approximation will be given here; a comprehensive description of this theory can be found in one of the standard references [36].

Substituting eq (3.13) into the expression for the energy eq (3.9) gives the following expression for the energy of a closed shell system:

$$E=2\sum_{u=1}^{n}H_u+\sum_{u=1}^{n}\sum_{k>u}^{n}(2J_{uk}-K_{uk}) \tag{3.14}$$

where

$$H_u=\left\langle \varphi_u(r_1) \mid h(r_1) \mid \varphi_u(r_1)\right\rangle \tag{3.15}$$

$$J_{uk}=\left\langle \varphi_u(r_1)\varphi_k(r_2) \mid \frac{1}{r_{12}} \mid \varphi_u(r_1)\varphi_k(r_2)\right\rangle=\left\langle \varphi_u(r_1) \mid J_k \mid \varphi_u(r_1)\right\rangle \tag{3.16}$$

$$K_{uk}=\left\langle \varphi_u(r_1)\varphi_k(r_2) \mid \frac{1}{r_{12}} \mid \varphi_u(r_2)\varphi_k(r_1)\right\rangle=\left\langle \varphi_u(r_1) \mid K_k \mid \varphi_u(r_1)\right\rangle \tag{3.17}$$

$$h(r_1) = \frac{1}{2}\nabla_1^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{1A}}$$

(3.18)

(3.18) is the Hamiltonian operator for an electron in the field of the nuclei alone. J_k is called the Coulomb operator and is defined by :

$$J_k(r_1) = \left\langle \varphi_k(r_2) \left| \frac{1}{r_{12}} \right| \varphi_k(r_2) \right\rangle$$

(3.19)

K_k is called the exchange operator and is defined by:

$$K_k(r_1)\varphi_u(r_1) = \left\langle \varphi_k(r_2) \left| \frac{1}{r_{12}} \right| \varphi_u(r_2) \right\rangle \varphi_k(r_1)$$

(3.20)

J_{uk} is called the Coulomb integral and describes the repulsion between electrons in φ_u and φ_K with probability density $|\varphi_u|^2$ and $|\varphi_k|^2$. K_{uk} is called the exchange integral. It does not have a classical analog but can be related to the correlation between pairs of electrons with parallel spin [36a]. According to the variational principle, eq (3.14) is minimised by varying the MOs with the restriction that they form an orthonormal set as expressed by eq (3.12). This type of constrained variational problem is solved mathematically by the calculus of variations using the method of Lagrange multipliers [36(a)]. This leads directly to n differential equations

$$\left[H_u + \sum_u (2J_u - K_u) \right] \varphi_k = \sum_u \varepsilon_{uk}\varphi_k \qquad k=1,2,.....,n$$

(3.21)

The quantity in square brackets is called the Fock operator, and eq. (3.21) can be written in the form:

$$F\varphi_u = \sum_k \varepsilon_{uk}\varphi_k \qquad k = 1,2,.....,n$$

(3.22)

This is not an ordinary one-electron eigenvalue equation because it contains a whole set of constants ε_{uk} instead of a single eigenvalue. Eq. (3.22) is brought into the form analogous to a standard eigenvalue problem by applying a unitary transformation to the orbitals such that the matrix of the lagrangian multipliers assumes a diagonal form, that is, all ε_{uk}=0 unless u=k:

$$F\varphi_u = \varepsilon_u\varphi_u \qquad u = 1,2,.....,n$$

(3.23)

These are the Hartree-Fock equations and state that the best molecular orbitals are eigenfunctions of the Fock operator F which is in turn defined in terms of these orbitals through the Coulomb and exchange operators. These equations can be solved by a iterative process: 1) a set of trial molecular orbitals is chosen; 2) Coulomb and exchange integrals are computed (eqs. (3.19) and (3.20)) and, 3) the Fock operator is constructed; 4)The Hartree-Fock equations (3.23) are solved. 5) the solution of the (3.23) is used as a new trial function and the procedure from 2 until 5 is continued until self-consistency is achieved. This procedure is called the Self-Consistent Field (SCF) procedure. The eigenvalues ε_u are the Hartree-Fock orbital energies. The n eigenfunctions φ_u which correspond to the lowest values of ε_u are the ground state orbitals. The Slater determinant obtained from the n ground state orbitals is the Hartree-Fock wave function Ψ_{RHF}.

The RHF approximation provides a suitable wave function for closed shell systems but it cannot be used if one or more electrons are unpaired such as in radicals and triplet states. In such cases, the constraint of assigning two electrons with different spin to the same spatial orbital has to be removed and all the orbitals optimized. This procedure is called Unrestricted Hartree-Fock (UHF) and produces lower energies than those obtained with the RHF method. This is clearly due to the increased flexibility of the UHF wave function and the possibility to optimize more parameters in the variational procedure. A

further limitation common to the HF methods is due to the assumption of regarding the electrons as independent particles moving in the average field generated by the other electrons. This is of course an unrealistic picture because the electrons do interact with each other. In writing the wave function as a Slater determinant, the correlation between electrons having the same spin is taken into account and this results in the appearance of the exchange integral in the Hartree-Fock equations. However, correlation between electrons having opposite spin is completely neglected; this correlation energy would decrease the energy obtained with the HF approximation. A common way to take into account the electron correlation, is to write the total wave function as a linear combination of Slater determinants obtained by considering not only the electronic configuration of the ground state, but also all the other possible configurations that arise from excitations to virtual orbitals. This procedure is known as 'interaction of configurations' or CI [36(a)]. The computer resources required to perform CI calculations are very large. For the molecules reported in this work it was not possible to conduct such calculations.

3.1.7 The LCAO-SCF method

For atomic calculations, the Hartree-Fock equations can be solved numerically. For molecular calculation however, the procedure is not easy and further simplifications are needed. In 1951, Roothaan showed that, by expanding the molecular orbitals in terms of atomic orbitals (or any set of spatial basis functions centred on each atom), the Hartree-Fock equations could be converted to a set of algebraic equations and solved by standard matrix techniques [45]. This method has the further advantage that the results of the calculation can be interpreted in terms of the constituent atoms of the molecules. In this approach, each molecular orbital is written in the form:

$$\varphi_u = \sum_{p=1}^{k} \chi_p c_{pu} \qquad (3.24)$$

where $\{c_{pu}\}$ are the expansion coefficients and $\{\chi_p\}$ a set of normalized AOs such that:

$$<\chi_p \mid \chi_p> = 1$$

(3.25)

Applying the variational procedure with the constraint that the LCAO-MOs form an orthonormal set, gives a set of coefficients c_{pu} for which the energy of the corresponding Slater determinant is a minimum. H, J_u, K_u, G and F are defined in terms of RHF-LCAO MOs and the final equation is the Roothann equations (in matrix form):

$$\mathbb{F}\mathbb{C}=\mathfrak{E}\mathbb{S}\mathbb{C}$$

(3.26)

were $\mathfrak{E}$ is a diagonal matrix of the orbital energies, $\mathbb{C}$ is a square matrix of the coefficients c_{pu} and $\mathbb{S}$ is the overlap matrix which arises because the AOs are not orthogonal; its elements are

$$S_{pq}=\langle \chi_p \mid \chi_q \rangle$$

(3.27)

$\mathbb{F}$ is the matrix representation of the Fock operator in the basis of the AOs with elements:

$$F_{pq}=\langle \chi_p \mid F \mid \chi_q \rangle = H_{pq}^{core} + G_{pq}$$

(3.28)

The elements of the core-Hamiltonian matrix H_{pq} are integrals involving the one-electron operator (3.18) describing the kinetic energy T and nuclear attraction V of an electron:

$$H_{pq}^{core} = T_{pq} + V_{pq}^{nucl}$$

(3.29)

$$T_{pq} = \left\langle \chi_p(r_1) \mid -\frac{1}{2}\nabla_1^2 \mid \chi_q(r_1) \right\rangle$$

(3.30)

$$V_{pq}^{nucl} = \left\langle \chi_p(r_1) \left| -\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} \right| \chi_q(r_1) \right\rangle$$

(3.31)

G_{pq} is the two-electron part of the Fock matrix:

$$G_{pq} = \sum_{kl} P_{kl} \left[\left\langle \chi_p(r_1)\chi_l(r_2) \left| \frac{1}{r_{12}} \right| \chi_q(r_1)\chi_k(r_2) \right\rangle - \frac{1}{2} \left\langle \chi_p(r_1)\chi_l(r_2) \left| \frac{1}{r_{12}} \right| \chi_k(r_1)\chi_q(r_2) \right\rangle \right]$$

(3.32)

Its elements are written in terms of the one-electron density matrix $\mathbb{P}$ which in the RHF-LCAO formalism is defined by:

$$P_{kl} = 2 \sum_{u=1}^{n\_occ} c_{ku} c_{lu}^*$$

(3.33)

where c_{ku} are the expansion coefficients of the LCAO-MOs (Eq. (3.24)) and the sum runs over the number of occupied orbitals.

The eigenvalues of eq (3.26) are obtained by solving the secular equation

$$\mathrm{Det}(\mathbb{F} - \mathfrak{S}\mathbb{S}) = 0$$

(3.34)

The Roothaan equations (3.26) and (3.34) are not linear as the Fock operator depends on its solutions, χ, through (3.32) and they are solved using the Self-Consistent-Field (SCF) method: First, a set of coefficients $\mathbb{C}$, i.e. a density matrix $\mathbb{P}$ is chosen, the matrix $\mathbb{G}$ (3.32) (hence $\mathbb{F}$ (3.28)) is calculated; Eq (3.34) and (3.26) are then solved and, the resulting coefficients compared with the assumed ones. If they are different, the procedure is repeated with the new values of the coefficients until self-consistency is reached. The Slater determinant obtained with the n ground state orbitals is called Ψ_{SCF}.

3.1.8 Atomic orbitals and basis sets.

An exact expansion of the molecular orbitals in (3.24) would require an infinite number of basis functions but, this is impossible to achieve and only a finite number of functions are used. A large basis set gives a lower energy than a smaller one but it also requires greater computational resources. Therefore, the basis set has to be carefully chosen in order to give results that are an acceptable compromise between accuracy and use of available resources. In early calculations, the basis sets used in the expansion of the MOs were of Slater Type Orbitals (STOs) [46] which are exponential functions with an exponent ξ. Slater functions were developed from hydrogenic atomic orbitals and describe correctly qualitative features of AOs for other systems. A minimal or single-zeta basis set is defined as a basis set which includes one Slater function for each occupied AO with distinct n and l quantum numbers; double-zeta basis sets include two STOs for each occupied AO. Any basis larger than double-zeta is referred to as an extended basis set. Of particular importance is the addition of functions with higher l values, (i.e. d, f, ...) on first row atoms and p on H which are termed polarization functions. The advantage of minimal basis sets is that they give results that can be easily interpreted in terms of the atomic orbitals of the constituent atoms of the molecule. Double-zeta or extended basis sets on the other hand, by using two or more functions to describe each atomic orbital, take into account the fact that within the molecular environment the atomic orbitals differ from those in the isolated atom (in particular the valence orbitals) and therefore they give a better wave function than the one obtained with a minimal basis set.

The main problem in evaluating the matrix elements F_{uv}, eq.(3.28), is the evaluation of the electron repulsion integrals. In theory, there are ~ $L^4/8$ such integrals were L is the number of basis functions used. The evaluation of these integrals with Slater-type functions are extremely time consuming and, most of the electronic calculations are done using gaussian-type functions (GTOs) as proposed initially by Boys [47]. The cartesian gaussian-type functions centred on the nuclei have the form:

$$B\, x^p\, y^q\, z^s\, e^{-\alpha r^2}$$

(3.35)

were B is a normalization factor, and α is the exponent. p, q, and s are integers which describe the angular dependence of the function, the radial dependence is described by $e^{-\alpha r^2}$. The main advantage of gaussian functions is that the product of two gaussians is a third gaussian which is centred between them. This property is used to simplify the two-electron integral calculation. Also, with cartesian GTOs (3.35) the evaluation of the two-electron integrals can be simplified (15-fold) by calculating them as a sum of one-dimensional integrals over the cartesian component that can be evaluated analytically [103]. However, GTOs do not describe the functional behaviour of AOs as well as STOs, especially near the nucleus and at a large distance from it. In order to obtain results equivalent to those using Slater functions more than one gaussian on each centre is required. The enormous number of integrals which have to be handled is dramatically reduced by using as a basis set a linear combination of primitive gaussians, g_k, with fixed coefficients d_k called Contracted Gaussian Type Orbitals (CGTOs) [42]:

$$\chi_u = \sum_{k=1}^{C} d_k g_k$$

(3.36)

where C is the degree of contraction. In solving the SCF equations, then, only the coefficients of each of the contracted functions χ_u must be determined.

A large number of gaussians basis sets of several sizes have been reported in the literature [48] and several of them have been used in this work including those proposed by Pople [49-53], Huzinaga [54,55], Dunning [56] and, Clementi [57].

The basis sets introduced by Pople and co-workers may be classified into three main groups: minimal, split-valence and, split-valence plus polarization functions. The STO-NG basis [49] belongs to the first group; in these basis sets, N GTOs are used to approximate each Slater-type function; the orbital exponents α (Eq. (3.35)) and the contraction coefficients d_k (Eq. (3.36)) which best reproduce a Slater function, were

determined by the least-squares method. A characteristic of all the basis sets proposed by Pople is that s and p GTOs with the same quantum number n share a common exponent. This reduces the computational cost [49-53]. The split-valence basis sets, of which 6-31G [51] and 3-21G [53] are examples, use minimal basis sets for the inner shell (core) and a double-zeta basis set for the valence orbitals. In the 6-31G (3-21G) basis set, the core is described by a CGTO which is a linear combination of 6 (3) GTOs while the two valence orbitals are the contraction of 3 (2) and 1 GTOs. Exponents and coefficients were optimized to give the lowest HF atomic energy of the ground state. This procedure is problematic because if the core part does not contain a large number of functions then, the valence functions tend to "fall inward" toward the nucleus [53]. For this reason, in order to construct the 3-21G basis set, the 6-21G basis had to be prepared first; then, the 6 inner shell functions were replaced by 3 gaussians; this was done without reoptimizing the valence-functions [53]. The third group belong the 6-31G* and the 6-31G** [52] basis sets. In the 6-31G* basis set, a set of d polarization functions is added to 6-31G; The 6-31G** also includes p functions for the hydrogen atoms.

Huzinaga MINI and MIDI basis sets [54] are very similar to the STO-3G and the 3-21G basis sets respectively; the main difference is that there is no exponent sharing among the valence s and p orbitals. Hence MINI-1 is somewhat more expensive to use than its STO-3G counterpart, but the properties obtained are far better than those obtained with the STO-3G basis set as indicated in chapter 9.

In the geometrical basis sets proposed by Clementi [57], the exponents of the single gaussians are terms of a geometrical progression and all the atoms from H to Sr are represented by the same set of exponents. These basis sets are specifically designed for use in large molecular calculations. The minimal basis set is called GEOSMALL; GEOSPV has a minimal basis for the core orbitals and a split one for the valence orbitals; GEOMEDIUM is contracted to a double-zeta basis set. The two large basis sets of Clementi GEOLARGE and GEOTRIPLEZ, have also been used. GEOTRIPLEZ uses a triple-zeta contraction scheme. GEOSMALL is a very promising basis set for calculation of large molecules; it has all the advantages of being a minimal basis set but it gives

energies and other properties which are far superior to those obtained with STO-3G or MINI-1. The energies obtained with GEOSMALL are in fact of the order of those obtained with 3-21G (see chapter 9).

The basis sets discussed so far do not perform well for calculations involving anions. For stable anions (i.e. that lie energetically below the ground state of their parent neutral molecules) [58], use of these basis sets yield positive energies for the highest-occupied molecular orbitals, indicating erroneously that the outermost valence orbitals are unbound. A way to overcome this problem is to include in the basis set one or more sets of diffuse functions, that is functions with a low exponent α. The 6-31+G basis [59] is an example of a basis set with diffuse functions.

3.2 Semi empirical methods

The approximations discussed so far are valid both for *ab-initio* and semi empirical methods.

In *ab-initio* methods, all one and two-electron integrals are retained and calculated, unless they are predicted to be zero by symmetry considerations and pre-screening. Efficient algorithms for computation of the integrals have been developed and single point electronic calculations can now been done even for large molecules using the 'SCF=direct' procedure [60] (in which the integrals are calculated only when they are needed and therefore the disk space is not longer a limiting factor in the calculation). Nevertheless, geometry optimization or studies of the potential energy surface (see later) for molecules of real biological interest are still too expensive to be attempted by *ab-initio* methods because of the large number of basis functions (hence the large number of integrals) that have to be handled.

In semi empirical methods, not all the electrons are explicitly considered. The Huckel theory, for instance, considers only the π electrons explicitly [61] and can be successfully used to study unsaturated and aromatic molecules. In the semi empirical

methods used in this work, only the valence electrons are explicitly included (core approximation); some integrals are neglected and others are approximated by parameters derived from experiments [63]. Despite the various approximations that are made, semi empirical methods can, by careful parametrization, provide a reliable framework for calculating properties of molecules that are similar to those used in the parametrization procedure; furthermore, they are simple enough to be applied to moderately large molecules using only modest amounts of computer time.

Among the various semi empirical methods which have been developed [63,64] in recent years, the AM1 (Austin Model 1) [65] and PM3 (Parametric Method 3) [66] have been extensively used in this work. Both are variations of the original MNDO (Modified Neglect of Diatomic Overlap) [67] method proposed by Dewar in 1976. MNDO uses the same assumptions made in the NDDO approximation (Neglect of Diatomic Differential Overlap) of Pople [61,68], in which the number of electron repulsion integrals is greatly reduced by using the core approximation, together with a minimal basis set of valence shell AOs, and by neglecting all integrals involving overlap of atomic orbitals between different atoms.

The next paragraph briefly describes the form of the matrix elements common to all the NDDO methods and the differences between AM1 and PM3.

3.2.1 The MNDO approximation

In the MNDO approximation, all inner electrons are treated as part of an unpolarizable core, the charge of which is set equal to that of the nuclei, minus that of the inner electrons. The molecular orbitals in eq. (3.24) are expanded in terms of a minimal basis set of Slater type functions. A further approximation is made (NDO approximation) in treating the atomic orbitals χ_p as if they form an orthonormal set; that is, the overlap matrix (3.27) is set equal to a unit matrix and the Roothaan equations (3.26) can be written (in matrix form):

$$\mathbb{FC}=\mathbb{SC}\varepsilon$$

(3.37)

were $\mathbb{F}$, $\mathbb{C}$, and $\mathbb{S}$ have the same meaning as in (3.26) except for the fact that only the valence electrons are considered. When this approximation is applied together with the neglect of the two-electron integrals between orbitals on different atoms, the following expressions for the elements of the Fock operator (for atoms of the first row only) are obtained:

$$F_{\mu\mu}=U_{\mu\mu}+\sum_{B}V_{\mu\mu,B}+\sum_{\upsilon}^{A}P_{\upsilon\upsilon}[(\mu\mu|\upsilon\upsilon)-\frac{1}{2}(\mu\upsilon|\mu\upsilon)]+\sum_{B}\sum_{\lambda,\sigma}^{B}P_{\lambda\sigma}(\mu\mu|\lambda\sigma)$$

(3.38)

$$F_{\mu\upsilon}=\sum_{B}V_{\mu\upsilon,B}+\frac{1}{2}P_{\mu\upsilon}[3(\mu\upsilon|\mu\upsilon)-(\mu\mu|\upsilon\upsilon)]+\sum_{B}\sum_{\lambda,\sigma}^{B}P_{\lambda\sigma}(\mu\upsilon|\lambda\sigma)$$

(3.39)

$$F_{\mu\lambda}=\beta_{\mu\lambda}-\frac{1}{2}\sum_{\upsilon}^{A}\sum_{\sigma}^{B}P_{\upsilon\sigma}(\mu\upsilon|\lambda\sigma)$$

(3.40)

Here, the chemists' notation for the two-electron integrals has been adopted, where $(\mu\upsilon|\mu\upsilon) = \langle\varphi_\mu\varphi_\mu|\varphi_\upsilon\varphi_\upsilon\rangle$. The atomic orbitals φ_μ and φ_υ are centred on atom A while φ_λ and φ_σ are centred on atom B.

$U_{\mu\mu}$ (the one-centre one-electron integrals) is the sum of the kinetic energy of an electron in φ_μ on atom A and the attraction to its own core; the values of these integrals are parametrized by using spectroscopic data [75].

$(\mu\mu|\nu\nu)$ and $(\mu\nu|\mu\nu)$ are the one-centre two-electron repulsion integrals; They are evaluated using experimental parameters which are chosen so that calculated properties fit experimental data for isolated atoms in AM1 whereas, in PM3 the parametrization is

designed to reproduce molecular properties. By deriving the one-centre repulsion integrals from experimental parameters, some allowance is made for correlation effects.

In the original NDDO method [61,68], the two-centre repulsion integrals $(\mu\nu|\lambda\sigma)$ were evaluated analytically. In all the MNDO methods they are evaluated with a semi empirical model in which the electron density distribution is approximated by a series of point charges and the integrals are evaluated as a series of multipole-multipole interactions.

$V_{\mu\sigma,B}$ is the core electron attraction, i.e. the interaction between the valence electrons and the core. It is expressed in terms of the two-centre repulsion integrals and an adjustable parameter.

$\beta_{\mu\lambda}$ are the two-centre one electron core resonance integrals which provide the main contribution to the bonding energy of a molecule. For this reason, the NDO approximation is violated and these integrals are assumed to be proportional to the corresponding overlap integrals $S_{\mu\lambda}$. The last are evaluated analytically, the orbital exponents being treated as adjustable parameters.

The total energy E_{tot} of the molecule is the sum of the electronic energy E_{el} and the repulsion between the cores of atom A and B.

$$E_{tot} = E_{el} + \sum_{A<} \sum_{B} E_{AB}^{core} \qquad (3.41)$$

In the original MNDO, the repulsion between the cores (CRF) was expressed in terms of the two-centre repulsion integrals plus an adjustable parameter. In AM1 and PM3 additional gaussian functions have been added to the CRF expression. These functions were added to the AM1 in order to correct some weakness of MNDO such as its underestimation of hydrogen bond energies and, the overestimation of energy obtained for sterically crowded molecules; These problems were caused by the tendency of MNDO

to overestimate the core-core repulsion between atoms when they are at their Van der Waals distance apart.

The heat of formation of the molecule ΔH_f, is obtained from its total energy by subtracting the electronic energies of the atoms (calculated from atomic NDDO calculations) and adding the experimental heat of formation ΔH_f^A of the atoms in the molecule:

$$\Delta H_f^{mol} = E_{tot}^{mol} - \sum_A E_{el}^A + \sum_A \Delta H_f^A \qquad (3.42)$$

Applications of PM3, at this time, have only been carried out in limited studies [69-71, 96-102]. PM3 was found to give, in general, better energies and heats of formations than AM1. However, the reliability of PM3 in calculating other molecular properties, is still under discussion [96,99] (see also chapter 5).

3.3 Properties derived from the wave function

For a molecule with a wave function ϕ, all the physical properties O, can in principle, be calculated from ϕ as the expectation value of the appropriate operator $\underline{\mathbf{O}}$. If the wave function is written as a single Slater determinant and, each occupied molecular orbital φ_u contains two electrons then, within the LCAO approximation, the expectation value of an operator can be written in terms of the density matrix:

$$O = \langle\phi|\mathbf{O}|\phi\rangle = \sum_u^{occ} \langle\varphi_u|\mathbf{O}|\varphi_u\rangle = \sum_u^{occ}\sum_{pk} c_{pu}c_{ku} \langle\chi_p|\mathbf{O}|\chi_k\rangle = \sum_{pk} P_{pk}\langle\chi_p|\mathbf{O}|\chi_k\rangle \qquad (3.43)$$

The density matrix defines the charge density $\rho(\mathbf{r})=|\phi(\mathbf{r})|^2$ given in terms of the basis set functions χ used in the calculation. The charge density is defined such that $\rho(\mathbf{r})d\mathbf{r}$ is the probability of finding an electron at co-ordinates $\mathbf{r}$ in a small volume of space $d\mathbf{r}$. The integral of the charge density gives the total number of electrons.

The first part of this section is dedicated to the main properties that have been calculated in this work, that is, the total energy, the dipole moment and, the electrostatic potential.

The second part describes the Mulliken population analysis. This is a method of describing the electron distribution in the molecular space and calculating atomic charges. Although the quantities which are derived from it are not physical observables, the Mulliken atomic charges correlate with the anti-cancer activity of the series of molecules studied in this work.

Finally, the third part describes the methods of geometry optimization and conformational studies.

3.3.1.1 Energy calculations

The total energy is the expectation value of the Hamiltonian operator. Within the Born Oppenheimer approximation it is obtained as sum of the electronic energy and the constant nuclear repulsion term for a given nuclear configuration. The electronic energy is given by:

$$E_{elect} = \frac{1}{2}\sum_{p}\sum_{k} P_{pk}(H_{pk}^{core} + F_{pk}) \qquad (3.44)$$

Where P_{pk} are the elements of the density matrix (3.33), H_{pk} and, F_{pk} are the elements of the core-Hamiltonian (3.29) and the Fock operator (3.28) respectively.

The occupied orbital energy ε_{υ} obtained as eigenfunctions of the Fock operator (3.26) represents the energy of an electron in the spin orbital ψ_{υ}. According to Koopmans' theorem [36(a)], ε_{υ} is a reasonable approximation to the negative value of the ionization potential of the molecule if the electron is removed from ψ_{υ}. This result is obtained from the calculation of the ionization potential assuming that the spin-orbitals of

the cation can be considered identical to those of the parent molecule. However, the electron affinity is not well approximated by the energy of the LUMO because the spin orbitals of the anion (obtained by adding an extra electron) cannot be considered identical to those of the parent molecule [36a].

3.3.1.2 The dipole moment

Neutral molecules often have permanent dipole moments because the centre of the negative charge, due to the electron distribution, and the centre of positive charges, due to the nuclei, do not coincide. An external electric field, due for instance to the charge distribution of a receptor, causes the molecule to orientate according to its dipole moment. The value of the dipole moment is very sensitive to the quality of the wave function. Nevertheless if it is calculated within the same level of accuracy for a series of related molecules, it can be used for comparative purposes.

From a computational point of view, the dipole moment is very easy to calculate once the wave function has been obtained. The expectation value is given by:

$$\vec{\mu} = -\sum_{\mu}\sum_{\upsilon} P_{\mu\upsilon}\left\langle \chi_{\upsilon} \left| \sum_{i=1}^{N} \mathbf{r}_i \right| \chi_{\mu} \right\rangle + \sum_{A} Z_A \mathbf{R}_A \qquad (3.45)$$

The electronic dipole operator is written as the sum of one-electron operators of the position vector $\mathbf{r}_i$. The second term is the contribution of the nuclei; $\mathbf{R}_A$ is the position vector of the nucleus A and Z_A its atomic charge. The others symbols have the usual meaning.

3.3.1.3 The electrostatic potential

The use of the molecular electrostatic potential to study chemical properties, molecular interactions and biological activities has been documented and reviewed

critically in Ref. [72]. In particular, the work of Pullman [72, 73], Weinstein [72, 74-76], and Tomasi [72, 77] showed how powerful the study of the MEPs can be in predicting the chemical reactivity of molecules of biological interest. Many other studies have been conducted in this area in recent years [78-81].

The electrostatic potential at the point **r** can be written as:

$$V(\mathbf{r}) = \sum_{A} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}|} - \sum_{\mu} \sum_{\upsilon} P_{\mu\upsilon} \left\langle \chi_\mu \left| \frac{1}{|\mathbf{r}' - \mathbf{r}|} \right| \chi_\upsilon \right\rangle \tag{3.46}$$

where $\mathbf{R}_A$ is the position vector of the nucleus A and Z_A its atomic charge. The first term represents the nuclear contribution to the electrostatic potential while the second is the electronic term. All the symbols have the usual meaning. V(r) can be obtained from either an *ab-initio* or semi empirical wave function. When semi empirical wave functions are used it is however, necessary to deorthogonalize first the basis set in order to obtain reliable results [82, 83].

The electrostatic potential is generally calculated on a plane or on a three-dimensional surface (MEPs) and displayed as equipotential contour diagrams or as three-dimensional surfaces in which for each range of potentials is associated a different colour. No single viewpoint can provide a complete inspection of the MEP on the molecular surface. Generally it is necessary to calculate the MEP in different planes of the molecule or, if the MEP has been calculated on a surface, different portions of the surface have to be analyzed in order to find the most useful information.

The analysis of the MEP can give information about the primary interaction between the drug and the relevant biological receptor. The interaction between two molecules can be divided into components by methods such as the Morokuma decomposition analysis [72]. At large distances, this interaction is primarily electrostatic; polarization, charge transfer, and exchange energies can be neglected if one of the species has a net charge or a high dipole moment. The electrostatic potential, being the electrostatic energy between the unperturbed charge density $\rho(r)$ of a molecule and a

positive unit charge located at a point **r**, provides a useful tool to help to understand the interaction with the charge distribution of another molecule which is far enough away, to be approximated by a point charge.

In molecular drug design, the approach is to compare the MEPs of active and inactive species to identify the characteristics that might be responsible for the activity. Once the characteristics of the active molecules have been found new active drugs may be rationally designed and the synthesis of molecules likely to be inactive can be avoided. In this work extensive studies of the electrostatic potential have been conducted although no definite correlation with the anti-cancer activity of the FAAs has be observed.

The electrostatic potential depends on the quality of the basis set, as do all the other properties calculated from the wave function. In order to obtain quantitatively good results, the use of a large basis set with the inclusion of polarization functions would be advisable. However, the calculation of the electronic part of the electrostatic potential increases with the number m of the basis functions as $(m^2+m)/2$ and it is necessary to find approximate expressions of V(r). Different kinds of approximations have been reported in the literature [72,84]. The most common expresses V(r) in terms of point charges located at the nuclei [84], or in terms of a multipole expansion [72]. Within the point charge àpproximation, the delocalized electron density is replaced by the localized atomic charges

$$V(r) = \sum_A \frac{q_A}{|\mathbf{r}' - \mathbf{r}|} \qquad (3.47)$$

where q_A is the atomic charge on nucleous A centred at $\mathbf{r}'$ while $\mathbf{r}$ is the point where the MEP is to be evaluated. The potential calculated with (3.47) become closer and closer to that calculated with (3.46) at large distances and fails to predict the right value of the potential at a distance smaller than the Van der Waals radius of the atom [108]. However, if Mulliken charges are used to calculate the potential, the MEP obtained differs significantly from that calculated directly from the wave function as can be seen in chapter

9. A study of the basis set dependence of the electrostatic potential has been conducted as part of this work in an attempt to find a minimal basis set able to reproduce high quality results.

The position and depth of the local minima in the MEP can be related to the reactivity of the molecule toward electrophilic attacks [85]. An index of the global similarity between MEPs of different molecules can be obtained by the similarity index, first proposed by Carbo et al [86]. The Hodgkin similarity index [87] has been used in this work, it is defined by:

$$H_{AB} = \frac{2\int \rho_A \rho_B dv}{\int \rho_A^2 dv + \int \rho_B^2 dv} \qquad (3.47)$$

were ρ_A is the electron density of molecule A. H_{AB} can take values between 0 and 1 (or between -1 and 1 in the case of the electrostatic potential), an index of 1 indicates complete similarity between molecules A and B.

3.3.2 The Mulliken population analysis

Within the LCAO approximation, the molecular orbitals are delocalized over the whole molecule and there is no quantum mechanical operator which can be used to represent the delocalised molecular charge distribution in terms of point charges associated with the atoms from which the molecule is built up. The common way to obtain atomic charges is to divide the molecular space into domains each containing an atom. This procedure is called 'population analysis' and the commonest one was introduced in 1955 by Mulliken [88 (a)].

The molecular orbital $\varphi_u = \sum c_{upk} \chi_{pk}$, written as a linear combination of normalized functions χ_p (of atoms k and l), contains N(i) electrons (2 in all the RHF calculations in

this work). In the Mulliken population analysis the space of these electrons is divided such that the number of electrons can be written as:

$$N(u) = N(u)\sum_{kp} c_{ukp}^2 + 2N(u)\sum_{l>k} c_{upk} c_{uql} S_{pkql} \qquad (3.48)$$

were S is the overlap integral between basis functions in the two atoms k and l. The term $2N(u)c_{upk}c_{uql}S_{pkql}$ is called *overlap population* and the sum of these elements over all the molecular orbitals between the atoms k and l gives the *total overlap population* n(k,l) between them. If the value of n(k,l) is positive, the atoms k and l are bonded, while if the value is negative they are antibonded [88 (b)]. The quantity $N(u)\ c_{ukp}^2$ is called *net atomic population*..

The *atomic charge* of an atom k is obtained by subtracting the *gross atomic population* from the total number of electrons in the ground state of the free neutral atom k; where the *gross atomic population* is obtained from eq. 3.48 by assigning to each centre the appropriate *net atomic population* plus half of the overlap population. The use of these charges as indices of reactivity should be conducted with caution mainly because the charge between two nuclei is divided equally between the two, even if the atoms have very different electronegativities; for some systems containing atoms with very different electronegativities the charges can lead to wrong predictions of chemical reactivity such as protonation sites. Protonation sites are generally better predicted by the minima in the electrostatic potential than by the most negative atomic charge. Also, Mulliken atomic charges are very dependent on the basis set and a large basis set does not necessarily give better results than a small one.

Atomic charges can also be obtained by another approach that fits the molecular electrostatic potential to a series of point charges placed at the atomic centres [89, 90]. Atomic charges calculated in this way, reproduce precisely defined molecular properties, such as dipole and quadrupole moments, better than Mulliken charges [91, 92].

In structure/activity studies it is possible to obtain very useful information from the use of atomic charges, in particular by considering differences between atomic charges of a series of related molecules. Extensive use of atomic charges has been made throughout this work and correlations between the anti-cancer activity of the FAAs and the charges of specific atoms have been observed.

3.3.3 Geometry optimization and potential energy surfaces

The total energy calculated within the Born-Oppenheimer approximation, as a function of the nuclear position, defines a potential energy surface. Given a molecular system containing M atoms, this surface will be 3M-6 dimensional (3M-5 dimensional for linear molecules). A complete knowledge of this surface becomes impossible as the number of atoms in the molecule increases. The most interesting points on the potential energy surface are, however, the stationary points that correspond to equilibrium geometries or transition states. In this work, equilibrium geometries have been determined mainly using semi empirical wave functions although a few ab-initio geometry optimizations have been carried out using the 3-21G basis set [123].

Beginning with a reasonable starting geometry, the geometry optimization proceeds by calculating the forces on each atom in the molecule and then moving the atoms so that these forces are reduced to zero and the energy of the system is minimised The general approach to locate minima, used both for *ab-initio* and semi empirical wave functions, utilizes gradient methods [93]. The gradient **g** of the potential energy function is a vector, the components of which are the partial derivatives of the energy with respect to the coordinates of each atom in the system. The norm of the gradient is the magnitude of the internal forces in the molecule. The matrix of the second derivatives of the potential energy with respect to the coordinates of each atom in the system is called "the Hessian". In a local minima the gradient must vanish and all the eigenvalues of the Hessian must be positive.

For the ab-initio geometry optimizations, the Berny procedure [94] was used. In this procedure the gradients are calculated each time that the energy is computed. A suitable estimate of the Hessian at the beginning of the optimization process is required. The gradient and the second derivatives are used to predict the changes in any individual coordinate, x_i, that will yield the stationary point; linear and quadratic searches are used. The coordinate changes predicted in the search are only applied if they fall below the specified threshold. A big change in the coordinates could, for example, break a bond, which is clearly not acceptable. The norm of the gradient and the absolute value of its largest component must also fall below their respective thresholds. At each step of the optimization, the predicted changes in the coordinates and the gradient are evaluated, compared to the threshold values and, if the conditions are not met, the Hessian is updated and the process is continued unless the maximum number of steps is exceeded.

The evaluation of the gradients is computationally expensive for ab-initio methods, due to the large number of additional integrals required. Semi empirical methods on the other hand, due to the approximations described previously, require only a relatively small number of such integrals and the time required to calculate them is negligible compared to the time required to calculate a single configuration. For the semi empirical optimizations, the BFGS (Broyden-Fletcher-Goldfarb-Shanno) method [93, 95] has been used. This method uses the inverse of the Hessian, which at the beginning of the optimization is normally chosen to be the unit matrix. The changes in the geometry are predicted by moving the atoms in a direction determined by the gradient so as to lower the energy of the system. The optimization is concluded when the gradient norm is smaller than a specified threshold, and the predicted change in energy, falls into a specific limit.

For the flavone acetic acids molecules studied in this work a range of conformations are possible because of the free rotation about the single bond between the phenyl ring and the chromone system, and between the flavone system and the acetic acid group. Semi empirical methods have therefore been used to investigate a part of the potential energy surface determined as a function of certain selected dihedral angles.

When the rotation only occurs about one bond, the results have been presented in the form of a curve of energy against angle. In order to establish the conformation of the acetic acid group, two rotational angles were required to specify the geometry, involving three dimensional surfaces, and these have been presented as contour diagrams. Such calculations are rather expensive since the computation must be performed at a large number of points. The rotation of each individual bond has to be performed through 360° (180°) in a series of m steps of step-size 360/m (180/m) degrees. When two variables are considered the number of points at which the energy is evaluated requires m separate steps for each angle leading to a total of m^2 points. At each point, the geometry of the FAAs have been fully optimized with the exception of the dihedral angles that were changing.

CHAPTER 4:

Computational Aspects

Introduction

In this chapter the computer facilities and the programs used in this thesis are reported. Over the last three years in which this work was conducted, there have been several changes both in the computer hardware and in the software available.

Section 4.1 includes a brief discussion of the computer hardware, whilst the computer software used in the work is described in section 4.2.

4.1 Computer hardware

Until spring 1991, the University of St-Andrews provided all the students and members of staff with two VAX 11/785 computers running the VMS operating system. Their use in this work has not been very extensive, mainly because the resources available in terms of computer time and disk space were too small for most of the calculations reported here. However, the VAX machines were licensed to run the SURFACE 2 program (see later) and without them it would have been very difficult to produce the contour diagrams reported in chapter 6.

The VAX computers have now been superseded by a network of SUN workstations that started to be operative in the Autumn 1990. In the Chemistry department, there are eight of these workstations, one of which is in colour. They run in a UNIX operating system and have been extensively used to perform semi empirical calculations. The colour SUN workstation is also the most powerful within the chemistry department, it has 16 MBytes of RAM while the others have 8 MBytes.

The computer facilities of the University of St. Andrews also includes microcomputers such as IBM PC 15 and several Apple Macintoshes. Recently, two Macintosh LCs have been installed having 12" colour monitors capable of displaying up to 256 colours. Apart for word processing, the Macintosh proved to be very useful during the course of this work; some of the programs it contains such as Cricket Graph and

Chem-3D, have been extensively used to analyze the results of the quantum mechanical calculations. Furthermore, because some of the computers are networked (either the X.25 network or the Ethernet), hard copies of graphic programs such as CHEM-X [112], SURFACE 2 [124], etc have been obtained with the laser printer connected to the Macintosh; this proved to be very useful in analysing the results and writing up the thesis.

Most of the calculations reported in this work have been performed on an FPS-500 (Floating Point System) mini-supercomputer of the research group. It replaced the SCS40 mini-supercomputer, in Autumn 1989. The FPS-500 runs the UNIX operating system and it is a two processor system, one scalar and one vector. This means that a number of different operations can be performed simultaneously rather than sequentially (as happens in a normal scalar processor) and this greatly increases the computational speed of the programs which run on the vector processor. The FPS-500 has 128-MBytes of RAM. The ab-initio calculations for the FAAs could only be run on the FPS-500.

The other machine available to the research group was a MicroVAX II/GPX colour workstation. It runs the VMS operating system and has 13MBytes of RAM. It is linked to the FPS-500 by Ethernet network. Its use in this work was limited to storage of files and graphics display. The colour workstation is provided with a Digital LJ250 colour printer to obtain hard copies. A Tektronix 4109 and a Digital VT340 could also be used to display graphics; hard copies from the Tektronic terminal could be obtained by a colour Tektronic printer.

4.2 Computer software

The program available to the research group to perform semi empirical calculations is MOPAC [104, 95]. Version 5 of this program has been used for all the semi empirical calculations reported here. It was installed in all the machines, the MicroVAX, the FPS-500 and the SUN, but only the UNIX machines were used for the calculations of the FAAs. With MOPAC5 it is possible to perform MNDO, MINDO 3, AM1 and PM3

calculations. The input to MOPAC used in this work was in the form of a Z-matrix although MOPAC also makes provision for an input in Cartesian coordinates.

FIG. 4.1: Example of input to MOPAC as Z-matrix

```
AM1 GNORM=0.05 PRECISE
Example of MOPAC input
Ethylene
C
C       1.35    1
H       1.08    1       120.0   1
H       1.08    1       120.0   1       180.0   1       2   1   3
H       1.08    1       120.0   1       0.0     1       1   2   4
H       1.08    1       120.0   1       0.0     1       1   2   3
```

Fig. 4.1 shows a typical Z-matrix: line 1 contains simple key-words that indicate the type of calculation that has to be performed; lines 2 and 3 are comments. The Z-matrix starts from line 4 and the bond lengths are in Å whilst bond and dihedral angles are in degrees. In the example of fig. 4.1, an AM1 full optimization of ethylene is required. The key-word GNORM=0.05 means that the geometry optimization has to be continued until the gradient norm drops below 0.05; the PRECISE option requires that the default SCF criterion (the heat of formation and the density matrix) is tightened by a factor of 100. Both these key-words have been used for all the geometry optimization calculations and there use is recommended in the design of new FAAs on the basis of the structure/activity correlations reported later in this thesis. An introduction to the MOPAC program is given in reference [105]. The calculation of the AM1 electrostatic potential from the wave function was performed using a modified version of MOPAC called MOPACQ obtained from Chris Reynolds at Oxford University.

