

**STUDIES TOWARDS THE
DEVELOPMENT
OF A
NOVEL ANTI-CANCER AGENT**

A thesis presented in part fulfilment of the requirements for the Degree of

Doctor of Philosophy

by

Jacqueline Coll

Department of Chemistry

The University

Glasgow

August 1993

ProQuest Number: 13832065

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 13832065

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

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

*Thesis
9757
copy 1*

GLASGOW
UNIVERSITY
LIBRARY

ACKNOWLEDGEMENTS

A big thank you to Dr. Bob Hill for his help and support. Thanks also to Graham and Mycology staff, especially Pearl.

Making the whole thing easier also thanks to Fran, Debs, Gail, Aileen, Janette, Steph and everyone else whose support I appreciated.

Thanks mum and dad for being there and being brilliant throughout everything.

Financial support from S.E.R.C. is acknowledged, as is C.R.C. funding of oncology and biochemical work.

Finally, thank you Duncan, for being understanding, patient and knowing just a wee bit about Chemistry. Lots of love.

To Mum and Dad.

SUMMARY

This project has been a collaborative effort involving the design, synthesis and testing of analogues of the fungal metabolites **viridin**, **demethoxyviridin** and **wortmannin**. Extraction methods of viridin, from *Trichoderma viride* were developed and initially provided a lead compound for testing. Data relating to the anti-inflammatory activity of wortmannin analogues was available in the literature and was used in a computer modelling program SYBYL to build a structure activity relationship from which activity of new analogues could be predicted. The crystal structure determination of demethoxyviridin provided detailed structural information about these molecules. This was used in modelling studies to guide the synthesis of new compounds, leading to the synthesis of an analogue several hundred times more active than the original lead compound. A patent has been applied for this compound and it is presently being forwarded for clinical trials in animals for toxicity testing. The compounds were shown to be cytotoxic *in vitro* and inhibit specific enzymes in the signal transduction pathway. Preferential inhibition of mitogen stimulated cells was shown, indicating that some degree of selectivity was built into the molecules, suggesting that future compounds may have relatively few side effects due to their novel mechanism of action.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
CHAPTER 2: POSSIBLE ENZYME INHIBITORS	25
CHAPTER 3: COMPUTER MODELLING	54
CHAPTER 4 : CRYSTALLOGRAPHY	120
CHAPTER 5 : ORGANIC SYNTHESIS	131
CHAPTER 6: DISCUSSION	150
APPENDICES	157
REFERENCES	172

ABBREVIATIONS

- ADA - adenosine deaminase
- AI - anti-inflammatory
- AMP - adenosine monophosphate
- ATP - adenosine triphosphate
- CoMFA - Comparative Molecular Field Analysis
- DAG - diacylglycerol
- EGF - epidermal growth factor
- g-csf - granulocyte colony stimulating factor
- GnRH - gonadotrophin releasing hormone
- HL-60 - human leukaemic cells
- HPLC - high performance liquid chromatography
- IP₃ - inositol-1,4,5-triphosphate
- IR - infra-red
- Leu - leucine
- LH - luteinising hormone
- LHRH - luteinising hormone releasing hormone
- Met - methionine
- MMS - Molecular Modelling System
- MR - multiple regression
- MS - molecular shape
- NMR - Nuclear magnetic resonance
- PA - phosphatidic acid
- PAP - phosphatidic acid phosphohydrolase
- PC - phosphatidylcholine
- PDGF - platelet derived growth factor
- Phe - phenylalanine
- PIP₂ - phosphatidyl inositol-4,5-biphosphate

ABBREVIATIONS (cont.)

PLC - phospholipase C, phosphoinositidase

PLD - phospholipase D

PLS - partial least squares

PRESS - predictive sum of squares

QSAR - quantitative structure activity relationship

TK - tyrosine kinase

TLC - thin layer chromatography

TPA - tetraphorbol acetate

UV - ultra-violet

CHAPTER 1
INTRODUCTION

CANCER - PRESENT STAGE

Cancer is now one of the leading causes of death world-wide¹. The emerging role for cancer as a cause of mortality is due not so much to continuing increases in cancer mortality but to the remarkable and consistent decline in heart disease mortality over the past four decades, due perhaps to a combination of a reduced prevalence of major risk factors (altered lifestyle, better detection and treatment of hypertension, etc.) and improved treatment of the various clinical manifestations of heart disease. As such, heart disease can serve as a model for the impact of a combination of altered lifestyle and improved treatment on mortality, which will hopefully soon be applied to cancer.

Since 1973 increases in mortality of greater than 15% have occurred for lung cancer, melanoma, non-Hodgkin's lymphoma, and multiple myeloma. Increases in cigarette smoking from 1900 until the early 1960s transformed lung cancer from a rare disease at the turn of the century to the current leading cause of cancer death. The annual mortality rate of lung cancer in men has finally levelled off after more than 50 years of unabated increase. In women, lung cancer surpassed breast cancer as the leading cause of cancer death in 1986, these rates are expected to continue to increase for at least another 10 years. The increase in melanoma mortality parallels a larger increase in the incidence of this disease, mainly caused by increased sunburning in fair skinned populations, attributable to changing fashions and recreational habits and also the recent much publicised depletion of the ozone layer.

The increase in mortality from, and the increased incidence of, non-Hodgkin's lymphoma and multiple myeloma remain largely unexplained. Improved diagnostic procedures may be part of the reason for this increase in lymphoproliferative diseases. Immunocytochemistry can distinguish genuine

lymphomas from undifferentiated epithelial tumours that have metastasised to lymph nodes from unknown primary sites, and distinguish more definitely extra nodal primary lymphomas from cancers of other histologies that occur at the same site. In younger age groups, however, much of the recent increased incidence of certain types of lymphomas is due to the increased prevalence of human immunodeficiency virus (HIV) infection, which is an established cause of these cancers.

WHAT CAUSES CANCER?

Most recognised cancer causing substances have either direct genotoxic effects on DNA (*e.g.* radiation) or increase cell proliferation (*e.g.* hormones) or a combination of both (*e.g.* tobacco)².

Increased cell proliferation appears to be the most important. Molecular genetics has provided evidence that cell division is essential for the genesis of human cancer and that an increased rate of cell division will increase cancer risk. 'Increased' may imply mitotic activity above the baseline rate or a division of a subset of cells that would ordinarily not be dividing. Cell division increases the risk of genetic errors of various kinds. Adducts or other single-stranded DNA damage may be converted to gaps or mutations, and mitotic recombination may result in more profound changes than those from a single mutation. The development of a fully malignant tumour appears to involve multiple stages - the activation or altered expression of proto-oncogenes and the loss or inactivation of tumour suppresser genes, which control normal cellular activity. The activation of proto-oncogenes, whether by mutation, translocation, or amplification, requires cell division. Inactivation of a tumour suppresser gene appears to involve first a mutational event that inactivates one allele, followed by a deletion during mitosis that

results in homozygosity. Both the fixation of the initial mutagenic event and the loss of the wild-type allele of the tumour suppresser (e.g. a hormone) can act at all stages in the pathogenesis of a malignant phenotype. Removal of the causative agent (except very late) can prevent or delay full development of the cancer.

Chemical and physical carcinogens leave traces of their activities on DNA because of the specific patterns of base changes they induce. Knowledge of the mutation patterns in genes believed to be involved in human cancer might be able to be used to predict the likelihood that an exogenous DNA damaging agent may be involved. It may ultimately be possible to predict what the agent might be. For example, liver mutations typically are found at one nucleotide pair of codon 249 of the tumour suppresser gene p53 in liver cancers occurring in individuals who live in geographic areas where exposure to both aflatoxin and hepatitis B (a probable and an established cause of liver cancer, respectively) is common.

CURRENT TREATMENT

Despite a vast array of drugs now available to treat cancer, one main drawback exists in a lack of selectivity, i.e. toxic side effects. Most drugs interfere with the role of nucleic acids, blocking cell division. The synthetic phase of the cell cycle is affected thus cell populations with a high mitotic index are most susceptible, e.g. rapidly dividing malignant cells, bone marrow, gonads, gastro-intestinal tract lining, hair follicles etc. The primary aim of chemotherapy is to prolong life, but treatment should be as free as possible from pain and distressful side effects, thus optimum chemotherapy should encompass a drug, or a range of mutually beneficial drugs which are potent, so that doses can be low, and free from toxicity, i.e. as selective as possible. The

particular type of cancer would also dictate the chemotherapeutic approach, e.g. lung cancer, the commonest carcinoma, is relatively slow growing but acute leukaemia cells are rapidly dividing.

Two main classes of product are used for the chemotherapy of neoplastic disorders, the **cytotoxic drugs** and the **sex hormones**³.