The *ab-initio* calculations on the FAAs reported here have been performed with an FPS-500 version of the GAUSSIAN 90 program [106]. This program is able to perform direct SCF calculations [60] and therefore, it has been possible to perform ab-initio calculations on molecules as large as FAAs, with split-valence basis sets. Furthermore, this version of the program is is very fast because it is adapted to run on a vector process machine such as the FPS-500. The input for GAUSSIAN used is in the form of a Z-matrix as shown in fig. 4.2. for a full optimization calculation on ethylene with the 3-21G basis set. An introduction to GAUSSIAN may be found in reference [105].

FIG 4.2: Example of input to GAUSSIAN as Z-matrix

```
$rungauss
%chk=/dir/filename
#RHF/3-21G  OPT  SCF=DIRECT

Example of GAUSSIAN input  for Ethylene

0  1
C
C          1          DB
H          2          HCB          1          ANG
H          2          HCB          1          ANG          3          DIH
H          1          HCB          2          ANG          4          DIH
H          1          HCB          2          ANG          3          DIH

DB=              1.35
HCB=             1.08
ANG=           120.0
DIH=           180.0
```

The GAUSSIAN 90 version for the FPS-500 became available only in the last few months. For most of the time in which this research was conducted the ab-initio program available in the mini-supercomputer was a modified version of GAUSSIAN80* [107] This program does not permit the possibility to perform direct SCF calculations and therefore it was not suitable for the study of large systems. The calculations reported in

chapter 10 were performed with this program. The input to the program to perform an *ab-initio* electrostatic potential calculation consist of two data files; One of them contains the molecular geometry along with control cards for the program options (Fig. 4.3), the other contains the coordinates (in atomic units) of each point at which the MEP is to be calculated (Fig. 4.4). In the example shown here the MEP is calculated over 5 points.

FIG. 4.3: Example of the first input file to GAUSSIAN80* for the calculation of the MEP of water with the STO-3G basis set.

```
99999.9       0       0     1     0
$N 1 1 0 4 3 0 0 0 0 2 0 0 2 0 0 0

0  1
 8
 1    1   0.96127
 1    1   0.96127       2    103.5315
```

FIG. 4.4: The second data file containing the coordinates of the surface at which the MEP has to be calculated.

1	EPOT	0.857000	0.857000	0.7000
1	EPOT	0.464000	1.120000	0.7000
1	EPOT	0.000000	1.212000	0.7000
1	EPOT	-0.464000	1.120000	0.7000
1	EPOT	0.857000	0.857000	0.7000

The output consists of a logfile and a second file containing the coordinates of the surface at which the potential has been calculated (in Å), and the value of the electrostatic potential in atomic units. This file can then be processed to give a suitable file for the display program 3D2 [108].

3D2 [108] is a program for the calculation and the display of MEPs. The MEP at each point is displayed by a coloured polygon perpendicular to the molecular surface. The colours of the polygons represent the different values of the potential. This program can also be used to calculate MEPs on a Connolly surface obtained by running the MS program [109]. The potential is calculated using the point charge approximation method (described in chapter 3) .The input to 3D2 for the calculation of MEPs contains the atomic coordinates and the atomic charges.

ASP [110] is a program released by Oxford Molecular Ltd for the calculation of the potential and electric similarity indexes. The potential is calculated with the point charge approximation over a grid common to all the molecules to be compared. The optimum fitting of the potential can also be obtained by translating or rotating the molecules. To compare the MEPs of the FAAs (chapter 6) a modified version of this program for the FPS-500 [111] has been used.

Other graphics programs that have been extensively used are CHEM-X [112] and MHCDRAW [111]. The latter was written by M. Charlton, a member of this research group, and can read any of the output produced by MOPAC and GAUSSIAN. It is a very useful and flexible display program that has been used extensively in this work.

To simplify the analysis of the quantum mechanical calculations in this work, it has often been necessary to write simple FORTRAN programs. The GRID program, for example, reads the MOPAC output file containing the result of a potential energy scan, identifies the local minima and writes them down with the corresponding value of the two variables (they are not shown explicitly in the MOPAC output file). With this program it also became very easy to identify the different minima in the potential energy surfaces

reported in chapter 6 without having to display the surface. Another program called SIM was written to calculate the Hodgkin similarity index [87] between two electrostatic potential maps calculated over the same surface, using the following formula:

$$SI_{RT}=\frac{2\sum_{i=1}^{N} V_R^i V_T^i}{\sum_{i=1}^{N} (V_R^i)^2+\sum_{i=1}^{N} (V_T^i)^2} \qquad (4.1)$$

The sum runs over the N points in which the electrostatic potential has been calculated; V_R^i and V_T^i are the EPM at the point i of the reference and test EPM respectively. This program has been very useful to analyse the *ab-initio* MEPs of the pyron-2-one obtained with the different basis sets (see chapter 9).

CHAPTER 5:

Comparison Between Calculated and Experimental Properties of FAA Analogues.

Introduction

The molecules used in the parametrization of AM1 and PM3 were mainly small molecules [65,66] therefore they are quite different from the molecules related to FAA. To see if the AM1 and PM3 methods were suitable methods to use for the study of this class of molecules, a comparison between the calculated and experimental properties has been made for molecules analogous to FAA in section 5.1. Subsection 5.1.1 reports the results of the comparison between calculated and experimental x-ray structures, while in 5.1.2 the atomic charges are compared with ^{13}C NMR chemical shift data.

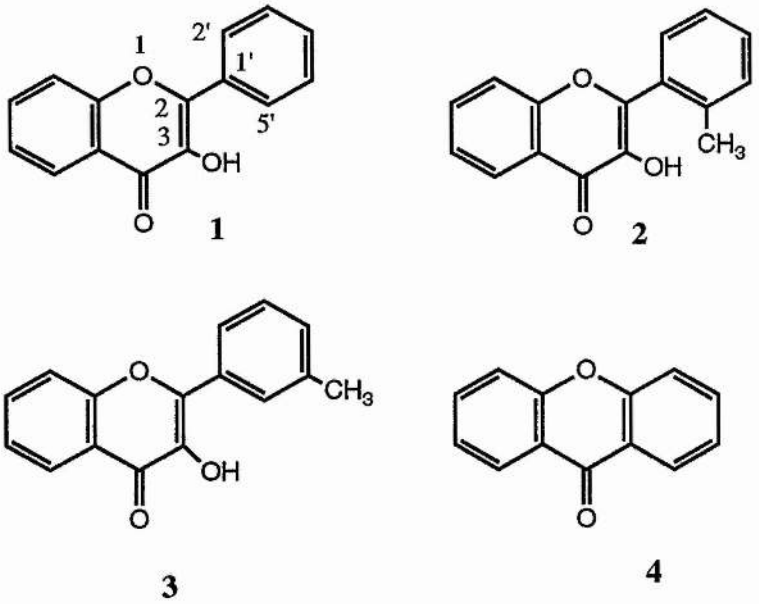
5.1 Comparison between calculated and experimental properties for analogues of FAA

5.1.1 Molecular geometries

The x-ray structure of the molecules shown in fig.5.1 have been compared with the optimised structures obtained using the AM1 and PM3 methods. Care has been taken to use the same strategy as used to optimize the structures of the FAAs and XAAs (see later). Therefore the results obtained from this study should be applicable to the FAAs and XAAs given the similarity of their structure with the molecules tested.

The x-ray structures were obtained by the Cambridge Crystallographic Data Base [113]. These molecules contain the basic ring structure found in the molecules of FAAs and XAAs. In particular, the three flavones were particularly interesting because of the differences in the position of the phenyl ring shown by their x-ray structures. The torsional angle τ_1 about the C2-C1' (C_2-C_1'-C_2-C_1) is of -4.34° for molecule **1**, while it is -57.51° for molecule **2** and -21° for molecule **3** (Fig. 5.1).

FIG. 5.1: Molecules for which the comparison between experimental and calculated structures have been made.



It was particularly important to see if the semi empirical methods AM1 and PM3 could predict the preferred conformation of this group because extensive conformational studies were performed in the FAAs molecules. Previous studies on flavonoids have shown that conformational preferences are important factors for recognition by responsive enzymes [115]. An attempt was made to see also if for flavone acetic acid derivatives, any correlation exists between experimental anticancer activity and conformation. The results of the rotation of the phenyl group for compounds **1**,**2** and **3** (fig. 5.1) respectively are summarized in Figures 5.2-5.4. The rotational curves obtained using the MNDO method have also been reported for comparative purposes.

The AM1 method gives similar rotational profiles for the compounds **1** and **3** (fig. 5.1) with the torsion angle $\tau_1 = \pm 30°$. To achieve a planar configuration (that is a torsion angle of 0° or 180°) these molecules must overcome a very small energy barrier of about 0.6 kcal/mole which is in the same order of the thermal energy. A configuration in which the phenyl group is at 90° is a maximum in the potential energy curves although the rotational barrier is not extremely high (2.2 kcal/mole). For compound **2** (fig. 5.1), the minimum is shifted at 40° and in order to achieve a planar configuration, a barrier of 2.5

kcal/mole has to be overcome. The energy of the planar configuration, τ_1=180°, is higher (Rotational barrier 7.8 kcal/mole) probably because the methyl group in position 5' and the hydroxyl group in position 3 are at a very short distance and repel each other. These results are in agreement with the experimental data for which the compound **2** has the phenyl group more rotated than compounds **1** and **3** (fig. 5.1).

The rotational curves obtained with the PM3 method are qualitatively similar for the three molecules, with the minima being predicted to be in between ±60° and ±120°. However, to achieve a planar conformation, compounds **1** and **3** must overcome a barrier of about 1.8 kcal/mole while for compound **2** a barrier of 5.5 kcal/mole or 11.7 kcal/mole (torsion angles 0° and 180° respectively) is request. The PM3 data suggests that all the three molecules prefer a twisted conformation of the phenyl ring and this is in contrast with the experimental evidence. The almost flat structure of compound **1** (experimental torsion angle -4.34°) could be due to an attractive interaction between the hydrogen in position 5' and the oxygen in position 3, their distance being only 2.2 Å. These results suggest that PM3 may still suffer the deficiency of the MNDO method in overestimating the repulsion between the cores of non-bonded atoms. A comparison with the rotational curves obtained with MNDO shows that both methods predict the same minima: however, the rotational barriers obtained with MNDO are far higher than those obtained with PM3 (Figs 5.2 - 5.4).

In order to check which one of the two optimized structures (the AM1 or PM3) was more accurate, a single point *ab-initio* calculation at the minimum predicted by AM1 and that predicted by PM3 was performed on molecule **1** with the 3-21G basis set. *Ab-initio* calculations obey the variational principle, and therefore the more accurate structure (with the 3-21G basis set) is that which gives the lowest energy. The energy of the 3-21G//AM1 structure was lower than that of 3-21G//PM3 by 0.01 au (about 6 kcal/mole). The force constants have also been analyzed; The maximum force constant for the 3-21G//AM1 geometry was -0.13061 Hartree/Rad for the angle O_1-C_2-C_3 (Fig. 5.1) (=121.60°) while for 3-21G//PM3 it was 0.11608 on the same angle (=123.48°) that is, the optimized 3-21G structure would have the angle O_1-C_2-C_3 (fig. 5.1) bigger than

121.6° but smaller than 123.5°. All the other force constants were very small for both methods; the gradient of the torsion angle 2'-1'-2-1 was -0.007 for 3-21G//AM1 and 0.008 for 3-21G//PM3 indicating that the optimized 3-21G torsional angle 2'-1'-2-1 would be smaller than 90° and bigger than 30°.

The geometry of the three flavones **1-3** (fig. 5.1) were optimised using as a starting point the geometry of the first minima from the 0° (see figs. 5.2-5.4); for the xanthone the starting point was that with standard bond lengths and bond angles. The comparison with the experimental data are presented as a graph of the difference (experimental value - calculated value) of bond lengths (Fig. 5.5), bond angles (Fig. 5.6) against the frequency value, that is the number of times a given value of the above difference is obtained. As can be noted in fig. 5.5, both AM1 and PM3 tend to give bond lengths which are slightly longer than the experimental ones (0.01 Å); very few bond lengths are longer than 0.05 Å and they all derive from the fact that the x-ray structure of xanthone, compound **4** (fig. 5.1) gives a molecule which is not symmetric, for instance the C-O_1 distances are 1.360 Å and 1.329 Å [114] while the calculation gives a symmetric molecule with these two bonds having the same length (1.386Å for AM1 and 1.378 Å for PM3). In general, both AM1 and PM3 predict bond lengths which are in agreement with the experimental values, PM3 performed slightly better than AM1 having a bigger number of bond lengths that differ less than 0.01 Å (in absolute value) from the experimental bond lengths (Fig. 5.5). The bond angles, for this class of molecules, are reproduced very well by both methods AM1 and PM3 as can be seen in fig. 5.6; both graphs show a symmetric distribution of differences in bond angles centred at 0°.

In conclusion, both AM1 and PM3 can be used with confidence to predict bond lengths and bond angles for this class of molecules. However, in predicting the conformation of groups with rotational freedom, AM1 gives results which agree better than PM3, with experimental values.

FIG. 5.2: Potential energy profiles of the compound **1** of Fig. 5.1, with AM1, PM3 and MNDO methods

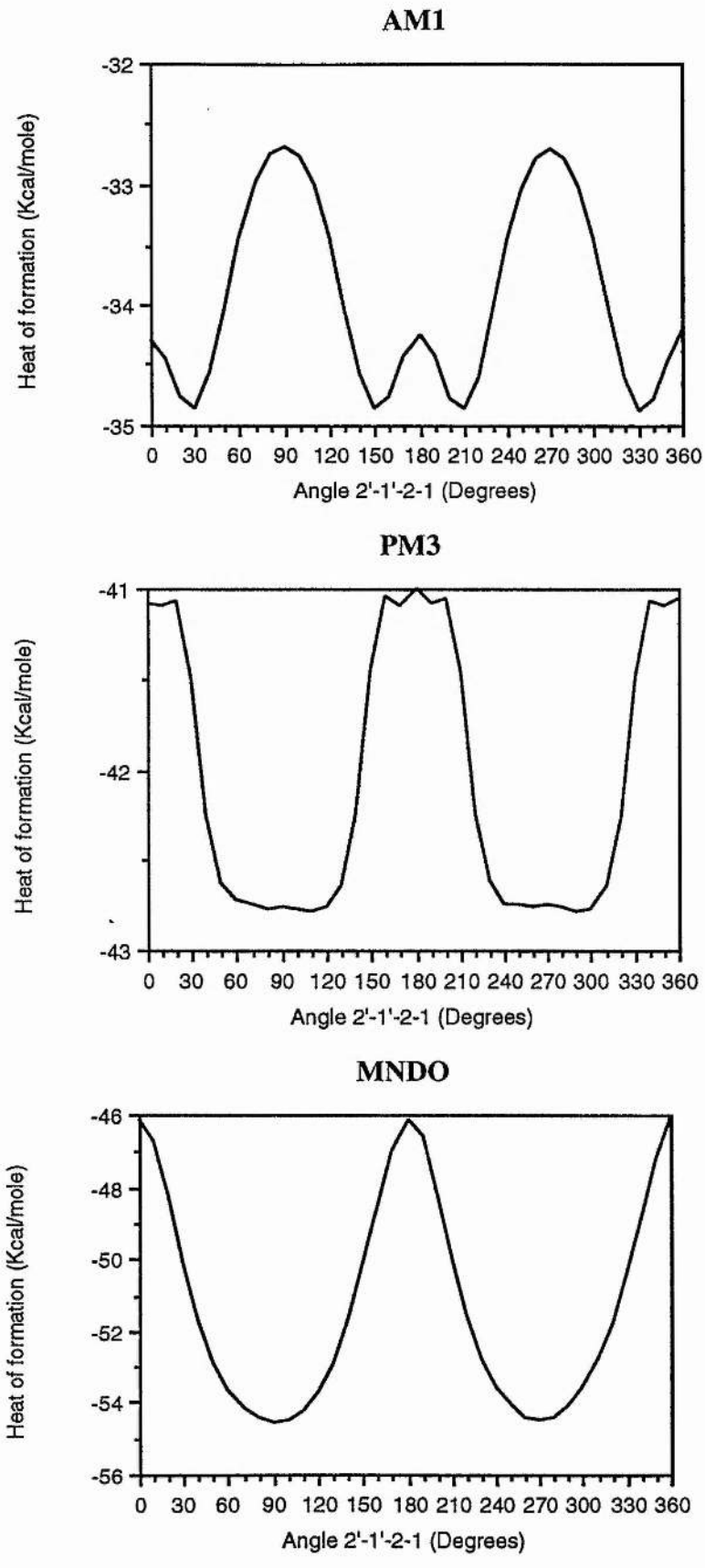


FIG. 5.3: Potential energy profiles of the compound **2** of Fig. 5.1, with AM1, PM3 and MNDO methods

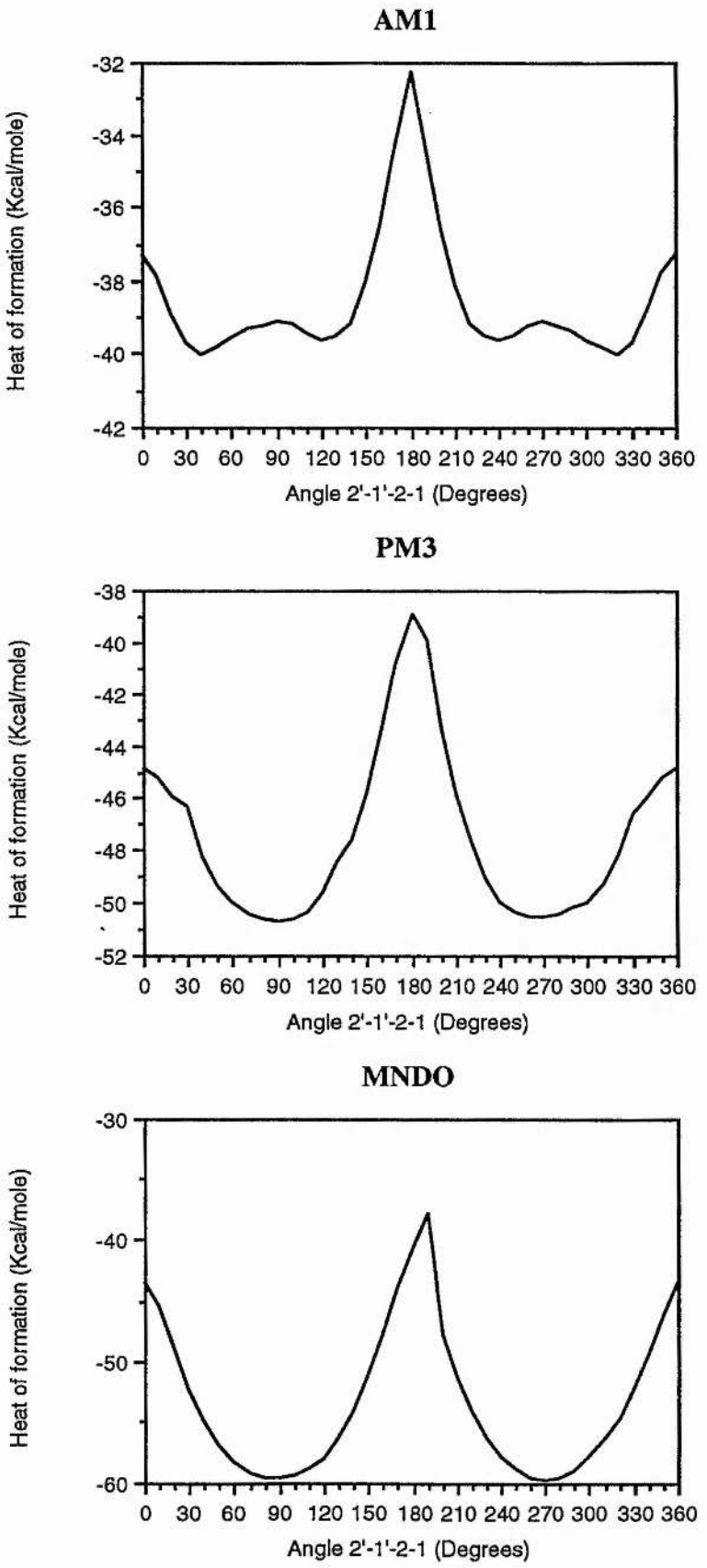


FIG. 5.4: Potential energy profiles for compound 3 in Fig. 5.1 with AM1, PM3 and MNDO method.

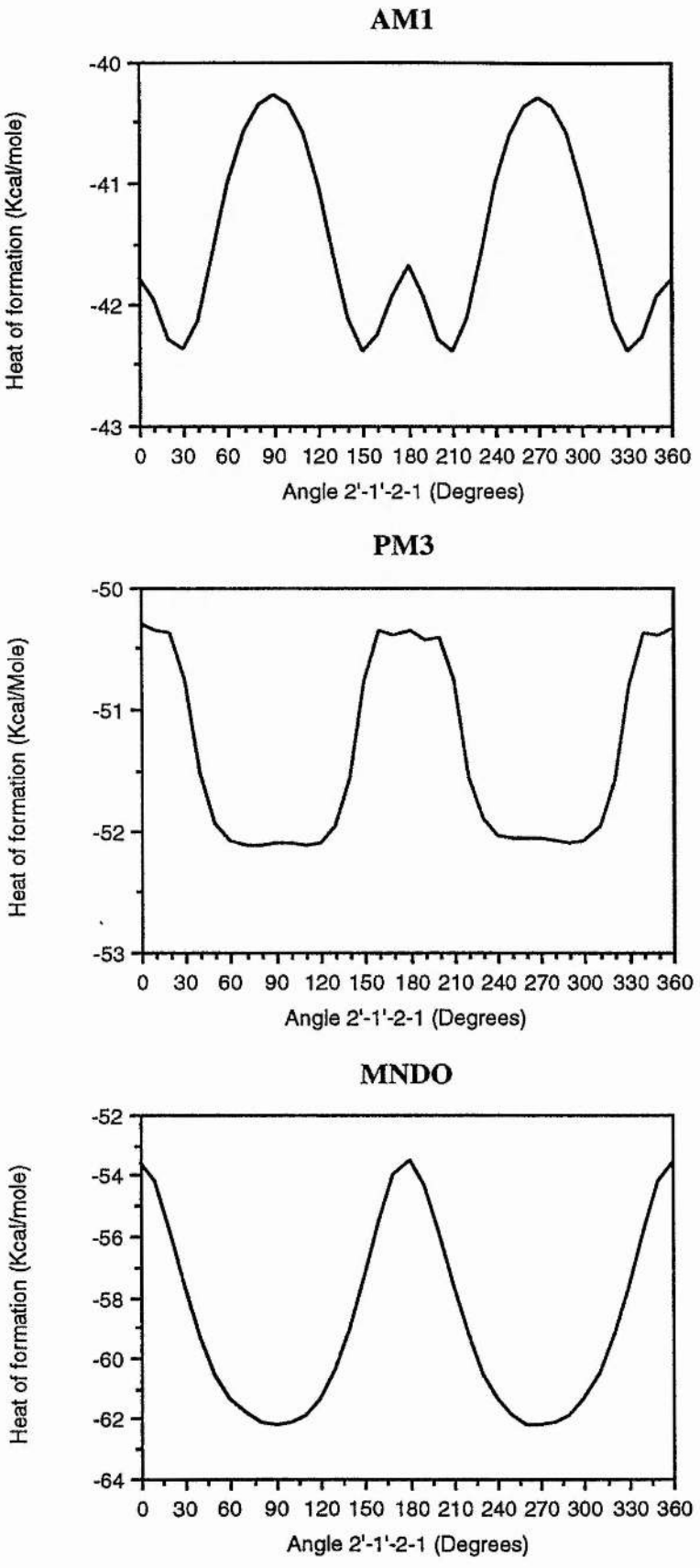
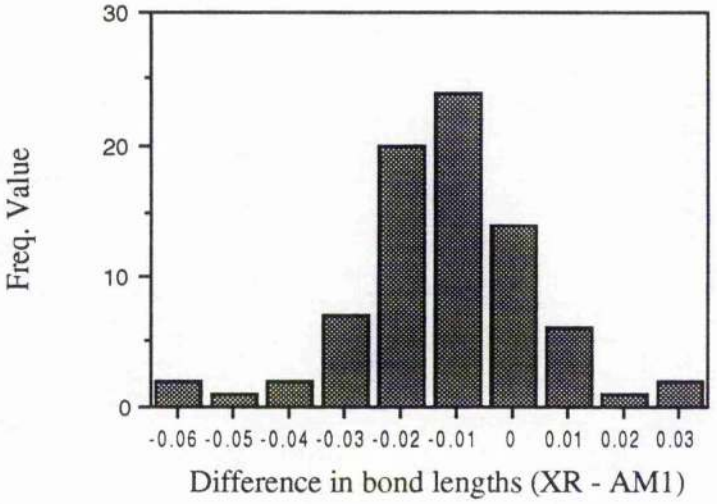


FIG. 5.5: Comparison between experimental and calculated bond lengths for the molecules reported in fig. 5.1.

AM1



PM3

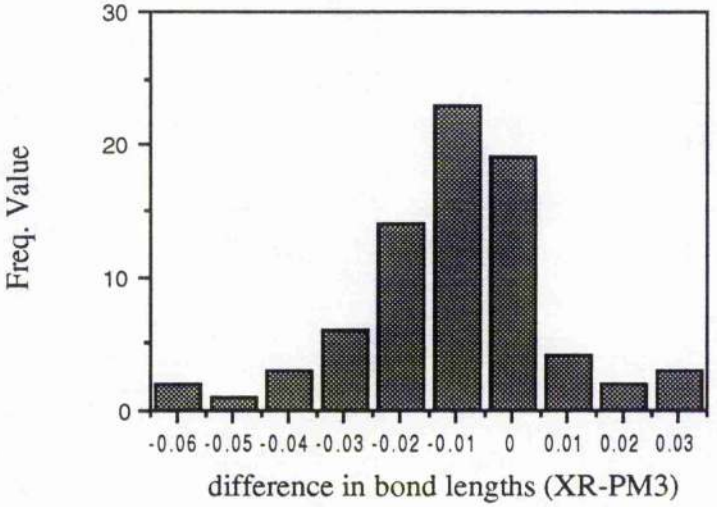
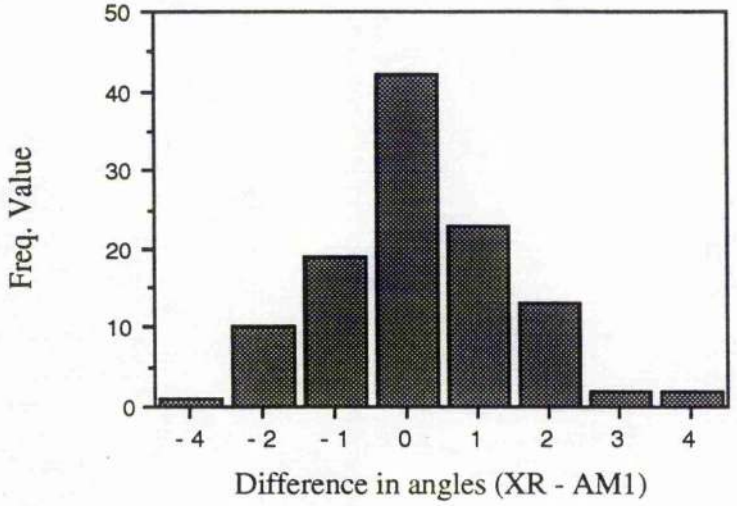
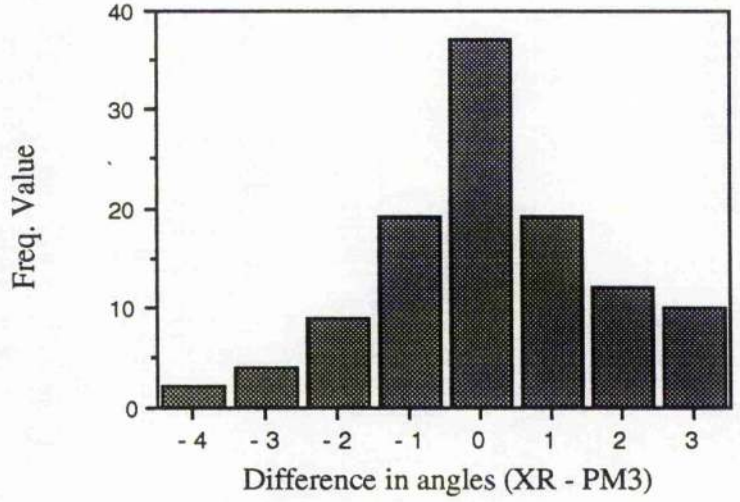


FIG. 5.6: Comparison between experimental and calculated bond angles for the molecules reported in fig. 5.1.

AM1



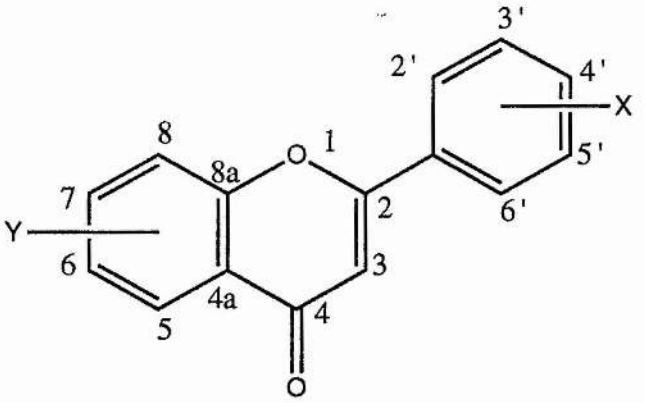
PM3



5.1.2 Atomic charges and ^{13}C chemical shifts

The chemical shift is a chemical observable and therefore can be calculated directly from the wave function. It is also very easy to measure and experimental values for the flavonoids shown in fig. 5.7 were available in the literature [116]. Unfortunately its calculation is not straightforward and it has not been attempted in this work. Instead, the experimental chemical shifts of the molecules shown in fig. 5.7 have been compared with the calculated atomic charges, both with AM1 and PM3. Atomic charges are easy to calculate and can been related to chemical shifts because it depends largely on the electron density around the nucleus in question [117-120].

FIG. 5.7: Structure and numbering scheme of the Methoxyflavones for which the chemical shifts have been compared with the calculated atomic charges.



1: X=H, Y=H; **2**: X=2'-OCH_3, Y=H; **3**: X=3'-OCH_3, Y=H; **4**: X=4'-OCH_3, Y=H; **5**: Y=5-OCH_3, X=H; **6**: Y=6-OCH_3, X=H; **7**: Y=7-OCH_3, X=H; **8**: Y=8-OCH_3, X=H;

The carbon-13 chemical shift data for the molecules shown in fig. 5.7 have been measured by Kingsbury et al. [116] and are reported in Table 5.2. This table also reports the atomic charges calculated using Mulliken population analysis with AM1 and PM3 semi empirical methods on the fully optimized structures; the heats of formation are reported in table 5.1. The comparison is presented as graphs of charges versus chemical

shifts in fig. 5.8 (with AM1 charges) and fig. 5.9 (with PM3 charges). These graphs show a good correlation between the measured chemical shifts and the atomic charges obtained from the two methods (AM1 and PM3) for all but molecule **6** (Fig. 5.7). The reason why for molecule **6** the correlation between charges and chemical shift is not linear is not very clear, initial investigations point to a misassignment of the NMR spectra for this molecule. This hypothesis is supported by the fact that for this particular molecule four chemical shifts are not reported in the paper (see table 5.2) and the concentration of the solution from which this spectra was obtained is far smaller than that of the other molecules [116].

In conclusion, AM1 and PM3 give atomic charges which correlate with experimental carbon-13 chemical shifts and this gives confidence that these methods describe correctly, the electron distribution for the class of molecules related to flavones.

TABLE 5.1: AM1 and PM3 heats of formation (kcal/mole) of the optimized structures of molecules in fig. 5.7

Ref. (see Fig.5.8)	$\Delta H_{f\ AM1}$	$\Delta H_{f\ PM3}$
1	6.7436	-1.6341
2	-28.1236	-37.4577
3	-30.5312	-39.2816
4	-31.6164	-39.9834
5	-27.3041	-35.7176
6	-29.7068	-38.5659
7	-31.7559	-40.3233
8	-27.7584	-36.2763

TABLE 5.2: ^{13}C chemical shift, AM1 and PM3 atomic charges of isomeric Methoxyflavones.

Compd.[b]		2'	3'	4'	5'	6'	3	5	6	7	8	4a	8a	CO
1	c	126.000	128.800	131.300	128.800	126.000	107.300	125.400	124.900	133.500	117.900	123.700	156.000	178.000
	d	-0.087	-0.137	-0.108	-0.136	-0.097	-0.300	-0.047	-0.156	-0.078	-0.159	-0.194	0.097	0.312
	e	-0.063	-0.107	-0.082	-0.105	-0.072	-0.301	-0.008	-0.144	-0.045	-0.147	-0.212	0.105	0.380
2	c	157.800	116.600	132.200	120.500	129.100	112.500	125.400	124.600	133.300	117.800	a	156.200	178.700
	d	0.123	-0.214	-0.074	-0.171	-0.066	-0.309	-0.047	-0.168	-0.081	-0.159		0.099	0.313
	e	0.128	-0.195	-0.042	-0.144	-0.036	-0.294	-0.009	-0.147	-0.045	-0.148		0.107	0.379
3	c	111.500	159.700	116.900	129.800	118.500	107.500	125.400	124.900	133.200	117.900	123.700	155.900	178.000
	d	-0.115	0.073	-0.184	-0.100	-0.134	-0.297	-0.047	-0.166	-0.078	-0.158	-0.195	0.097	0.311
	e	-0.101	0.082	-0.168	-0.064	-0.111	-0.299	-0.008	-0.145	-0.045	-0.146	-0.213	0.105	0.380
4	c	127.700	114.200	162.100	114.200	127.700	105.900	125.300	124.700	133.000	117.700	123.700	155.800	177.900
	d	-0.056	-0.165	0.103	-0.213	-0.059	-0.308	-0.047	-0.166	-0.079	-0.160	-0.194	0.098	0.312
	e	-0.028	-0.148	0.111	-0.190	-0.031	-0.308	-0.008	-0.145	-0.045	-0.147	-0.212	0.105	0.381
5	c	125.600	128.600	131.000	128.600	125.600	108.700	159.400	109.800	133.400	106.200	114.000	157.900	177.800
	d	-0.089	-0.137	-0.110	-0.135	-0.098	-0.293	0.163	-0.252	-0.041	-0.200	-0.223	0.130	0.315
	e	-0.064	-0.107	-0.084	-0.105	-0.073	-0.297	0.182	-0.245	-0.001	-0.192	-0.244	0.137	0.385
6	c	126.100	128.900	131.300	128.900	126.100	106.700	104.800	a	123.600	119.400	a	a	a
	d	-0.087	-0.137	-0.108	-0.136	-0.097	-0.304	-0.076		-0.151	-0.124			
	e	-0.064	-0.107	-0.083	-0.105	-0.072	-0.305	-0.046		-0.125	-0.107			
7	c	125.800	128.700	131.100	128.700	125.800	107.200	126.700	114.100	163.700	100.200	117.600	157.700	177.400
	d	-0.088	-0.137	-0.108	-0.135	-0.097	-0.294	-0.019	-0.190	0.133	-0.245	-0.230	0.136	0.315
	e	-0.064	-0.107	-0.083	-0.105	-0.072	-0.296	0.023	-0.182	0.148	-0.238	-0.252	0.146	0.385
8	c	126.100	128.700	131.200	128.700	126.100	107.100	114.200	124.600	116.100	148.800	124.000	146.000	178.000
	d	-0.084	-0.138	-0.108	-0.138	-0.097	-0.307	-0.083	-0.129	-0.160	0.056	-0.166	0.072	0.311
	e	-0.060	-0.107	-0.083	-0.106	-0.172	-0.308	-0.045	-0.104	-0.135	0.047	-0.186	0.076	0.379

[a] Values of chemical shifts not reported in Ref. [116] .[b] See Fig. 5.8. [c] ∂(ppm). [d] Charges AM1.[e] Charges PM3

FIG. 5.8: AM1 Atomic charges of molecules shown in Fig. 5.7 versus Carbon-13 chemical shifts (ppm)

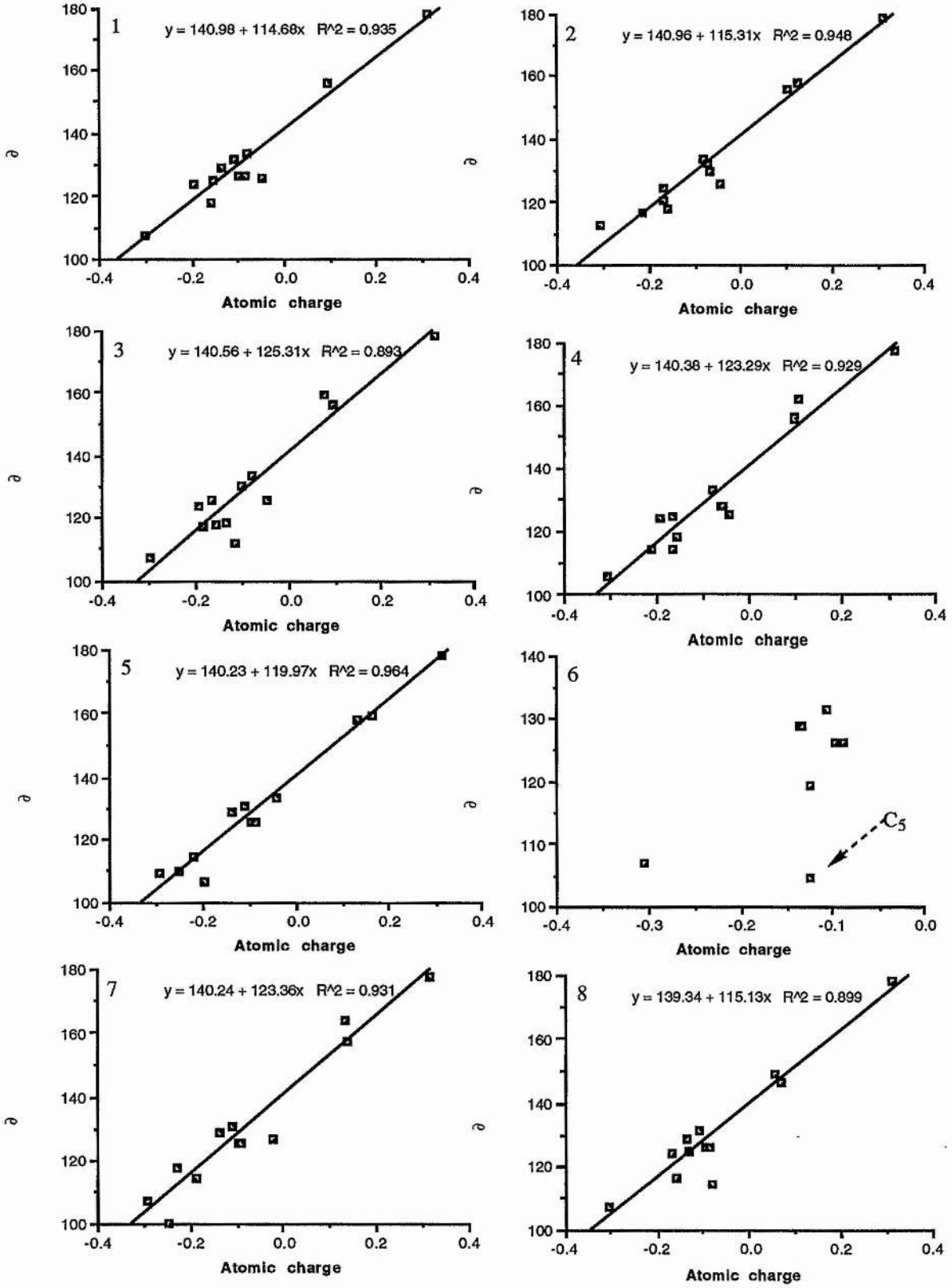
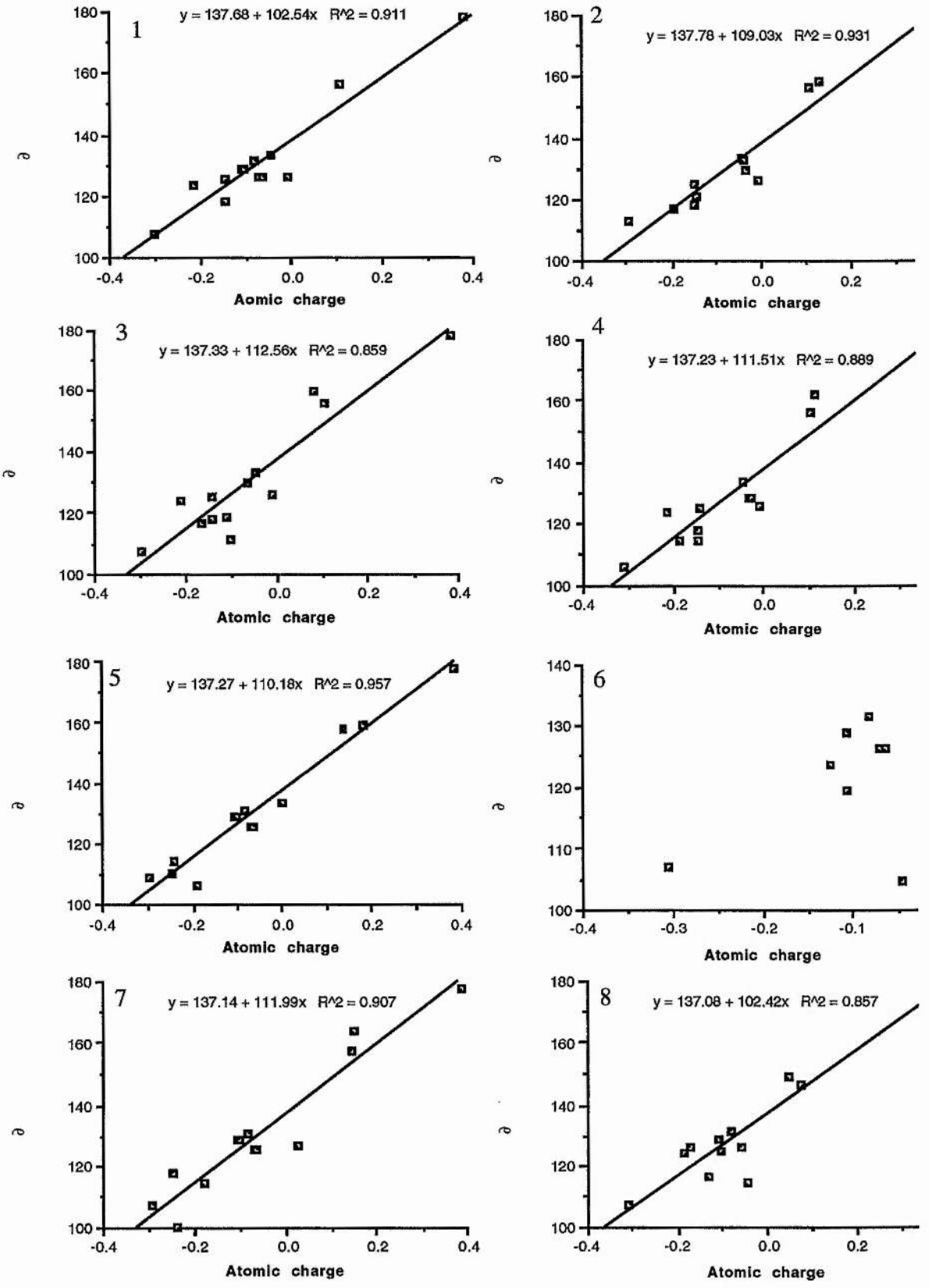


FIG. 5.9: PM3 Atomic charges of molecules shown in Fig. 5.7 versus Carbon-13 chemical shifts (ppm).



CHAPTER 6:

Semi empirical Results for FAAs

Introduction

This chapter reports the results of the structure-activity relationship studies of the derivatives of flavone acetic acid using the semi empirical methods AM1 and PM3 [65,66]. Among the data reported in chapter 2, the ten molecules shown in table 2.3 have been chosen for the SAR studies because of the large variation in the activity (from 0% to 100% TGI) (see chapter 2 for the definition of %TGI) that they show for small variations in the nature and position of the substituents.

Section 6.1 contains the results of the geometry optimization of the molecules in table 2.3. For the molecule studied in this work, no experimental structures were available. The geometry optimization process was complicated by the fact that all the FAAs contain three (or more) single bonds and therefore there will be several relatively close energy minima. Extensive conformational studies have been performed on all the molecules of table 2.3 and the geometry of all the molecules has been optimized using the strategy described in paragraphs 6.1.1-6.1.3. The advantages of having a standard procedure, is that it is possible to reproduce the results of a calculation at any time and the different molecules can be optimized to a similar structure of the flexible groups. This was particularly important because some electronic properties, such as the dipole moment and the electrostatic potential, are very dependent on the conformation and in order to be able to compare these properties between the different molecules, it was imperative to have all the molecules in a similar conformation, particularly for the acetic acid group.

In section 6.2 the electronic properties of FAAs are reported in an attempt to establish correlations between the anticancer activity and factors related to the electronic distribution.

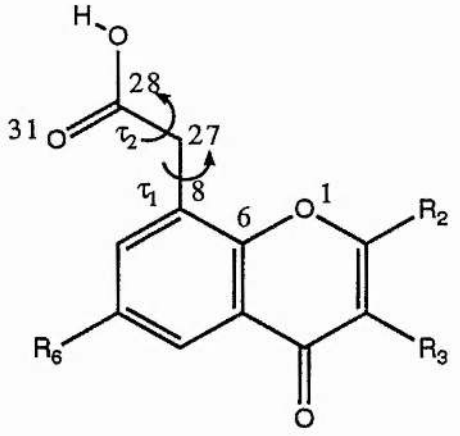
Section 6.3 contains a discussion of the results on the use of AM1 and PM3 in SAR studies.

6.1 Geometry optimization

As mentioned previously, the geometry optimization has been conducted by using a standard strategy. The initial configuration was constructed from standard bond lengths and bond angles [64]. The molecules were initially assumed to be flat and rotations around the flexible groups were performed for each of the molecules in order to determine the dihedral angles to be used for the final optimization. The conformation of the acetic acid groups was determined first, followed by that of the other groups. In this way, a number of possible conformations for each molecule have been selected and all of them used for the structure-activity relationships studies.

6.1.1 Conformation of the acetic acid group

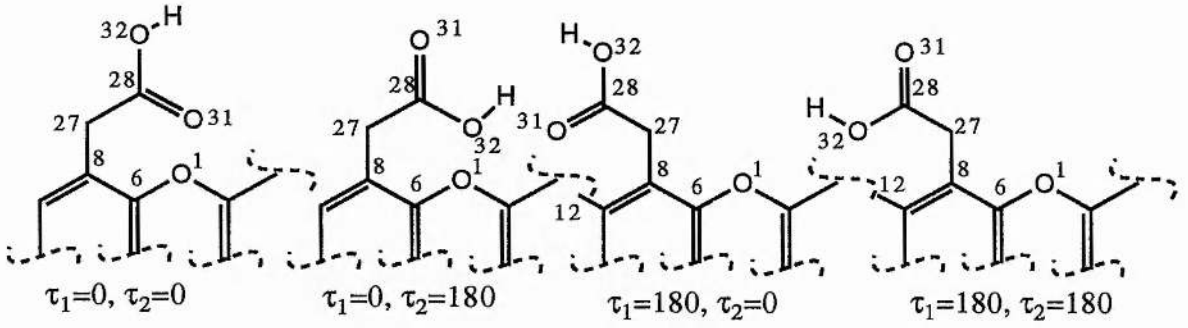
The acetic acid group is common to all the molecules studied in this work. To specify its conformation two torsional angles are required: $\tau_1 = C_{28}\text{-}C_{27}\text{-}C_8\text{-}C_6$ and $\tau_2 = O_{31}\text{-}C_{28}\text{-}C_{27}\text{-}C_8$:



The portion of the potential energy surface necessary to specify the conformation of this group has been obtained for $0 \leq \tau_1 \leq 180$ and $0 \leq \tau_2 \leq 180$ on a grid of 121 points. The value of all the other variables have been optimized at each point. The results for the FAAs of table 2.3 are presented in figs. 6.1 (AM1) and 6.2 (PM3) as contour maps τ_1 versus τ_2,

with the energy contours in kcal/mole. The difference in energy between two consecutive contours is 1kcal/mole.

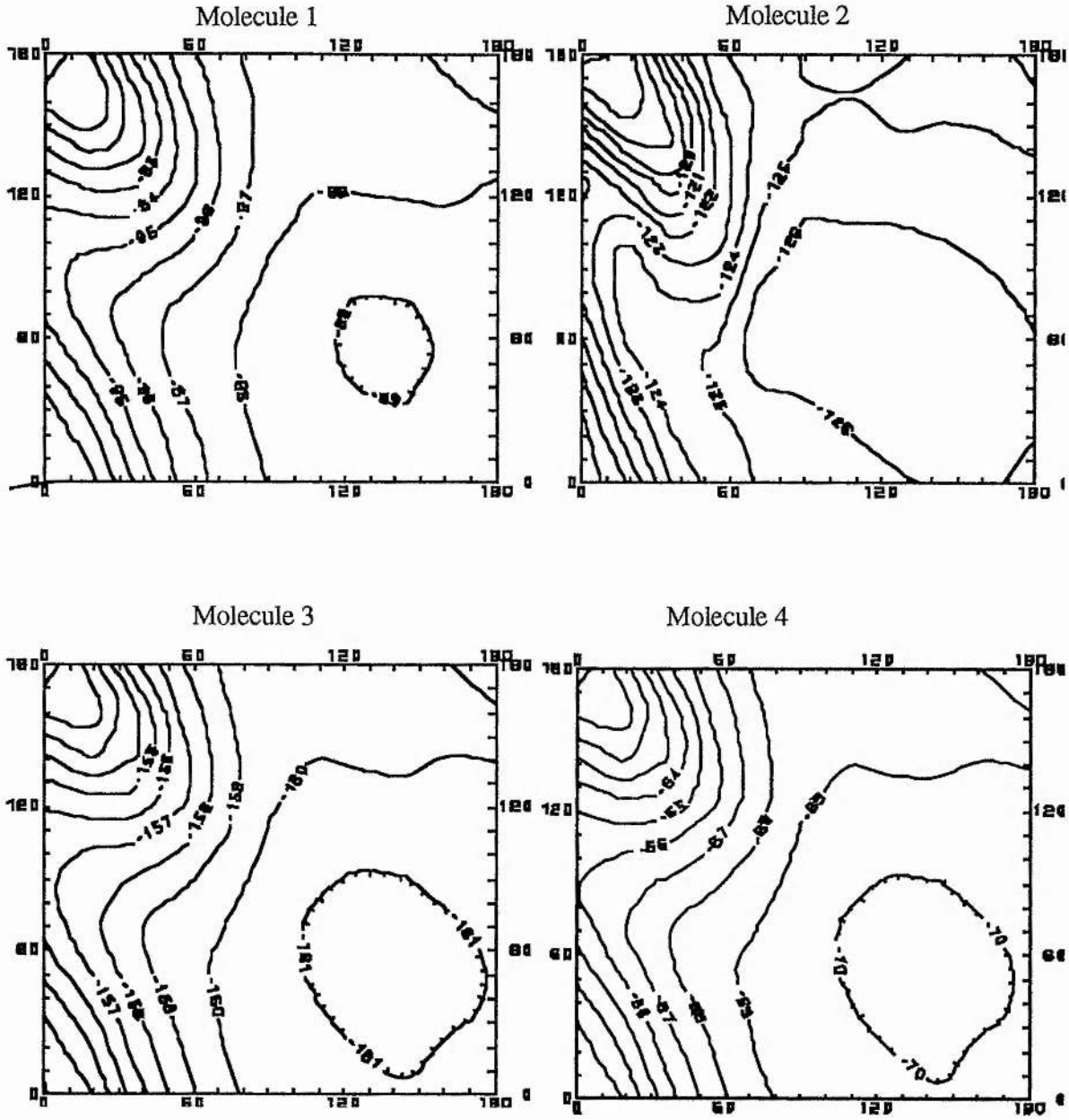
Both AM1 and PM3 predict that for these molecules, the rotation around τ_1 is not a free rotation but conformations for which $\tau_1 \sim 0°$ is at least 7 kcal/mol above the minima (Figs 6.1 and 6.2). The reason for such an impediment to free rotation may be due to the strong repulsive interaction between the lone pairs of the two oxygen's atoms O_1 and $O_{31/32}$. If we make the assumption that all the conformations whose energy is within 2kcal/mole above the minimum are possible, then the minimum predicted by AM1 is wide and shallow for all the molecules and it is achieved for $60° \leq \tau_1 \leq 300°$ and for every value of τ_2 (Fig. 6.1).



On the other hand, PM3 gives a more localized minimum and a energy barrier has to be overcome for τ_1=180, suggesting again the fact that PM3 overestimates the repulsion between the cores of nonbonded atoms; in this case H_{12} and $O_{32/33}$. There is no indication of any conformational requirement of the acetic acid group for activity with AM1 or PM3.

Within the AM1 approximation, the conformation for which τ_1=140° and τ_2=60° is a minima for all the molecules and it has been choosen as the initial configuration for the conformational studies of the other flexible groups (see paragraphs 6.1.2 and 6.1.3) or for the final geometry optimization. With PM3 instead, the initial configuration of the acetic acid group was τ_1=90° and τ_2=72° for the same reason.

FIG. 6.1: AM1 conformational energy maps of the FAAs in table 2.3, τ_1 versus τ_2. See table 2.3 for reference and text for the definition of τ_1 and τ_2.



Cont....

...Fig. 6.1 cont.

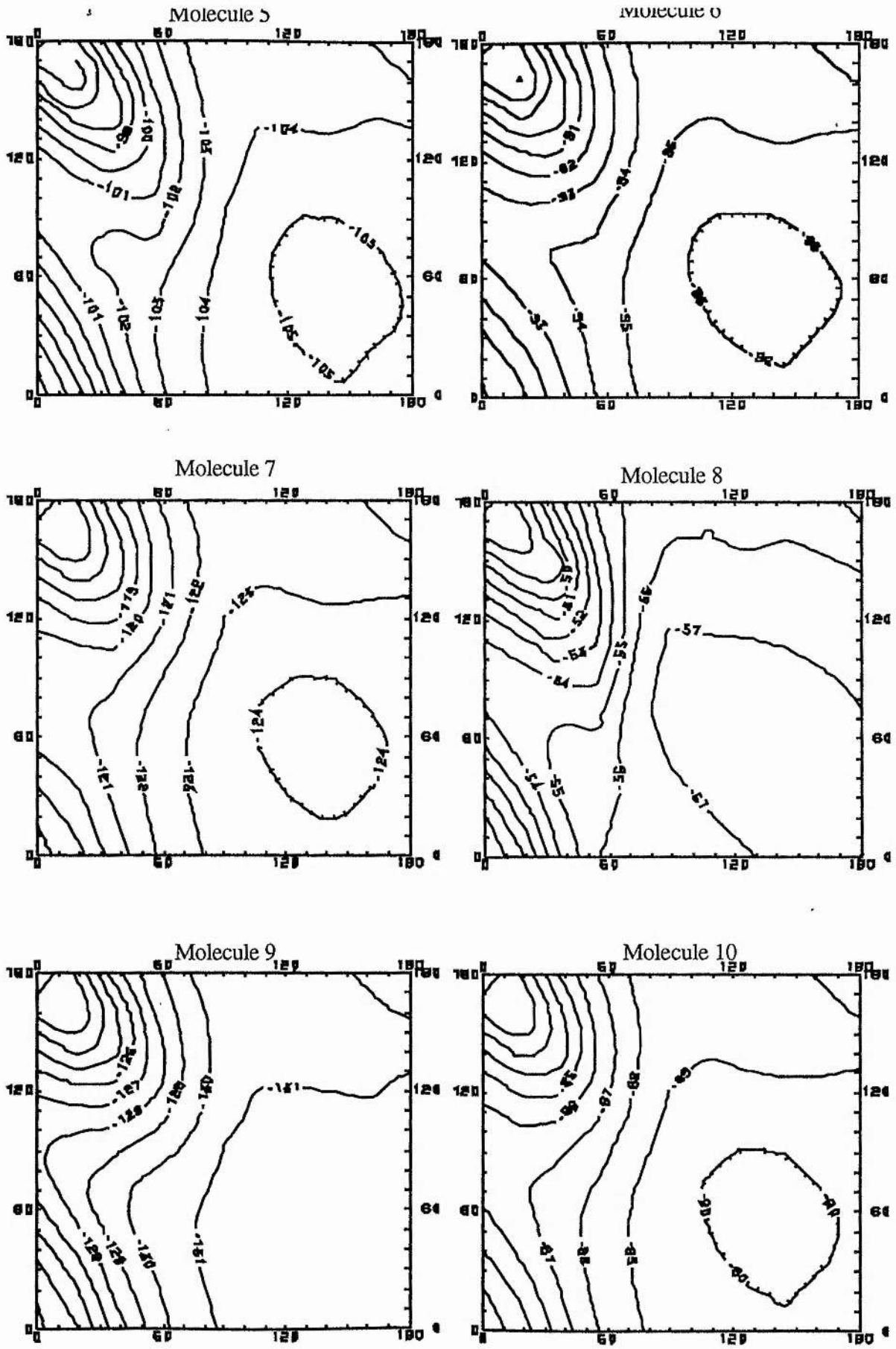
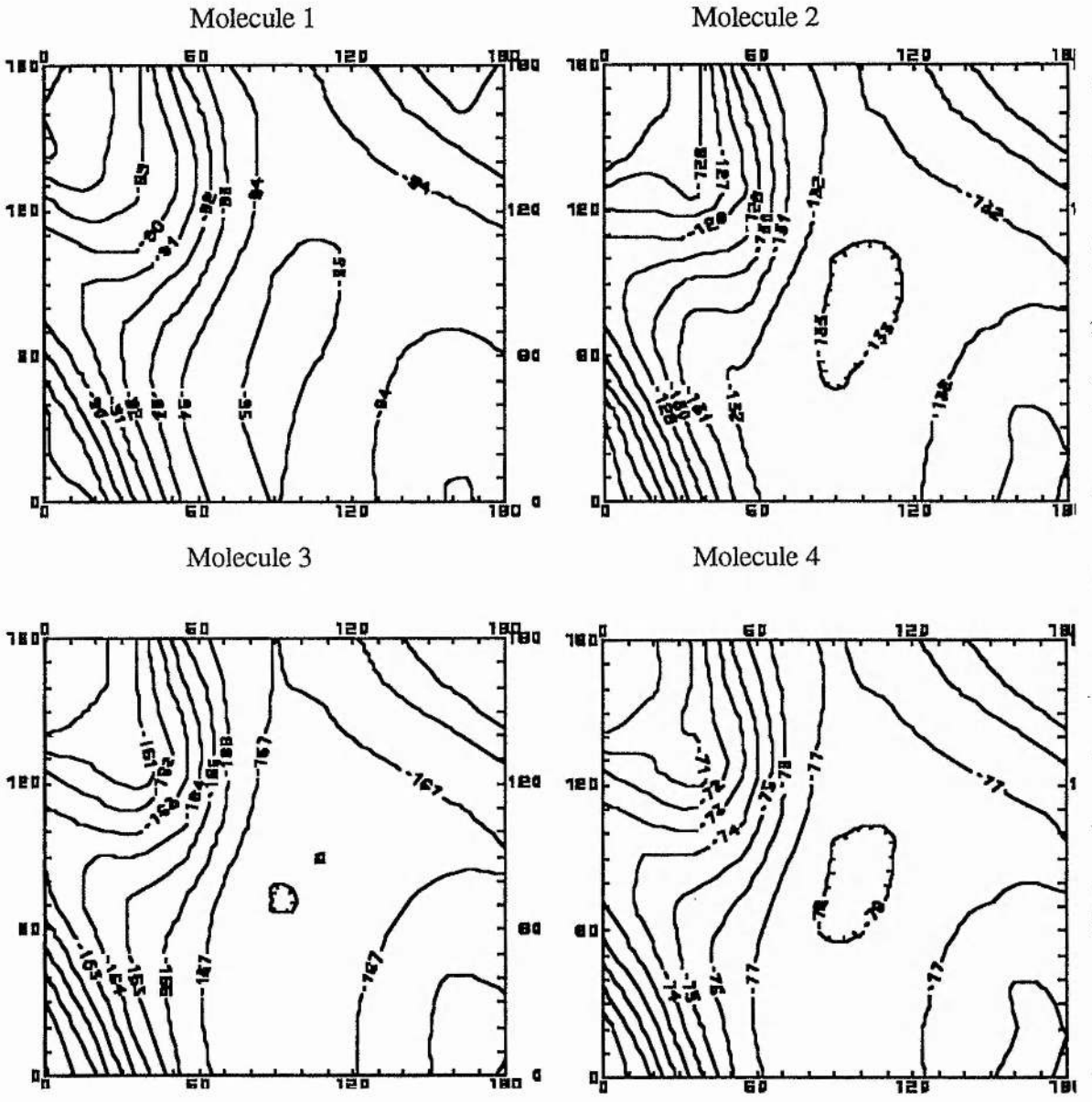

Molecule 5
Molecule 6
Molecule 7
Molecule 8
Molecule 9
Molecule 10

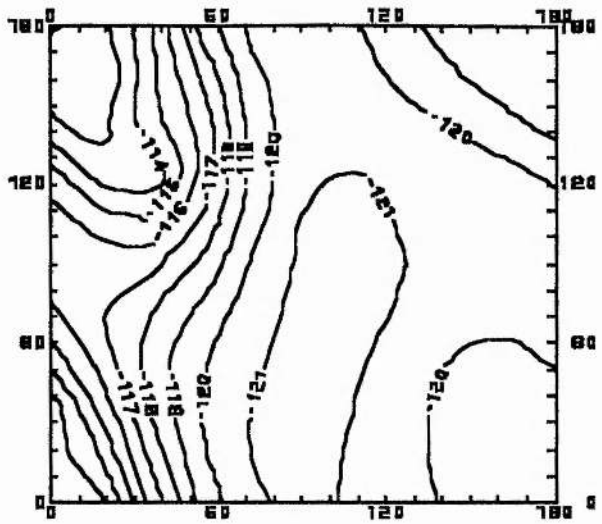
FIG. 6.2: PM3 conformational energy maps of the FAAs in table 2.3, τ_1 versus τ_2. See table 2.3 for reference and text for the definition of τ_1 and τ_2.



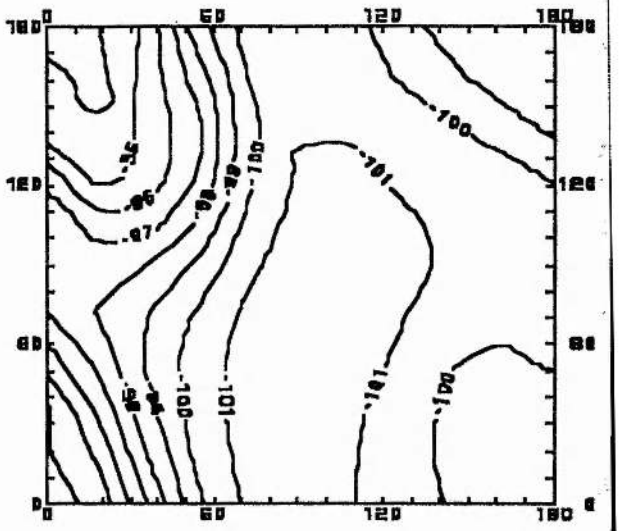
Cont....

...Fig. 6.2 cont.

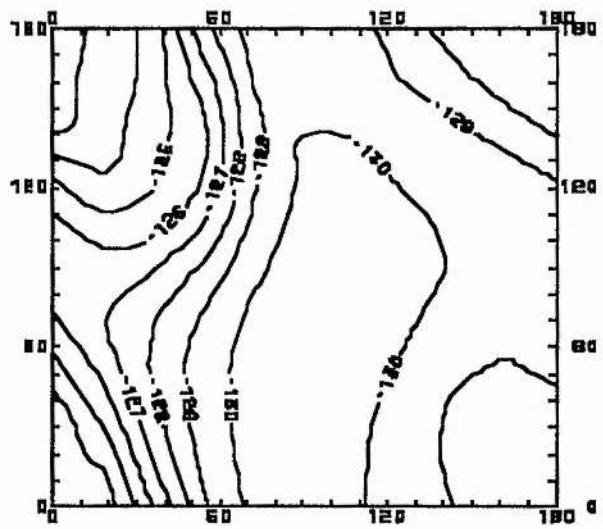
Molecule 5



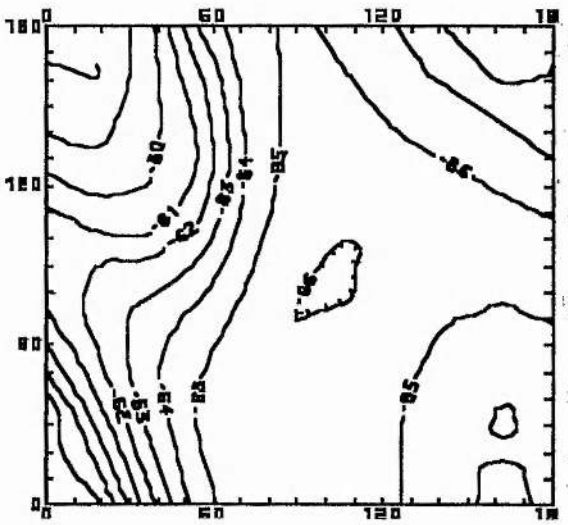
Molecule 6



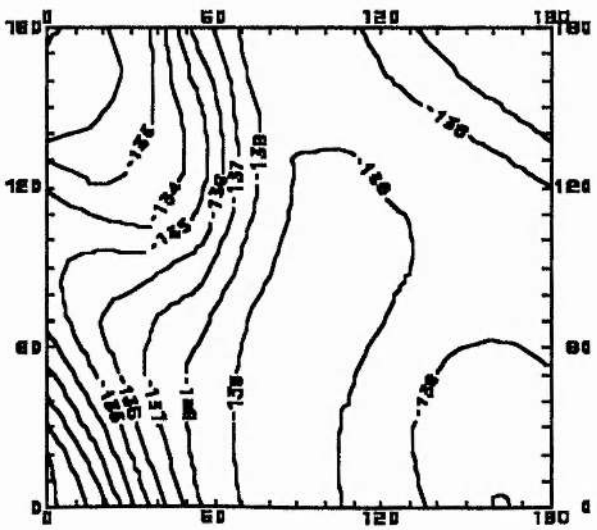
Molecule 7



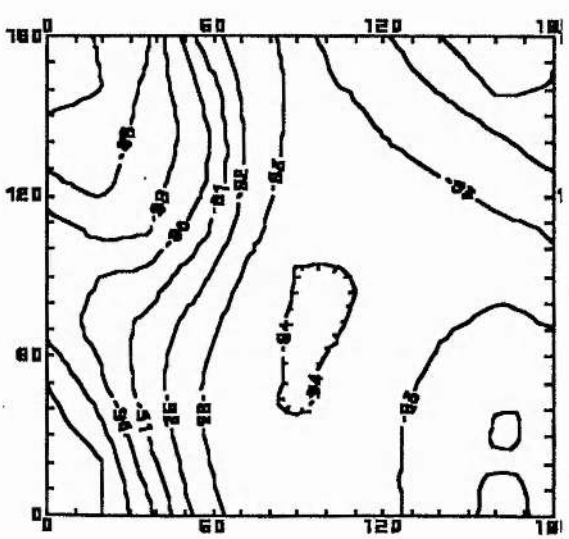
Molecule 8



Molecule 9

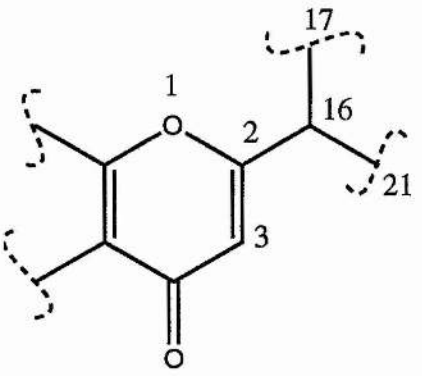


Molecule 10



6.1.2 Conformation of the other flexible groups

A glance at table 2.3, shows that the specification of the conformation of the substituent in position R_2 requires knowledge of one (molecules **1, 4, 5, 9,** and **10**) or more torsional angles (molecules **2, 3, 6** and, **8**). For molecules **1, 3, 4, 5** and, **9,** the rotational energy profiles were obtained for $0° \leq \tau_3 \leq 360°$ with a step of 10°; τ_3 is defined as the torsion angle C_{17}-C_{16}-C_2-O_1:

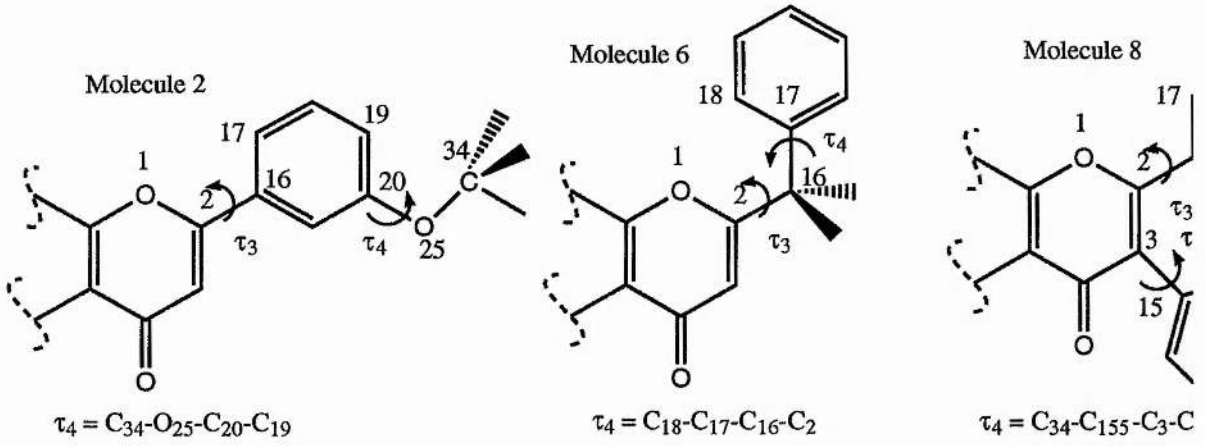


The initial conformation of the acetic acid group was chosen to be the same for all the molecules that is, τ_1=140° and τ_2=60° for AM1 and, τ_1=90° and τ_2=72° for PM3; these values of the torsional angles form a minimum in the potential energy surfaces τ_1 v. τ_2 (see figs 6.1 and 6.2). Except for the torsional angle τ_3, the molecule was allowed to relax in all calculations. Fig 6.3 shows the rotational energy profiles obtained with AM1, the same profiles for PM3 are shown in fig. 6.4.

The shape of the potential energy surface along τ_3 depends mainly on the following factors: 1) The stabilizing effect due to the delocalization of the double bounds on the whole molecule, this would favour a planar configuration. 2) the repulsive steric interaction between the hydrogen atoms in position 3 and 21, this would favour a torsional angle of 90° (maximum distances of the two H's atoms). For all the molecules under consideration AM1 predicts two minima at 30° and 210°; to achieve a planar configuration it is necessary to supply 1 kcal/mole of extra energy, except for molecule **5** which prefers an almost planar configuration (the furyl group is smaller than the phenyl group therefore the steric interaction between the hydrogen atoms in position 3 and 21 is

negligible); To achieve the configuration with τ_3=90° requires at least 3 kcal/mole extra and this is not likely to occur (although not impossible) (Fig. 6.3). According to PM3 (fig. 6.4), the rotation around τ_3 is almost free, the barrier to rotation being in the order of less than 1 kcal/mole; the planar configuration is predicted to be a minima, τ_3=0° or τ_3=180°.

For molecules **2, 6** and **9**, conformational energy maps have been obtained on a grid of 121 points τ_3 v. τ_4 (from 0° to 360°) using the same initial configuration of the acetic acid group as before and allowing the molecule to relax in all calculations, except for τ_3 and τ_4 where τ_3 has already been defined and τ_4 is defined as:



The conformational maps are shown in fig. 6.5 for AM1. The same calculations have also been performed with PM3 but they are not reported here.

FIG. 6.3: AM1 rotational energy profiles (see table 2.3 for reference and text for the definition of the torsion angle)

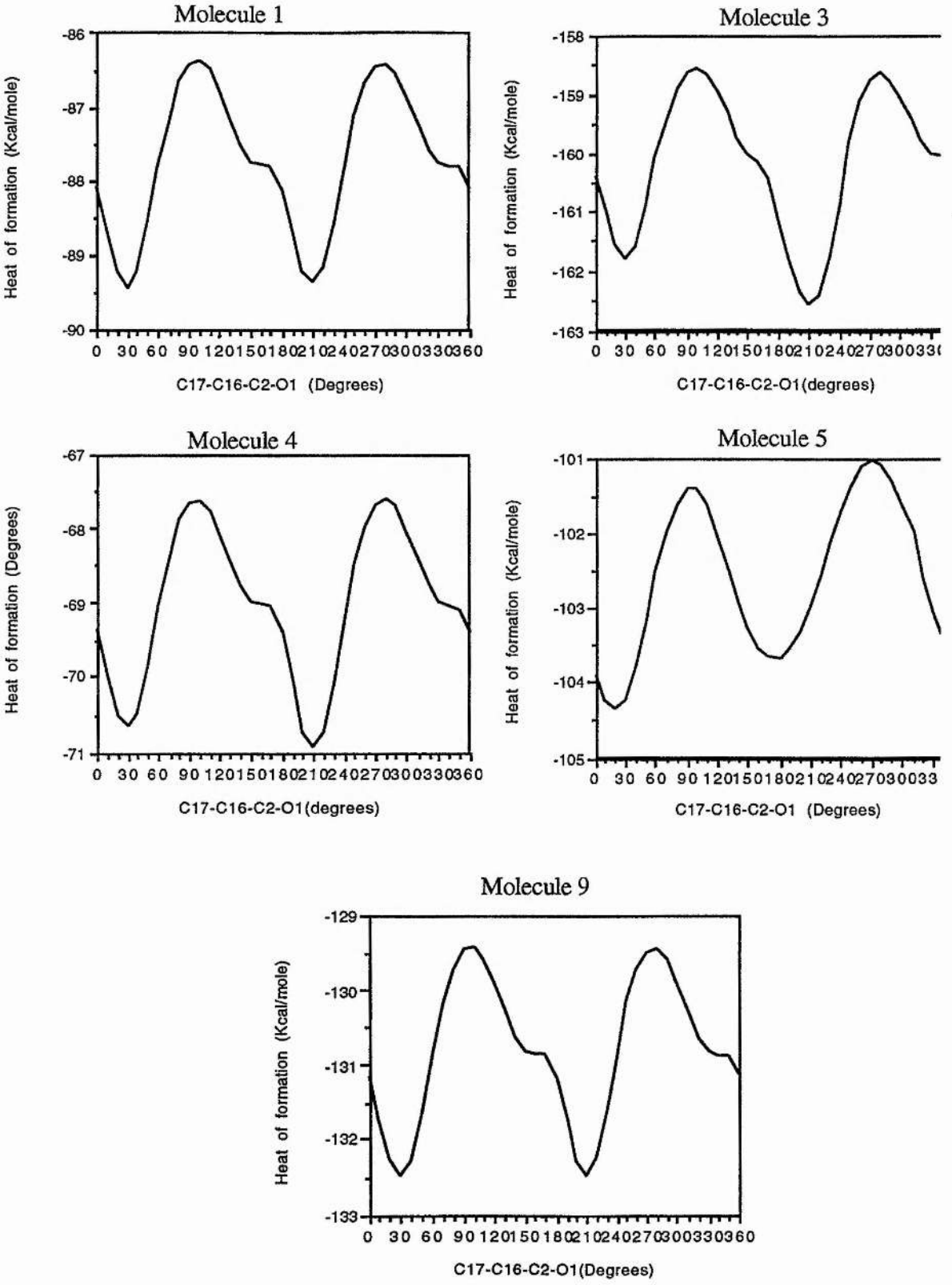


FIG. 6.4: PM3 rotational energy profiles (see table 2.3 for reference and text for the definition of the torsion angle)

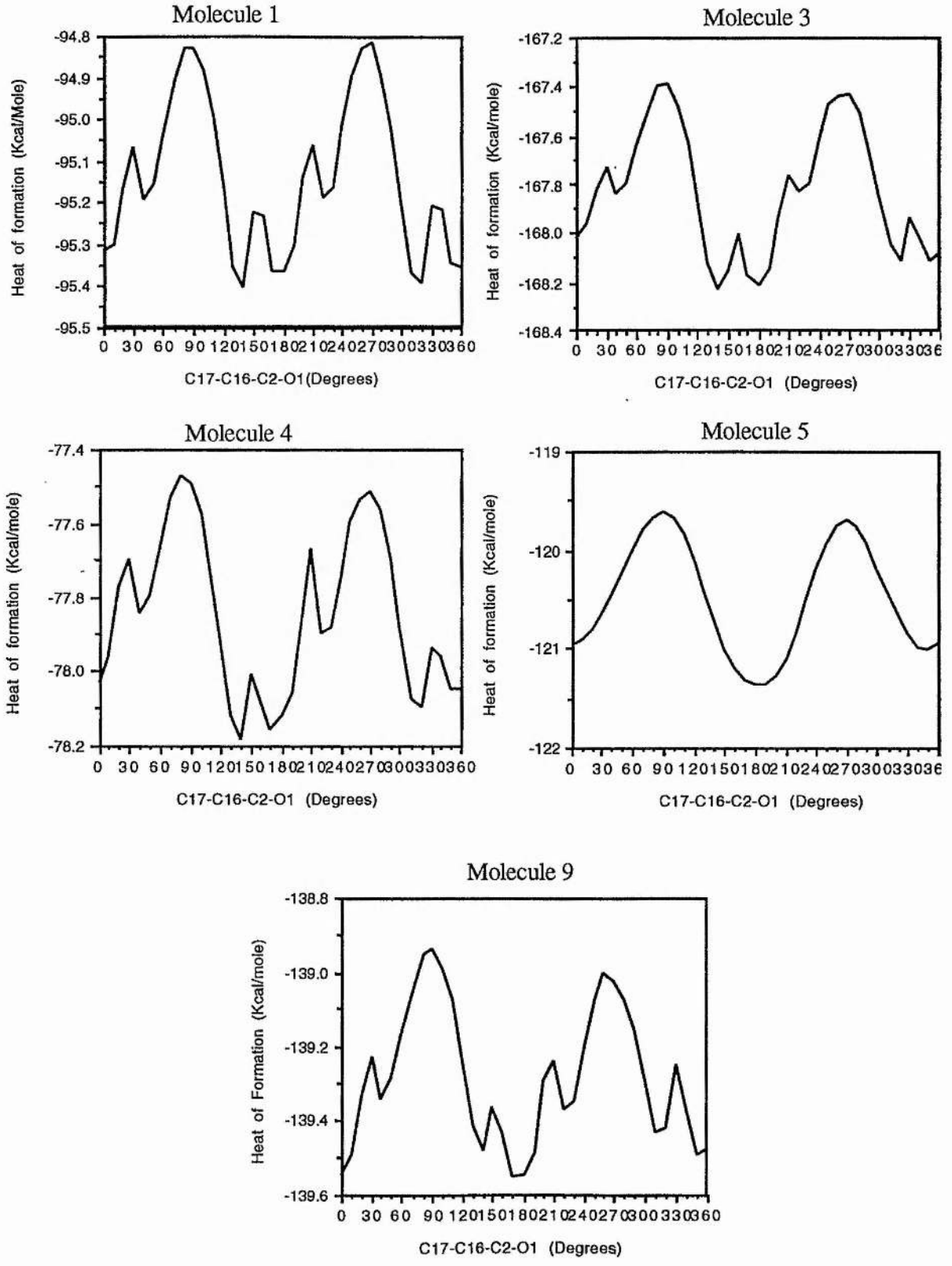
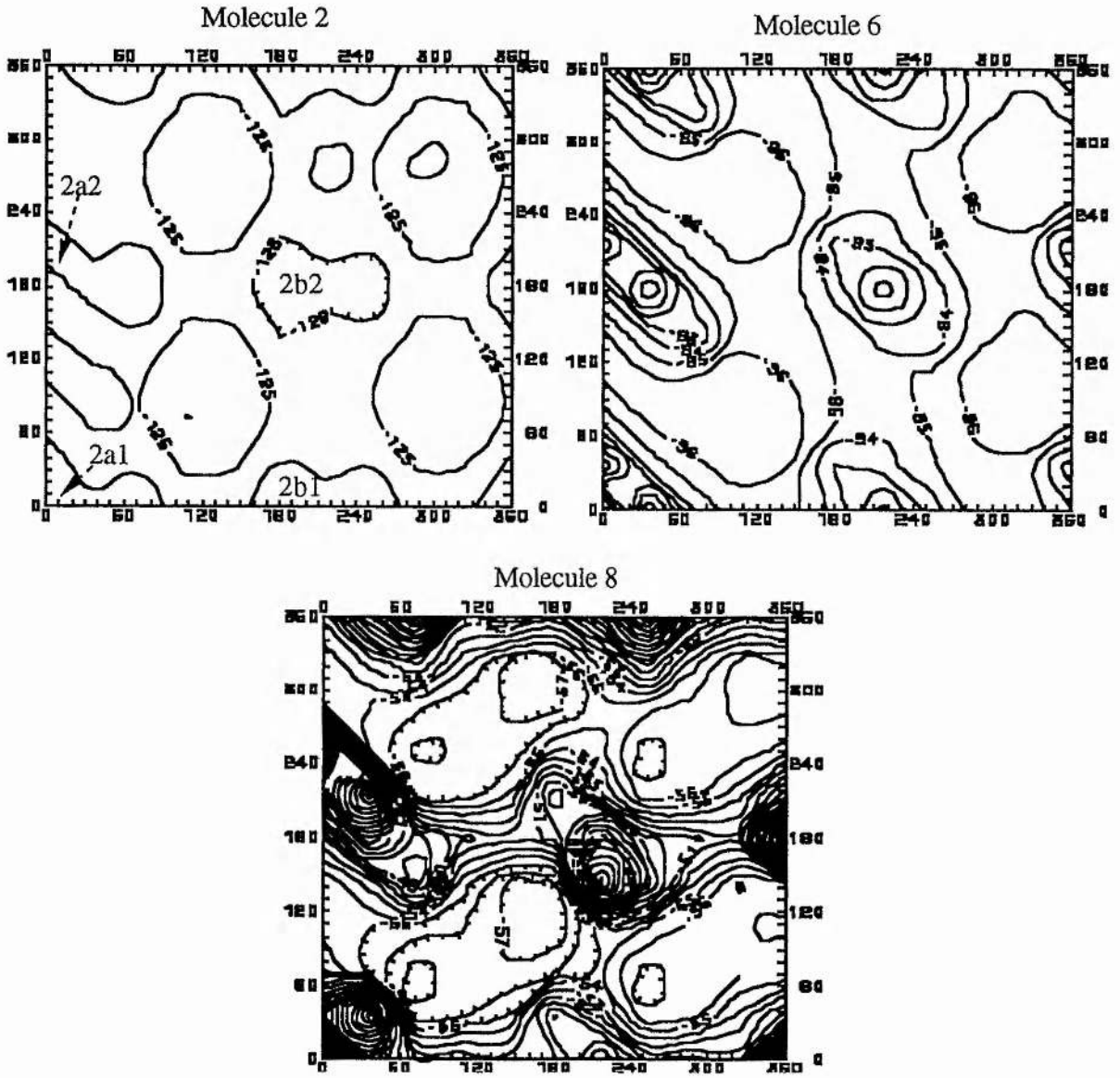


FIG. 6.5: AM1 conformational energy maps τ_3 v. τ_4. (see table 2.3 for reference and text for the definition of τ_3 and τ_4)



6.1.3 Final geometry

Figs 6.3 - 6.5 have been used to define the conformations of the molecules reported in table 2.3 that have actually been used for the structure activity correlation studies. The conformation for which τ_3=30° or τ_3=210° is a minima on the AM1 potential energy surfaces for all the FAAs, while on the PM3 potential energy surfaces the minima is found for τ_3=0° or τ_3=180° (fig. 6.3 - 6.5). These have been used as initial values of the torsional angle τ_3 for the final geometry optimization of the FAAs with AM1 and PM3 respectively. The geometry optimization was straightforward for most of the molecules and the condition of a gradient norm ≤ 0.05 was imposed. The heats of formation obtained for all the molecules in the different conformations are reported in table 6.1 and 6.2 for the AM1 and PM3 method respectively. For convention, the conformation for which τ_3=30° (or 0° for PM3) will be called 'a' while that for τ_3=210° (or 180 for PM3) 'b'. For molecule **2**, four conformations are possible and the indexes '1' and '2' refer to the position of the methoxy group: τ_4=0° or τ_4=180° respectively (See fig 6.5). Force calculations were performed and these conformations were found to be minima in the potential energy surface, all having the second derivatives of the energy (force constant) positive.