CYTOTOXIC DRUGS

Cytotoxic drugs can be divided into 5 main groups;

1. Alkylating drugs
2. Cytotoxic metabolites
3. Antimetabolites
4. Vinca alkaloids and etoposide
5. 'Others'

Malignant disease is treated by surgery, radiotherapy, and/or chemotherapy. Certain tumours are highly sensitive to chemotherapy but many are not, and inappropriate drug administration in these circumstances can only increase morbidity.

All of these drugs are designed to stop cell division in cancer cells. However, normal cells, particularly those with a rapid rate of mitosis, can also be affected. Toxic effects commonly observed are severe nausea and vomiting, alopecia and bone marrow suppression. Myelosuppression necessitates checking blood counts prior to each stage of treatment, with the dosage modified or treatment being delayed accordingly.

1. The alkylating drugs are among the most extensively used in cancer chemotherapy. They act by damaging the DNA, thus interfering with cell replication. They are not entirely selective for the cancer cells thus have many side effects. These include **extravasation**, *i.e.* leakage of the drug into the extravascular compartment causing severe tissue necrosis, **hypercalcaemia**, **hyperuricaemia** which can result in uric acid crystal formation in the urinary tract and associated renal dysfunction, **nausea and vomiting**, **bone marrow suppression**, **alopecia** and interference in **reproductive function**. Although many of these side effects are associated with most of the cytotoxic drugs, there are two main problems associated with long term usage of alkylating drugs. Firstly, gametogenesis is often severely affected and secondly, prolonged use of these drugs, particularly when combined with extensive radiation, is associated with a marked increase in the incidence of certain types of cancer, *i.e.* of acute non-lymphocytic leukaemia. These drugs include cyclophosphamide, chlorambucil and busulphan.

2. Cytotoxic antibiotics are also widely used. Many act as radiomimetics so simultaneous use of radiotherapy is avoided as it can result in markedly enhanced tissue toxicity. **Doxorubicin** is one of the most successful drugs, and is used to treat the acute leukaemias, lymphomas and a variety of solid tumours. It can be **cardiotoxic** at high doses so care is needed when treating patients with previous history of heart disease. Epirubicin is structurally related to doxorubicin and is thought to be slightly less cardiotoxic. Bleomycin is used to treat the lymphomas, certain solid tumours and malignant effusions. **Dermatological toxicity** is common, as is **mucositis** and an association with **Raynaud's** phenomenon - an attack of marked vasoconstriction of the extremities, affecting the hands and feet giving rise to tingling, pain and sometimes ischaemic damage. Hypersensitivity reactions are shown by chills and fevers commonly occurring a few hours after drug

administration, but which may be prevented by concurrent administration of a steroid, e.g. hydrocortisone.

3. Antimetabolites are incorporated into new nuclear material or combine irreversibly with vital cellular enzymes, preventing normal cell division. **Methotrexate** inhibits the enzyme dihydrofolate reductase, essential for the synthesis of purines and pyrimidines. It can cause **myelodepression**, **mucositis** and rarely **pneumonitis**. It is contra-indicated if significant renal impairment is present as the kidney is the route of excretion. Folinic acid ('rescue') is given to speed recovery from methotrexate toxicity. Other antimetabolites include fluorouracil, and mercaptopurine.

4. Vinca alkaloids and etoposide interfere with microtubule assembly, causing metaphase arrest. All vinca alkaloids have similar activity but vary in the site of toxicity. Vincristine and vinblastine are commonly associated with **peripheral and autonomic neuropathy** shown by peripheral paraesthesia and loss of deep tendon reflexes. Recovery of the nervous system in these cases is generally slow but complete. Etoposide is used to treat small cell carcinoma of the bronchus, the lymphomas and testicular teratoma.

5. 'Others' include carboplatin, cisplatin and procarbazine. Carboplatin is a derivative of cisplatin with similar activity although it appears to be better tolerated in some ways, *i.e.* less nausea, vomiting, nephrotoxicity and neurotoxicity although it is more myelosuppressive and for this reason is not used more frequently than 4 week intervals. Cisplatin is an alkylating type drug but is extremely toxic. Procarbazine is a first line treatment in Hodgkin's disease. Toxic effects include rash. It is also a mild monoamine oxidase inhibitor.

SEX HORMONES

The **sex hormones** and antagonists are mainly used for the treatment of prostatic and breast cancer. Hormonal treatment of advanced prostatic cancer is based on the concept that hormone-dependent neoplasms will regress when deprived of hormonal stimulation. In prostatic carcinoma, which in general is androgen dependent, **oestrogens** are often effective but their use is limited by cerebral and cardiovascular side effects. The anti-androgen **cyproterone acetate** blocks adrenal and testicular androgens which provide the androgenic stimulus to the prostate. This is a more rational form of treatment for prostatic cancer and has produced responses in patients who have relapsed after oestrogen therapy and orchidectomy. Gynaecomastia is less of a problem and the cardiovascular risk is decreased. **Flutamide** is also an anti-androgen which acts by inhibiting androgen uptake and/or by inhibiting nuclear binding of androgens in target tissue. In advanced prostatic cancer, flutamide produces a response which is comparable to that achieved with oestrogen therapy and when used in combination with a LHRH (luteinising hormone releasing hormone) agonist, survival rate is prolonged. GnRH (gonadotrophin releasing hormone) analogues such as buserelin, goserelin and leuprorelin are also effective and are also without the limiting side effects of the oestrogens. They produce a down-regulation of the receptors in the pituitary, preventing LH (luteinising hormone) release therefore decreasing serum testosterone levels. Anti-androgens, *e.g.* **tamoxifen** have largely replaced oestrogens for the treatment of advanced breast cancer especially in post menopausal women. Progestogens are second or third line treatment in breast cancer, and are also used to treat endometrial carcinoma.

TOXICITY - RECENT ADVANCES

Selectivity for rapidly dividing cells results in many side effects, one of the most commonly recognised by the general public is perhaps loss of hair, alopecia, as a result of direct action at the hair follicles. This is often temporary. Some of the more serious side effects include myelodepression, i.e. damage to the blood forming elements of the bone marrow. Sterility and teratogenicity can also arise as a result of permanent damage to the gonads.

Many clinical supportive measures are used to combat such serious side effects. These range from simple measures such as placing ice packs on the patients' heads, to more sophisticated measures such as bone marrow transplant or use of haematopoietic growth factors.

Since 1987, a variety of these growth factors have been used to stimulate white blood cell growth within the bone marrow before or after treatment, *e.g.* the genetically engineered version of the natural colony stimulating factor 'Neupogen' (g-csf) which also particularly stimulates neutrophil count. If bone marrow rescue becomes possible in this way, treatment can become bolder in terms of frequency and dosage of drug regimens. Transplants can become either autologous i.e. the patient donates his own bone marrow which is stored while he undergoes treatment, then transplanted back into the patient, or it can be allogenic where the tissue is a close match, *e.g.* from the patient's sibling. This is to minimise the chance of rejection or 'host-versus-graft' reaction. Other supportive measures such as barrier nursing, which ranges from the use of barrier creams to laminar air flow bed and positive pressure ventilation rooms, can also contribute to a successful treatment.

One of the latest anti-cancer drugs to be released in this country is **Nipent** (containing pentostatin and deoxycoformycin). The launch was announced by Lederle Laboratories⁴ and its intended use is for first line treatment of hairy cell leukaemia. This is a rare, chronic leukaemia in which leukaemic cells have characteristic projections or 'hairs'. The disease accounts for about 2% of all leukaemias and mainly affects middle aged men.

Speaking at a Lederle meeting on April 15th 1993, Professor Daniel Catovsky (honorary consultant physician, Royal Marsden hospital, London) said that pentostatin was a potent inhibitor of **adenosine deaminase (ADA)**, an enzyme known to play a key role in lymphocyte maturation and function. Although the exact mechanism of action of the drug was unclear, it was possible that in addition to ADA inhibition, it exerted its effects through inhibiting RNA synthesis and increasing DNA damage.

Another current advance is the use of high dose cytotoxics with stem cell transplants recently reported⁵. 'Is more drug better?' is a controversial issue currently facing oncologists. Now that there are improved ways of reducing myelotoxicity, the usual dose limiting toxicity of anticancer drugs, doctors are able to give more intensive cytotoxic treatment, although so far there is no proof that high dose therapy improves survival as randomised trials have generally not yet been carried out.

High dose, intensive chemotherapy does, however, look very promising in certain types of cancer. The best results seen so far have been with diseases known to be responsive to chemotherapy, such as lymphoma, leukaemia, testicular cancer and Hodgkin's disease. This type of chemotherapy has become possible with the advent of improved supportive techniques -

antibiotics, haematopoietic growth factors, bone marrow transplantation and, more recently, peripheral stem cell transplantation - all of whom have been designed to reduce myelotoxicity.