Fig. 6.6 shows the optimized bond lengths and bond angles for the lead molecule flavone acetic acid; The value of the C_2-C_3 bond length (1.35 Å) suggest that the pyrone ring has not a significant aromatic character in the FAA; for a delocalization of the π system over the molecule, the bond C_2-C_3 would be expected to be longer. This result was expected from the similarity with the parent molecule 4H-pyran-4-one, whose low aromatic character was suggested by Frieman and Allen [121], Thomson and Edge [125]. Norris et al. also suggested this as a result of Zeeman studies on the molecule [122].

TABLE 6.1: AM1 Heat of formation of FAAs (see fig. 6.6)

Ref.[a]	τ_1[b]	τ_2[b]	τ_3[b]	τ_4[b]	ΔHf (kcal/mole)
1a	134.35	57.89	26.61	-	-89.29
2a1	134.53	58.04	27.03	-0.46	-126.56
2a2	134.59	57.95	28.68	-179.44	-126.72
2b1	135.18	58.90	-152.21	0.54	-126.63
2b2	133.88	56.37	-152.53	178.79	-126.59
3a	134.36	57.87	27.17	-	-161.60
3b	133.70	56.94	-155.39	-	-161.59
4a	134.35	57.80	25.40	-	-70.60
4b	134.69	57.91	-152.18	-	-70.58
5a	133.97	57.99	-3.37	-	-105.07
5b	137.04	57.52	-179.76	-	-105.54
6a	133.22	59.75	91.35	119.97	-96.57
7	134.40	59.40	-	-	-124.52
8a	135.97	58.27	44.67	55.15	-57.78
9a	135.16	60.37	26.82	-	-132.03
10	134.89	58.52	-	38.40	-90.54

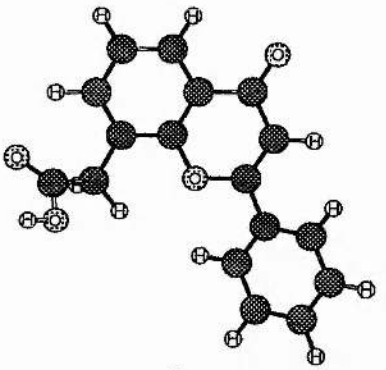
[a] See fig.6.6 and table 2.3 - [b] Optimized values. See text for definition

TABLE 6.2: PM3 Heat of formation of FAAs (fig. 6.6)

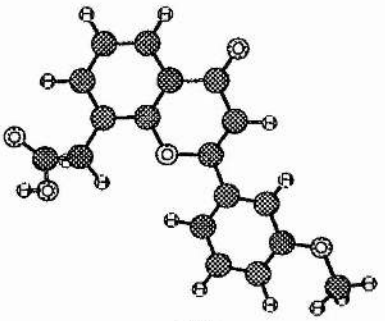
Ref.[a]	τ_1[b]	τ_2[b]	τ_3[b]	τ_4[b]	ΔHf (kcal/mole)
1a	93.45	73.12	-8.56	-	-95.42
2a1	93.79	73.32	-9.12	18.28	-132.99
2a2	94.37	74.15	-8.81	165.24	-133.08
2b1	97.23	79.04	174.02	-18.12	-132.94
2b2	88.81	66.85	169.96	-164.82	-133.32
3a	93.11	72.61	-9.21	-	-168.16
3b	91.90	69.55	171.87	-	-168.34
4a	94.20	73.16	-6.70	-	-78.11
4b	91.95	72.25	168.92	-	-78.22
5a	91.76	70.12	-9.43	-	-121.02
5b	100.26	80.63	-176.04	-	-121.40
6a	94.85	75.68	93.44	81.39	-101.76
7	95.90	77.80	-	-	-130.77
8a	100.16	85.69	-62.10	95.15	-66.14
9a	88.32	66.24	-7.74	-	-139.60
10	95.70	77.52	-	59.70	-95.49

[a] See fig.6.6 and table 2.3 - [b] Optimized values. See text for definition

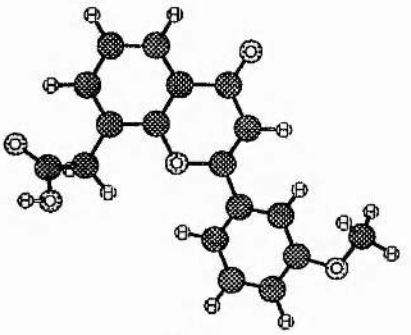
FIG. 6.6: AM1 Optimized structures of FAAs used for SAR



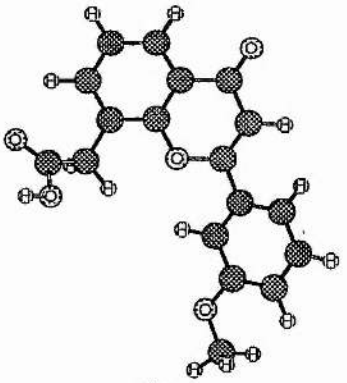
1 a



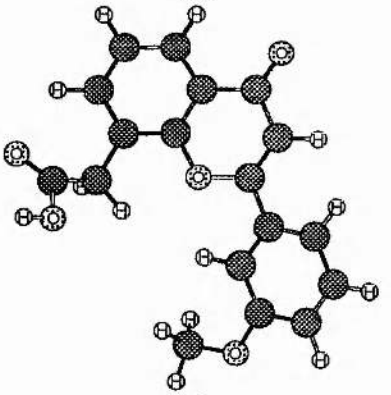
2a1



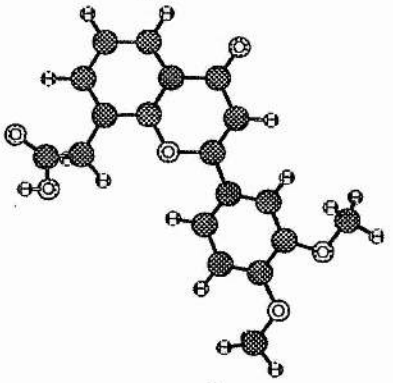
2a2



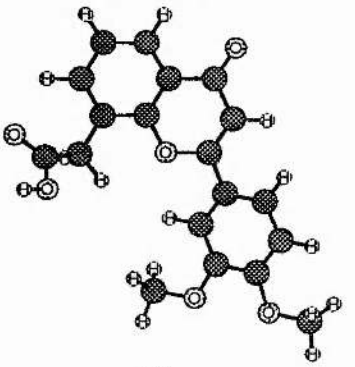
2b1



2b2



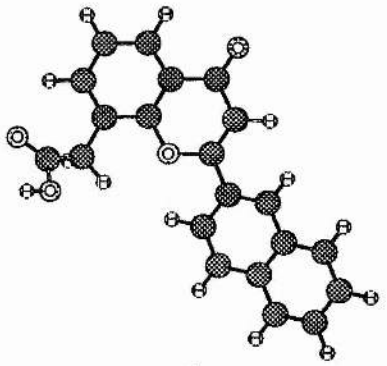
3a



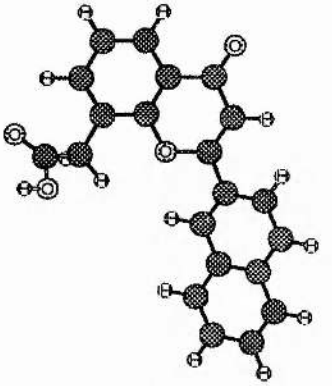
3b

Cont.....

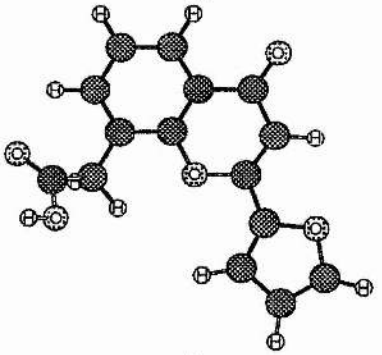
...Fig. 6.6 cont.



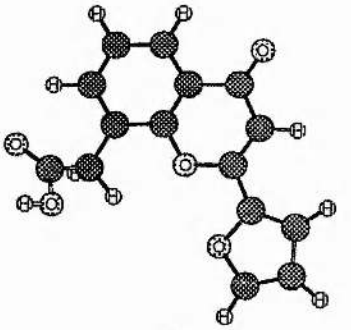
4a



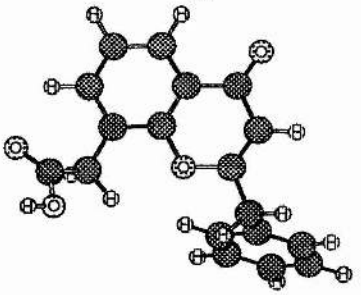
4b



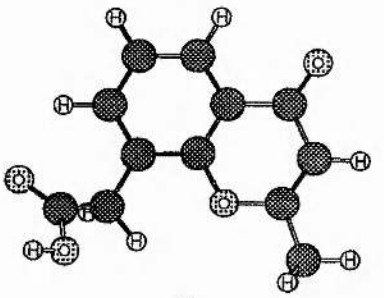
5a



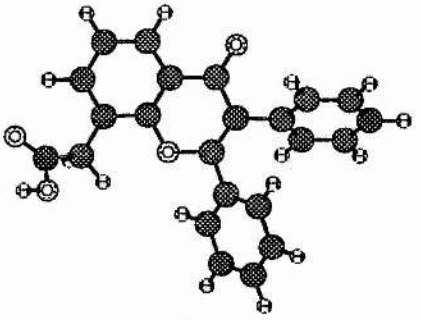
5b



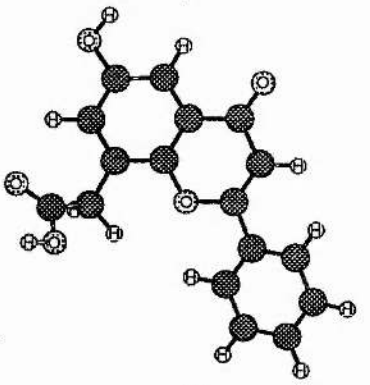
6 a



7



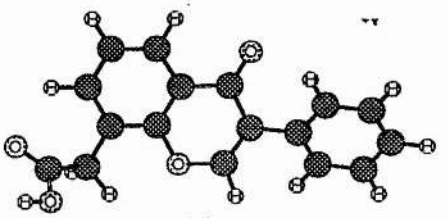
8 a



9a

Cont.....

...Fig. 6.6 cont.



10

The AM1 and PM3 optimized structures for the others molecules are available but for reason of space are not reported here. In table 6.3 the differences between selected bond lengths in the pyrone ring of the test molecules and the same bond lengths in FAA are reported.

FIG. 6.7: AM1 and PM3 optimized bond lengths and bond angles of Flavone Acetic Acid

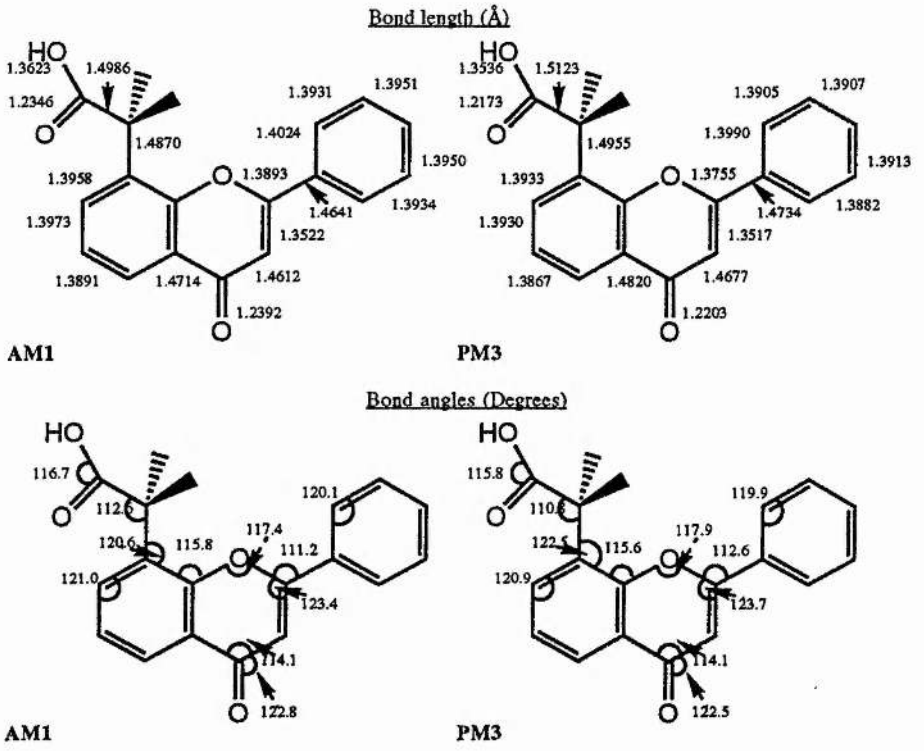


TABLE 6.3: Differences between bond lengths of test molecule and the corresponding bond length of FAA (**1a**) with AM1 and PM3

Ref	TGI%	Δ1-2 (Å)		Δ2-3 (Å)		Δ3-4 (Å)		Δ4-7 (Å)		Δ16-2 (Å)	
		AM1	PM3	AM1	PM3	AM1	PM3	AM1	PM3	AM1	PM3
7	0	-0.0031	-0.0002	-0.0033	-0.0044	0.0006	0.0024	-0.0002	-0.0001	0.0197	0.0164
9a	0	-0.0016	-0.0018	0.0007	0.0007	-0.0016	-0.0013	-0.0000	0.0003	0.0002	0.0003
10	24	-0.0165	-0.0173	0.0033	-0.0015	0.0121	0.0128	-0.0009	-0.0018	-	-
4a	41	-0.0002	-0.0004	0.0004	0.0005	-0.0001	-0.0002	0.0000	0.0001	0.0001	-0.0001
4b	41	0.0002	0.0004	-0.0002	-0.0002	0.0000	0.0001	0.0000	0.0000	0.0002	0.0003
3a	70	0.0001	-0.0001	0.0003	0.0005	-0.0006	-0.0006	0.0002	0.0002	-0.0008	-0.0009
3b	70	0.0005	0.0002	0.0005	0.0004	-0.0006	-0.0004	0.0001	0.0001	-0.0009	-0.0011
1a	96	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
6a	97	-0.0032	-0.0020	-0.0038	-0.0044	0.0010	0.0024	-0.0003	-0.0002	0.0314	0.0296
2a1	100	-0.0000	0.0002	-0.0003	-0.0002	0.0005	0.0005	-0.0001	-0.0001	0.0015	0.0016
2a2	100	-0.0005	-0.0005	-0.0002	0.0001	0.0002	0.0001	0.0000	0.0001	0.0015	0.0011
2b1	100	-0.0005	-0.0003	-0.0003	0.0000	0.0004	0.0010	-0.0000	0.0001	0.0015	0.0021
2b2	100	0.0003	0.0002	-0.0003	-0.0003	0.0003	0.0005	-0.0002	-0.0001	0.0013	0.0008
5a	100	0.0015	0.0034	-0.0003	-0.0013	0.0004	0.0026	-0.0005	-0.0003	-0.0203	-0.0166
5b	100	-0.0023	0.0002	0.0013	-0.0001	0.0003	0.0016	-0.0002	-0.0001	-0.0216	-0.0173
8a	100	-0.0009	0.0004	0.0084	0.0029	0.0138	0.0014	-0.0006	-0.0015	0.0039	0.0038

As can be seen from table 6.3, the bond lengths of the FAAs do not differ much from those of the the parent molecule FAA and this is not surprising given the similarity between all the molecules. The bond lengths in this table are reported with four significant figures, this is most unusual because it is known that the (average) experimental error of

the calculated bond lengths involving carbon is about 0.002 Å [69]. However it is interesting to note that some of these differences are positive while other are negative and some correlation with the antitumour activity can be observed. For example, for molecules more active than FAA, the difference in the bond C_3-C_4 is positive, that is in molecules more active than FAA the bond C_3-C_4 is "longer" (If we can ascribe physical significance) than that in FAA itself. This result is common to both AM1 and PM3 (table 6.3). Molecules less active than FAA on the other hand show the opposite trend except for molecules **7** and **10** which have a positive difference for this bond although they are inactive. It has to be noted however, that both molecules **7** and **10** do not have the phenyl group in position 2 and this topographical difference together with some other factors may explain their inactivity. For molecule **4** (41% active) the difference in the bond C_3-C_4 depend on the conformation: the difference is negative for conformation 'a' (Δ_{3-4} = -0.0001Å for AM1 and -0.0002Å for PM3) while it is positive for conformation 'b' (Δ_{3-4} = 0.0000Å for AM1 and 0.0001Å for PM3). This double behaviour of the naphtyl derivative **4** is common with other correlations found (see later). From now on, FAAs less active than FAA will be called 'inactive' whereas, those more active than FAA will be called 'active'.

The differences in the bond lengths between the FAAs are indeed very small and doubtless have no physical significance, however, it has to be pointed out that these calculations were conducted under the same level of accuracy and the geometry optimization process was conducted in the same way for all the molecules. Furthermore all the molecules under study are very similar therefore it is likely that all of them are affected by the same errors. The trend with the antitumour activity that has been observed may then arise from small differences between the active and inactive FAAs. In energetic terms, the results in table 6.3 suggest that the C_3-C_4 bond of active molecules is weaker than that of the inactive ones. This comes from the correlation between bond length and energy [117]. The correlation is better described by the bond orders and therefore they have been calculated and are reported in the next section.

6.1.4: Bond orders

The bond order between a pair of atoms is defined as the sum of the squares of the density matrix elements connecting the two atoms and is a measure of the covalent bond energy of that bond [88 (b)]. For example, the C-C bond orders of ethane, ethylene and benzene are roughly 1.00, 2.00 and, 1.40 reflecting the single and the double bond of ethane and ethylene, and the delocalised π electron system of benzene. In the case of single polar bonds, the bond orders also reflect the degree of polarity. In the extreme case of an ionic bond, the bond order would be zero because the coefficients of the less electronegative atom would also be zero. In view of the correlation between the bond length C_3-C_4 and the anticancer activity of the series of FAAs which have been reported in the previous section, it was interesting to calculate the bond orders which can be related better to the strength of the bonds. In order to obtain more information from the bond energies, it was important to also look at some of the atomic charges reported in section 6.2.1.

Fig. 6.8 shows the selected AM1 and PM3 bond orders of flavone acetic acid. As previously predicted by the analysis of the bond lengths, the π electron system or the pyran ring is quite localized: the O_1-C_2 and O_1-C_6 are single bonds with very little double bond character probably due to a partial transition sp3 -> sp2 of the oxygen. This is consistent with the value of the bond angle C_6-O_1-C_2 in fig. 6.7, (117.4°) far larger than a tetrahedral angle (109°). This transition allows some of the π-electron charge to be drawn from the carbon atoms toward the more electronegative oxygen; the later is in fact negative while the neighbouring carbon atoms are positive (see fig. 6.9). C_2-C_3 is a quite localized double bond, while C_4-C_5 is delocalized over the condensed benzene ring. C_3-C_4 and C_4-C_5 are single bonds with an ionic character; here the σ-electron charge is probably drawn from the carbon C_4 to carbons C_3 and C_5 which, therefore, become very negative (see fig. 6.9).

FIG. 6.8: Selected AM1 and PM3 bond orders of flavone acetic acid

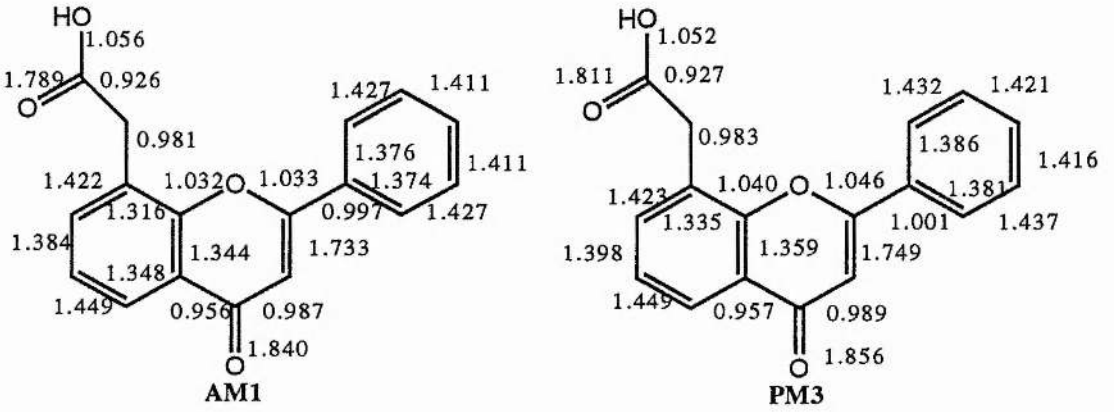


Table 6.4 reports the differences between the bond orders of the test molecules and those of flavone acetic acid. The differences in the C_3-C_4 bond order are negative for all the molecules more active than FAA with both methods AM1 and PM3. This indicates that the C_3-C_4 bond is weaker in the active molecules than in the inactive ones. These results are in agreement with those obtained with the comparison of the bond lengths in the previous section; as can be noted from tables 6.4 and 6.5 the difference in bond orders C_2-C_3 and C_4-O_7 is positive for most of the active molecules, this indicates a stronger double bond character of the bonds C_2-C_3 and C_4-O_7 and therefore more localization of the π-electron system over the pyrone ring of the molecules more active than FAA.

TABLE 6.4 : AM1 difference (bond orders of test molecule - bond order in FAA) of the pyran ring of the FAAs

Ref.	TGI%	Δ1-2	Δ2-3	Δ3-4	Δ4-7	Δ2-16
7	0	0.003	0.024	-0.001	0.002	-0.014
9a	0	0.006	-0.008	0.004	0.000	0.000
10	24	0.031	-0.003	-0.025	0.012	-
4a	41	0.001	-0.003	0.000	0.000	0.002
4b	41	-0.001	0.001	0.000	0.000	-0.001
3a	70	-0.000	-0.004	0.002	-0.002	0.003
3b	70	-0.001	-0.005	0.002	-0.001	0.004
1a	96	0.000	0.000	0.000	0.000	0.000
6a	97	0.002	0.031	-0.003	0.003	-0.038
2a1	100	-0.000	0.004	-0.001	0.002	-0.002
2a2	100	0.001	0.001	-0.001	0.000	-0.003
2b1	100	0.001	0.003	-0.001	0.000	-0.003
2b2	100	-0.001	0.003	-0.001	0.002	-0.002
5a	100	-0.004	-0.006	-0.002	0.004	0.014
5b	100	0.008	-0.020	-0.001	0.001	0.015
8a	100	0.001	-0.032	-0.026	0.011	-0.011

TABLE 6.5 : PM3 difference (bond orders of test molecule - bond order in FAA) of the pyran ring of the FAAs

Ref.	TGI%	Δ1-2	Δ2-3	Δ3-4	Δ4-7	Δ2-16
7	0	0.006	-0.007	0.004	-0.002	-0.000
9a	0	0.006	-0.007	0.004	-0.002	-0.000
10	24	-0.001	0.033	-0.003	0.003	-0.020
4a	41	0.001	-0.004	0.000	-0.000	0.002
4b	41	-0.001	0.001	0.000	0.000	-0.001
3a	70	0.000	-0.005	0.002	-0.002	0.002
3b	70	-0.001	-0.003	0.001	-0.001	0.003
1a	96	0.000	0.000	0.000	0.000	0.000
6a	97	0.003	0.036	-0.004	0.003	-0.044
2a1	100	-0.000	0.003	-0.001	0.001	-0.002
2a2	100	0.001	0.000	-0.000	-0.000	-0.002
2b1	100	0.001	0.002	-0.001	-0.000	-0.002
2b2	100	-0.001	0.003	-0.001	0.002	-0.002
5a	100	-0.006	0.001	-0.003	0.004	0.006
5b	100	0.002	-0.009	-0.002	0.002	0.008
8a	100	-0.002	0.001	-0.027	0.016	-0.026

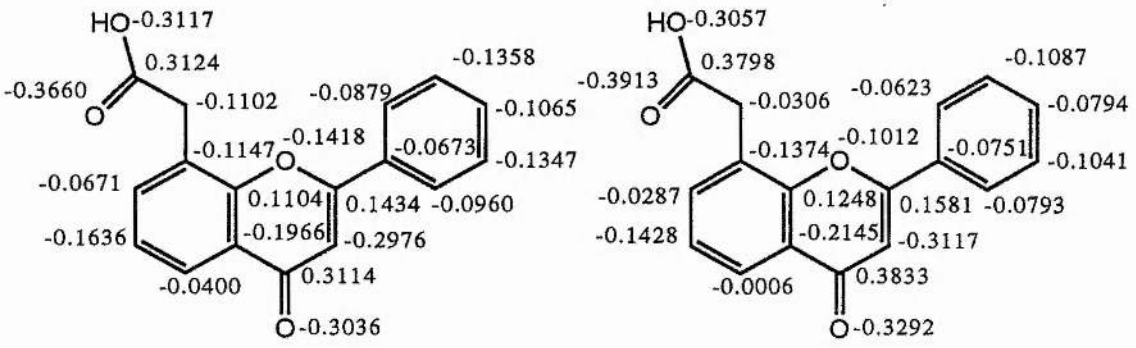
6.2 Electronic properties of FAAs

6.2.1 Charge distribution

Selected atomic charges obtained with the Mulliken population analysis with AM1 and PM3 methods are reported for FAA in fig 6.9. Table 6.6 shows some of the differences between the atomic charges of the test molecules and the charge at the same position in FAA sorted in order of increasing activity. Only the charges at positions C_3, O_7, C_{16} (or $C_{1'}$) and $C_{2'}$ have been reported because they are the most interesting in the context of looking for structure/activity correlations of the FAAs reported in table 2.3. The atomic charges at position C_{16} and $C_{2'}$ are those that change most along the series of molecules considered (to the second decimal place) but also the charge at position C_3 can vary considerably such as in molecules **10**, **5** and **8** (table 6.6). Trends of the differences in the atomic charges with the antitumour activity can be observed both with the AM1 and PM3 methods.

Within the AM1 method, correlation with the activity is observed with the charges at positions C_3 and O_7. For molecules more active than FAA, the atomic charges at C_3 and O_7 are both less negative than in FAA, i.e the difference "atomic charge in the test molecule minus atomic charge in FAA" is positive as shown in table 6.6 and fig. 6.9. Exceptions are molecule **10** and molecule **4** in the **b** conformation, conformation **4a** however shows the expected sign in the charge difference. The PM3 method on the other hand, shows a correlation between the activity and the charges at atoms C_3, C_{16} and, $C_{2'}$: active molecules have a more positive charge than FAA at C_3 and C_{16} but a more negative one at $C_{2'}$. The inactive molecules (or less active than FAA) violate at least one of these conditions.

FIG. 6.9: Selected atomic charges of FAA



AM1 **PM3**

Calculations on the anions of the FAAs were also conducted as part of this study (but the data is not reported here) and similar structure/activity correlations were obtained. These results are encouraging particularly in view of the correlations described previously with the bond lengths. It is clear that the nature of the substituents in position C_2 has an influence on the antitumour activity of these molecules and the correlations found so far point to an important involvement of the pyrone ring in this activity.

TABLE 6.6: Differences between selected atomic charges of test molecules and the atomic charge at the same position in FAA (**1a**) with AM1 and PM3

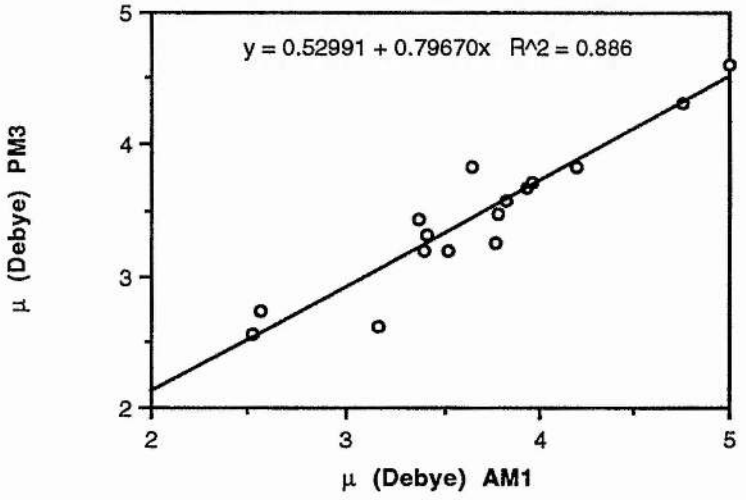
Ref	%TGI	C_3-C_{3FAA}		O_7-O_{7FAA}		C_{16}-C_{16FAA}		[a]$C_{2'}$-$C_{2'FAA}$	
		AM1	PM3	AM1	PM3	AM1	PM3	AM1	PM3
7	0	-0.007	0.007	0.000	0.001	-0.121	-0.005	-	-
9a	0	-0.004	-0.005	-0.000	-0.002	-0.001	-0.001	0.000	-0.000
10	24	0.105	0.087	0.006	0.012	-	-	-	-
4a	41	-0.001	-0.002	-0.000	-0.000	0.004	0.003	-0.010	-0.011
4b	41	0.001	0.001	0.000	0.000	0.001	-0.001	0.022	0.024
3a	70	-0.005	-0.006	-0.002	-0.002	-0.004	-0.001	0.007	0.010
3b	70	-0.004	-0.004	-0.001	-0.001	-0.002	-0.001	-0.054	-0.057
1a	96	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6a	97	0.002	0.010	0.001	0.001	-0.022	0.059	0.005	-0.030
2a1	100	0.005	0.006	0.002	0.001	0.033	0.035	-0.038	-0.040
2a2	100	0.001	0.001	0.000	-0.001	0.035	0.040	-0.045	-0.037
2b1	100	0.004	0.004	0.000	-0.000	0.033	0.035	-0.028	-0.038
2b2	100	0.004	0.005	0.002	0.002	0.036	0.040	-0.079	-0.090
5a	100	0.022	0.022	0.006	0.005	0.022	0.019	-0.036	-0.039
5b	100	0.007	0.013	0.003	0.002	0.023	0.023	-	-
8a	100	0.107	0.093	0.008	0.013	-0.000	0.002	-0.004	-0.000

[a] Atomic charge at C_{17} or C_{21}

6.2.2 Dipole moments

Table 6.7 shows the magnitude (μ) and the direction of the calculated dipole moment of the FAAs with both AM1 and PM3. In the table, the dipole moments have been calculated as a vector with the oxygen atom O1 as origin. α and β give the direction of the dipole moment as D-O_1-C_2 and D-O_1-C_2-C_{16} respectively, where D is the point of the arrow which describes the vector. In general AM1 values of μ are slightly bigger than those calculated with PM3, however the differences are not very big and the graph of AM1 versus PM3 dipole moments gives a straight line with a slope 0.8 (Fig. 6.10).

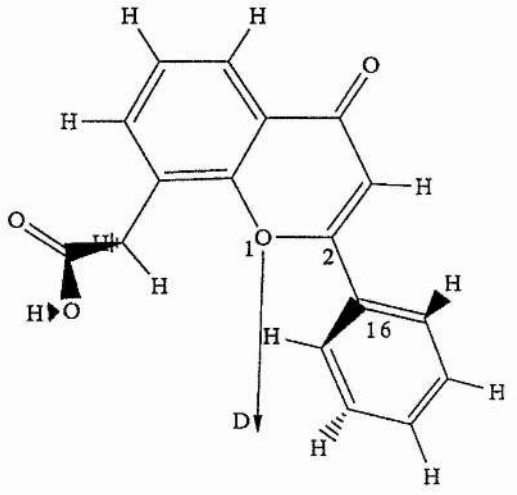
FIG 6.10: AM1 v. PM3 dipole moments of FAAs



The main difference between AM1 and PM3 dipole moments of the FAAs is in the direction of the vector. As indicated in table 6.7, the values of β predicted by PM3 are bigger (in absolute value) than those of AM1; this means that PM3 vectors are further away from the pyrone (pseudo) plane than the AM1 vectors. This is probably due to the fact that the acetic acid group is more twisted in the PM3 than in the AM1 structures.

No correlations between the dipole moments and the anticancer activity of the FAAs have been observed.

TABLE 6.7: Magnitude (Debye) and direction of AM1 and PM3 dipole moments of FAAs



Ref	Activity	μ_{AM1}	μ_{PM3}	α_{AM1}	α_{PM3}	β_{AM1}	β_{PM3}
7	0	3.395	3.205	94.95	98.14	-7.91	-20.93
9a	0	3.770	3.247	74.40	84.93	-6.57	-23.91
10	24	2.519	2.572	95.99	100.39	-8.93	-22.76
4a	41	3.938	3.676	87.14	92.09	-5.68	-19.20
4b	41	4.193	3.818	92.95	99.10	-8.36	-17.92
3a	70	3.648	3.815	101.41	109.16	-10.78	-30.06
3b	70	3.162	2.630	88.70	94.73	-4.47	-12.63
1a	96	3.826	3.571	91.87	97.13	-7.58	-19.74
6a	97	3.413	3.311	94.29	97.76	-13.39	-25.98
2a1	100	4.994	4.607	90.54	95.49	-7.95	-17.98
2a2	100	2.565	2.746	87.88	95.94	-1.95	-34.26
2b1	100	4.750	4.317	81.54	83.58	-2.99	-19.37
2b2	100	3.370	3.435	109.21	119.28	-18.63	-28.78
5a	100	3.957	3.714	97.61	101.18	-6.62	-20.14
5b	100	3.784	3.474	85.94	89.52	-6.23	-16.96
8a	100	3.524	3.193	86.94	90.04	-7.91	-16.34

6.2.3 Frontier orbitals

In order to study the ability of FAAs to lose or accept electrons, the HOMO and LUMO eigenvectors have been analysed and reported in this section.

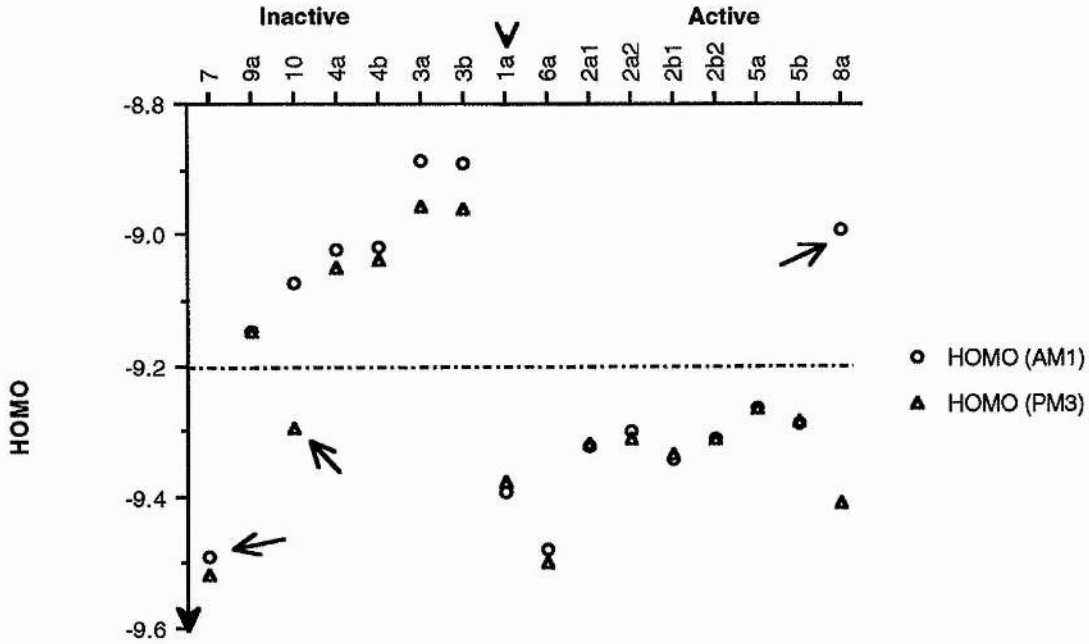
Table 6.8 shows the energies of the HOMO and LUMO of the FAAs calculated with AM1 and PM3.

TABLE 6.8: AM1 and PM3 energies (in eV) of HOMO and LUMO of FAAs

Ref.	Activity	HOMO (AM1)	LUMO (AM1)	HOMO (PM3)	LUMO (PM3)
7	0	-9.49	-2.05	-9.52	-2.12
9a	0	-9.15	-2.36	-9.15	-2.42
10	24	-9.07	-2.13	-9.30	-2.16
4a	41	-9.02	-2.45	-9.05	-2.53
4b	41	-9.02	-2.44	-9.04	-2.52
3a	70	-8.88	-2.28	-8.95	-2.37
3b	70	-8.89	-2.29	-8.96	-2.37
1a	96	-9.39	-2.30	-9.38	-2.39
6a	97	-9.48	-2.05	-9.50	-2.11
2a1	100	-9.32	-2.30	-9.32	-2.38
2a2	100	-9.30	-2.33	-9.31	-2.42
2b1	100	-9.34	-2.30	-9.33	-1.39
2b2	100	-9.31	-2.34	-9.31	-2.41
5a	100	-9.26	-2.41	-9.26	-2.40
5b	100	-9.29	-2.43	-9.28	-2.43
8a	100	-8.99	-2.23	-9.41	-2.16

The data in table 6.8 is reported in order of increasing activity and some correlation between the energy of the HOMO (with both AM1 and PM3) and the activity can be observed. This may be easier to visualize in the graph of fig. 6.11 which contains the energies of the HOMO (in eV) on the Y-axes and the reference number of the FAAs (in order of increasing activity) in the abscissa. For all the active molecules the energy of the HOMO is lower than -9.3 a.u. with the PM3 method; AM1 gives a similar result apart from molecule **8** whose HOMO has an energy of -8.99 eV. For most of the inactive molecules, the energy of the HOMO is higher by at least 0.2 eV, with the exceptions of molecule **7** (with both methods), and molecule **10** (only with PM3).

FIG. 6.11: AM1 and PM3 energies of highest occupied molecular orbital (eV) of FAAs.



According to Koopmans' theorem, the energy of the HOMO of a Hartree-Fock wavefunction is a good approximation of the negative value of the ionization potential [36a]. The results of table 6.8 and fig. 6.11 suggest therefore that the inactive molecules

they have a lower ionization potential. In order to study in more detail if there was any other substantial differences between the HOMO of active and inactive molecules, the AM1 coefficients of the HOMO have been calculated.

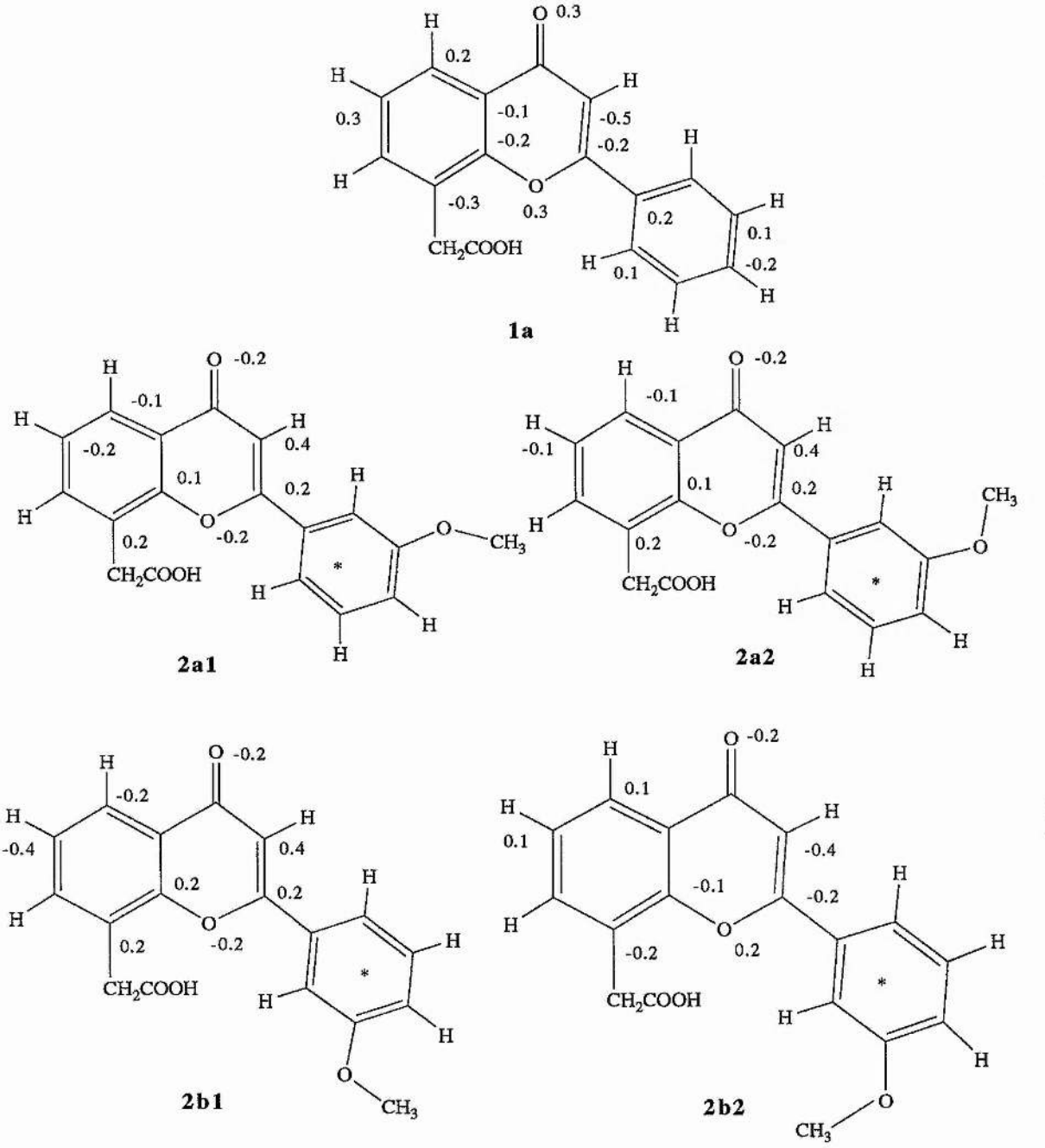
As can be seen in figs. 6.12 and 6.13, the HOMO of all the FAAs is a π-type orbital in which the p_z orbitals of the constituents atoms are mainly involved with only small contributions of p_x and p_y orbitals in the out-of-plane phenyl (or naphtyl) rings of molecules **2**, **8**, **3** and, **4**. The HOMOs of the active molecules present some common features:

1) The p_z orbitals in C_2, C_3, C_6 and C_8 are all always involved and they have the same sign. (In molecule **6** and FAA the p_z orbital of C_5 is also involved and it has the same sign).

2) The p_z orbitals in O_1, O_7, C_{10} and C_{11} are all always involved and they have the same sign opposite to that of C_2 such that there is a nodal plane in the bonds between O_1-C_2 and O_1-C_6.

A pictorial representation of the condition 1) and 2) is given in fig. 6.14.

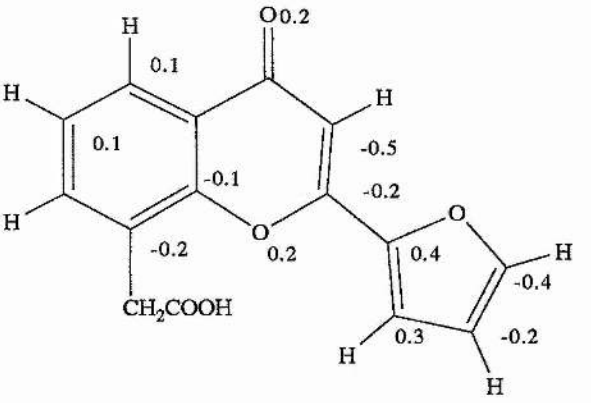
FIG. 6.12: Coefficients of the HOMO (>0.1) of active FAAs



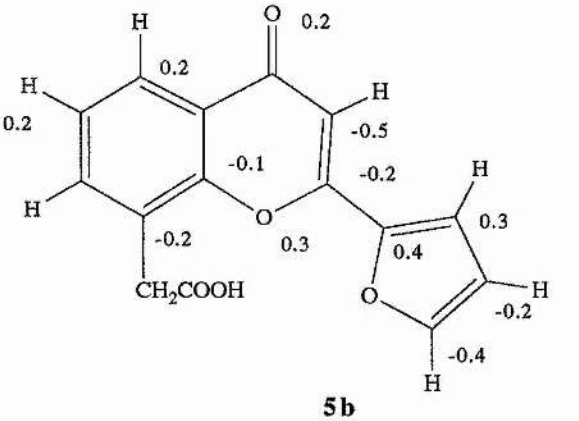
* Pz and Px delocalized orbitals over the ring.
** Pz and Py delocalized orbitals over the ring.

Cont....

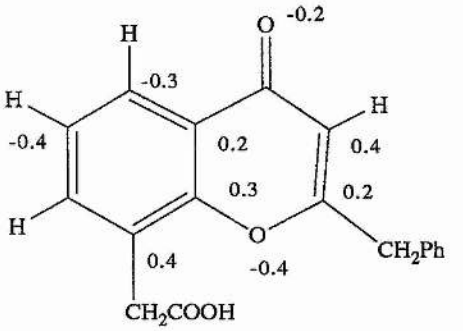
...Fig. 6.12 cont.



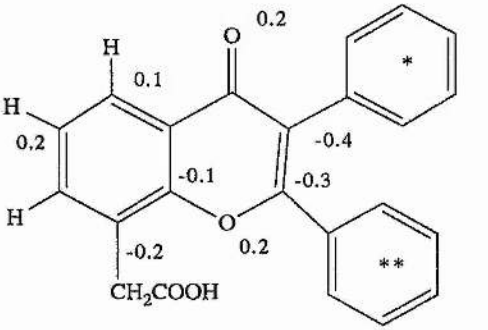
5a



5b



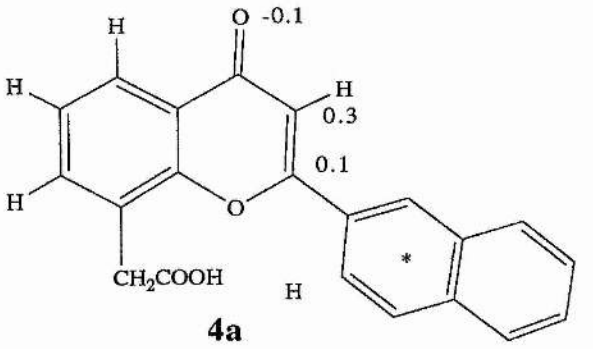
6a



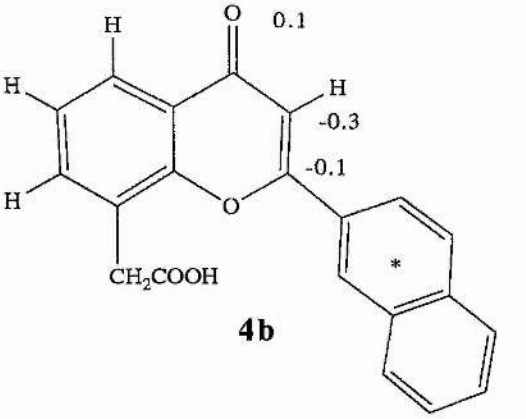
8a

* Pz and Px delocalized orbitals over the ring.
** Pz and Py delocalized orbitals over the ring.

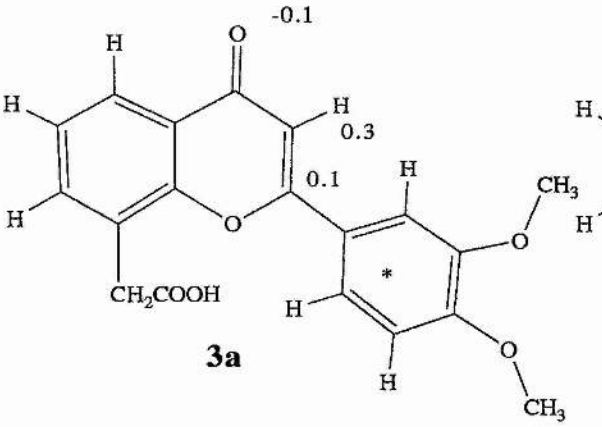
FIG. 6.13: Coefficients of the HOMO (>0.1) of inactive FAAs.



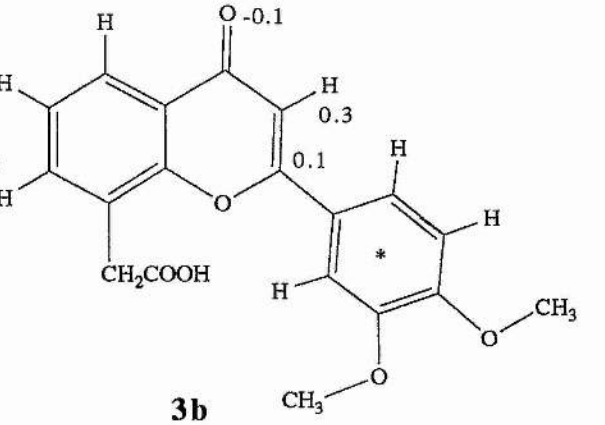
4a



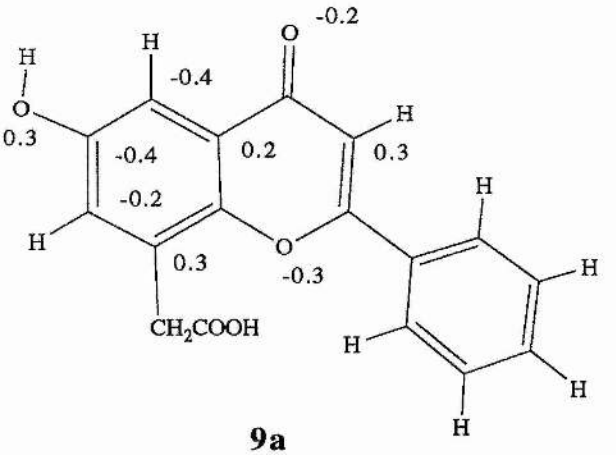
4b



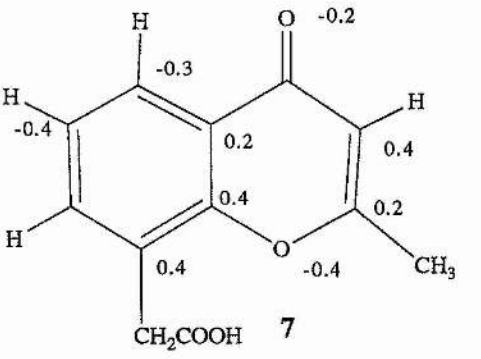
3a



3b

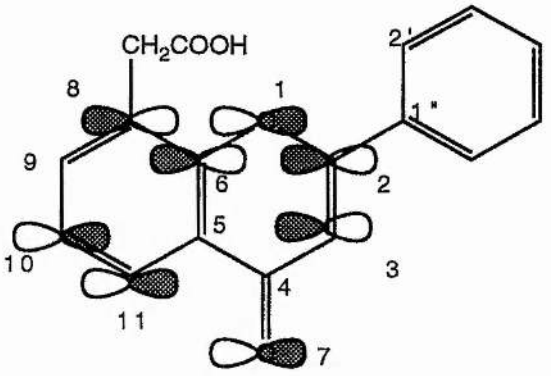


9a



7

FIG. 6.14: Pictorial representation of the orbitals involved in the HOMO of molecules more active than FAA.



As can be seen in fig.6.13, the HOMO of the inactive molecules are not similar to the HOMO of active molecules; the only exception being molecule **7** whose HOMO is very similar to that of the active molecule **6a**.

To summarize, inactive molecules have a lower ionization potential than active molecules and also, the orbitals involved in the formation of their HOMO are not the same as those used in active molecules; the latter show a precise topology common to all the active molecules. These results suggest that one of the reasons why molecules **3, 4, 9, 7** are inactive may be because, given their low ionization potential, they are oxidised before they even reach the target. Alternatively, it may be that part of the antitumor mode of action involves the donation of an electron to a specific region of a receptor and only the HOMO of the active molecules present the right topology to achieve this.

6.2.4 Electrostatic potential

The electrostatic potential gives a global picture of the forces acting on the molecules being a function of both the electron density and the geometry of the molecule. The correlations between the activity of the FAAs and the atomic charges, bond lengths and HOMO, discussed in the previous sections, point to existing differences in the

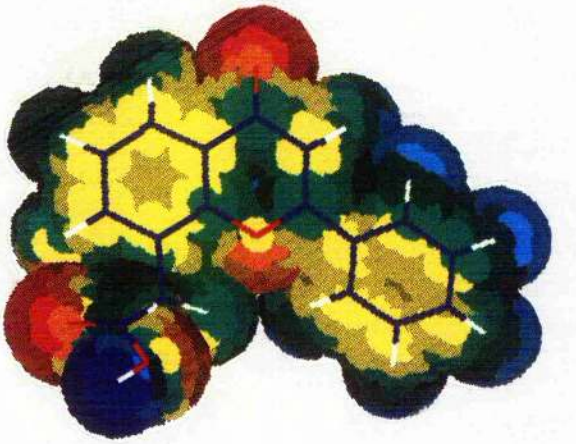
electron distribution of the series of active and inactive molecules under study, that shoud be highlighted by the study of the MEP. The MEPs of the FAAs have been calculated and many attempts to look for structure/activity correlations have been made.

Unfortunately, most of the information contained within the electrostatic potential map (MEP) can be lost if a suitable way of displaying it is not available. The first problem that arose in the study of the MEP of the FAAs was connected with the difficulty in dealing with molecules with many degrees of rotational freedom. Many different conformations may be achieved in a very short amount of time because of the rotations about the flexible bonds, but it is very difficult to compare the MEPs of different molecules in the different conformations and it is necessary to choose one or two structures for each molecule to use in the S/A studies. Initially, the structure of the molecules to be used for the SA correlations was choosen on the basis of the lowest heat of formation [27]. The results obtained in this way were not satisfactory because for some of the FAAs the acetic acid group was found to be on one side of the chromone molecular plane, whereas for others it was on the opposite side. The position of the acetic acid group makes a big difference to the aspect of the MEP, but the differences in energy between the acetic acid group being on one side rather than the other, were often very small [123]. The main reason that lead to the development of the strategy for the geometry optimization described in section 6.1 was to obtain comparable MEPs of the different molecules.

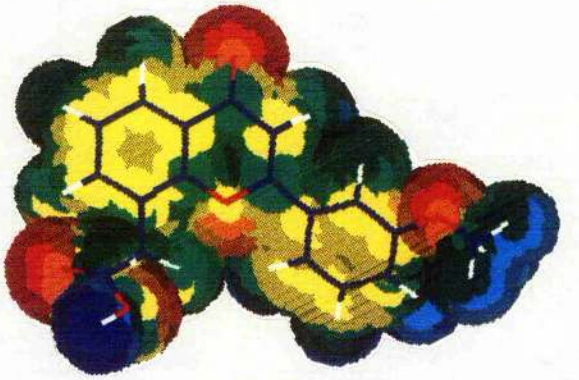
Fig. 6.15 shows the MEPs of the active FAAs calculated from the AM1 wavefunction, the MEPs of the inactive molecules are in fig.6.16. No substantial differences in the MEPs of active and inactive molecules can be observed. This may be due to an unsuitable methods of displaying the MEPs or to the fact that the activity of FAAs is not related to their MEPs. The 3D2 program used to display the MEPs only allows seven contour levels to be used. If correlation with the activity exists in the MEPs of FAAs they are probably very small and limited to a small region of the molecular space. The calculation of the MEPs on the plane of the atoms $C_3C_4O_7$ may highlight these differences.

The MEPs have also been compared with the Carbo and Hodgkin similarity indexes (SIs) in which FAA was used as the lead molecule. The Carbo and Hodgkin methods gave very similar values of SI as it is evident from the graph in fig. 6.17 which show the straight line obtained by plotting the Carbo SI versus the Hodgkin SI of the FAA. In Fig. 6.18 is reported a plot of the activity of the FAAs and the Carbo SI: no correlations between SI and activity can be observed. With the ASP program used to calculate the SI, it is also possible to optimize the SI in order to have the best fit of the MEPs; this can be done by moving the test molecule with respect to the lead molecule either by a rigid translational movement or by also rotating the molecule about the flexible bonds. The SI optimized, both with rigid and flexible bonds, versus the activity of the FAAs are shown in figs. 6.19 and 6.20. No correlation with the antitumour activity has been observed neither with AM1 or PM3 charges.

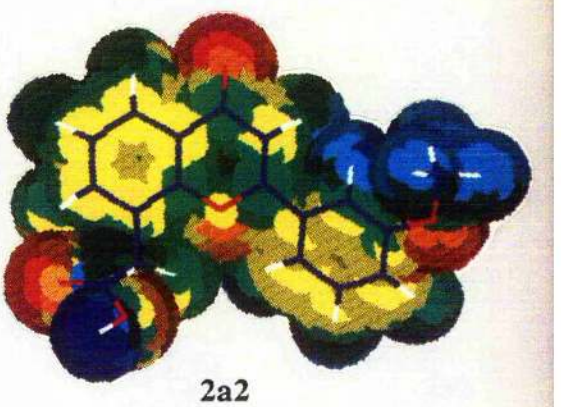
FIG. 6.15: AM1 MEPs of active FAAs calculated directly from the wavefunction



1a



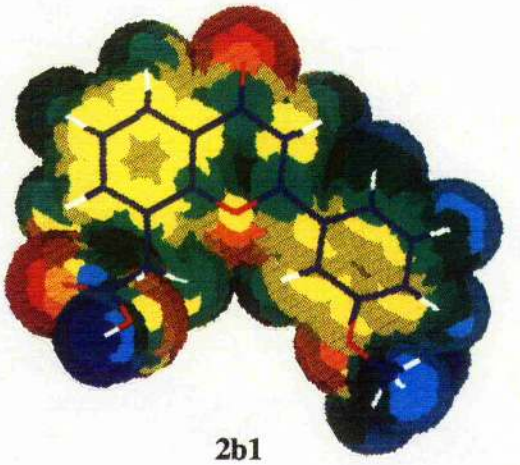
2a1



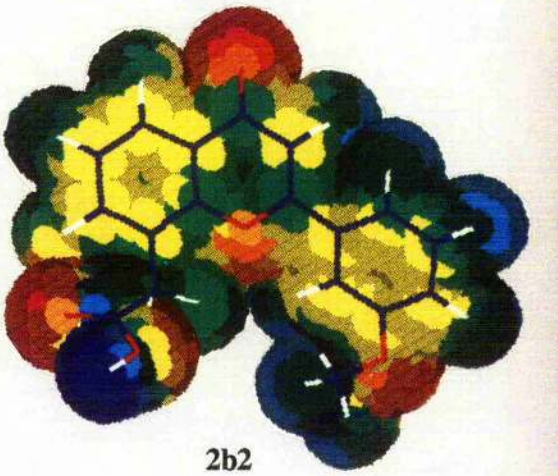
2a2

Cont.....

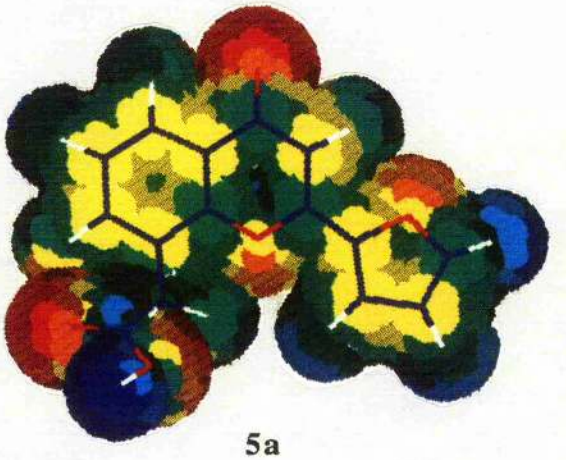
...Fig 6.15 cont.



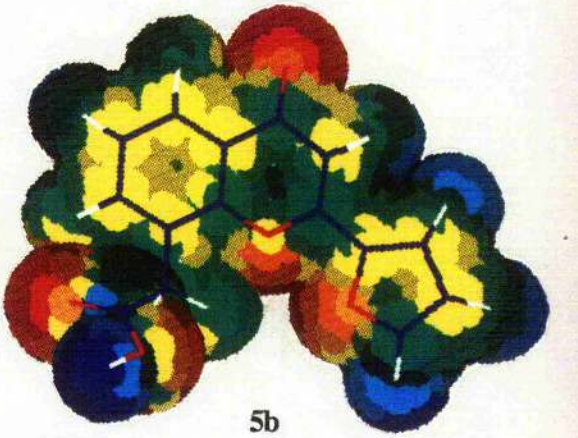
2b1



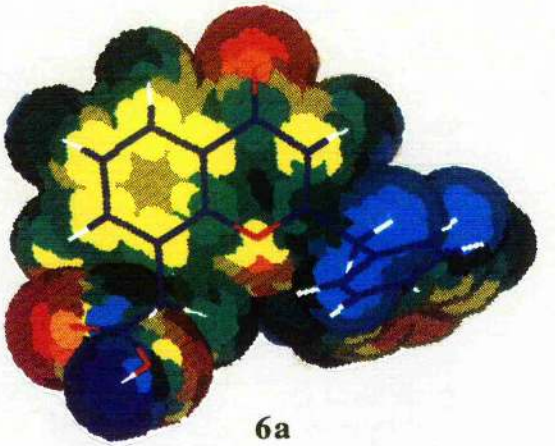
2b2



5a



5b

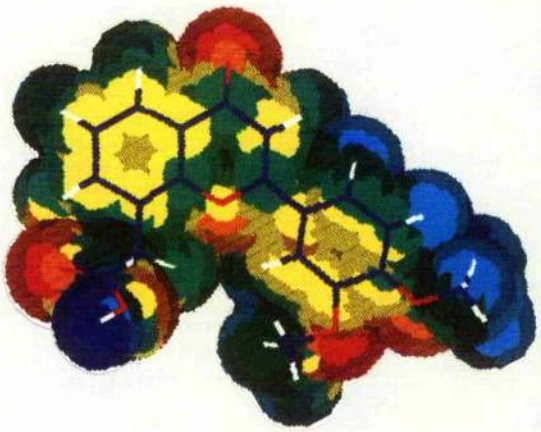


6a

8a

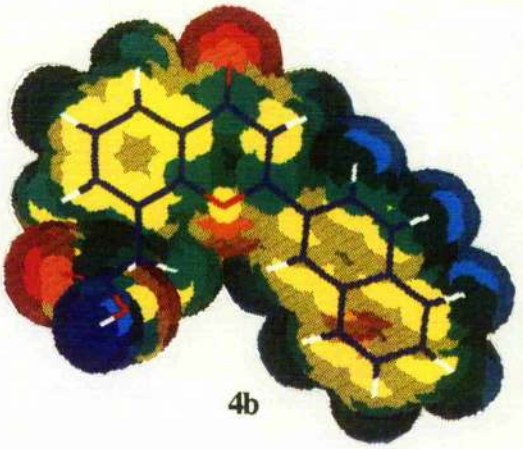
FIG. 6.16: AM1 MEPs of inactive FAAs calculated directly from the wavefunction

3a



3b

4a



4b

Cont.....

...Fig 6.16 cont.

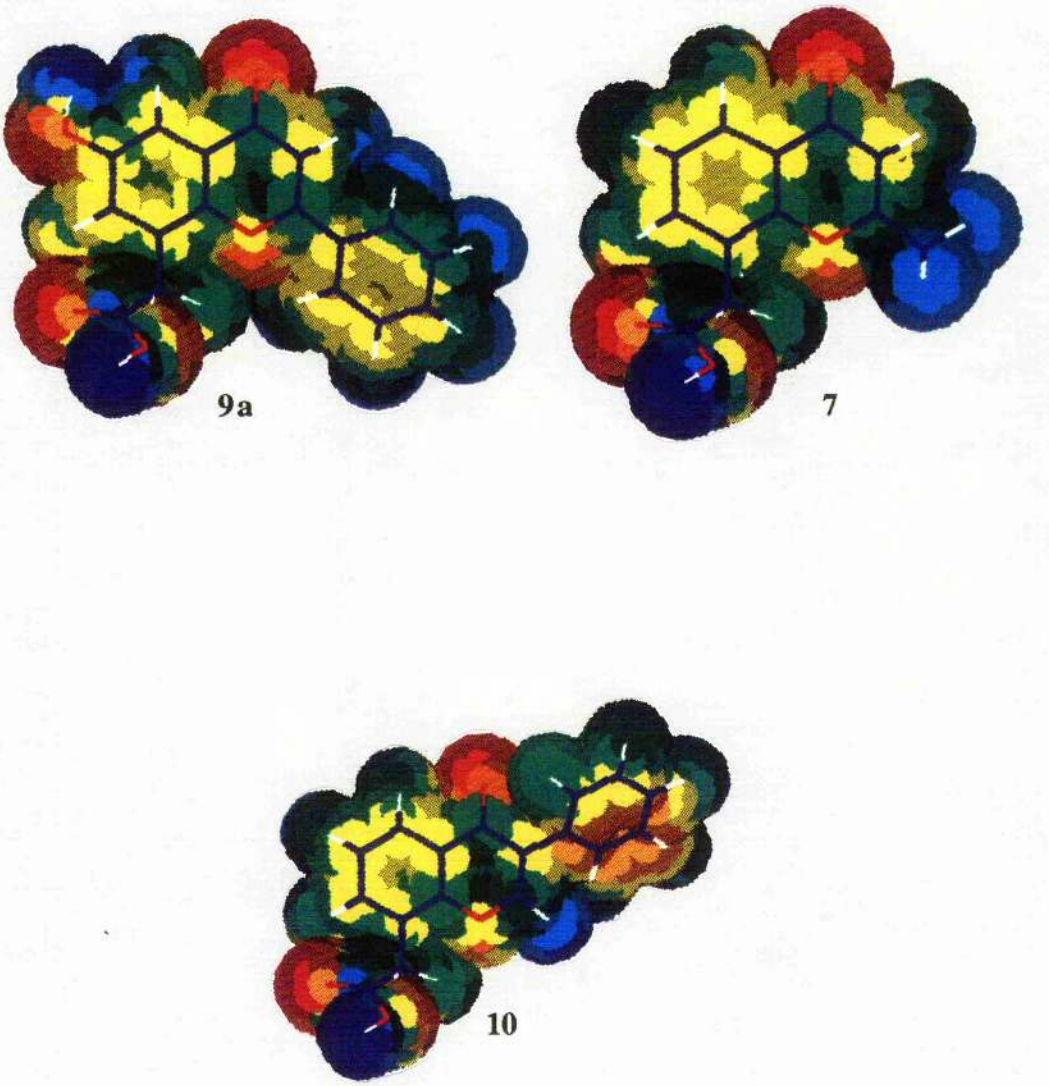


FIG. 6.17: Carbo versus Hodgkin Similarity indexes of the FAA.

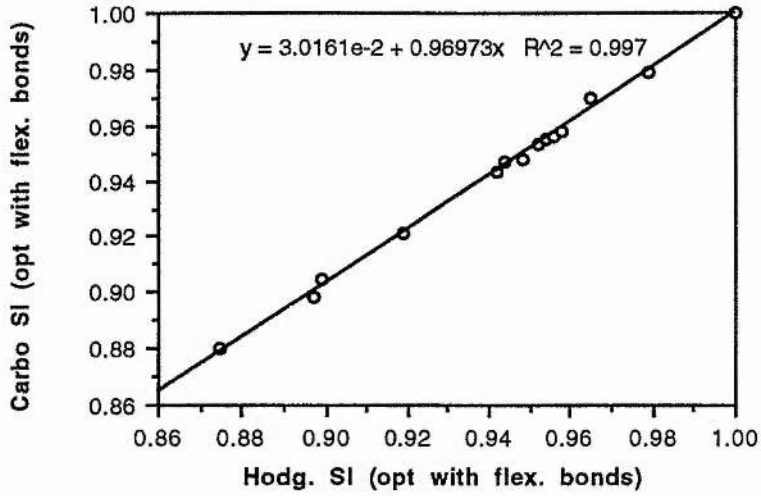


FIG. 6.18: Carbo similarity index of the AM1 MEPs of FAAs.

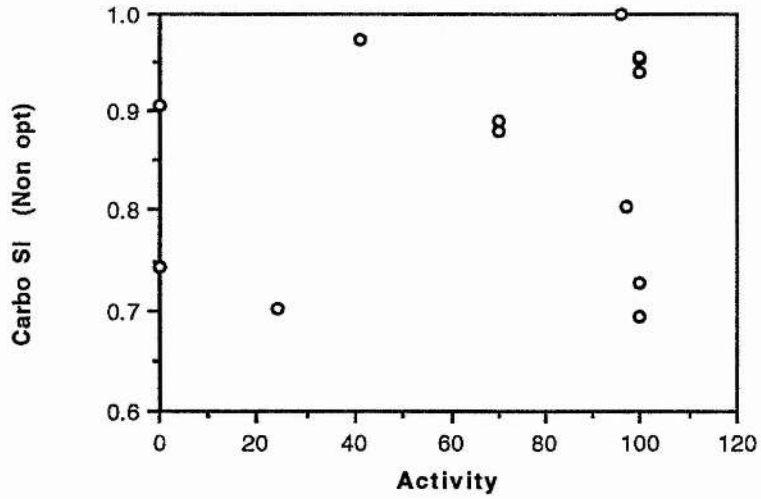


FIG. 6.19: Carbo similarity index optimized with rigid bonds of the AM1 MEPs of FAAs.

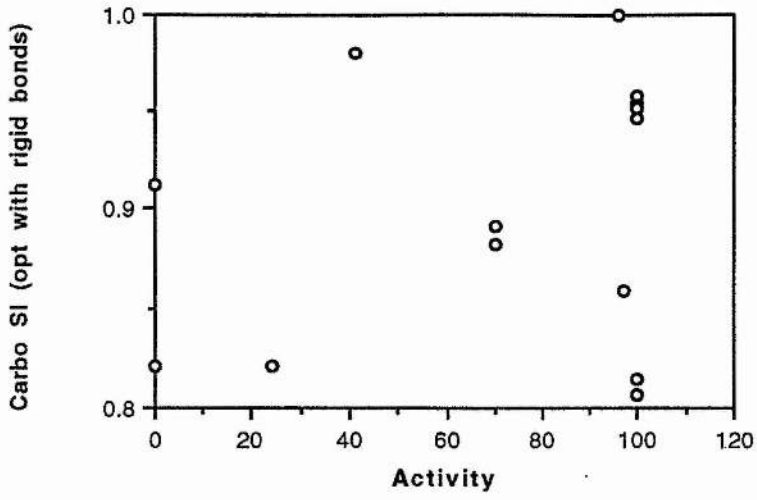
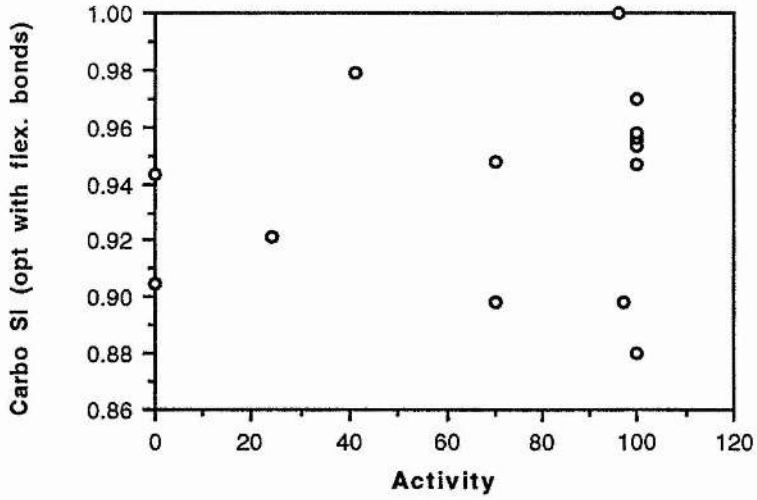


FIG. 6.20: Carbo similarity index optimized with flexible bonds of the AM1 MEPs of FAAs.



6.3 Summary and discussion

AM1 and PM3 semi empirical calculations on ten derivatives of Flavone acetic acid, whose activity against colon 38 solid tumours was known, have been performed

and correlation between the antitumour activity and calculated properties have been obtained.

With the AM1 method the following conditions are common to all the active molecules:

Condition 1) Atomic charge at C_3 minus the same charge in FAA ≥ 0

Condition 2) Atomic charge at O_7 minus the same charge in FAA ≥ 0

Condition 3) Bond length C_3-C_4 minus the same bond length in FAA ≥ 0

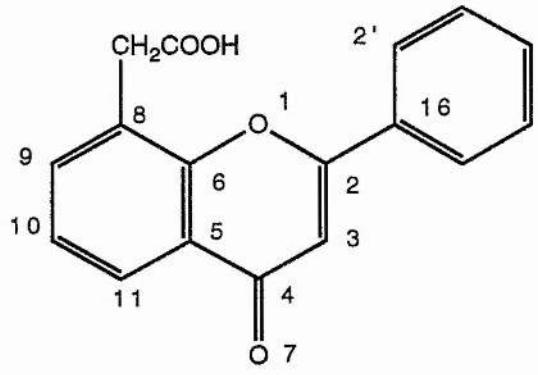
Condition 4) Bond length C_4-O_7 minus the same bond length in FAA ≤ 0

Condition 5) Bond order C3-C4 minus the same bond order in FAA ≤ 0

Condition 6) In the formation of the HOMO, the p_z orbitals in C_2, C_3, C_6 and C_8 are all always involved and they have the same sign. (the p_z orbital of C_5 may also be involved and it has the same sign).

Condition 7) In the formation of the HOMO, the p_z orbitals in O_1, O_7, C_{10} and C_{11} are (fig. 6.21) all always involved and they have the same sign opposite to that of C_2 such that there is a nodal plane in the bonds between O_1-C_2 and O_1-C_6 (See fig. 6.13).

FIG. 6.21: Structure and numbering scheme of FAA



Furthermore, the energy of the HOMO for active molecules is smaller than -9.3 eV but there is an exception (molecule **8**) and therefore it has not be included as one of the conditions for activity. Conditions 1-4 can be obtained from a single calculation and they are sufficient to explain the activity of all the FAAs studied so far; these are used in chapter 8 to predicte the activity of others FAA analogues and xanthones.

The PM3 set of conditions that all the active molecule respect are as follow:

Condition 1) Atomic charge at C_3 minus the same charge in FAA ≥ 0

Condition 2) Atomic charge at C_{16} minus the same charge in FAA ≥ 0

Condition 3) Atomic charge at $C_{2'}$ minus the same charge in FAA ≤ 0

Condition 4) Bond length C_3-C_4 minus the same bond length in FAA ≥ 0

Condition 5) Energy of the HOMO $\leq$ -9.3 eV.

The orbitals involved in the formation of the HOMO have not been calculated with the PM3 method; this study is suggested for future work.

The correlation found suggests that the pyrone ring may be directly involved in the antitumour effect of flavone acetic acid, because here is where all the differences between active and inactive molecules have been observed. The AM1 and PM3 results are consistent with the π-electron system of the pyrone ring being more localized in the active molecules than in the inactive ones; as a consequence, the bond C_3-C_4 becomes weaker and may break more easily than in the inactive molecules. It may also be that part of the antitumour mode of action involves the donation of an electron to a specific region of a receptor and only the HOMO of the active molecules presents the right topology to achieve this. Another possibility is that molecules **3, 4, 9, & 7** are inactive because, given their low ionization potential, they are oxidised before they reach the target.

These correlations obtained with AM1 and PM3 are able to distinguish between the active and inactive series of FAAs used in the S/A study and there is a good

probability that the FAAs which fail to respond to the conditions of activity do not exert an antitumour action similar to FAA. These S/A correlations have been tested on the other 38 molecules analogous to FAA whose activity was also known and the results are reported in chapter 8.

CHAPTER 7:

Ab-initio Results

Introduction

Ab-initio calculation with 3-21G and 6-31G basis sets have been performed on the AM1 optimized structures of the FAAs reported in chapter 6; Table 7.1 shows the energy obtained (in atomic unit). Some of the data is missing because at times the memory or disk-space were insufficient to successfully complete the ab-initio calculation. As expected, 6-31G gives lower energies than 3-21G; the differences in energy are about 4 au for molecule **7**, 5 a.u. for molecules **5**, **10**, **1** and, **6** and finally about 6 a.u. for the largest molecules **2**, **4** and, **3**.

TABLE 7.1: 3-21G and 6-31G energies of FAAs

Ref[a]	TGI%[b]	Energy 3-21G[c]	Energy 6-31G[c]
1a	96	-944.915	-949.829
2a1	100	-1058.162	-1063.650
2a2	100	-1058.162	-1063.650
2b1	100	-1058.162	-1063.650
2b2	100	-1058.162	-1063.649
3a	70	-1171.404	-1177.464
3b	70	-1171.404	-1177.464
4a	41	-1096.710	-1102.426
4b	41	-1096.710	d
5a	100	-942.850	-947.733
5b	100	-942.843	-947.728
6a	97	-983.730	-988.842
7	0	-755.470	-759.383
8a	100	-1173.176	d
9a	0	d	-1024.647
10	24	-944.909	-949.820

a) See table 2.3 for reference. b) Tumour Growth inhibition % (See chapter 2) c) Energies in a. u.. d) Data not available.

7.1 Frontier orbitals

The energies of the HOMO and LUMO of the FAAs calculated with 3-21G and 6-31G basis sets are reported in table 7.2. It is interesting to note that the energy of the HOMO obtained with 3-21G is very similar to that obtained with the bigger 6-31G basis set. The maximum difference is 0.013 a.u. for molecule **9**, however all the other values differ by no more than ± 0.003 a.u..

TABLE 7.2: 3-21G and 6-31G Energies of highest occupied and lowest unoccupied molecular orbitals (in a.u.).

Ref[a]	TGI%[b]	3-21G HOMO[c]	6-31G HOMO[c]	3-21G LUMO[c]	6-31G LUMO[c]
7	0	-0.335	-0.337	0.076	0.068
9a	0	-0.312	-0.325	0.070	0.049
10	24	d	-0.312	d	0.062
4a	41	-0.302	-0.299	0.050	0.047
4b	41	-0.303	d	0.051	d
3a	70	-0.304	-0.307	0.061	0.055
3b	70	-0.303	-0.306	0.061	0.054
1a	96	-0.329	-0.329	0.059	0.053
6a	97	-0.333	-0.334	0.076	0.068
2a1	100	-0.326	-0.327	0.059	0.053
2a2	100	-0.323	-0.323	0.057	0.051
2b1	100	-0.325	-0.326	0.059	0.054
2b2	100	-0.321	-0.322	0.056	0.050
5a	100	-0.317	-0.318	0.052	0.045
5b	100	-0.317	-0.318	0.051	0.044
8a	100	-0.305	d	0.064	d

a) See table 2.3 for reference. b) Tumour Growth inhibition % (See chapter 2) c) Energies in a. u.. d) Data not available.

The energies of the LUMOs obtained with the two basis sets are also very similar apart from molecule **9**, for which the value obtained with the 3-21G basis set is 0.021 a.u higher than that with 6-31G. The correlation between the values obtained with 3-21G and 6-31G basis sets can be better visualized in Figs. 7.1 and 7.2. The reason why molecule **9** behaves differently from the other FAAs is not clear at this time.

FIG. 7.1: 3-21G versus 6-31G energies (in a.u.) of the HOMO of the FAAs.