Peripheral blood stem cell transplantation is an alternative to bone marrow transplantation for 'rescuing' patients after bone marrow has been destroyed by high dose chemotherapy. The technique, which is still being evaluated, is known as 'off the peg' bone marrow transplantation. Blood is removed before the high dose chemotherapy and the stem cells are separated out, stored then re-infused after treatment to repopulate the bone marrow. The normal low yield of stem cells in the circulation is boosted by priming with mild chemotherapy and haematopoietic growth factors, which stimulate migration of stem cells into the blood. After reinfusion, growth factors are generally given again, to boost recovery.

These recent clinical advances are an invaluable advantage for the patient, but the problem of developing selectively toxic anti-cancer agents remains.

DRUG DESIGN

The past century has witnessed the virtual realisation of one of man's ancient dreams - the conquest of disease through the use of effective drugs. Much of the drama of these discoveries arose from the opportunistic exploitation of unexpected observations, that combination of chance and sagacity known as serendipity. Recently, however, major new drugs have been developed through the application of scientific thought.

In 1940 Charles Huggins⁶ was responsible for the first ever demonstration of a synthetic drug having induced an unquestionable improvement in a malignant disease. As a professor of surgery at the University of Chicago, Huggins had been collecting prostatic fluid from dogs for several weeks as part of an enquiry into the effect of testosterone injections on prostatic fluid production. It was during one such study that he noticed that castration consistently caused regression of tumours which had spontaneously arisen in elderly dogs. Realising these tumours were hormone dependent, he then administered stilboestrol, a synthetic hormone with high oestrogenic potency, to produce 'chemical' castration. His success encouraged him to conduct a successful clinical trial in patients with prostatic cancer which had spread into their bones.

These pioneering investigations were later confirmed by others and Charles Huggins was awarded the Nobel prize in 1966 for physiology and medicine.

More recent advances have been made using a more rational approach to drug design. The work described in this thesis attempts to adopt this approach.

INHIBITION OF SIGNAL TRANSDUCTION ENZYMES

The aim of this project was to design, synthesise and test various structural analogues of the fungal metabolites which have been shown to have phospholipase D inhibitory capacity. This enzyme is critical to signal transduction in cells and its activation is a key step in the chain of events leading to cell division. The project presents a novel mechanism of action of a potential anti-neoplastic agent in the form of a signal transduction enzyme inhibitor and represents a rational approach to drug design.

The majority of currently available cytotoxic agents interact with DNA as their principal molecular source of action. Their general toxicity, as well as a feeling that clinical activity of these agents has reached a plateau has impelled cancer pharmacologists to consider novel targets for anti-cancer drug development^{7,8}.

The stimulation of cell proliferation by growth factors involves a series of biochemical events which are regulated by the generation of intracellular messengers in response to the binding of the mitogen to the cell surface receptor. The signals implicated in the onset of mitogenesis are tyrosine kinase activity and changes in the levels of the second messengers cyclic AMP, DAG and inositol triphosphate. Transformation of cells by a number of oncogenes has been shown to subvert the regulation of pathways controlling second messenger concentration. Thus it is critical to the inhibition of the growth of both normal and a range of transformed cell types to identify key sites of regulation of these pathways which might serve as targets for therapeutic intervention⁹.

THE PHOSPHOINOSITIDE CASCADE

The phosphoinositide cascade evokes a wide variety of responses in many kinds of cells¹⁰. Like the adenylate cyclase cascade, it converts extracellular signals into intracellular ones. The intracellular messengers formed by activation of this pathway arise from **phosphatidyl inositol 4,5-biphosphate (PIP₂)**, a phospholipid in the plasma membrane. PIP₂ is formed by the phosphorylation of phosphatidyl inositol (PI), a membrane constituent synthesised from CDP-glycerol and inositol (see Fig. 1).

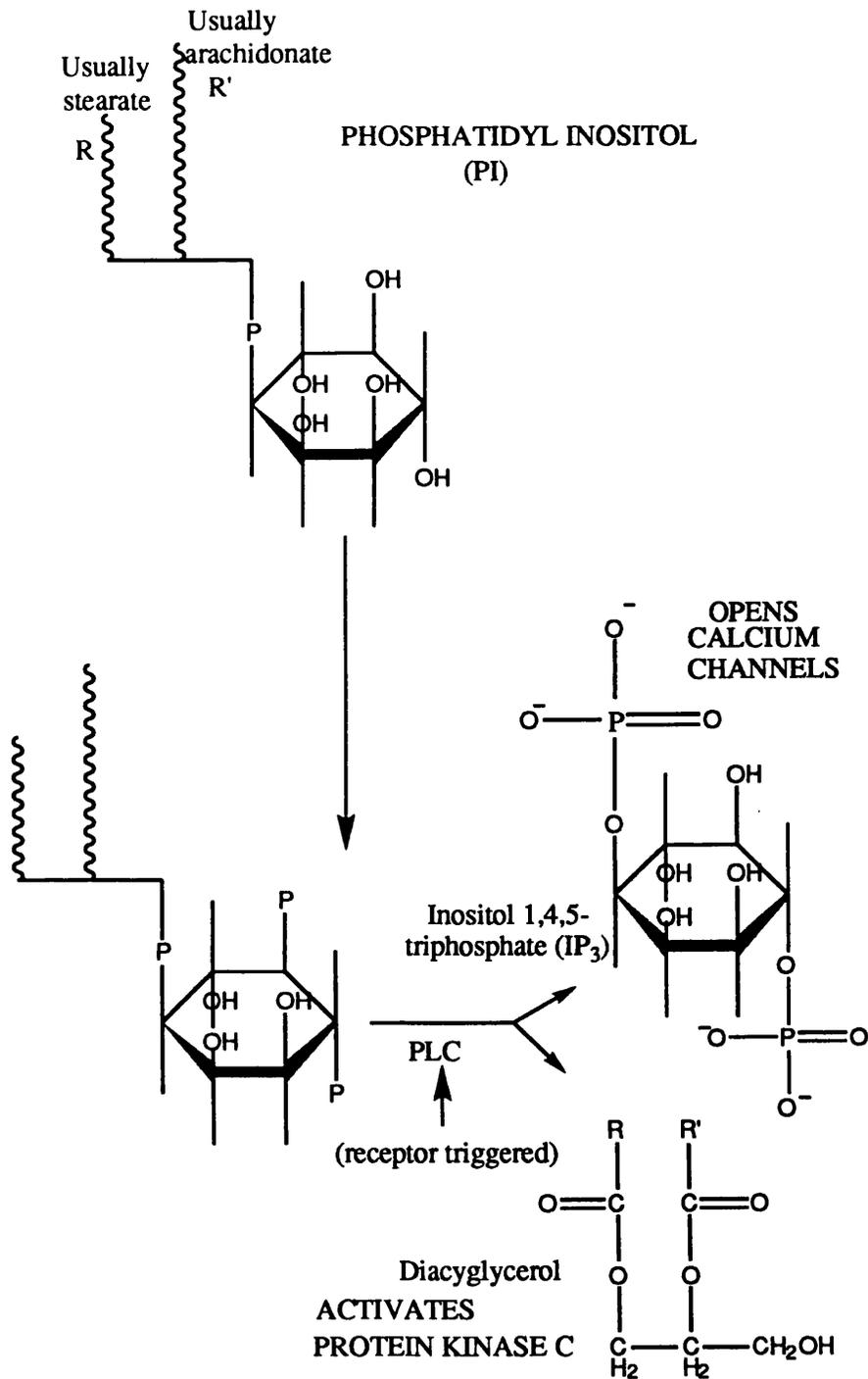


FIGURE 1. Synthesis and receptor triggered hydrolysis of phosphatidylinositol-4,5-bisphosphate (PIP₂).

The binding of a hormone such as serotonin to a cell surface receptor leads to the activation of **phosphoinositidase** (or **phospholipase C**). This membrane-bound enzyme hydrolyses the phosphodiester bond linking the phosphorylated inositol unit to the acylated glycerol moiety. Two messengers, **inositol 1,4,5-triphosphate (IP₃)** and **diacylglycerol** are formed by the cleavage of PIP₂. This hydrolysis is markedly stimulated by hydrolysis resistant analogues of GTP and by AlF₄⁻, suggesting that a G-protein carries the excitation signal from the activated receptor to phosphoinositidase. The inhibition of this cascade by pertussis toxin also points to the participation of a G-protein.