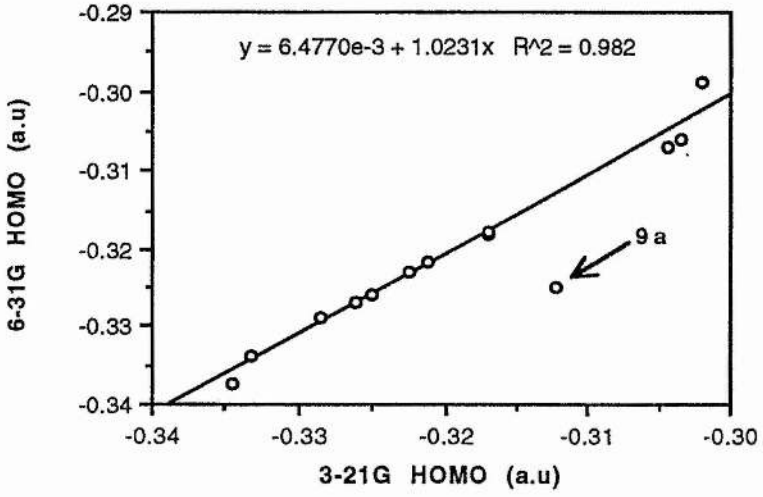
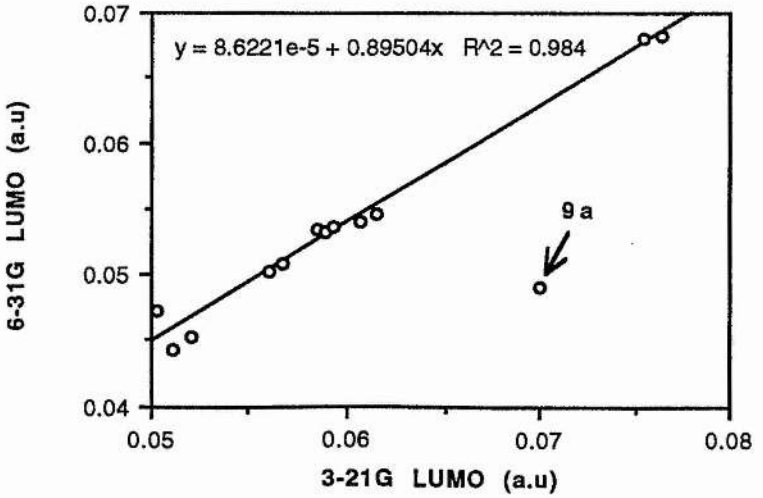


FIG. 7.2: 3-21G versus 6-31G energies (in a.u.) of the LUMO of the FAAs



The *ab-initio* energies of the HOMO of the FAAs broadly correlate with the antitumour activity in a similar fashion to the correlation obtained with the semi empirical methods AM1 and PM3 described in the previous chapter. Figs. 7.3 and 7.4, show plots of the energy of the HOMO (in a.u) (Y-axes) and the reference numbers of the FAA (X-axes), sorted in order of increasing activity. It is interesting to compare these plots with Fig. 6.11 obtained with the semi empirical methods: AM1 and 3-21G give very similar plots, while the PM3 plot is very similar to that obtained with 6-31G.

FIG 7.3: 3-21G energies of HOMO

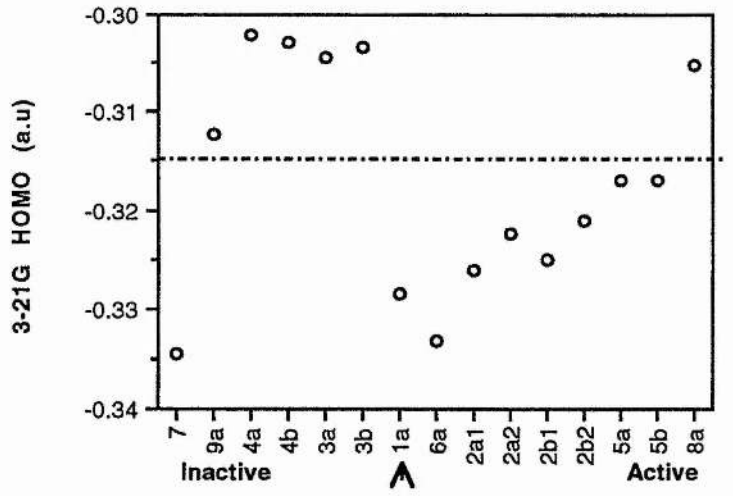
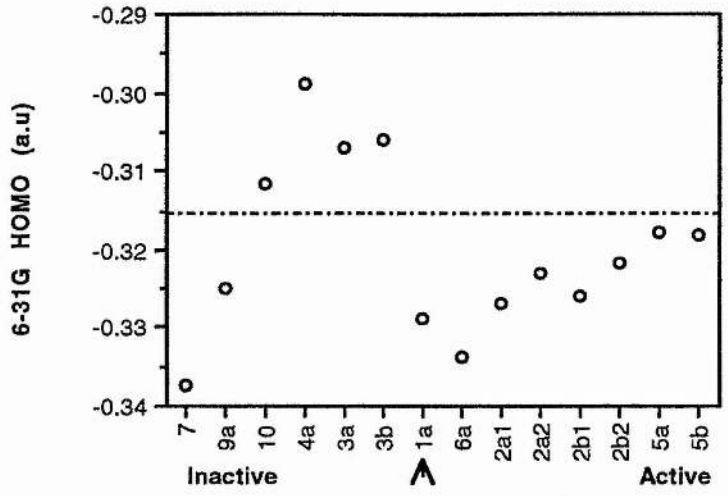


FIG 7.4: 6-31G energies of HOMO



It should be noted however that the 6-31G value for molecule **8** is missing (due to the lack of sufficient memory in the FPS-500) therefore the similarity of results between PM3 and 6-31G is not very sure.

To summarise, although the 3-21G basis set is smaller than the 6-31G, the energies of the HOMO obtained with it, are very similar to those obtained with the bigger basis set. These energies broadly correlated with the antitumour activity of the FAAs in a similar way to the semi empirical methods AM1 and PM3, therefore the considerations reported in section 6.2.3 are also valid for these basis sets.

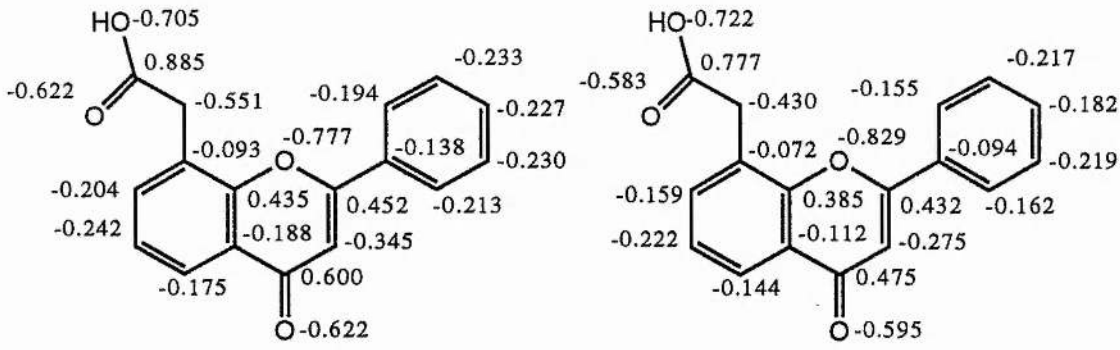
7.2 Charge distribution and MEP

The Mulliken atomic charges of FAA obtained with 3-21G and 6-31G basis sets are shown in fig.7.5. The charges obtained with the two basis sets are very similar, the maximum difference being 0.1 a.u.for the charge at C_4. The 3-21G basis set overestimates the negative and positive charges i.e. it gives more negative charges for atoms with a negative charges and more positive charges for positive atoms, than 6-31G.

This behaviour can be seen in the graph of fig. 7.6 were the atoms in the x-axes are sorted in order of increasing (6-31G) atomic charge. The only exception is given by the charge at O_1 for which 6-31G gives a more negative value (-0.83 a.u.) than 3-21G (-0.78 a.u.).

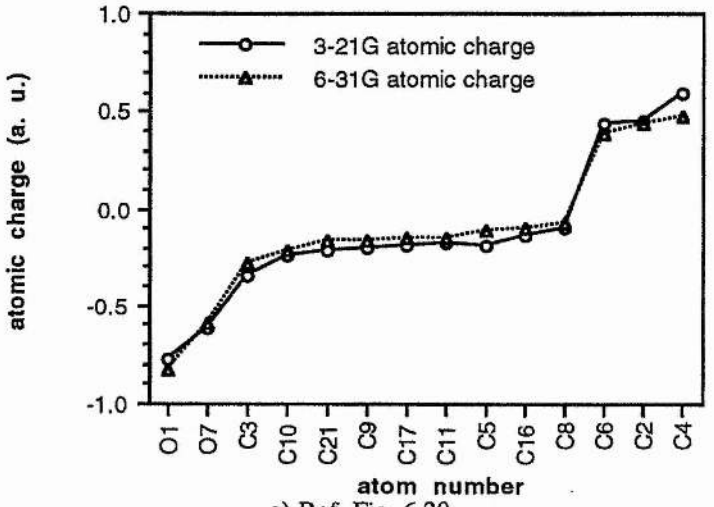
It is clear from fig.7.6 (see also figs. 7.1 and 7.2) that although 3-21G is a basis set of modest size, it gives calculated molecular properties which are in reasonable agreement with those calculated with the 6-31G basis set. This similarity can be extended to MEPs calculated with 3-21G Mulliken charges which are expected to contain the same information as those obtained with 6-31G charges in a quantitative way. As it will be shown later (chapter 9), the MEP calculated directly from the wavefunction with 3-21G and 6-31G basis sets are also very similar.

FIG. 7.5: Selected 3-21G and 6-31G atomic charges of FAA



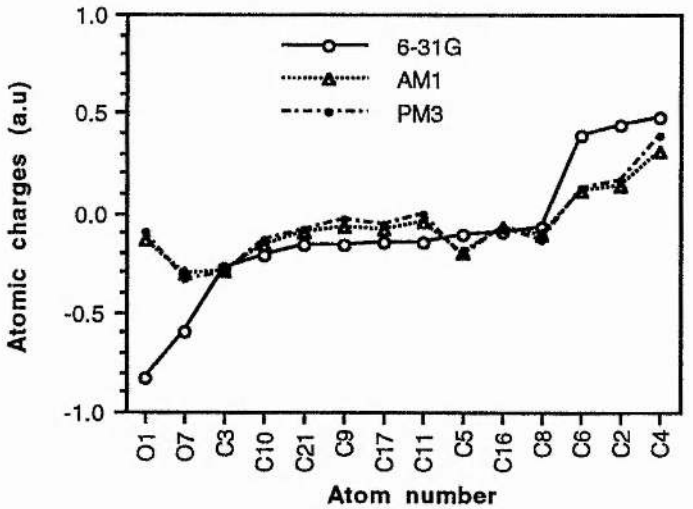
3-21G **6-31G**

FIG. 7.6: Atom number [a] versus 3-21G and 6-31G atomic charges of FAA



a) Ref. Fig. 6.20

FIG. 7.7: Atom number [a] versus 6-31G, AM1 and PM3 atomic charges of FAA

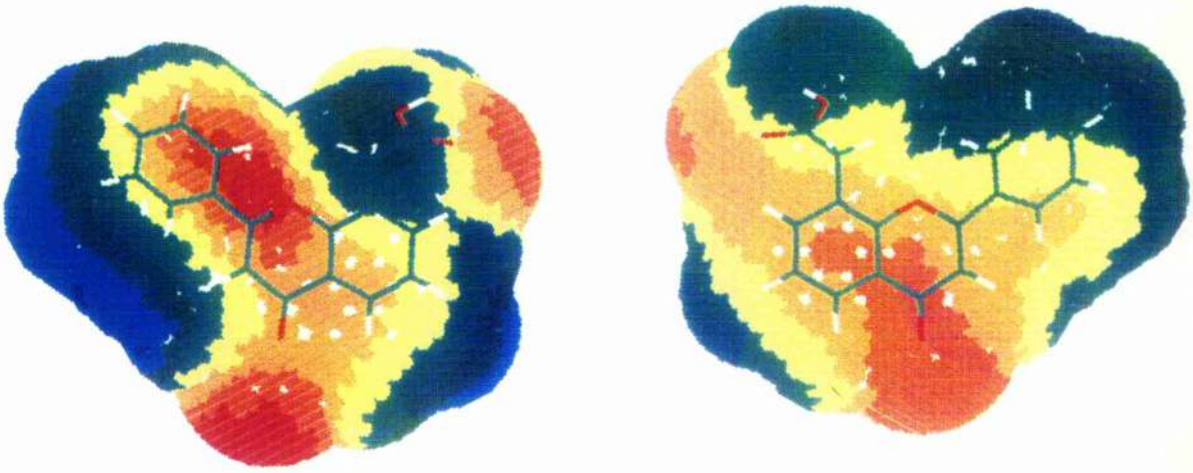


a) Ref. Fig. 6.20

The similarity between the charges obtained with the semi empirical methods AM1 and PM3 (see fig. 6.9) and the ab-initio charges is, on the other hand, not very good as can be seen from the plot in fig. 7.7. The main differences between the two appear in

atoms which are either very negative or very positive. What is important, is that these differences are not constant and therefore the interpretation of the charge distribution leads to different conclusions depending which method is considered. For example, according to the semi empirical methods, the most negative atoms in the pyrone ring are O_7 and C_3 while 3-21G and 6-31G basis sets assign the most negative charge to O_1. Intuitively, an electrophilic reaction would be expected to occur at the the most negative atom; Looking at the semi empirical results, the attack by electrophiles would occur preferentially at O_7 or at C_3 while the ab-initio results suggest that it occurs at O_1. This later result is very unlikely from a chemical point of view. As discussed by Tomasi [77], the site of an electrophilic attack is better represented by the minima of the electrostatic potential rather than by the atomic charges. However if the MEPs are calculated from these charges the same result has to be expected as it is clear from Fig. 7.8 which shows the MEP of FAA calculated with the point charge approximation method with (a) 3-21G and (b) AM1 atomic charges. According to the 3-21G MEP the attack by electrophiles occurs at O_1 or even C_2 because they are surround by a (red) negative region. The AM1 MEP on the other hand, matches more closely the expectation that an electrophilic reaction would occur at O_7. The match between atomic charges (and MEP derived from it) and the chemical intuition observed here, are closer with semi empirical than *ab-initio* methods.

FIG. 7.8: MEP of FAA calculated with a) 3-21G and b) AM1 Mulliken charges. Regions of negative potential are coloured in red.



(a) (b)

The differences between selected atomic charges and the charge at the same position in FAA obtained with 3-21G and 6-31G basis sets respectively, have been calculated and are reported in Tables 7.3 and 7.4. No correlation with the antitumour activity and 3-21G atomic charges can be observed in table 7.3, while the results of table 7.4 are more encouraging. As can be seen from this table, the 6-31G charges at C_5, C_6, and O_7 are ≥ 0, ≤ 0 and, ≥ 0 respectively, only for molecules more active than FAA. For inactive molecules, at least one of these conditions is not respected. The correlation with charges found here is similar to those obtained with the semi empirical methods, however the atoms involved are different (see chapter 6).

To summarize, the Mulliken charge distribution obtained with 3-21G is very similar to that obtained with 6-31G basis sets. The set of charges obtained with such basis sets should however not be used to construct MEPs (using the point charge approximation) because they can give false information about possible sites of electrophilic reactions. 3-21G charges do not seem to correlate with the antitumour

activity of the FAAs while with the 6-31G basis set the following conditions have to be respected in order to have active molecules:

1) Atomic charges at C_5 of test molecule minus charge of FAA at the same position ≥ 0

2) Atomic charges at C_6 of test molecule minus charge of FAA at the same position ≤ 0

3) Atomic charges at O_7 of test molecule minus charge of FAA at the same position ≥ 0

Failure to respect one of these condition leads to inactivity.

TABLE 7.3: 3-21G differences between selected atomic charges of test molecules and the charge at the same position of FAA

Ref[a]	%TGI[b]	ΔO_{1FAA}	ΔC_{2FAA}	ΔC_{3FAA}	ΔC_{4FAA}	ΔC_{5FAA}	ΔC_{6FAA}	ΔO_{7FAA}	ΔC_{8FAA}
9a	0	0.023	-0.215	0.105	0.034	-0.002	0.002	0.003	0.099
7	0	0.011	-0.048	-0.020	0.001	-0.007	0.001	0.001	-0.475
4a	41	-0.002	0.004	0.002	-0.000	0.001	-0.000	-0.000	-0.005
4b	41	0.001	0.000	-0.002	0.000	-0.000	-0.000	0.000	-0.002
3a	70	0.000	0.005	-0.006	0.001	-0.000	-0.000	-0.002	-0.010
3b	70	-0.004	0.006	-0.001	-0.000	0.000	0.000	-0.002	-0.014
1a	96	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6a	97	0.006	-0.011	0.005	-0.001	-0.003	-0.000	0.001	-0.366
2a1	100	0.001	-0.003	0.003	-0.000	0.000	0.000	0.002	0.015
2a2	100	0.003	-0.001	-0.001	0.000	-0.000	-0.001	0.000	0.018
2b1	100	0.003	-0.002	0.002	-0.000	0.000	-0.000	0.001	0.015
2b2	100	0.002	-0.005	0.003	-0.000	0.000	-0.002	0.002	0.017
5a	100	-0.002	-0.108	0.036	-0.003	0.004	-0.002	0.003	0.470
5b	100	0.013	-0.088	0.022	0.001	0.003	-0.004	-0.001	0.464
8a	100	0.003	0.010	0.136	0.042	0.001	0.005	0.006	0.018

a) See table 2.3 for reference. b) Tumour Growth inhibition % (See chapter 2)

TABLE 7.4: 6-31G differences between selected atomic charges of test molecules and the charge at the same position of FAA

Ref[a]	%TGI[b]	ΔO_{1FAA}	ΔC_{2FAA}	ΔC_{3FAA}	ΔC_{4FAA}	ΔC_{5FAA}	ΔC_{6FAA}	ΔO_{7FAA}	ΔC_{8FAA}
9a	0	0.001	0.003	-0.002	-0.002	0.018	-0.016	-0.001	-0.001
7	0	0.020	0.011	-0.025	0.009	-0.003	-0.007	0.001	-0.403
10	24	0.043	-0.207	0.116	0.035	-0.002	-0.010	0.010	-
4a	41	-0.003	0.001	0.002	-0.001	0.001	0.001	-0.001	-0.011
3a	70	0.000	-0.003	-0.004	0.001	-0.000	-0.000	-0.002	0.000
3b	70	-0.006	0.003	-0.001	-0.000	0.000	0.002	-0.001	-0.008
1a	96	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6a	97	0.008	0.026	-0.004	0.006	0.000	-0.005	0.000	-0.330
2a1	100	0.001	-0.010	0.005	-0.001	0.000	0.000	0.002	0.018
2a2	100	0.004	-0.010	0.000	0.000	0.000	-0.001	0.001	0.020
2b1	100	0.004	-0.012	0.002	-0.000	0.000	-0.001	0.001	0.018
2b2	100	0.002	-0.010	0.002	-0.000	0.001	-0.002	0.003	0.017
5a	100	-0.004	0.012	0.014	0.004	0.001	-0.002	0.004	0.327
5b	100	0.024	0.004	0.027	-0.005	0.004	-0.008	0.000	0.301

a) See table 2.3 for reference. b) Tumour Growth inhibition % (See chapter 2)

7.3 Dipole moments

Table 7.5 reports the dipole moment of the FAAs obtained with 3-21G and 6-31G basis sets. Again, as can be seen from fig. 7.9, the dipole moments calculated with 3-31G correlate closely with those obtained with 6-31G. AM1 dipole moments of FAAs (table 6.7) are lower than the 6-31G values, however the plot of 6-31G versus AM1 dipole moments gives a straight line (Fig. 7.10). No correlation with the antitumour activity could be observed.

TABLE 7.5: Magnitude (Debye) of 3-21G and 6-31G dipole moments of FAAs

Ref[a]	Activity[b]	μ3-21G	μ6-31G
9a	0	c	4.738
7	0	4.002	4.190
10	24	2.984	3.044
4a	41	4.609	4.924
4b	41	4.846	c
3a	70	4.222	4.364
3b	70	3.414	3.299
1a	96	4.443	4.730
6a	97	3.940	4.214
2a1	100	5.880	6.237
2a2	100	2.895	3.042
2b1	100	5.466	5.682
2b2	100	3.867	3.966
5a	100	5.236	5.669
5b	100	4.653	4.952
8a	100	4.045	c

a) See table 2.3 for reference. b) Tumour Growth inhibition % (See chapter 2). c) Data not available

FIG. 7.9: 3-21G versus 6-31G dipole moments of FAAs (Debye).

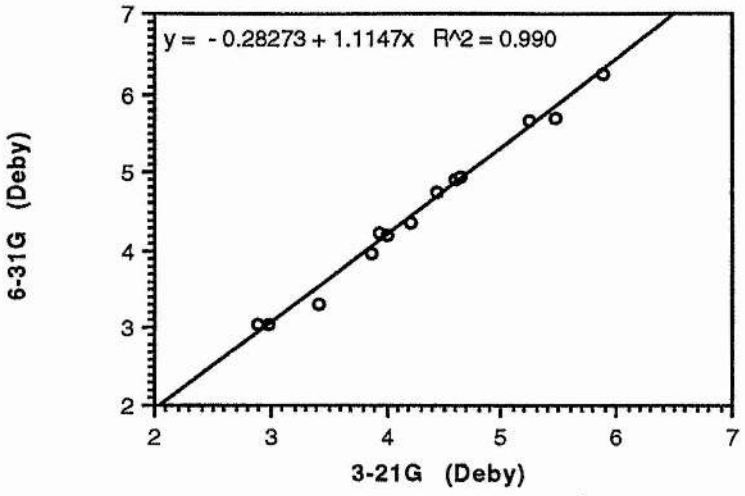
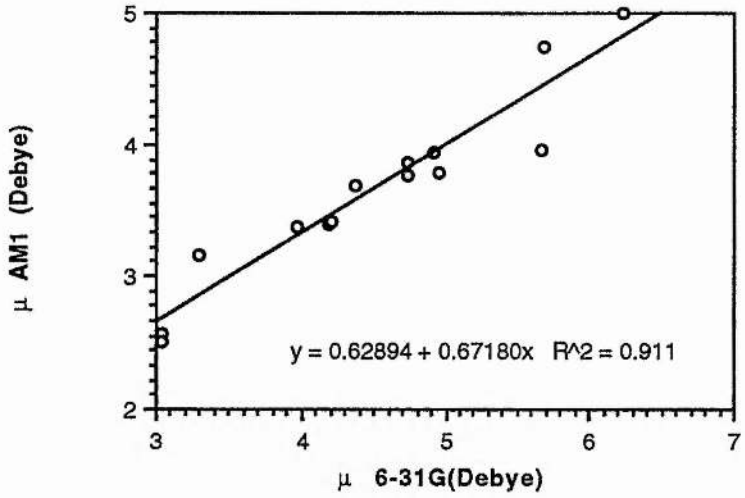


FIG. 7.10: 6-31G versus AM1 dipole moments of FAAs (Debye).



In conclusion, the 3-21G basis set predicts molecular properties such as energies of the HOMO, dipole moments and atomic charges with the same accuracy as 6-31G. It can be used instead of the 6-31G basis set on large molecules and the extra computational cost can therefore be avoided. The Mulliken charge distribution with these split-valence

basis sets cannot be associated with reactivity toward electrophiles, as well as the MEP derived from them. Correlation with activity has been observed between the energies of the HOMO and atomic charges at specific atoms in the molecules. Similar correlations have been found with AM1 and PM3 and therefore the extra computational time necessary for the *ab-initio* calculations can be avoided. Instead, the ab-initio electrostatic potential maps calculated directly from the wavefunction may give further insight towards the reactivity and it is suggested as future work.

CHAPTER 8:

Predictive Value of the Semi Empirical Structure/Activity Correlations

Introduction

Some factors related to the antitumour activity of the FAAs have been identified and reported in the previous chapters. In order to establish if the correlations found have a general validity for other related molecules, semi empirical AM1 and PM3 calculation have been performed on some of the analogues of FAA and XAAs. The activity data of these, reported in chapter 2, has been compared to the calculated properties, to see if the S/A correlations found for FAAs also hold for these other molecules. A positive answer consolidates the structure/activity relationships found, adds generality to them and gives confidence in their use to predict the activity of new FAA analogues.

The antitumour activity for most of the compounds reported in this chapter was measured by Atwell et all. in % of hemorrhagic necrosis of the tumour caused by the drug [29-33]. The activity of the FAAs used for the SAR studies in the previous chapter were on the other hand, measure by Atassi et al. in % of tumour growth inhibition. In using the Atwell data we are assuming that a linear correlation exists between the activity data measured by Atassi and that measured by Atwell. We are also assuming that the key features of the drug action of FAA and XAAs are the same. Experimental evidences support this hypothesis [34].

Section 8.1 contains the AM1 and PM3 results of the FAA analogous reported in fig.2 of Chapter 2; it also contains the results for the 5-methyl-FAA, a FAA analogue synthesized in St-Andrews by Dr. Aitken et al. [138] and found to be inactive by the group of Bibby et al. in Bradford [126]. The AM1 results from the XAAs reported in table 4 and table 5 of Chapter 2 are in section 8.2.1 while the results from some of the those XAAs obtained with PM3 are in section 8.2.2. Section 8.3 contains a summary and discussion of the results while in section 8.4 suggestions for future work are reported.

8.1 Analogues of FAA

The molecules studied are reported in fig. 8.1. The geometry optimization was conducted using the same method used for the FAAs described in section 6.3. The heats of formation obtained are shown in table 8.1.

FIG. 8.1: Analogues of FAA

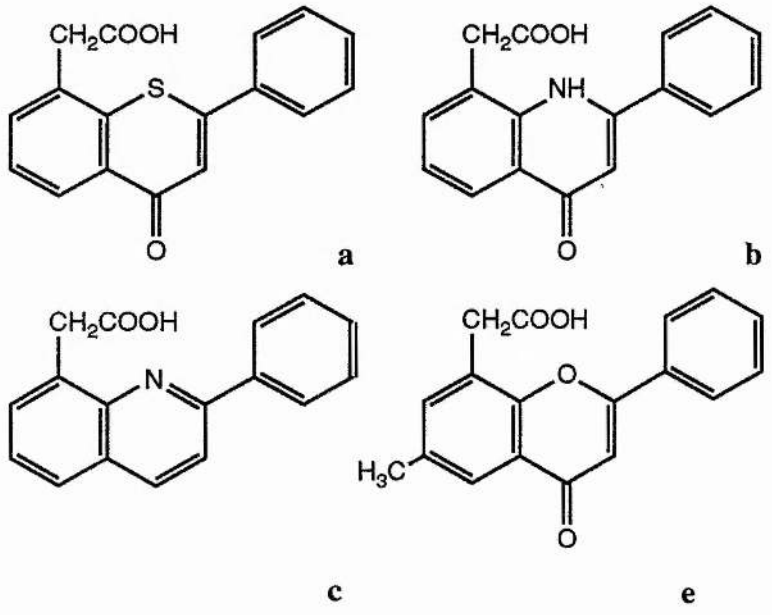


TABLE 8.1: AM1 and PM3 Heat of formation of the optimized structures of the analogues of FAA reported in Fig. 8.1

Ref.[a]	ΔHf_{AM1} (Kcal/mole)	ΔHf_{PM3} (Kcal/mole)
a	-57.8033[b]	-49.7819
b	-62.4854	-66.7370
c	-16.2488	-21.3033
e	-96.8570	-104.5714

[a] See fig. 8.1. [b] The AM1 parameters for the Sulphur atom were taken from reference [127]

The version 5 of the MOPAC program used to perform these calculations did not contain the AM1 parameters for the sulphur atom, these have been taken from ref. [127]. It is interesting to note that for molecule **a**, which contains a sulphur atom, the heat of formation obtained with PM3 is higher than that obtained with AM1; this is in contrast to the general trend which has been observed, i.e. PM3 giving lower heats of formation than AM1 (see for instance other molecules in table 8.1 and tables 6.2, 6.3 and, 6.4).

According to the AM1 structure/activity correlation discussed in chapter 6, active molecules obey the following criteria:

Condition 1) Atomic charge at C_3 minus the same charge in FAA ≥ 0 (C_3-C_{3FAA})

Condition 2) Atomic charge at O_7 minus the same charge in FAA ≥ 0 (O_7-O_{7FAA})

Condition 3) Bond length C_3-C_4 minus the same bond length in FAA ≥ 0 (Δ3-4)

Condition 4) Bond length C_4-O_7 minus the same bond length in FAA ≤ 0 (Δ4-7)

While according to the PM3 method the criteria for activity obtained with the series of FAAs in chapter 6 are as follows:

Condition 1) Atomic charge at C_3 minus the same charge in FAA ≥ 0 (C_3-C_{3FAA})

Condition 2) Atomic charge at C_{16} minus the same charge in FAA ≥ 0 (C_{16}-C_{16FAA})

Condition 3) Atomic charge at $C_{2'}$ minus the same charge in FAA ≤ 0 ($C_{2'}$-$C_{2'FAA}$)

Condition 4) Bond length C_3-C_4 minus the same bond length in FAA ≥ 0 (Δ3-4)

Condition 5) Energy of the HOMO $\leq$ -9.3 eV

Table 8.2 show the differences of the AM1 atomic charges at C_3 and O_7 and of the bond lengths C_3-C_4 and C_4-O_7 of the test molecules **a**, **b**, **c** and, **e** with those in FAA. The PM3 structure/activity correlations are in table 8.3.

TABLE 8.2: AM1 structure/activity correlations for molecules in fig. 8.1 (see text).

Ref.	C_3-C_{3FAA}	O_7-O_{7FAA}	Δ3-4 (Å)	Δ4-7 (Å)	I.P.(eV)
a	0.078	-0.003	0.002	0.004	8.643
b	0.003	-0.027	-0.004	0.004	8.708
c	0.123	-	-0.092	-	9.056
e	-0.001	-0.001	0.000	0.000	9.254

TABLE 8.3: PM3 structure/activity correlations for molecules in fig. 8.1 (see text).

Ref.	C_3-C_{3FAA}	C_{16}-C_{16FAA}	$C_{2'}$-$C_{2'FAA}$	Δ3-4 (Å)	I.P.(eV)
a	0.114	0.017	-0.012	0.010	9.012
b	0.044	0.008	-0.028	0.002	8.691
c	0.169	0.027	-0.001	-0.103	9.149
e	0.010	0.001	0.001	0.002	9.301

The activity of molecules **a** - **c** have been measured Atwell et al [29] and they were found to be inactive. Molecule **e** was synthesized in St. Andrews by Dr. Aitken et al [138], its *in-vitro* activity was found to be similar to that of FAA whereas it is completely inactive *in-vivo* [126]. As can be seen from table 8.2 and table 8.3, the condition for activity is not respected and all these molecules are indeed expected to be inactive. Table 8.2 also reports the value of the ionization potential that, although is not included in the AM1 conditions for activity, was found to be lower than 9.3 eV. for most of the inactive

molecules (See Fig. 6.11). For all the molecules in fig. 8.2, the ionization potential is indeed lower than 9.3 eV in agreement with the results in section 6.2.3.

These new results obtained add confidence to the AM1 and PM3 structure/activity correlations discussed in chapter 6; not only are they valid for the series of molecules used to establish the correlations but also for other molecules topologically similar to FAA but with different hetro-atoms than oxygen such as molecules **a-c** in fig.8.1.

8.2 XAAs

This section contains the results of the structure/activity correlation studies on some of the xanthones studied by Atwell and reported in chapter 2. AM1 and PM3 results are reported separately in two different sections because they lead to different conclusions.

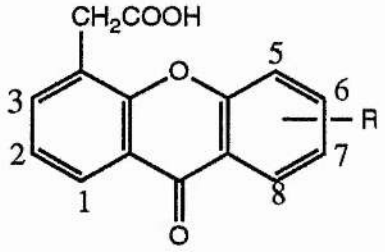
8.2.1 AM1 Results

The geometry optimization of the XAAs in table 8.4 was conducted using the strategy reported in section 6.2. The potential energy surface τ_1 v. τ_2 (the notation is the same as the one used for FAAs) was very similar to that of the FAAs (see Fig. 6.11) and the minima was obtained for τ_1=140° and τ_2=60°. The heats of formation of the optimized structures are shown in table 8.4.

Table 8.5 shows the AM1 structure activity correlations applied to the 35 XAAs of table 8.4. It is encouraging to note that for all but molecule **13** of the 21 active molecules, the differences C_3-C_{3FAA}, O_7-O_{7FAA}, Δ3-4 and, Δ4-7 have the right sign expected for active molecules. However, of the 13 inactive molecules of table 8.5, only 5 of them are predicted to be inactive by the structure/activity correlations the other 8 molecules, according to these correlations, are also expected to be active. It may be due to the fact that the activity data used by Atwell are very strict and as mentioned in chapter 2,

all the molecule less active than 50% are considered inactive; or perhaps there are other factors than those described by these correlations, that determine the activity of the XAAs. For instance, previous structure/activity correlation studies conducted on these molecules by Atwell and co-workers, showed that the activity broadly correlates with the lipophilicity and the potency of these molecules, was increased by having small lipophilic substituents (see section 2.4). They also performed quantum mechanical calculations on some of the XAAs and were able to correlate the activity with the direction of the dipole moment [33]. The latter, was not an important factor for the activity of the FAAs in Chapter 5, however the topology of the XAAs is different from that of the FAAs. In the FAAs, the direction of the dipole moment is changed to a large extent, by the rotation of the phenyl group, XAAs instead are less flexible molecules therefore the direction of their dipole moment will be strictly defined by the nature and position of the substituents. Molecules whose dipole moment is not in the right orientation will not be able to achieve the orientation required for the reaction with the 'receptor' and, even if they are potentially active, they will not react, being unable to move close enough to the active site of the 'receptor'

FIG. 8.2: Numbering scheme of XAA



In conclusion, the AM1 structure/activity correlations described in Chapter 6 are also valid for XAA. They highlight some necessary electronic conditions that analogues to FAA must have in order to show a significant antitumour activity. For XAAs however, other conditions may also be important such as the lipophilicity and dipole moment.

TABLE 8.4: AM1 Heats of Formation of XAAs in Kcal/mole

Ref.	R[a]	ΔH_f (Kcal/mole)	Ref.	R[a]	ΔH_f (Kcal/mole)
11	H	-102.86	**29**	5-CF3	-255.74
12	1-Cl	-103.72	**30**	5-Et	-115.03
13	1-Me	-108.86	**31**	5-Ph	-74.45
14	1-OH	-149.40	**32**	5-NH_2	-103.40
15	1-OMe	-135.56	**33**	5-NHCOMe	-139.19
16	2-Cl	-108.63	**34**	5-NO_2	-95.45
17	2-Me	-110.44	**35**	6-OH	-147.62
18	2-OH	-145.47	**36**	6-Cl	-109.06
19	2-OMe	-139.41	**37**	6-Me	-110.70
20	3-Me	-109.40	**38**	7-OMe	-139.28
21	3-OH	-145.78	**39**	7-Cl	-109.01
22	3-OMe	-139.32	**40**	7-Me	-110.47
23	5-Cl	-106.49	**41**	7-NO_2	-97.55
24	5-OH	-145.57	**42**	7-OH	-145.47
25	5-OMe	-136.96	**43**	8-OMe	-137.06
26	5-OEt	-142.98	**44**	8-Cl	-104.03
27	5-OPr	-149.74	**45**	8-Me	-108.94
28	5-Me	-109.60			

a) See Fig. 8.2 or Chapter 2

TABLE 8.5: AM1 structure/activity correlations for XAAs of table 8.4

Ref.[a]	Activity[b]	C_3-C_{3FAA}	O_7-O_{7FAA}	Δ3-4 (Å)	Δ4-7 (Å)
14	-	0.100	-0.038	0.006	0.005
15	-	0.101	0.014	0.010	-0.001
18	-	0.098	0.004	0.007	-0.000
21	-	0.103	-0.002	0.009	0.000
22	-	0.103	-0.003	0.009	0.000
24	-	0.140	0.010	0.010	-0.001
35	-	0.063	-0.000	0.005	0.000
41	-	0.085	0.021	0.011	-0.002
42	-	0.136	0.010	0.011	-0.001
45	-	0.100	-0.003	0.009	0.001
43	-	0.072	0.025	0.007	-0.002
33	-	0.126	0.007	0.011	-0.000
31	-	0.099	0.003	0.009	-0.000
19	+	0.098	0.002	0.007	-0.000
40	+	0.105	0.003	0.008	-0.000
13	+	0.098	-0.003	0.010	0.001
16	+	0.101	0.009	0.008	-0.001
17	+	0.100	0.003	0.008	-0.000
20	+	0.100	0.001	0.008	-0.000
34	+	0.082	0.012	0.013	-0.001
25	+	0.128	0.006	0.010	-0.000
36	+	0.098	0.007	0.009	-0.001
37	+	0.094	0.001	0.007	-0.000
39	+	0.107	0.009	0.010	-0.001
38	+	0.135	0.009	0.011	-0.001
44	+	0.115	0.025	0.014	-0.002
29	+	0.092	0.010	0.011	-0.001
32	+	0.138	0.005	0.011	-0.000
12	++	0.101	0.026	0.012	-0.002
23	++	0.106	0.007	0.011	-0.001
11	++	0.100	0.003	0.008	-0.000
30	++	0.104	0.003	0.009	-0.000
28	++	0.104	0.003	0.009	-0.000
26	++	0.129	0.004	0.010	-0.000

[a] See table 8.4. or chapter 2 [b] Refer to section 2.4.

8.2.2 PM3 results

The results obtained with the PM3 method are disappointing. In fact, the PM3 structure/activity correlation found with the 10 FAAs in chapter 6 are not valid when applied to the XAAs. The geometry of a small number of XAAs (smaller than that studied with the AM1 method) was optimized using the standard strategy described in chapter 6; the heats of formation obtained are reported in table 8.6 and as expected, are smaller than those obtained with AM1.

TABLE 8.6: Heats of formation of the XAAs studied with PM3.

Ref.[a]	ΔH_f (Kcal/mole)	Ref.[a]	ΔH_f (Kcal/mole)
11	-104.57	**34**	-114.47
13	-115.62	**24**	-152.11
14	-158.37	**25**	-143.72
15	-143.64	**36**	-115.91
16	-115.44	**47**	-148.40
18	-153.67	**39**	-115.70
19	-146.73	**40**	-118.94
46	-114.56	**41**	-117.43
20	-118.32	**38**	-146.52
22	-146.06	**44**	-113.12
23	-114.14	**45**	-115.84
28	-117.47	**43**	-151.08

[a] See table 8.4. or chapter 2

The PM3 structure/activity correlations are shown in table 8.7. $C_{2'}$-$C_{2'FAA}$ is not reported because $C_{2'}$ does not exist in XAAs. According to the results obtained in chapter 6 for the FAAs, the sign of all the differences C_3-C_{3FAA}, C_{16}-C_{16FAA} and, Δ3-4 should be positive for all the active molecules and the ionization potential larger than 9.3 (a.u). From table 8.7 however, it is evident that these conditions are not respected, for example the difference C_{16}-C_{16FAA} is negative for all active and inactive molecules and also the ionization potential of very active molecules such as **23** and **24** is smaller than 9.3 (a.u).

These results are clearly unsatisfactory and point to a failure of the PM3 methods in providing a simple tool for SAR studies of FAA analogues.

TABLE 8.7: PM3 structure/activity correlations for XAAs of table 8.6

Ref.[a]	Activity[b]	C_3-C_{3FAA}	C_{16}-C_{16FAA}	Δ3-4 (Å)	I.P. (a.u)
14	-	0.091	-0.140	0.004	8.663
15	-	0.096	-0.137	0.016	9.153
18	-	0.090	-0.142	0.011	9.057
22	-	0.096	-0.141	0.013	9.282
24	-	0.134	0.076	0.015	9.175
41	-	0.063	-0.163	0.015	9.909
45	-	0.101	-0.152	0.016	9.225
43	-	0.059	-0.186	0.014	9.159
13	+	0.090	-0.140	0.014	9.223
16	+	0.093	-0.139	0.011	9.224
19	+	0.090	-0.142	0.011	9.012
20	+	0.094	-0.141	0.012	9.269
34	+	0.074	-0.435	0.016	9.817
25	+	0.111	0.035	0.016	9.253
36	+	0.089	-0.146	0.012	9.423
47	+	0.053	-0.232	0.009	9.277
39	+	0.104	-0.127	0.013	9.224
40	+	0.098	-0.134	0.012	9.183
38	+	0.130	-0.101	0.014	8.989
44	+	0.097	-0.147	0.014	9.267
46	++	0.095	-0.140	0.012	9.394
23	++	0.102	-0.174	0.013	9.264
28	++	0.096	-0.104	0.013	9.233

[a] See table 8.4. or chapter 2 [b] Refer to section 2.4.

8.3 Summary and discussion

AM1 and PM3 quantum mechanical calculations on ten derivatives of flavone acetic acid, with known experimental anticancer activity, were conducted in order to look for structure/activity correlation. A standard strategy was used in the optimization of the geometry to ensure reproducibility of the results. As reported in chapter 6, both AM1 and PM3 methods showed that the activity correlates with the difference between selected atomic charges and bond lengths of test molecules minus those of the lead molecule FAA.

These differences were often very small nevertheless it was possible to differentiate between active and inactive molecules.

In order to determine whether these S/A correlations could be successful in predicting the activity of analogues of FAA, AM1 and PM3, quantum mechanical calculations were carried out on some XAAs and FAAs other than those used to identify the correlations. The experimental activity of these XAAs and FAAs was known although it was expressed in a different way than for FAAs (See chapter 2).

PM3 did not predict correctly the activity of the XAAs while AM1 succeeded in this goal and predicted correctly the activity of 25 of the 34 XAAs derivatives plus the activity of 4 FAA derivatives; this give a total of 76 % correct predictions. If only the active molecules are considered, then, the SA correlations show a 95% success rate (of 21 active XAAs 20 were predicted to be active by the correlations). It also emerged from this study that in the case of XAAs there may be other conditions that have to be fulfilled in order for them to be active, such as small lipophylic groups and a proper direction of the dipole moment as suggested by Atwell [33].

AM1 has emerged from this study as a suitable quantum mechanical method for structure/activity relationship studies on molecules related to FAA. About 50 molecules have been used to establish the fact that it can be used with confidence to predict the activity of FAA analogues. In doing so, it is important to remember that if these correlations are going to be used to predict new active derivatives, care has to be taken in the geometry optimization process, it should be conducted using the strategy described in section 6.2.