IP₃ is a short lived messenger, lasting only a few seconds. It can be converted to inositol by the sequential action of three phosphatases. The removal of the 5-phosphate group terminates its messenger role. Inositol 1,4-biphosphate is hydrolysed to inositol 4-phosphate and then to inositol. Alternatively, IP₃ can be phosphorylated to inositol 1,3,4,5-tetrakisphosphate, which is then hydrolysed to inositol 1,3,4-triphosphate, a different isomer from the one formed by cleavage of PIP₂. This 1,3,4-isomer can then be converted to inositol by successive dephosphorylations. The phosphatase that acts on inositol 1,3,4-phosphate is inhibited by millimolar concentrations of lithium ions. It is interesting to note that lithium is used to treat manic-depressive disorders; it may exert its action by inhibiting the recycling of 1,3,4-triphosphate. Some of the compounds formed in the conversion of IP₃ to inositol may have messenger roles.

The action of IP₃ has been investigated by microinjecting it into cells or by adding it to cells whose plasma membrane has already been made permeable. It has been found¹¹ that IP₃ causes rapid release of calcium from intracellular stores- the endoplasmic reticulum and, in smooth muscle cells, the

sarcoplasmic reticulum. The elevated level of calcium ions in the cytosol triggers processes such as smooth muscle contraction, glycogen breakdown, and exocytosis. Indeed, the injection of IP₃ into *Xenopus* oocytes suffices to activate many of the early events of fertilisation. Submicromolar levels of IP₃ mobilise calcium ions from intracellular stores by directly opening calcium channels in the membrane of the endoplasmic and sarcoplasmic reticulum.

For the purposes of this study, however, we are more interested in diacylglycerol (DAG), the other messenger formed by the receptor triggered hydrolysis of PIP₂ and which activates **protein kinase C (PKC)**. This 77kD enzyme phosphorylates serine and threonine residues in many target proteins¹⁰. For example, the phosphorylation of glycogen synthase by protein kinase C stops the synthesis of glycogen. This action of PKC nicely complements the IP₃-induced increase in the activity of glycogen phosphorylase which is mediated by an increase in the cytosolic level of calcium ions. Indeed, most effects of DAG and IP₃ are synergistic.

The phosphorylation of proteins has long been recognised as a pivotal step in the generation and transduction of signals in the cell¹⁰. Protein kinases, the enzymes which catalyse protein phosphorylations and their complementary protein phosphatases are able to modulate the biological activity of proteins in a rapid and reversible manner¹². With the advent of molecular cloning techniques the number of proven or putative receptors has risen, and the finding that a large number of receptors of growth factors and products of oncogenes are protein kinases underlines the important role that protein phosphorylation plays in the cell.

One of the protein kinases, protein kinase C (PKC), is known to have a crucial influence on cell proliferation and differentiation¹³. Research interest in PKC has been stimulated by the finding that the enzyme is a cellular receptor through which the tumour promoting phorbol esters such as 12-O-tetradecanoyl-13-acetate (TPA) mediate their pleiotropic effect on cells^{14,15}.

It has been shown that PKC is only active in the presence of calcium ions and phosphatidyl serine¹⁶. DAG greatly increases the affinity of PKC for calcium ions and thereby renders it active at physiological concentrations of this ion. The inactive enzyme is located mainly in the cytosol, whereas the active form is membrane-bound.

DAG can be phosphorylated to phosphatidic acid, which reacts with CTP to form CDP-diacylglycerol. Alternatively, it can be hydrolysed to glycerol and its constituent fatty acids. Arachidonate, the C₂₀-polyunsaturated fatty acid that usually occupies the 2-position of the glyceryl moiety of PIP₂, is the precursor of a series of eicosanoid hormones, thus the phosphoinositide cascade gives rise to many molecules that have signalling roles^{17,18}.

PKC ACTIVATORS

Protein kinase C (PKC) has been referred to as the Clapham Junction of cellular signalling traffic¹⁹. It is widely distributed in the tissues and organs of mammals and other organisms. The purified protein consists of a single polypeptide chain with a molecular weight of around 77kD¹⁹ and is composed of two separate domains. The enzyme can be cleaved by calpain, a calcium dependent protease, to produce two fragments with molecular weights of 51 and 26kD. The larger fragment is hydrophilic and contains the catalytic domain but is active in the presence of calcium, phospholipid, and DAG, while the smaller hydrophobic domain possesses the regulatory and membrane binding functions.

The enzymatic reaction catalysed by PKC is the transfer of the γ -phosphate group of ATP to seryl and threonyl, but not tyrosyl, residues of proteins. Activated by DAG, PKC is the site of action for tumour promoting phorbol esters¹⁵, of which TPA is the most potent. PKC phosphorylates many proteins including the EGF receptor, the insulin receptor and *ras*²¹.

Molecular cloning studies have revealed that PKC is in fact a family of enzymes of which several isoforms have been described, thus it seems feasible that the variety of PKC mediated cellular processes might be regulated by the differential expression of PKC isoforms. As an important integrator in signal transduction PKC is an especially attractive target for drug development.

Activation of protein kinase C seems to be absolutely essential for mitosis²⁰. Protein kinase C activity can be down regulated in some cell lines by chronic exposure to the tumour promoter TPA. The consequence of this is that the cell is unable to enter mitosis in response to growth factor stimulation and is also unable to divide in response to transfection by *v-Ha-ras*. Consequently inhibition of C-kinase activity has been regarded as a key site for therapeutic intervention of tumour states. However, despite some 10 years effort, no specific inhibitor has been developed. Protein kinase C is activated physiologically by DAG and is also the site of action of tumour promoting phorbol esters.

The importance of PKC in controlling cell division and proliferation is revealed by the action of the **phorbol esters**. These polycyclic alcohol derivatives from croton oil are carcinogenic; they are known tumour promoters. Phorbol esters activate PKC because they resemble DAG. The activation is persistent because, unlike DAG, they are not degraded.

A number of naturally occurring terpenoid compounds structurally unrelated to phorbol esters can activate PKC, e.g. mezerein, teleocidin, asplysiatoxin and ingenol 3-tetradecanoate, act as tumour promoters, mimic some of the other biological actions of these esters and compete with tritiated phorbol dibutyrate for binding to the phorbol ester receptor. In general the biological potency of the esters is paralleled by their PKC-activating activity. The propensity of phorbol esters to activate PKC is governed by a strict structure activity relationship. Computer molecular modelling of phorbol esters, related tumour promoters and DAG has shown a marked similarity in the relative position in space of certain heteroatoms and hydrophobic groups in these molecules^{21,22,23}.

For phorbol esters, the pharmacophoric moieties are the hydroxyl groups at C4, C9 and C20 (see Fig. 2) and the hydrophobic region filled by a long chain acyl functionality attached to either the C12 or C13 positions. DAGs also fit this model in a stereospecific fashion therefore one might expect that exogenously added synthetic DAGs induce the same biological responses observed with phorbol esters.

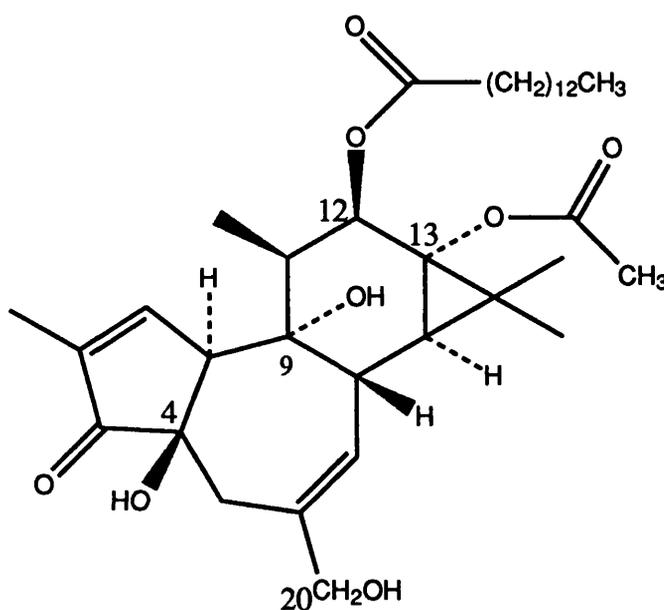


FIGURE 2 - A Phorbol Ester (TPA).

Certain studies support this hypothesis, although some studies show a slight discrepancy in results possibly due to the fact that DAGs are much weaker activators of PKC than phorbol esters since they possess a much lower affinity for the enzyme and unlike phorbol esters are metabolised rapidly by kinases and lipases which renders DAG-mediated activation transitory.

All DAGs able to bind to and activate PKC *in vitro* have a 1,2-*sn* configuration, and neither the *sn*-2,3-DAG enantiomer nor the 1,3-DAG

diastereomer are active⁸. The observed stereospecificity was also observed in intact cells, suggesting that a highly specific DAG-protein interaction is necessary for the activation of PKC.