8.4 Possibilities and suggestions for future work

FAA is active against solid tumours but not leukemias; it acts on the vascular system of solid tumours without affecting the normal cells, as proven by the fact that it has no substantial side effects. It means that FAA is able to differentiate between solid

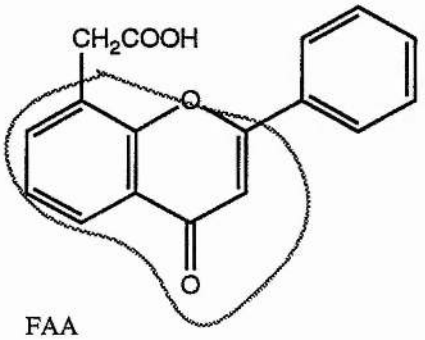
tumours cells and other cells. Bibby and Double suggested that because the pH within solid tumours is lower than that inside the normal cells, the action of FAA could be increased by adding to it a group that at low pH becomes cytotoxic [128]. An other characteristic of solid tumours is that the environment inside its cells is more reductive than that of the well oxygenated normal cells. This characteristic has already been used to introduce the idea of bioreductive antineoplastic agents [129]. Because FAA is able to differentiate between the normal cells and those of solid tumours, its activity could also be increased by adding chemical groups that become cytotoxic after a reductive reaction that can occur only after FAA reaches the solid tumour.

The choice of which groups should be added goes beyond the aim of this thesis because it requires an in-depth knowledge of the biochemistry of the tumour cells and pharmacology. Once a chemical group has been proposed however, AM1 quantum mechanical calculations can be performed on the new molecules obtained adding the group to different positions of FAA using the strategy described in 6.2 for the geometry optimization. If the calculated atomic charges and bond lengths match the requirement of the structure/activity correlations described previously, there is a good chance that the new molecule will be active. If these requirement are not respected, then it may not be necessary to synthesize the molecule at all, saving a large amount of time and of money.

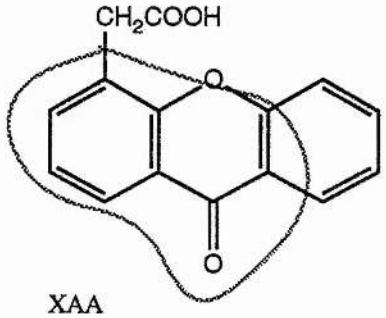
An understanding of the mode of action of FAA is puzzling numerous scientists all over the world (including myself during the course of this work). As mentioned in Chapter 2, FAA activates the blood clotting cascade [18] after 15 minute of drug administration [20] giving rise to thrombus formation in the blood vessels. This latter action of FAA is particularly interesting because it is well known that vitamin K is essential in the production of certain factors involved in the clotting cascade [18]. Vitamin K, and FAA show some similarity in their structure as can be seen in fig. 8.3. Dicoumarol and warfarin are two antagonists of vitamin K used clinically as anticoagulating agents and also as rodenticides (at large doses), because they cause excessive permeability of blood vessels [18]. They could also achieve a similar structure

to FAA after rearrangment (Fig. 8.4). On the other hand, FAA could achieve a similar structure to that of dicoumarol by protonation of the exocyclic oxygen.

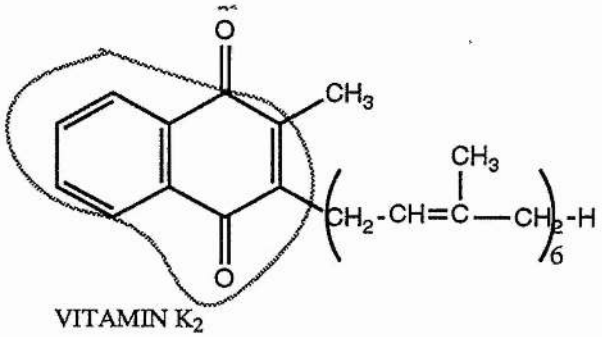
FIG 8.3: Formulas of FAA, XAA, vitamin K2, dicoumarol and warfarin.



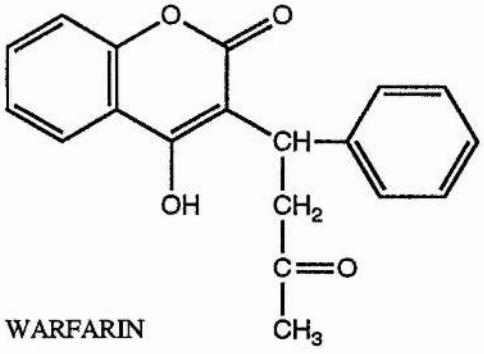
FAA



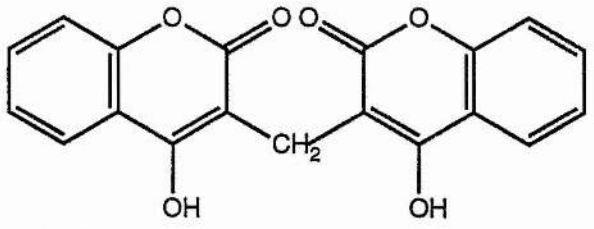
XAA



VITAMIN K_2

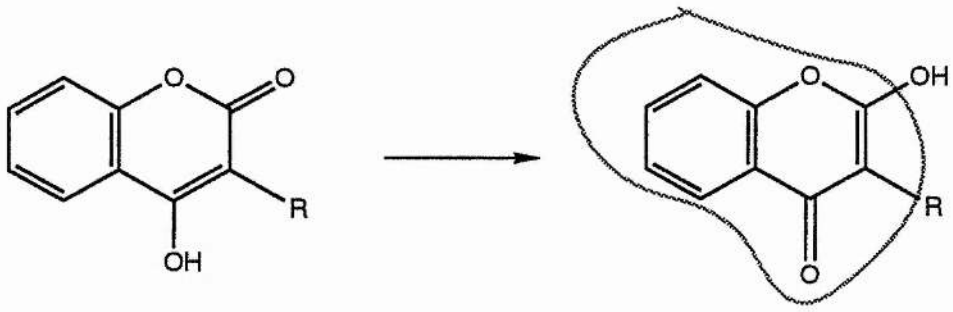


WARFARIN

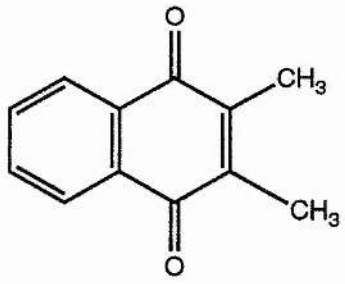


DICOUMAROL

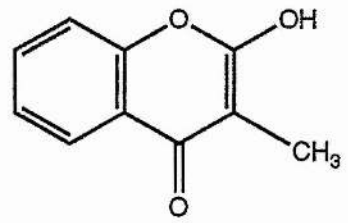
FIG. 8.4: Dicoumarol and warfarin can achieve a similar structure to that of FAA by rearrangement



AM1 quantum mechanical calculations were carried out on **f** and **g**. The first contains the basic ring structure of Vitamin K and the later that of dicoumarol and warfarin. The heats of formations of the optimized structures were -30.104 Kcal/mole and -77.072 Kcal/mole respectively.



f



g

The AM1 structure/activity correlations were applied to these molecules and they are reported in table 8.8. As the table shows, according to the S/A correlation, molecule **f** could show a similar activity to FAA.

That Vitamin K, or molecule **f**, may have a similar mode of action to FAA is only a speculation, but if true it would be a big contribution to cancer research because the way in which vitamin K works is well understood. In future work, it would be interesting to perform a set of experiments in which Vitamin K is tested for the anticancer activity of solid tumours. Alternatively, it (or its antagonists) could be administered with FAA to

see whether its anticancer activity is affected. Such experiments may bring further insight into the mode of action of FAA.

TABLE 8.8: AM1 structure/activity correlations of analogues of Vitamin K and dicoumarol.

Ref.[a]	C_3-C_{3FAA}[b]	O_7-O_{7FAA}[b]	Δ3-4 (Å)[b]	Δ4-7 (Å)[b]
f	0.188	0.026	0.021	-0.001
g	0.030	-0.002	0.004	0.000

[a] See Fig.8.2 . [b] See chapter 6

As concerns calculations, it will be interesting to perform *ab-initio* calculations with high-quality basis sets, perhaps with polarization functions, on the set of FAAs used in the S/A correlation studies including the study of the MEPs calculated directly from the wavefunction. Such studies would be useful in understanding the differences in the reactivity of the FAAs and therefore in understanding the mode of action of FAA.

To conclude, the results obtained from this study add confidence to the fact that thanks to the joint effort of biochemists, pharmacologists, medicinal chemists and, theoretical chemists, the rational design of anticancer drugs is no longer a dream but is achievable now.

CHAPTER 9:

Studies on the Dependence of the MEP on the Basis Set.

Introduction

The validity of the molecular electrostatic potential (MEP) as a tool for understanding the chemical reactivity of molecules, in particular towards electrophilic reactions has become, in recent years, a matter of fact [72-81]. The calculation of the MEP is however computational expensive and several approximations have been proposed [72]. The point charge approximation [72,84] allows the calculation of the MEP at a minimal computational cost and has been used extensively over the years. The MEP obtained with this approximation however, is not always very rigourous and may lead to erroneous interpretations. As discussed in chapter 7, one of the reason for this failure is that the charges used are generally Mulliken charges, which depend strongly on the basis set used. Furthermore, the accuracy of these charges does not necessarily increase with the basis set and it is often the case that minimal basis sets gives a set of charges (and therefore the MEP calculated with the point charge approximation) more credible than those obtained with a split valence basis set (see chapter 7). With the development of computer technology and quantum mechanical software today, it is possible to perform ab-initio calculations with large basis sets even on large molecules of biological interest and the calculation of the MEP directly from the wavefunction is no longer so difficult as it was, for example, only three years ago, when this project started. It is known that the MEP depends on the basis set and some of the small basis sets may lead to erroneous interpretation of the MEP [130]. It seems, however, that in existing studies of the dependence of the MEP over the basis set, a very small selection of basis sets were actually considered [130-133] and a systematic study of the dependence of the MEP on the basis set has not yet been performed.

In this chapter we report ab-initio calculations of the MEP of pyran-2-one with a large number of basis sets [49-57] including the geometrical basis sets of Clementi et al [57] whose ability to predict accurate MEP have not yet (as far as is known) been tested. The MEPs are compared by means of the Hodgkin similarity index calculated from equation 4.1; in this way a quantitative measure of the similarity between the MEPs is

obtained. Particular attention is given to minimal basis sets because they are suitable for the study of large molecules.

The charges which best reproduce the MEPs are also calculated and compared with the Mulliken charges. In this way, the differences between the MEP calculated from the wavefunction (as the expectation value of the operator 1/**r**) and that calculated with the Mulliken charges (using the point charge approximation) are also highlighted. This is because the charges obtained from the MEP are those that best fit the electrostatic potential calculated as expectation value from the wavefunction and, using these charges to calculate the MEP with the point charge approximation, reproduce the most similar MEP to that obtained directly from the wavefunction that can be obtained within the point charges approximation. The MEP generated from other charges will differ from the latter and a measure of these differences can be obtained by comparing these charges with those obtained from the potential.

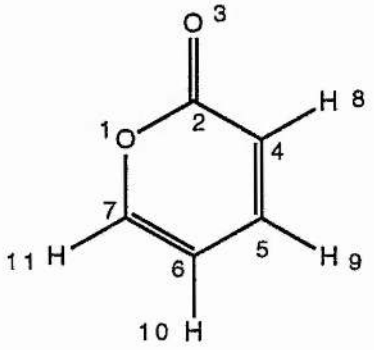
The study of pyran-2-one was of interest to the research group, because its basic ring structure is found in biological interesting molecules such as coumarin and fostriecin. For the kind of molecules reported in this work, the study of the γ-pyrone would have been more appropriate: however, both the molecules have the same functional groups and the results obtained with one should be transferable to the other. Furthermore, while the study of the MEP of γ-pyrone has been conducted by Thomson and Edge [125], the study of the MEP of pyran-2-one has, as far as is known, not been reported before.

Bond lengths and angles of pyran-2-one were taken from the experimental geometry measured by microwave spectroscopy [134]. The MEP was calculated on a Connolly surface at 1 Van der Waals radius from the molecule; the surface point density was set equal to 5 and the MEP was calculated at 492 points. The numbering scheme of this molecule used throughout this chapter is shown in fig. 9.1.

Section 9.1 contains four subsections, each of them reports the results obtained with the basis sets proposed by Pople [49-53], Huzinaga [54,55], Dunning [56] and,

Clementi [57] respectively. In section 9.2 is reported the comparison between minimal basis sets and, section 9.3 contains discussion and conclusions.

FIG. 9.1: Numbering scheme for pyran-2-one



9.1: The basis sets studied

9.1.1: The Pople's basis sets

Among the basis sets used in this study there are those developed by the Pople group such as the minimal STO-NG [49]; N-21G [53]; N-31G [51], N-31G* and N-31G** [52]. These basis sets are quite well known and most of the MEP calculations reported in literature are conducted using STO-3G, 3-21G, 6-31 and, 6-31G* basis sets [130-133]. Here, a systematic study of these basis sets has been conducted and the influence of the number of the gaussian functions N has also been analyzed. In table 9.1 the energies of α-pyrone with the number of primitives, shells and basis functions for each of the basis set used are reported. As expected, 6-31G** gives the lowest energy of all. In fig. 9.2 the MEPs obtained with these basis sets are compared in terms of the Hodgkin similarity index where the 6-31G** MEP is used as a lead. From the plot in fig 9.2 it can be possible to identify three main groups of basis sets for which the MEPs are very similar. The first group contains the basis sets with polarization functions; going from N-31G* to N-31G** i.e. adding polarization functions to the H atoms, do not change the MEPs sensibly, because for all of them the similarity index is equal to one. To

the second group, belong all the split valence basis sets N-31G and N-21G for which the similarity index is very close to one. It is interesting to note that 3-21G performs slightly better than 6-31G, the similarity indexes being 0.992 and 0.989 respectively. The third group contains the minimal basis sets STO-NG. As can be seen, the similarity between the MEP obtained with 6-31G** and with any of these minimal basis sets is very poor especially that of STO-3G.

In fig 9.3 are shown the atomic charges which best reproduce these MEPs. As for the similarity index, within the same group the charges do not change very much however, the charges produced by STO-NG differ from those obtained with N-21G and N-31G and further changes are observed when polarization functions are added. As can be seen in plot **d** of fig. 9.3, the main differences between the charges obtained from 6-31G** and the other basis sets such as 6-31G, 3-21G and STO-3G are on atoms which are either positive or negative; neutral atoms show more or less the same charge whatever the basis set. In comparison to 6-31G**, STO-3G tends to underestimate the charges obtained from the MEP while 3-21G and 6-31G overestimates them.

FIG. 9.2 Pople's basis sets versus Hodgking similarity index of the MEP

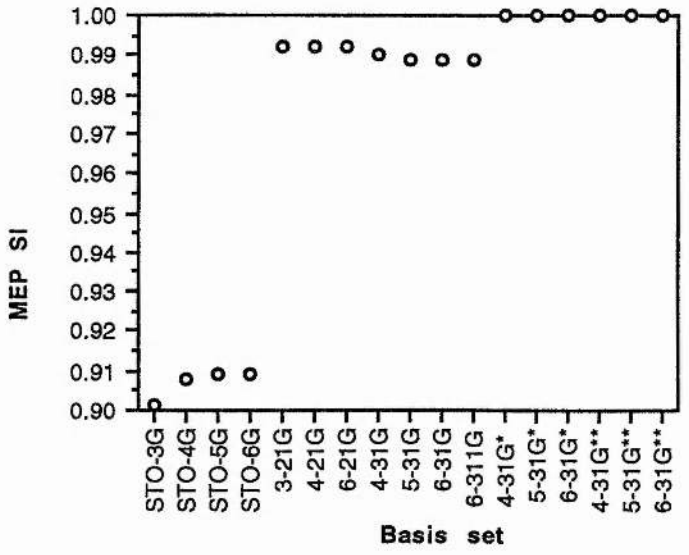
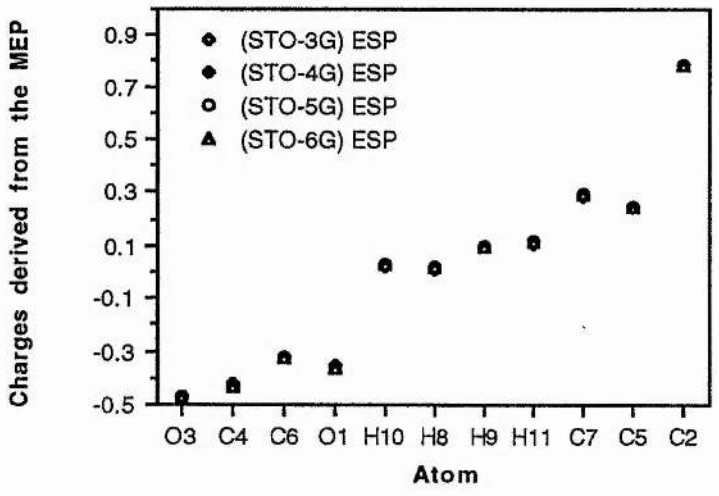


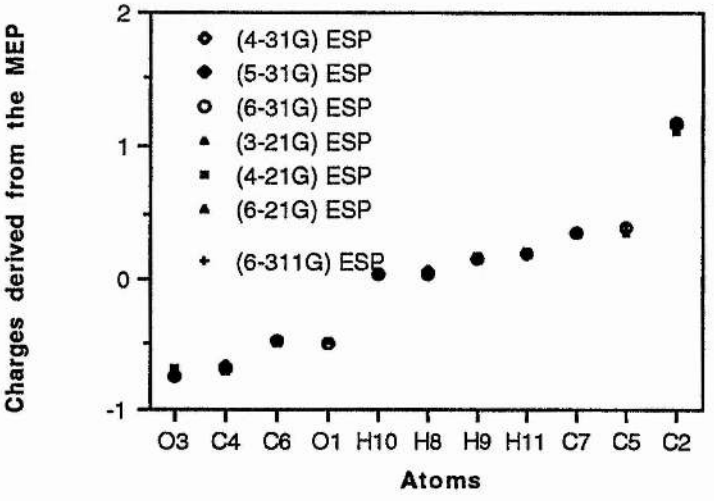
TABLE 9.1: Energies, number of primitives, shells and of basis functions for Pople's basis set for pyran-2-one.

Basis set	Energy (a.u)	N° primitives	N° shells	N°basis funct
STO-3G	-336.961	117	18	39
STO-4G	-339.393	156	18	39
STO-5G	-340.043	195	18	39
STO-6G	-340.244	234	18	39
3-21G	-339.435	117	29	71
4-21G	-340.576	124	29	71
4-31G	-340.834	156	29	71
5-31G	-341.115	163	29	71
6-21G	-340.902	138	29	71
6-31G	-341.187	170	29	71
6-311G	-341.264	202	40	103
4-31G*	-341.010	198	36	113
5-31G*	-341.275	205	36	113
6-31G*	-341.344	212	36	113
4-31G**	-341.018	210	40	125
5-31G**	-341.282	217	40	125
6-31G**	-341.351	224	40	125

FIG 9.3: Charges which best reproduce the MEP calculated with: (a) minimal basis STO-NG, (b) Split-valence basis sets N-21G, N-31G and, 6-311G; (c) N-31G* and N-31G**. (d) Comparison between the basis sets.



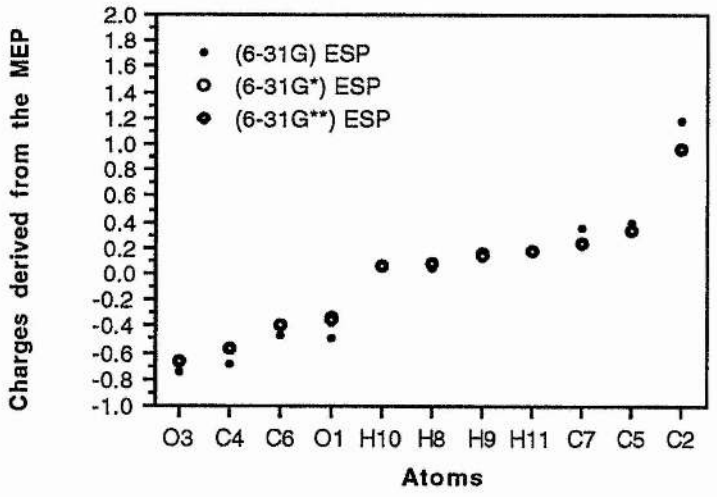
a



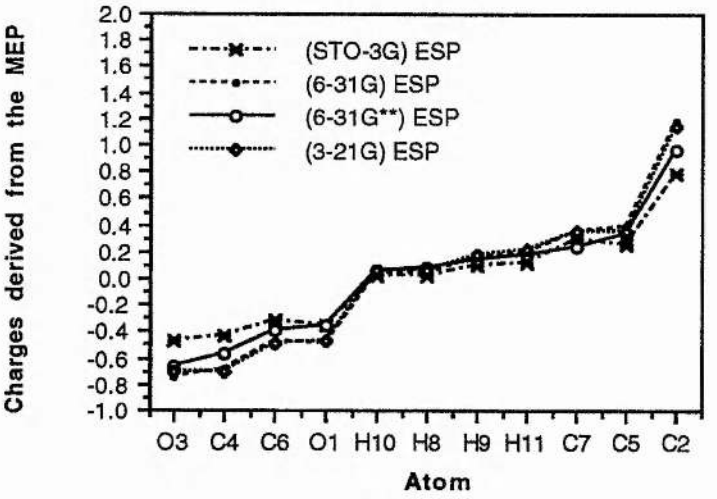
b

Cont...

...Fig. 9.3 cont.



c

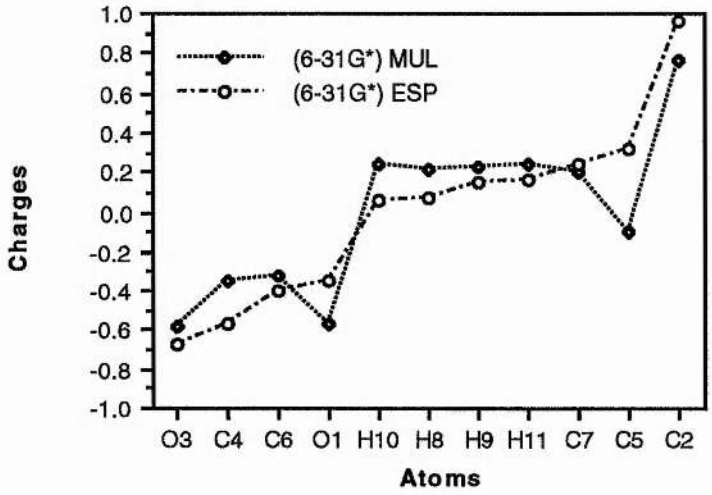


d

Fig. 9.4 shows the differences between the Mulliken charges and those obtained from the MEP both with the 6-31G* basis set. From this plot is clear that MEP obtained from Mulliken charges will be very different from those obtained as expectation value of the 1/**r** operator. For example, the Mulliken charge at C_5 is -0.068, while the electrostatic potential derived charge at the same position is 0.324 that is 0.392 a.u bigger! On the other hand, the electrostatic potential derived charge at C_2 is only 0.193 a.u. bigger while

it is 0.22 a.u. smaller at C_4. It is also interesting to note that O_1 and O_3 have the same Mulliken charges, while with the charges derived from the potential O_3 is about 0.3 a.u. more negative than O_1; this is what one would expect given the fact that protonation reaction of pyran-2-one occur at O_3 [134].

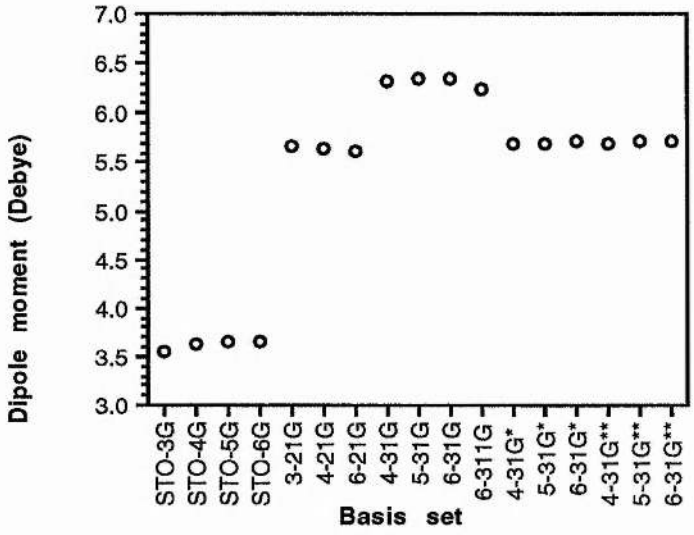
FIG. 9.4: Mulliken charges and charges obtained from the MEP.



The value of the dipole moment predicted from the STO-NG basis sets is about 2 Debye lower than the value predicted by 6-31G** (5.7 Debye), as can be seen in fig. 9.5 which shows the magnitude of the dipole moment as a function of the basis set. The N-21G basis sets, in particular the 3-21G, perform better than N-31G basis sets in predicting the dipole moment. According to the latter, the dipole moment is between 6.2 and 6.4 Debye, that is higher than that predicted by 6-31G** and the experimental value which, for pyran-2-one, is known to be about 5 Debye [134]. The superiority of the 3-21G basis set over the 6-31G has also been observed for the similarity of the MEP with that obtained with 6-31G** (fig. 9.2) and for the charges derived from this MEP (fig. 9.3). In one way this result is unexpected because the 6-31G is a larger basis set than 3-21G and from a variational point of view, it is a 'better' basis set; however, the fact that it

produces a lower energy does not necessary means that it is also going to produce better properties and the example of the pyran-2-one is one which clearly is not.

FIG. 9.5: Dipole moment of pyran-2-one calculated with Pople's basis sets



9.1.2: The Dunning basis sets

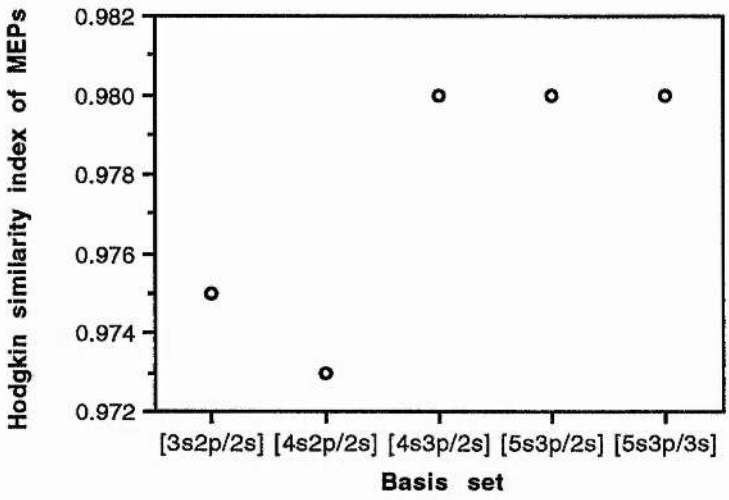
The influence of the degree of contractions of the (9s5p/4s) in the Dunning basis sets [56] have been studied. MEPs have been calculated with the [3s2p/2s], [4s2p/2s], [4s3p/2s], [5s3p/2s] and [5s3p/3s] basis sets. These basis sets are quite large, the smallest [3s2p/2s] for the molecule under study has 71 basis functions, the same as 6-31G. The energy and details of the basis set for pyran-2-one with these basis sets is in table 9.2. The Hodgkin similarity indexes of the MEPs has been calculated and reported in fig. 9.6. The similarity index of the MEP is affected in particular by the degree of flexibility of the p orbitals which are also the more external orbitals for carbon and oxygen. In fact, for all the basis sets in which the 5p gaussians have been contracted to 2p functions, the similarity index is smaller than 9.75, while where they are contracted to 3p functions, the similarity index increases. The degree of contraction of the s orbitals

both in heavy and hydrogen atoms do not appear to be of particular importance (See fig. 9.6).

TABLE 9.2: Energies of the pyran-2-one with Dunning basis sets.

Basis set	Energy (a.u)	N° primitives	N° shells	N°basis funct
[3s2p/2s]	-341.217	184	43	71
[4s2p/2s]	-341.236	184	50	78
[4s3p/2s]	-341.256	184	57	99
[5s3p/2s]	-341.259	184	64	106
[5s3p/3s]	-341.260	184	68	110

FIG. 9.6 Dunning's basis sets versus Hodgking similarity index of the MEP (6-31G** MEP is used as reference).



The differences in the similarity indices within the Dunning basis sets are however very small and it can be stated that MEPs obtained with these basis are almost independent from the degree of contraction of the basis set. This is perhaps easier to see

in fig. 9.7 where the atomic charges derived from the MEPs with the different basis sets, are plotted against the atom's number: the charges are practically the same for all the Dunning's basis sets. When compared to the charges derived from the 6-31G** basis set, it can be seen that the Dunning charges are slightly overestimated (Fig.9.7). Fig 9.8 shows the comparison between Mulliken charges and charges derived from the MEP for the [5s3p/3s] basis sets. As can be seen, Mulliken charges underestimate very positive charges, for example, the Mulliken charge at C5 is -0.045 a.u. while the electrostatic potential fitted charge it is +0.405. The use of Mulliken charges to obtain MEP is therefore not advisable because these differences will also appear in the electrostatic potential calculated with these charges.

FIG. 9.7 Atomic charges derived from the MEP

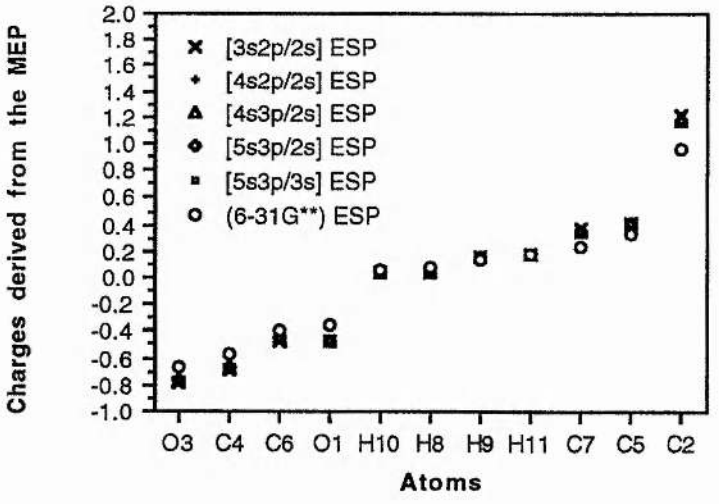
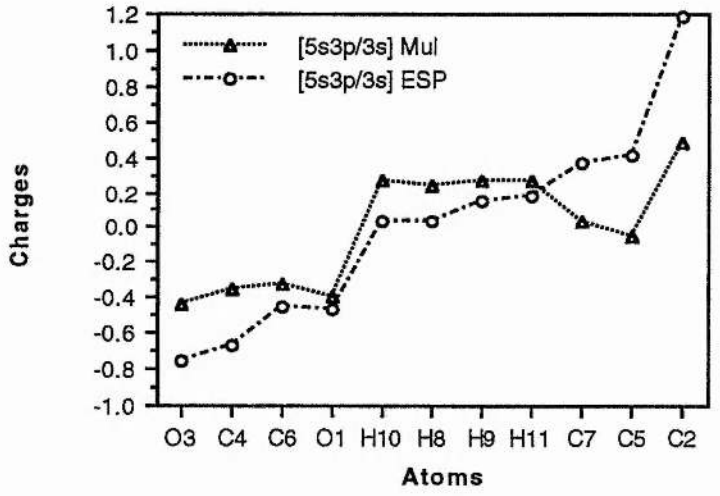
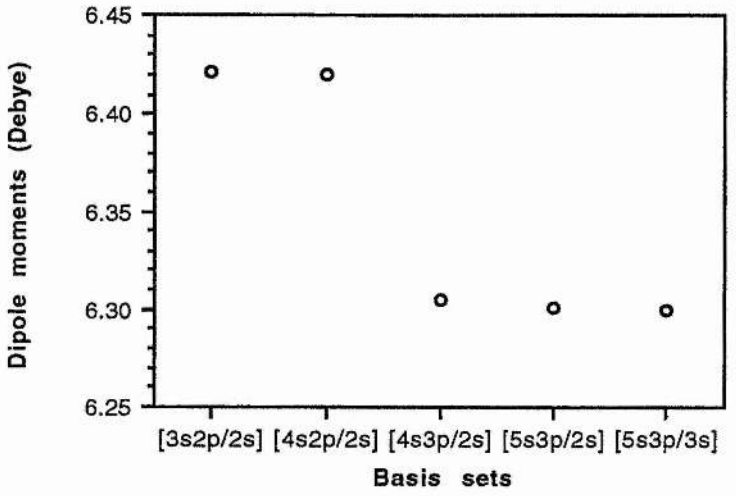


FIG. 9.8 Comparison between atomic charges derived from the MEP and Mulliken charges



As for the MEP, the dipole moment obtained with Dunning basis sets also seems to depend on the flexibility of the external p orbitals of the heavy atoms; in fact the dipole moment obtained with the basis sets [3s2p/2s] and [4s2p/2s] is about 0.2 Debye higher than the value obtained when the p orbitals are described by 3 functions instead of by two (See Fig. 9.9). These values are also higher than those obtained with the 6-31G** basis sets and are comparable to those predicted with the N-31G basis sets (Fig. 9.5).

FIG. 9.9 Dipole moments of pyran-2-one with Dunning basis sets.



9.1.3: The MINI-i and MIDI-i basis sets

In 1980, Tatewaki and Huzinaga published a collection of minimal and split valence basis sets called MINI-i and MIDI-i (i=1-4) respectively [54]. Some of these basis sets have already been used by some researchers (see ref. [135-137]) and their use will probably increase over the next few years. Energies of pyran-2-one and details of the basis sets can be found in Table 9.3. MINI-1 and MIDI-1 contains the same number of basis functions and primitives as STO-3G and 3-21G (Table 9.1) althought they differ in the number of shells. In Pople basis sets the number of shells is decreased by the constraint that 2s and 2p orbitals share the same exponent while this constraint is released in the MINI and MIDI basis sets. The MINI-1 basis set is gives lower energies than STO-3G, for pyran-2-one the difference is about 2 a.u. .

TABLE 9.3: Energies of the pyran-2-one with Huzinaga's MINI-i and MIDI-i basis sets.

Basis set	Energy (a.u)	N° primitives	N° shells	N°basis funct
MINI-1	-338.931	117	25	39
MINI-3	-340.284	128	25	39
MIDI-1	-339.288	117	43	71
MIDI-2	-339.520	138	43	71
MIDI-3	-340.575	128	43	71
MINI-1*	-339.267	171	36	93
MINI-3*	-340.560	182	36	93
MIDI-1*	-339.504	171	54	125
MIDI-2*	-339.722	192	54	125
MIDI-3*	-340.756	182	54	125

The MEP of pyran-2-one with these basis sets has been calculated and compared to that obtained with 6-31G**, the results of this comparison is shown in fig. 9.10. The inclusion of polarization functions on the minimal basis sets produce worst electrostatic potential maps that those without extra functions. The production of such polarized basis set by Tatewaki and Huzinaga is surprising because it was expected that these basis sets would be affected by a large superimposition error. Within the MIDI-i family, the MIDI-1 produces the best MEP and its similarity index is close to that obtained with 3-21G basis set. Inclusion of polarization functions to the MIDI-i basis sets improve the MEP in particular MIDI-2* has a similarity index very close to 1.

FIG. 9.10: MINI- and MIDI-i basis sets versus Hodgking similarity index of the MEP

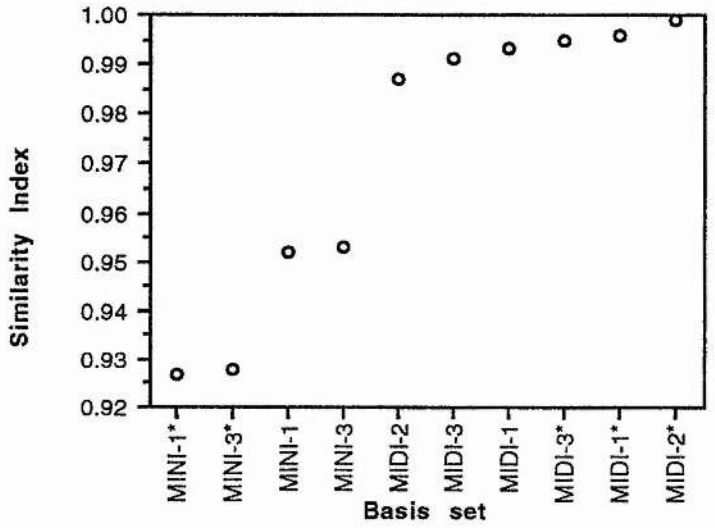
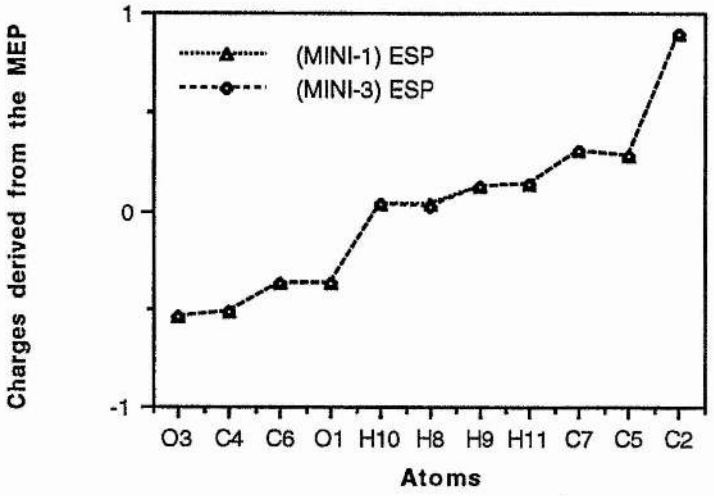
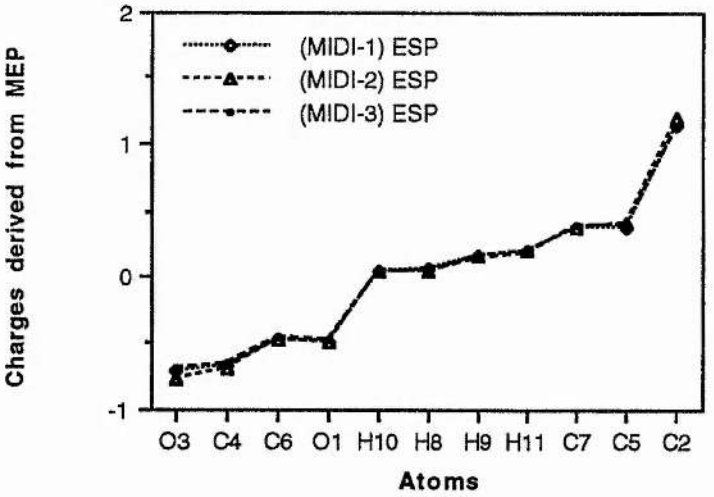


Fig. 9.11 shows the charges which best reproduce these MEPs. Adding primitive functions to MINI-1 or MIDI-1 basis sets does not affect the charges obtained from the MEPs (Fig. 9.11 **a**, **b** and, **c**). As can be seen from plot **d** of fig. 9.11, MIDI-2* and 6-31G** produce almost the same charges, both MINI-1 and MIDI-1 tend to overestimate the positive charges while negative charges are overestimated by MIDI-1 and underestimated by MINI-1.

FIG. 9.11: Atomic charges derived from the MEP



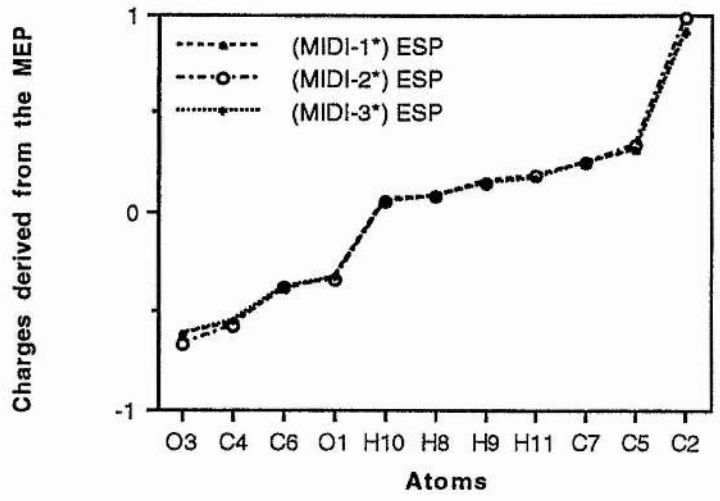
a



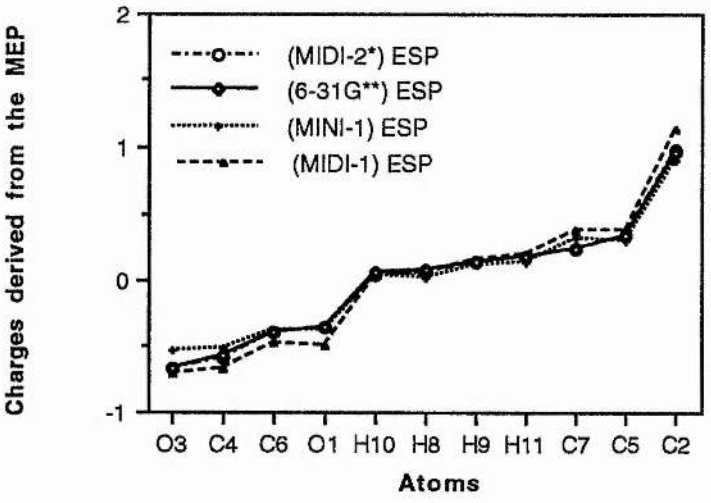
b

Cont...