Structure activity studies have shown that virtually any modification of the glycerol backbone of DAG results in a strong decrease in activity²³. Examples of such modifications are the addition of a single carbon atom in the glycerol backbone or modifications of the hydrophilic residues of the DAG.

The structures of those compounds which inhibit PKC are seemingly as diverse as those exhibited by PKC activators²³. Agents with PKC inhibiting properties have been found among ether phospholipids, dialkyl and alkylacylglycerols, lipoidal amines and an array of heterocyclic compounds such as aminoacridines, staurosporine and sangivamycin. In general these kinase inhibitors are neither selective for PKC nor are they, with the exception of staurosporine and certain related microbial products, very potent inhibitors. The mechanisms by which they inhibit PKC activity are manifold and are not completely understood.

It was originally thought that the source of agonist stimulated DAG was the phospholipase C-catalysed breakdown of PIP₂, however it is now clear that this only provides a minor component of the elevation in this second messenger system in response to mitogen stimulation in a range of cell types²⁴, e.g. fibroblasts, smooth muscle cells and neutrophils. The major and sustained source of DAG (and indeed that generated in response to mitogens that do not stimulate PIP₂ breakdown e.g. EGF), is PC²⁴ (see Fig. 3).

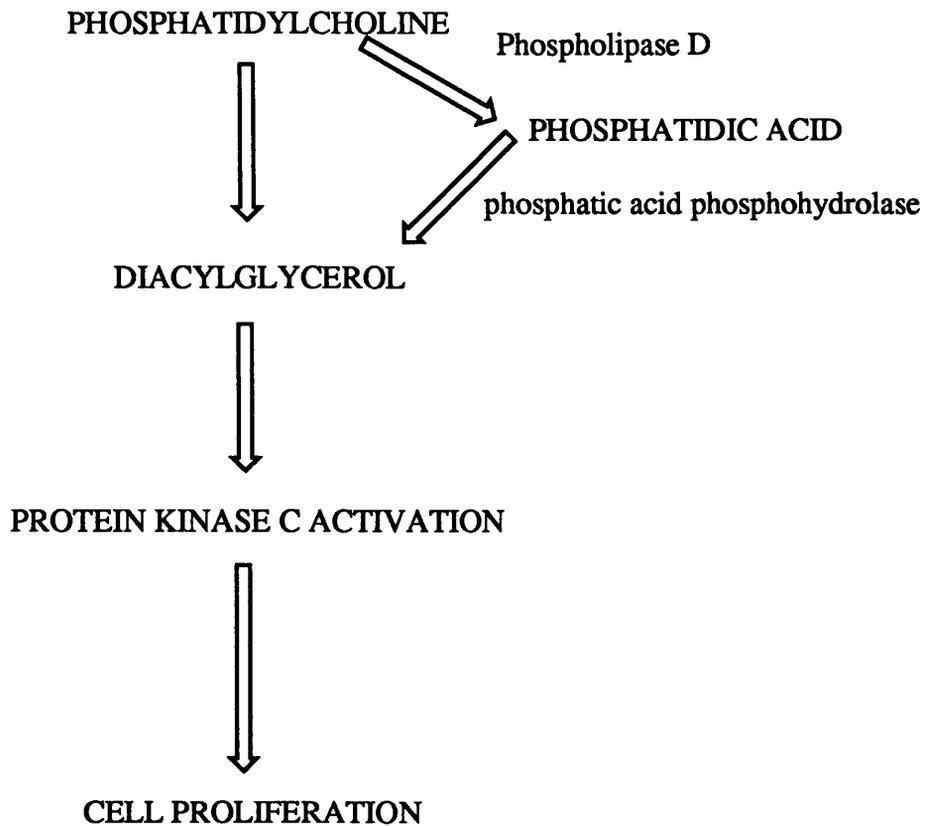


FIGURE 3 - The phospholipase D pathway

Stimulated PC breakdown has been observed in a number of cell types in response to a range of mitogens, i.e. bombesin, vasopressin, prostaglandin F₂ alpha, PDGF, EGF, thrombin, endothelin and serum¹⁸. It has also been observed in response to stimulation with tetraphorbol acetate (TPA). In contrast to the breakdown of PIP₂, PC hydrolysis is catalysed by a phospholipase D type enzyme. The lipid product of this reaction is phosphatidic acid (PA). PA has been demonstrated to be mitogenic in its own right, but it can be converted to DAG by the action of phosphatidate phosphohydrolase. The regulation of PLD activity appears to be accomplished in both a receptor-dependent and independent manner²⁵.

Phosphatidylcholine (PC) is found almost exclusively in eukaryotic cell membranes and can account for up to 50% of the total cellular phospholipid content¹⁰.

PC from mammalian tissues normally contains a saturated fatty acid residue at the C1 position and an unsaturated fatty acid residue at the C2 position of the glyceryl moiety. Different species of PC have varying proportions of fatty acids e.g. stearic, arachidonic, linoleic, oleic thus phosphodiesteric cleavage of PC can lead to multiple species of DAG with important ramifications in cell activation. The fact that PC is a widespread receptor-linked process suggests that its importance in signal transduction is critical.

In addition to being activated by a diverse range of growth factors, an increase in PC breakdown has been shown in cells transfected by *ras* and *src* oncogenes. This raises the possibility that PLD is a convergent point in the action of both growth factors and certain oncogenes and therefore pharmacological intervention at the level of PLD may lead to the inhibition of the proliferation of both normal and cells transformed by a range of oncogenes. Consequently the identification of inhibitors of PLD may lead to the generation of novel, specific anti-cancer drug.

Although phospholipase D (PLD) activities of plant origin have been known since 1948, the existence of mammalian PLD acting on PC was first detected in 1975 using a microsomal preparation of the brain¹⁰. Subsequent studies have demonstrated PC-preferring PLD in homogenates and membranes from various tissues and cells including lung, liver, adipose tissue, epithelial cells, HL-60 cells and spermatozoa, with lung and brain being the richest sources. Mammalian PLD from various systems exhibits a transphosphatidylolation

activity. Requirements of calcium, fatty acids and detergents for PLD expression vary from cell to cell, suggesting again that multiple forms exist.

It has been demonstrated that growth factors can stimulate PC breakdown through activation of PLD. This conclusion can be derived from a range of studies demonstrating the appearance of [^3H -choline] labelled cells before any choline phosphate is detectable. Since cells contain high levels of choline kinase and the resting level of choline phosphate in equilibrium labelled cells is 506 times higher than that of choline, it is likely that the rise in choline phosphate is due to the phosphorylation of the liberated choline (i.e. from choline hydrolysis) rather than a PLC activity being stimulated.

The most conclusive demonstration of agonist stimulated PLD activity has been by Pai *et al.*²⁶. This group labelled HL-60 cells with an alkyl-[^{32}P] lyso-PC, which was incorporated into the membrane of these cells and rapidly acetylated. They then demonstrated that stimulation with N-formyl-Met-Leu-Phe caused the generation of alkyl-[^{32}P] PA. Since the cellular pool of ATP is not labelled under these conditions, the alkyl-[^{32}P] PA must be formed through PLD catalysed hydrolysis of alkyl-[^{32}P] PC.

The lead compound for this study is the compound **viridin**, a fungal metabolite from *Trichoderma viride*²⁷ which has potent cytotoxic activity, at first thought to be through inhibition of PLD. It is now thought, however, that its activity may be mediated through an as yet undefined tyrosine kinase which may or may not have a direct action on levels of PLD in cells. Although it is not clear at which point in the second messenger system these compounds

are acting, it is obvious that they are potent and structurally novel compounds, and a potential lead into a new generation of effective anti-cancer drugs.

This project has therefore involved:

1. the growth, isolation, purification and characterisation of the fungal metabolite viridin using a modified version of known techniques;
2. identification of all structurally similar compounds using the search facilities by computer of the Crystallographic database CSSR;
3. determination of the crystal structure of demethoxyviridin;
4. partial synthesis of hibiscone C, a partial analogue of viridin;
5. similarity modelling of all compounds using HYDRA and MMS;
6. building a QSAR model of previously synthesised analogues of wortmannin with known anti-inflammatory activity data;
7. correlating this model with synthesised analogues of demethoxyviridin to predict activity of new compounds;
8. testing new compounds for cytotoxic and PLD/tyrosine kinase inhibitory activity;
9. using these results to predict which compound will be the most active, hence which should be synthesised next;
10. synthetically modifying the most active analogue in order to optimise stability and other pharmacochemical properties such as lipophilicity .