...Fig. 9.11 cont.



c



d

The Mulliken charges obtained with these basis sets (see fig. 9.12) are also very different from the fitted charges and the same considerations which were made for the other basis sets are also valid for the Huzinaga MINI and MIDI basis sets.

FIG. 9.12: Comparison between Mulliken charges and charges obtained from the potential

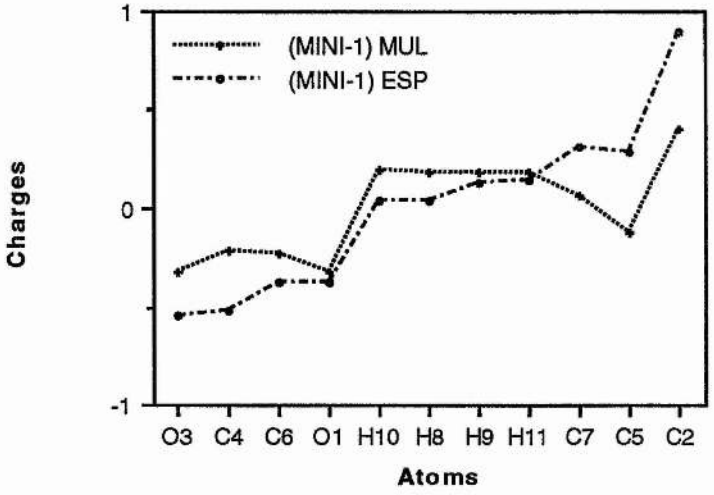


FIG. 9.13: The dipole moment of pyran-2-one with MINI and MIDI basis sets.

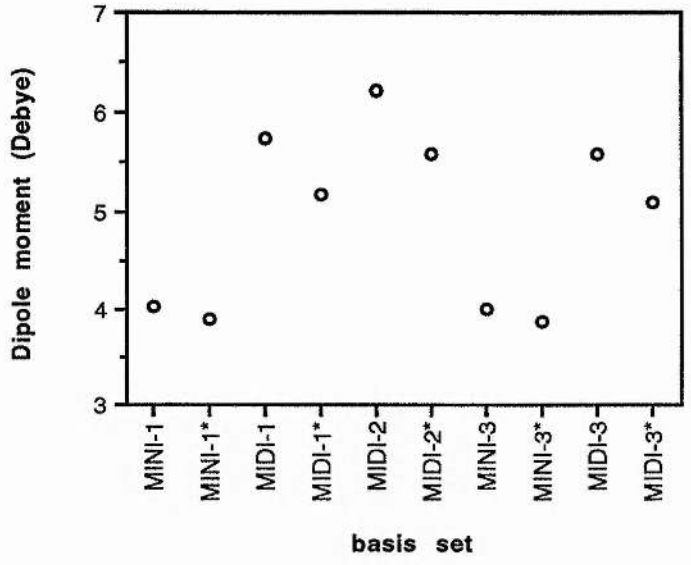


Fig. 9.13 shows the dipole moment of pyran-2-one obtained with the Huzinaga basis sets. Values of dipole moment similar to those obtained with 6-31G** basis sets are obtained with the split-valence MIDI-i basis sets; the MINI basis instead give values which are too low (~ 4 Debye).

9.1.4: Geometrical basis sets

The geometrical basis sets were introduced in 1982 by Clementi and Corongiu [57]. As mentioned in chapter 3, the term 'geometrical' is derived from the fact that the exponents of the single gaussians are terms of a geometrical progression. Energies and basis sets information for pyran-2-one are reported in table 9.4. It is interesting to note the low value of the energy obtained with the minimal basis set GEOSMALL. Although this basis set is slightly smaller in size than STO-5G, the energy it produces is only about 0.1 a.u. lower than the energy obtained with the 4-21G basis set (see table 9.1).

TABLE 9.4: Energies of the pyran-2-one with geometrical basis sets.

Basis set	Energy (a.u)	N° primitives	N° shells	N°basis funct
GEOSMALL	-340.395	192	25	39
GEOSPV	-340.826	192	43	71
GEOMEDIUM	-341.178	236	50	78
GEOSPTCV	-341.037	192	50	78
GEOLARGE	-341.292	299	68	110
GEOTRIPLEZ	-341.297	299	75	117

Fig. 9.14 shows the similarity indexes of the MEPs, obtained with geometrical basis sets, compared to that obtained with 6-31G** basis set. The similarity indexes obtained with basis sets larger than GEOSPV is within the average obtained with the Pople and Huzinaga basis sets of similar size (See also figs. 9.2 and 9.10). Unfortunately

the geometrical basis sets with inclusion of polarization functions are not available for comparison. The similarity index obtained with the minimal basis set, GEOSMALL, is however far higher than that obtained with the other minimal basis sets as can also be seen in figs. 9.2 and 9.10.

FIG. 9.14: Hodgkin similarity index of MEPs with geometrical basis sets.

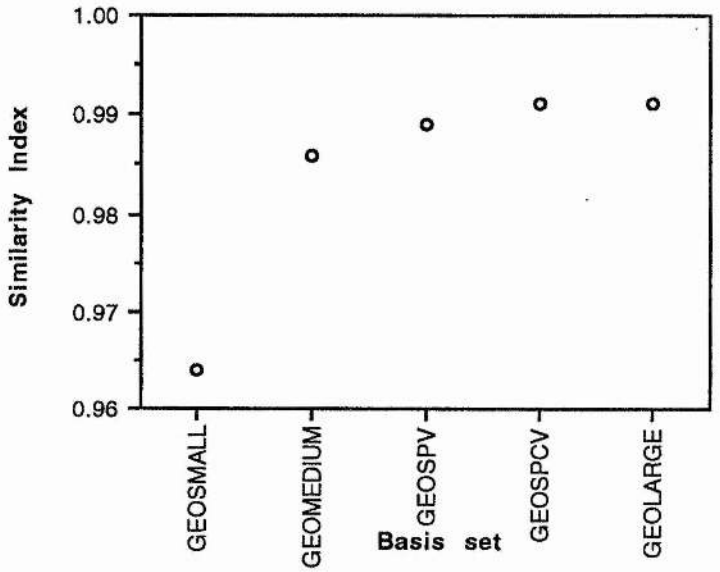
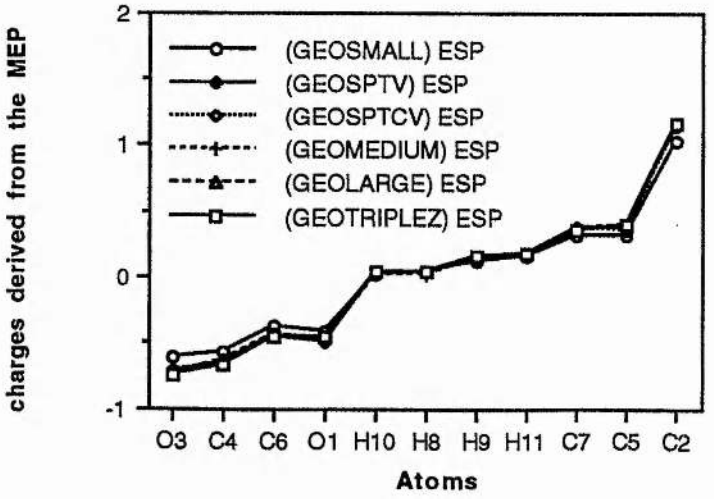


FIG. 9.15: Charges derived from the MEPs with geometrical basis sets.



The charges obtained from the MEP with the geometrical basis sets are all very similar, only GEOSMALL tends to slightly underestimate negative and positive charges as can be seen in fig. 9.15. As expected, Mulliken charges are not similar to the charges obtained from the potential (See fig. 9.16) and are not recommended for the calculation of the MEP using the point charge approximation.

Of the geometrical basis sets only GEOSPV and GEOMEDIUM give a value of the dipole moment similar to that obtained with 6-31G**, GEOSMALL underestimates it whereas GEOSPCV, GEOLARGE and, GEOTRIPLEZ overestimates it (Fig. 9.17).

FIG. 9.16: Comparison between Mulliken charges and charges derived from the MEP with geometrical basis sets.

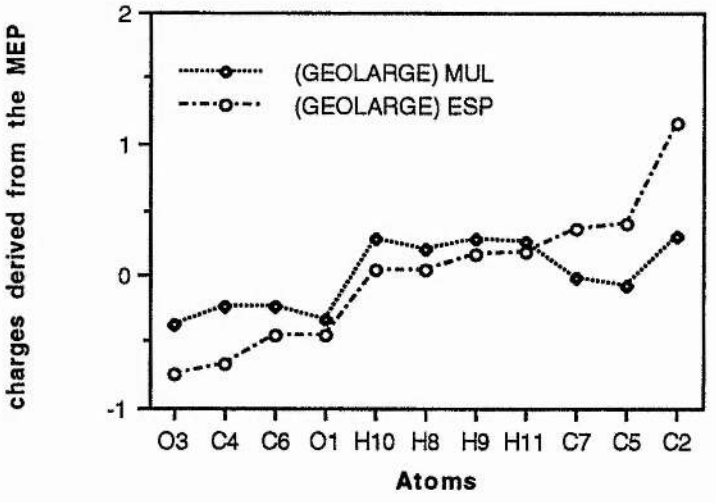
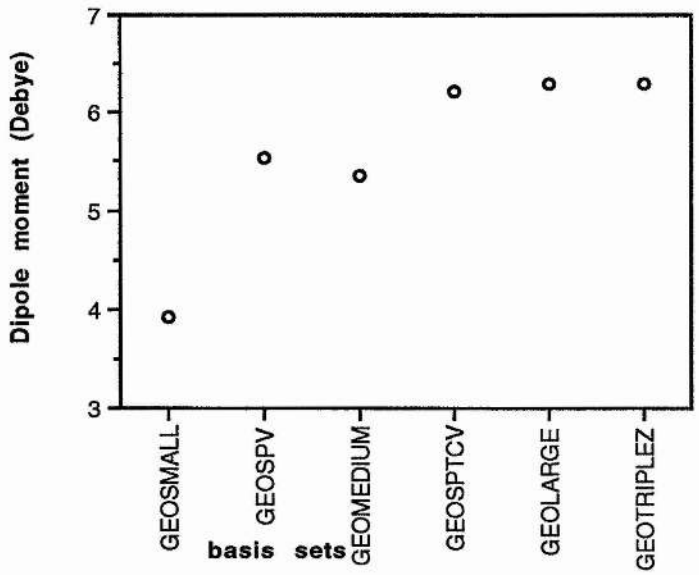


FIG. 9.17: Dipole moment of pyran-2-one obtained with geometrical basis sets.



9.2: Minimal basis sets

As mentioned in the introductory chapter, the main goal of this study was to identify small basis sets able to produce accurate MEP for large molecules of biological interest. Of the basis sets studied in the previous section, those of the series STO-NG, MINI-N and, the GEOSMALL are minimal basis set, all having the same number of basis functions (i.e. 39 for the pyran-2-one). In this section, the total energy, the electrostatic potential, the charges derived from the electrostatic potential and the dipole moments predicted by these basis sets are compared to those obtained using the 6-31G** basis set, because it gives the lowest energy and it is known to give properties comparable with the experiment.

The total energy (in atomic units) of pyran-2-one as a function of the basis sets is shown in Fig. 9.18. GEOSMALL gives the lowest energy of all the basis sets studied here, it is 0.9566 a.u. higher than the energy obtained with 6-31G** but it requires the

computation of only 1.5% two-electrons integrals. The number of two electron integrals to be calculated with GEOSMALL are 163843, those with 6-31G** are 11388332.

FIG. 9.18: Total energy of pyran-2-one with minimal basis sets.

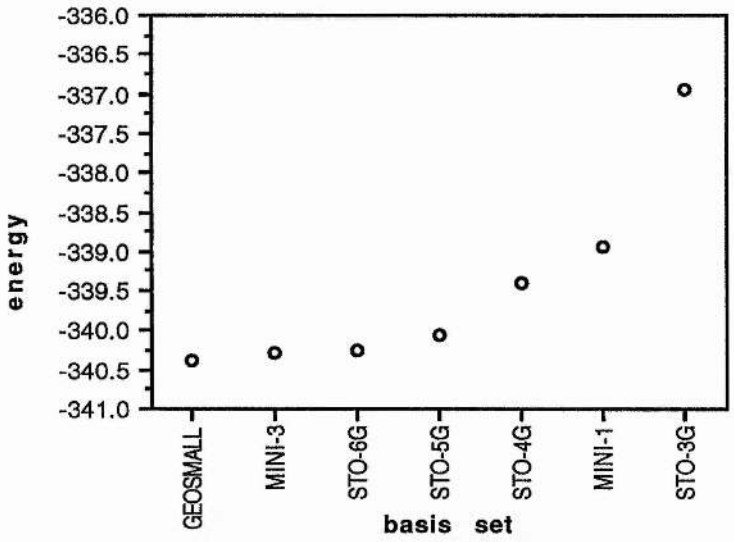
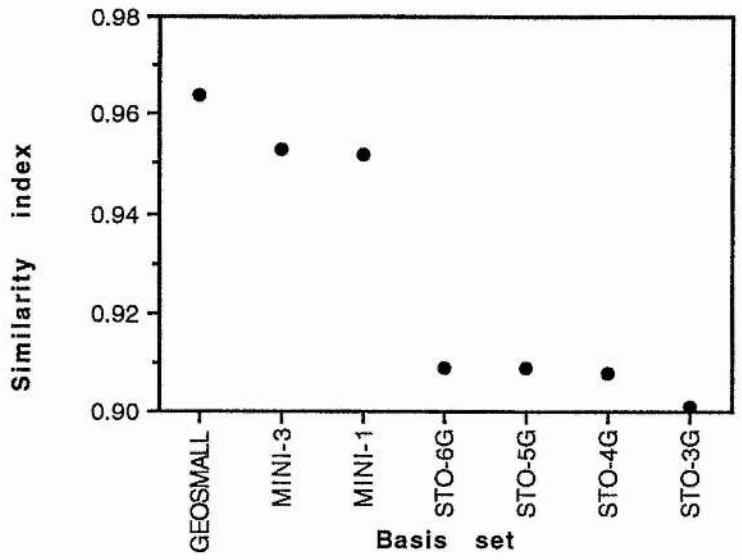


Fig. 9.19 shows the similarity indexes of the EPM as function of the basis set. Again, GEOSMALL is the basis set which best reproduces the EPM map obtained with 6-31G** basis set, its similarity index is 0.964, the similarity index with MINI-i basis sets is only slightly lower, at 0.953. The MEP obtained with STO-3G on the other hand is the most different, the similarity index being 0.901.

FIG 9.19: Hodgkin similarity indexes for the EPMs obtained using minimal basis sets.



In figs 9.20 - 9.22 the charges derived from the 6-31G** MEP are compared with those derived with the minimal basis sets. All the STO-NG (N=3,4,5,6) basis sets give the same electrostatic potential fitted atomic charges, no matter the value of N (Fig. 9.20); the same can be said for the MINI-i (i=1,3) basis sets (Fig. 9.21). From these plots, both MINI-i and GEOSMALL basis sets emerge as suitable minimal basis sets for the calculation of the MEP and the charges derived from them. As can be seen in fig 9.21 and 9.22 the differences in the charges derived from the potential with GEOSMALL and MINI-i basis sets are very similar to those obtained with 6-31G** whereas STO-NG basis sets tend to underestimate both negative and positive charges (Fig. 9.20).

FIG 9.20: Comparison between 6-31G** and STO-3G potential derived charges

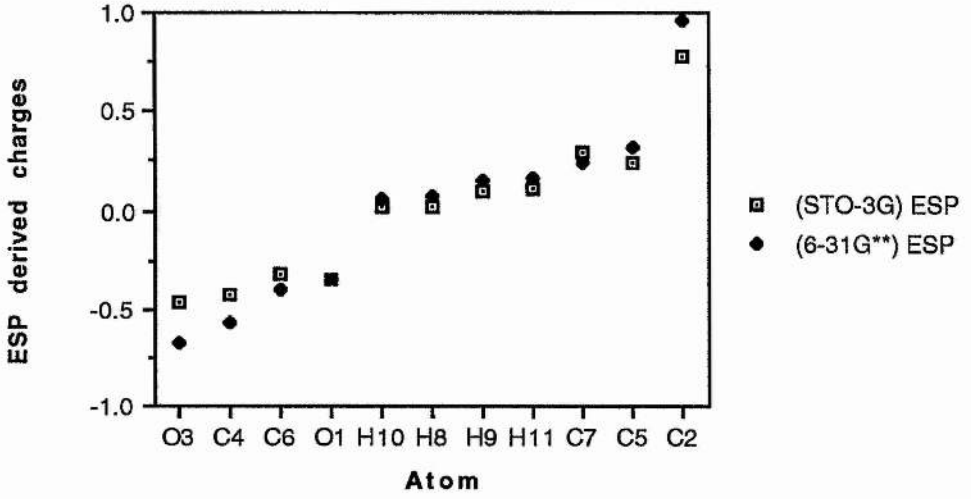
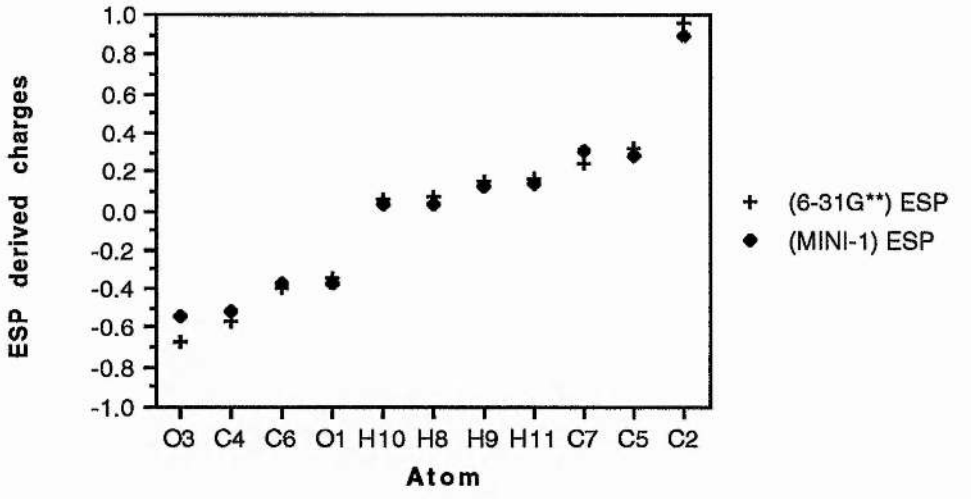


FIG 9.21: Comparison between 6-31G** and MINI-N potential derived charges



In fig. 9.23 are reported the dipole moments of pyran-2-one with the minimal basis sets, and the value obtained with 6-31G** is also reported for comparison. As can be seen, all of these basis sets underestimate the dipole moment, however, the largest differences are observed with the STO-NG basis sets.

FIG 9.22: Comparison between 6-31G** and GEOSMALL potential derived charges

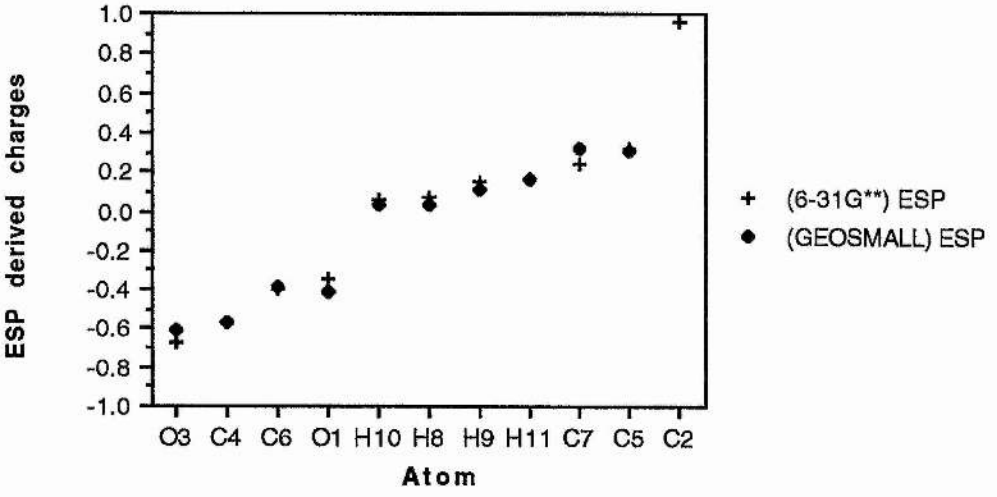
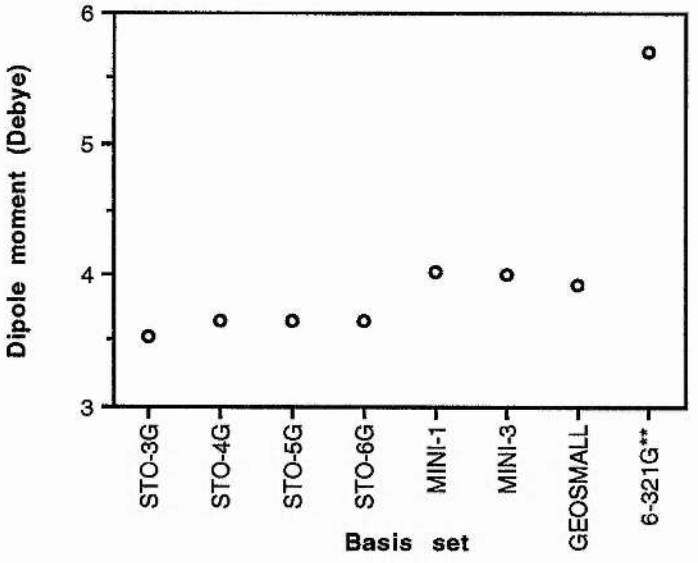


FIG 9.23: Dipole moments of pyran-2-one with minimal basis sets



9.3: Discussion and Conclusions

The electrostatic potential is a molecular property that can be associated with the chemical reactivity of the molecule towards electrophiles. Its use is of particular interest in

quantum pharmacology because the large size of the molecules studied do not generally allow the calculation of supermolecules. An accurate calculation of the MEP is therefore particularly interesting in order to obtain information about the reactivity of the molecules that would be otherwise very difficult to obtain for most of the molecules of biological interest. However, the calculation of the MEP for large molecules can be quite expensive in terms of computer resources and it is very often the case in which it can be obtained only with small basis sets or it can be obtained by using approximate expressions such as, the point charge approximation. If the information obtained from the MEP calculated with a small basis set can be misleading [130] then it appears that for the molecules of real biological interest, the usefulness of the MEP as an index of reactivity cannot be used because for these large molecules the MEP can only be evaluated with small basis sets. From a literature search however, it emerged that the number of basis sets generally used to calculate the MEP is limited to Pople's basis sets [130-133] and the STO-3G basis set seems to be the only minimal basis set used. Huzinaga and Clementi have introduced the MINI-i [54] and GEOSMALL [57] minimal basis sets respectively and they are rarely used for MEP calculations.

From this study of the MEP with a large selection of basis sets is emerged that the accuracy of the MEP is not necessarily increased by the size of the basis set used. For pyran-2-one, for example, 3-21G gives a better MEP than 6-31G and, the MEP obtained with 4-31G* and that with 6-31G** do not differ at all although the latter basis set is considerably larger than the former one (fig. 9.2). If the computer resources available are sufficiently powerful for the use of polarization functions, then a MEP similar to that obtained with 6-31G** can be obtained using the 4-31G* basis set; using the latter basis set gives rise to a saving in the number of two-electron integrals of about 25%. If more limited computer resources are available then the MIDI-1 basis set can be used to produce a MEP with a similarity index equal to 0.991 having to calculate only 12.5 % of the two-electron integrals. If the MEP can only be calculated with minimal basis set then the choice has to be GEOSMALL basis set, its similarity index is 0.953 but it requires the computation of only 1.5% two-electrons integrals compared to 6-31G**. Within the

minimal basis sets GEOSMALL and MINI-i give also better energies and better properties than STO-NG and their use is recommended when properties of large organic molecules are of interest.

From the comparison between charges derived from the MEP and Mulliken charges it is clear that the use of the later charges for the calculation of the MEP with the point charge approximation is not advisable.

References

1) a) La Fond R. R.; Cancer: the Outlaw Cell; American Chemical Society; Washington; (1988). b) De Vita V. T., Hellman S. J., Rosenberg S. A.; Eds.; Cancer: Principles and Practice of Oncology; 2nd ed.; Lippincott, Philadelphia; (1985). c) Symington T.; Carter R. L.; Scientific Foundations of Oncology; William Heinemann Medical books Ltd; London; (1976).

2) Pinedo H. M.; Ed; Cancer Chemotherapy 1979; Excerpta Medica Amsterdam, Oxford; (1979).

3) Venditti J. M.,Wesley R. A., Plowman J.; 'Current NCI preclinical Anti-tumour screening in vivo. Results of tumour panel screening, 1976 1982, and future directions'.; In Advances in Pharmacology and Chemotherapy; Garathini S., Goldin A., Hawking F.; Eds; Academic Press: Orlando, Florida.

4) Comprehnsiv emedicinal chemistry

5) Richards W. G.; Quantum Pharmacology; Butterworths; (1983); 2nd Edition

6) Briet P., Berthelon J.J., Collonges F.; Eur. Pat. Appl. EP 80943 AL; Nov. 24; p34; (1982).

7) Plowman J., Narayanan V. L., Dykes D., Szarvasi E., Briet P., Yoder O. C., Paull K. D.; Cancer Treat. Rep., **70**, 5, (1986).

8) Corbett T. H., Bissery M. C., Wozniak A., Plowman J., Polin L., Tapazoglou E., Dieckman J., Valeriote F.; Inv. New Drugs, **4**, 207, (1986).

9) Capolongo L. S., Balconi G., Ubezio P., Giavazzi R., Taraboletti G., Regonesi A., Yoder O. C., D'Incalci M.; Eur. J. Cancer Clin. Oncol; **23**, 1529, (1987).

10) Finlay G. J., Smith P., Fray L. M., Baguley B. C.; J. Nat. Cancer Inst.; **80**, 24, (1988).

11) Wiltrout R. H., Hornung R.L.; J. Nat. Cancer Inst.; **80**, 21, (1988).

12) Ching L., Baguley B. C.; Eur. J. Cancer Clin. Oncol.; **25**; 10; 1513; (1989)

13) Bibby M. C., Phillips R. M., Double J. A., Pratesi G.; Br. J. Cancer; **63**, 57, (1991)

14) Pratesi G., Rodolfo M., Rovetta G., Parmiani G.; Eur. J. Cancer; **26**; 10; 1079; (1990).

15) Ching L. M., Baguley B. C.; Eur. J. Cancer Clin. Oncol.; **24**, 1521, (1988).

16) Bibby M. C., Double J. A., Loadman P. M., Duke C. V.; J. Natl. Cancer Inst.; **81**; 216; (1989).

17) Mahadevan V., Malik S. T. A., Meager A., Fiers W., Lewis G. P., Hart I. R.; Cancer Research; **50**; 5537; (1990).

18) Principles of Biochemistry; White A., Handler P., Smith E. L., Hill R. L., Lehman R.; McGraw Hill; 6th Edition.

19) Atassi G., Briet P., Berthelon J., Collongenes F.; Eur. J. Med. Chem - Chim. Ther.; **20**; 5; 393; (1985).

20) Hill S., Williams K. B., Denekamp J.; Eur. J. Cancer Clin. Oncol.; **25**; 1419; (1989).

21) Weiss R.B., Greene R. F., Knight R. D., Collins J. M., Pelosi J. J., Sulkes A., Curt G. A.; Cancer Research; **48**; 5878; (1988).

22) Kerr D. J., Maughan T., Newlands E., Rustin G., Bleehen N. M., Lewis C., Kaye S. B.; Br. J. Cancer; **60**; 104; (1989).

23) Cassidy J., Kerr D. J., Setanoians A., Zaharko D. S., Kaye S. B.; Cancer Chemother. Pharmacol.; **23**; 397; (1989).

24) Bibby M. C.; Br. J. Cancer; **63**; 3; (1991).

25) Kerr D. J., Graham J., Cassidy J., Harding M., Setanoians A., McGrath J. C., Vezin W. R., Cunningham D., Forrest G., Soukop M.; Cancer Research; **46**; 3142; (1986).

26) Thomson C., Higgins D.; Int. J. Quantum Chem.: QC; **22**; 697; (1988).

27) Sedda P., Thomson C.; in 'Chemical Carcinogenesis 2; Columbano A. et al.; Eds; Plenum Press; New York; (1991).

28) Havlin K. A., Kuhn J. G., Craig J. B., Boldt D. H., Weiss G. R., Koeller J., Harman G., Schwartz R., Clark G. N., Von Hoff D. D.; J. Natl. Cancer Inst.; **83**; 124; (1991).

29) Atwell G. J., Rewcastle G. W., Baguley B. C., Denny W. A.; Anti-Cancer Drug Design; **4**; 161; (1989).

30) Rewcastle G. W., Atwell G. J., Baguley B. C., Calveley S. B., Denny W. A; J. Med. Chem.; **32**; 793; (1989).

31) Atwell G. J., Rewcastle G. W., Baguley B. C., Denny W. A.; J. Med. Chem.; **33**; 1375; (1990).

32) Rewcastle G. W., Atwell G. J., Zuan L., Baguley B. C., Denny W. A; J. Med. Chem.; **34**; 217; (1991).

33) Rewcastle G. W., Atwell G. J., Palmer B. D., Boyd D. W., Baguley B. C., Denny W. A; J. Med. Chem.; **34**; 491; (1991).

34) Ching L., Joseph W. R., Zhuang L., Atwell G. J., Rewcasle G. W., Denny W. A., Baguley B. C.; Eur. J. Cancer; **27**; 1; 79; (1991).

35) Zwi L. J., Baguley B. C., Gavin J. B., Wilson W. R.; Br. J. Cancer; **62**; 932; (1990).

36) (a) Szabo A. and Ostlund N.S.; Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory; Macmillan; New York; (1982). (b) Lowe J. P.; Quantum Chemistry; Academic Press, New York; (1978). (c) Daudel R., Leroy G., Peeters D., Sana M.; Quantum Chemistry; John Wiley, New York, (1983).

37) Schrodinger E.; Phys. Rev., **28**, 6, (1926)

38) Dirac P. A. M..; The Principles of Quantum Mechanics; Oxford University press, London, (1958); 4th Edition.

39) Born M., Oppenheimer R.; Ann. Phys.; **84**, 457, (1927).

40) Pauli W; Z. Physik; **31**; 765; (1925)

41) (a) Mulliken R. S.; Phys. Rev.; **32**, 186, (1928).(b) **32**, 761, (1928). (c) **33**, 730, (1930).

42) Clementi E, Davis D. R.; J Comp. Physics; **2**, 245, (1967).

43) Hartree D.R.; Proc. Cambridge Phil. Soc.; **24**, 89, (1928)

44) Slater J. C.; Phys. Rev.; **34**, 1293, (1929).

45) Roothaan C. C. J.; Rev. Mod. Phys.; **23**, 69, (1951).

46) Slater J. C.; Phys. Rev.; **36**, 57, (1930).

47) Boys S. F.; Proc. R. Soc. London Ser. A.; **200**, 542, (1950).

48) Davidson E. R., Feller D.; Chem. Rev.; **86**, 681, (1986).

49) Hehre W. J., Stewart R. F., Pople J. A.; J. Chem. Phys.; **51**, 2657, (1969).

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

51) Hehre W. J., Ditchfield R., Pople J. A.; J. Chem. Phys.; **56**, 2257, (1972).

52) Hariharan P. C., Pople J. A.; Theor. Chim. Acta; **28**, 213, (1973).

53) Binkley J. S., Pople J. A., Hehre W. J.; J. Am. Chem. Soc.; **102**, 939, (1980).

54) Tatewaki H., Huzinaga S.; J. Comp. Chem.; **1**, 205, (1980).

55) Huzinaga S.; Gaussian basis sets for molecular calculations; Elsevier, Amsterdam - Oxford- New York- Tokyo; (1984).

56) Dunning T. H.; J. Chem. Phys.; **53**, 2823, (1970).

57) Clementi E., Corongiu G.; IBM Tech. Rep.; POK-11, (April 27, 1982).

58) Simons J., Jordan K. D.; Chem. Rev.; **87**, 535, (1987).

59) Clark T.; Chandrasekhar J., Spitznagel G. W., Schleyer P. v. R.; J. Comp. Chem.; **4**, 294, (1983).

60) Almlof J., Faegri K., Korsell K.; J. Comp. Chem.; **3**, 385, (1982).

61) Miller W. H., Schaefer H. F., Berne B. J., Segal G. A.; Modern Theoretical Chemistry, Vol. 7; Plenum Press New York; (1977).

62) Ransil B. J.; Rev. Mod. Phys.; **32**, 239, 245, (1960).

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

64) Dewar M. J. S.; The Molecular Orbital Theory of Organic Molecules; McGraw-Hill, New York; (1969).

65) Dewar M. J. S., Zoebisch E. G., Healy E. F., Stewart J. J. P.; J. Am. Chem. Soc.; **107**, 3902, (1985).

66) Stewart J. J. P.; J. Am. Chem. Soc.; **111**, 209, (1989).

67) Dewar M. J. S., Thiel W.; J. Am. Chem. Soc.; **99**, 4899, (1977).

68) Pople J. A., Santry D. P., Segal G. A.; J. Chem. Phys.; **43**, 10, S129, (1965).

69) Stewart J. J. P.; J. Comp. Chem.; **10**, 221, (1989).

70) Thomson C., Scano P.; J. Comp. Chem.; **2**; 12; (1991).

71) (a) Dewar M.J.S., Healy E. F., Holder A. J., Yuan Y.C.; J. Comp. Chem.; **11**, 541, (1990). (b) Stewart J. J. P.; J Comp. Chem.; **11**, 543, (1990).

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

73) Bonnaccorsi R., Scrocco E., Tomasi J., Pullman A.; Theoret. Chim. Acta; **36**, 339, (1975).

74) Weinstein H., Osman R., Topiol S., Green J. P.; Ann. NY Acad.Sci.; **367**, 434, (1981).

75) Namboodiri K., Osman R., Weinstein H., Rabinowitz J. R.; Mol. Tox., **1**, 131, (1987).

76) Osman R., Namboodiri K., Weinstein H., Rabinowitz J. R.; J. Am. Chem. Soc.; **110**, 1701, (1988).

77) Petrongolo C., Tomasi J.; Int. J. Quantum Chem.: QB; **2**, 181, (1975).

78) Nakayama A., Richards W. G.; Quant. Struct.-Act. Relat.; **6**, 153, (1987).

79) Naray-Szabo G.; Int. J. Quantum Chem.: QB; **16**, 87, (1989).

80) Thomson C. Wilkie J.; Carcinogenesis, **10**, 531, (1989).

81) Murray J. S., Sukumar N., Ranganathan S., Politzer P.; Int. J. Quantum Chem.; **37**, 611, (1990).

82) Culberson J. C., Zerner M. C.; Chem. Phys. Lett.; **122**, 5, (1985).

83) Luque F. J., Orozco M.; Chem. Phys. Lett.; **168**, 3,4, (1990).

84) Williams D. E., Yan J. M.; Adv. Atomic Mol. Phys.; **23**, 87, (1988).

85) Scrocco E., Tomasi J.; Adv. Quantum Chem.; **11**, 115, (1978)

86) Carbo R., Leyda L., Arnau M.; Int. J. Quantum Chem.; **17**, 1185, (1980).

87) Hodgkin E. E., Richards W. G.; Int. J. Quantum Chem.: QB; **14**, 105, (1987).

88) Mulliken R. S.; J. Chem. Phys.; **23**, (a) 1833; (b) 1841; (1955)

89) Singh U. C., Kollman P. A.; J. Comp. Chem.; **5**, 129, (1984).

90) Besler B. H., Merz K. M., Kollman P. A.; J. Comp. Chem.; **11**, 431, (1990).

91) Orozco M., Luque F. J.; J. Comp. Chem.; **11**, 909, (1990).

92) Woods R. J., Khalil M., Pell W., Moffat S. H., Smith V. H.; J. Comp. Chem.; **11**, 297, (1990).

93) Fletcher R.; Practical Methods of Optimization; Vol. 1; John Wiley, New York (1985).

94) Schlegel H. B.; J. Comp. Chem.; **3**, 214, (1982).

95) Stewart J. J. P.; J. Comp.-Aided Mol. Des.; **4**, 1, (1990).

96) Gano J. E., Jacob E. J., Roesner R.; J. Comp. Chem.; **12**; 126; (1991).

97) Martin J. M., Francois J. P., Gijbels R.; J. Comp. Chem.; **12**; 52; (1991).

98) Fraser-Reid B., Wu Z., Andrews C.W., Skowronski E., Bowen J. P.; J. Am. Chem. Soc.; **113**; 1434; (1991).

99) Buemi G.; Theochem; **67**; 253; (1990).

100) Karelson M., Katritzky A. R., Zerner M. C., J. Org. Chem.; **56**; 134; (1991).

101) Broughton H. B., Woodward P. R.; Comput. -Aided Mol. Des.; **4**; 147; (1990).

102) Karaman R., Huang J. T. L., Fry J.; J. Comp. Chem., **11**; 1009; (1990).

103) Almlof J.; in : IMA, Vol 15: Mathematical frontiers in computational chemical physics; Donald G. Truhlar; Springer Verlag; (1988).

104) Stewart J. J. P.; MOPAC; Q.C.P.E. Bull; 3, 455; 43;. (1983).(The MOPAC version 5 was supplied by Dr Stewart).

105) Clark T.; A Handbook of Computational Chemistry; John Wiley and Sons; New York; (1985).

106) GAUSSIAN 90; Revision I; Frish M. J., Head-Gordon M., Trucks G. W., Foresman J. B., Schlegel H. B., Raghavachari K., Robb M., Binkley J. S.,

Gonzalez C., Defrees D. J., Fox D. J., Whiteside R. A., Seeger R., Melius C. F., Baker J., Martin R. L., Kahn L. R., Stewart J. J. P., Topiol S., and Pople J. A.; Gaussian; Inc; Pittsburgh PA; (1990).

107) Singh U. C., Kollman P.; QCPE Program N° 446; (1982).

108) Higgins D.; PhD Thesis; University of St. Andrews; (1988).

109) Connolly M. L.; Molecular Surface Program; QCPE Program 429; (1983).

110) ASP: Automatic Similarity Package; Oxford Molecular Terrapin House; South Parks Road; Oxford; Oxford OX1 3UB; U.K.

111) Charlton M. H..; PhD Thesis; University of St-Andrews (1991).

112) CHEMX: Developed and Distribuited by Chemical Design Ltd; Oxford

113) Allen F. H., Kennard O., Taylor R.; Cambridge Crystallographic Database; Acc. Chem. Res. **16**, 146; (1983)

114) Biswas S. C., Sen R. K.; Indian J Pure Appl Phys; **20**; 414; (1982)

115) Cody V.; in: 'Plant Flavanoids in Biology and Medicine II: Biochemecal, Cellular and Medicinal Properties'; Alan R. Liss, Inc; (1988).

116) Kingsbury C. A., Looker J. H.; J. Org. Chem.; **40**; 8; 1121; (1975).

117) McWeeny R.; Coulson's Valence; 3rd Edition; Oxford Science Publications; (1990).

118) Pople J. A.; J. Chem. Phys.; **37**; 53; (1962).

119) Pople J. A.; Mol. Phys.; **7**; 301; (1964).

120) Karplus M., Pople J. A.; J. Chem. Phys.; **38**; 2803; (1963).

121) Frieman P., Allen L. C.; Int. J. Quantum Chem.; **23**; 1053; (1986).

122) Norris C. L., Benson R. C., Beak P., Flygare W. H.; J. Am. Chem. Soc.; **95**, 2767, (1973).

123) Sedda P.; Thomson C.; unpublished data

124) Sampson R. J.; SURFACE II Graphics System; Kansas Geological Survey; Lawrence; KS; (1978).

125) Thomson C., Edge C.; J. Mol. Structure (Theochem); **121**; 173; (1985).

126) Bibby et al.; AICR Annual Report; (1990).

127) Dewar M. J. S., Y. Ching-Juan Inorg Chem; **29**, 3881, 1990

128) Bibby M. C., Double J. A; Personal communication.

129) Kirkpatrick D. L.; Drugs of Today; **26**; 2; 91; 1990.

130) Edwards W. D., Weinstein H.; Chem. Phys. Lett.; **56**; 3; 582; (1978).

131) Orozco M., Luque F. J.; Chem. Phys. Lett.; **160**; 3; 305; (1989).

132) Orozco M., Luque F. J.; Chem. Phys. Lett.; **168**; 3; 4; (1990)

133) Murray J. S., Sukumar N., Ranganathan S., Politzer P.; Int. J. Quantum Chem.; **37**, 271, (1990).

134) In "Comprehensive heterocylic chemistry"; Vol. 3; eds.Boulton A. J., Mc Killop A.; Pergamon Press.

135) Kerins M. C., Fitzpatrick N. J.; J. Mol. Structure (Theochem); **180**; 297; (1988).

136) Kurnig P. K.; Int. J. Quantum Chem.: QB; **14**; 47; (1987)

137) Barker J. D.; PhD Thesis; University of St-Andrews (1991)

138) Aitken R. A., Sharma S. K. ; AICR Annual Report (1990).