CHAPTER 2

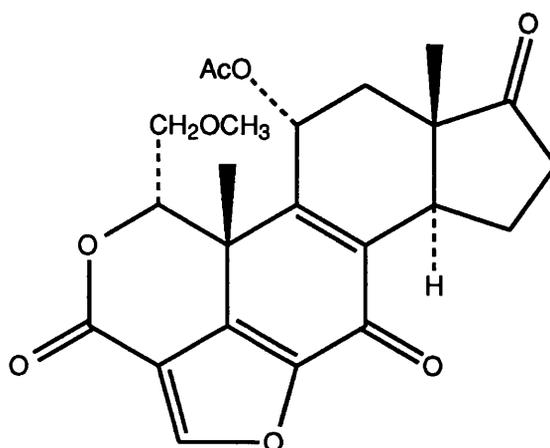
**POSSIBLE ENZYME
INHIBITORS**

LEAD COMPOUNDS

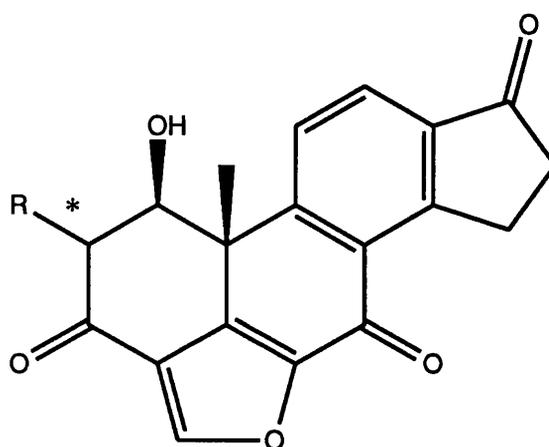
In neutrophils the respiratory burst induced by *f*-met-leu-phe, a chemotactic peptide, was shown to be dependent on PLD activation²⁸. The main consequence of phagocytosis in granulocytes and mononuclear phagocytes is this sudden, large increase in non-mitochondrial oxygen consumption commonly called the *respiratory burst*. In this process, oxygen is reduced to superoxide by an oxidase located in the plasma membrane while cytosolic NADPH is oxidised. The nature of the respiratory burst is still unknown. A flavoprotein and a b-type cytochrome were proposed as the components of what could function as an electron transport acceptor with oxygen as the final acceptor.

The triggering of the phagocyte by particles depends on the interaction of opsonic molecules and other surface determinants with plasma membrane receptors. The transduction sequence leading to the activation of NADPH oxidase is likely to involve an enhancement of cytosolic calcium, the breakdown of inositol phospholipids with generation of inositol triphosphate and DAG, and activation of PKC.

This respiratory burst is inhibited by the fungal metabolite **wortmannin**²⁹, produced by *Penicillium dangeardii*. This compound is a structural relation of **viridin**, produced by *Trichoderma viride*³⁰ (see Fig. 4).



WORTMANNIN [1]

*  OCH₃ - VIRIDIN [2]*  OCH₃ - -VIRIDIN [2A]

R=H DEMETHOXYVIRIDIN [3]

FIGURE 4 - Lead compounds

PLD activity can be readily quantified in intact cells by measuring the formation of phosphatidylalcohols, which are produced in a unique PLD-dependent transphosphatidylation reaction. Recent work by Wellcome Research Laboratories has indicated that in cytochalasin B-primed neutrophils, superoxide production stimulated by *f*-met-leu-phe, is totally dependent on PLD activation²⁸. Wellcome have also shown that

demethoxyviridin and wortmannin block phospholipase C and D activation in the human neutrophil²⁵. Consequently, a modified viridin molecule (see Fig. 4) was synthesised, tested and found to have inhibited PLD activity in neutrophils, Swiss 3T3 and NIH 3T3 cells.

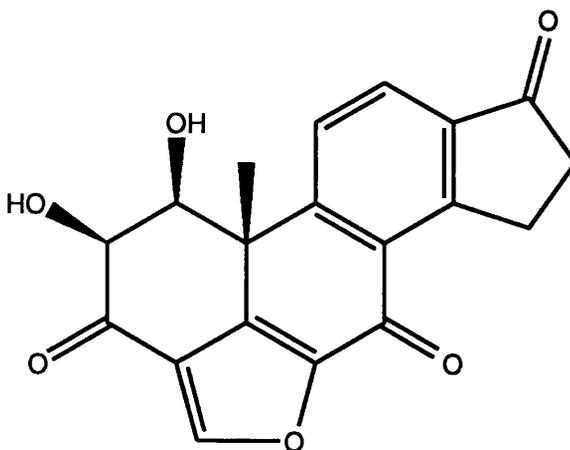


FIGURE 5 - A modified viridin structure [4].

Addition of this analogue to the medium of the cells inhibited the incorporation of [³H]- thymidine into the DNA of both control and transformed cells. There was no effect of the compound on agonist stimulated PIP₂ hydrolysis and it appeared to have no effect on cellular morphology over the 24 hour period examined. These results suggest that PLD activity is essential for proliferation and that its inhibition by these compounds can prevent the growth of normal and transformed cells.

This was therefore the basis of the project to design a novel anti-cancer agent. A wortmannin or viridin derivative with maximum PLD inhibitory activity was sought to have optimum physicochemical features such as stability, and lipophilicity. These compounds were shown to have a preferential activity on

malignant dividing cells suggesting a desired selectivity of action may be present.

MYCOLOGY

The two main lead compounds **viridin** and **wortmannin** are metabolites of the fungi *Trichoderma viride* and *Penicillium dangeardii* respectively.

EXPERIMENTAL

VIRIDIN

Trichoderma viride was obtained from Kew Mycological Institute in London as a freeze dried culture in test tubes. These were used to inoculate seed tubes that were then stored for 7 days at 25°C. A water suspension was produced by adding a few ml of water to each seed tube and gently scraping to rub spores from the growing fungus into the suspension. The solutions were bulked, homogenised and used to inoculate Roux bottles containing Raulin-Thom medium³¹ with extra tartaric acid added as in the literature method. The ingredients of this medium are shown below. The resulting suspension was green and soup-like.

Deionised water	H ₂ O	5l
Glucose	CH ₂ OH(CHOH) ₄ CHO	250g
Tartaric acid	(CHOHCOOH) ₂	13.3g
Ammonium tartrate	(CHOH.COONH ₄) ₂	13.3g
Di-ammonium hydrogen orthophosphate	(NH ₄) ₂ HPO ₄	2.0g
Potassium carbonate	K ₂ CO ₃	2.0g
Magnesium carbonate	MgCO ₃	1.35g
Ammonium sulphate	(NH ₄) ₂ SO ₄	0.85g
Zinc sulphate	ZnSO ₄ .7H ₂ O	0.25g
Ferrous sulphate	FeSO ₄ .7H ₂ O	0.45g

TABLE 1 - Raulin Thom medium.

This process was carried out under as sterile conditions as possible. These, along with two media control bottles, were then stored in the dark. These conditions were chosen after a trial batch was grown in the dark after this fact being mentioned but not explained in the literature, resulting in an increase in viridin yield from 23mg/l to 85mg/l with an increased purity of 91% compared to 65% when grown in the light. The pH was monitored daily. The fungus was cropped off after 21 days and extracted continuously with ethyl acetate. The pH at time = zero, and of the control bottles throughout the experiment was 2.9. The pH of the inoculated broth, however, fell gradually to a resulting value of 2.60.

The purified extract yield from the broth varied from batch to batch but was usually around 80mg/l. In the initial stages of this project the mycelium, left after filtering, was dried and subjected to continuous soxhlet extraction, but this gave no compounds of sufficient interest. The concentrated crude extract was purified by chromatography as follows:

One 1m silica plate was prepared for each 500mg crude sample. 50g of silica (GF254) was made into a slurry with approximately 200ml deionised water. This was then spread onto 1m glass plates and the concentrated sample painted onto the dried plate in layers. The plates were immersed in a solvent tank containing 70:30 light petrol:ethyl acetate, the solvent ratios being previously determined by TLC. The plates were run once and the appropriate bands scraped off and washed several times with ethyl acetate. Viridin ($R_f=0.47$) was positively identified by melting point (222-224°C), high field NMR and mass spectrometry³² (for experimental data see Appendix 5). Its purity was later determined by HPLC. The literature also revealed that β -viridin ($R_f=0.36$) could be separated from pure viridin. This was done by chromatography on a small silica plate with 10:1 toluene:methanol as solvent. β -viridin was positively identified by NMR: and melting point (140°C.).

WORTMANNIN

The first few batches of *Penicillium dangeardii* received from Kew were discovered to be contaminated. The fungus was eventually isolated and grown under similar conditions as *Trichoderma viride*. The growth was bright yellow and the texture was of rubber when scraping the seed tubes for scaling up and inoculation into larger batches. When isolated, however, the main compound was produced in extremely low yield and did not give convincing TLC, NMR or mass spectrometry results. After several unsuccessful attempts to rectify this, a sample of wortmannin was donated by Sandoz for analysis at the Beatson Institute by HPLC.

DEMETHOXYVIRIDIN

The viridin samples initially isolated and purified were used for initial testing. As the results looked promising, more compound was required for analogue synthesis. SmithKline Beecham kindly donated a large sample of demethoxyviridin, which enabled several analogues to be made and tested. The demethoxyviridin comes from the fungus *Nodulisporum hinnuleum*.

HPLC ANALYSIS

HPLC analysis of viridin, wortmannin and β -viridin (see Fig. 4) was carried out at the Cancer Research Campaign Beatson Laboratories, CRC Department

of Medical Oncology, University of Glasgow. Separate reverse phase HPLC assays were developed for the analysis of wortmannin and its breakdown products. Both assays used a C18 μ Bondpack column ($5\mu\text{m} \times 25\mu\text{m}$) and a photodiode array detector at 200-500nm. For wortmannin, the mobile phase was methanol: acetonitrile: water (1:1:2), pH7, flow rate of 1ml/minute. For the breakdown products the mobile phase consists of methanol: acetonitrile: 0.01M sodium acetate (1:1:2) at pH4. Fig. 6 shows resulting HPLC traces of wortmannin before and after 20 minutes had elapsed.

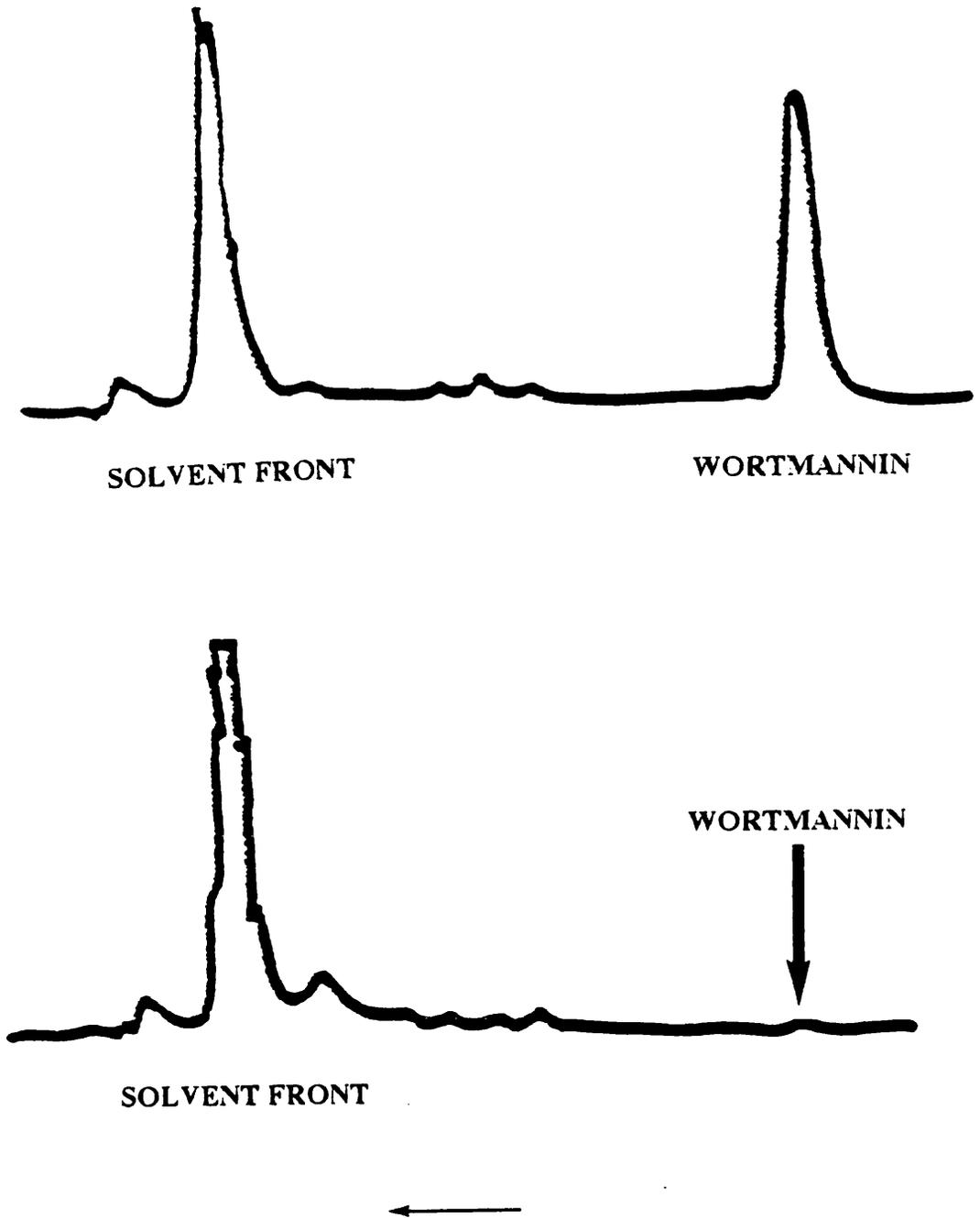


FIGURE 6 - HPLC analysis of wortmannin.

Above - wortmannin, below - no wortmannin is seen after 20 minutes, some traces of degradation products are seen.

PEAK	%AREA	%AREA
	t=0 minutes	t=20 minutes
1	3.04	3.84
2 (solvent peak)	62.44	88.95
3	1.21	6.17
4 (breakdown product)	1.53	1.04
5 (wortmannin)	31.77	-

TABLE 2 - Results from HPLC analysis of wortmannin.

Wortmannin was seen to degrade extremely rapidly, after 20 minutes there was no measurable peak 5. Peak 4, however, increased several times in area and was thought to be the major breakdown product. More of the substance represented by peak 3 was also seen, this was thought perhaps to be a minor degradation product. These breakdown products have been observed, unlike the parent compound, to show strong absorption at 440nm. This was used as the basis for an assay of the increase in degradation of wortmannin with time to determine half-life (see Fig. 7).

At physiological pH, wortmannin was found to degrade rapidly and completely with a half life of less than 10 minutes (figure 7). The breakdown of wortmannin is a pH dependent process since degradation can be prevented by acidifying aqueous solutions of the molecule (pH5). For cytotoxicity assays where dilutions are made at pH 7.4, wortmannin has been shown to be fully degraded by the time the drug is added to the cells. The breakdown product of wortmannin may therefore be crucial in the activity, and our understanding of the cytotoxic activity, of these compounds.

Unfortunately, sample size was not sufficient to allow isolation and analysis of these products.

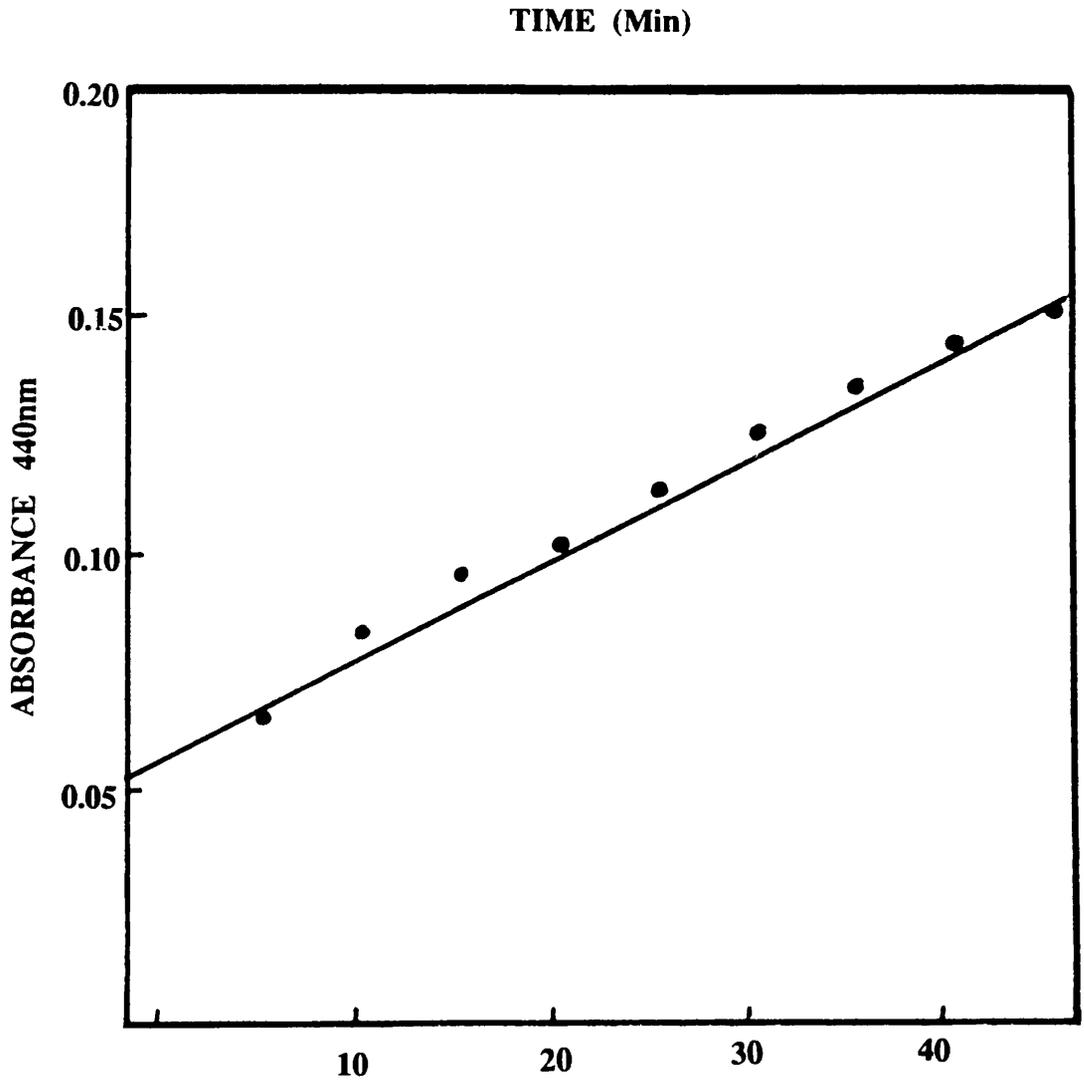


FIGURE 7 - Decomposition of wortmannin shown by increase in breakdown products.

Cytotoxicity studies have been carried out by co-workers at the Beatson Institute in Glasgow. It was found that wortmannin is cytotoxic to Swiss 3T3 cells with an ID_{50} of $41.3 \pm 1.9 \mu\text{M}$ and inhibits phospholipase D (figure 8). The inhibition of phospholipase D was measured by co-workers in the Biochemistry department at Glasgow. This test was based on a transphosphatidylolation assay. It is known that PLD catalyses the cleavage of the terminal diester bond of phospholipases (see Fig.9). Its preferred substrate is phosphatidylcholine, thus generating phosphatidic acid (PA) and choline (Cho). PA will then be hydrolysed by phosphatidic acid phosphohydrolase, and break down to generate diacylglycerol. If cells are labelled with [^3H]palmitic acid, which is taken up by the cells as R_2 of the PC, then stimulated in the presence of butanol, a butylated adduct is preferentially formed which cannot be hydrolysed by the PA specific enzymes, and can then therefore be counted by scintillation counting after extraction (see Fig.10).

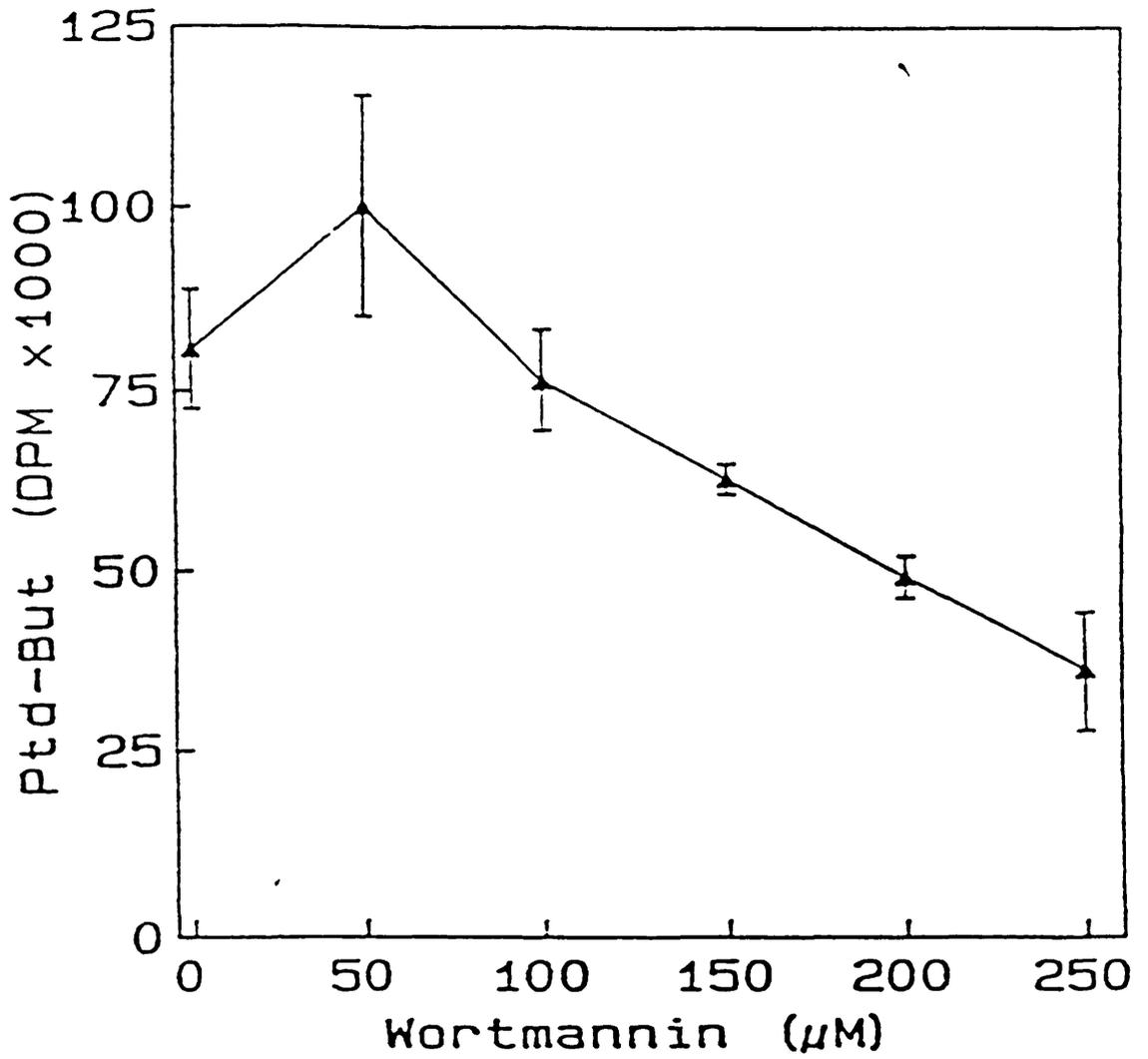


FIGURE 8 - Inhibition of phospholipase D by wortmannin

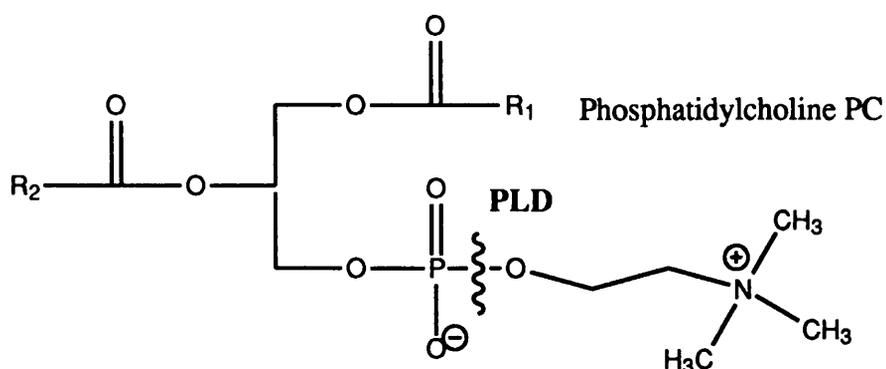


FIGURE 9 - Hydrolysis of phosphatidylcholine by PLD.

R_1 = usually saturated fatty acid, R_2 = *palmitic acid

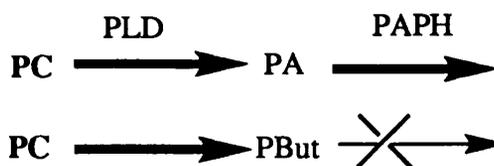


FIGURE 10 - PLD assay.

Demethoxyviridin showed significant cytotoxicity against a panel of cell lines including Swiss 3T3s, mink cell lines, ovarian and breast tumour cell lines (table 3)³⁴. It appeared to be most toxic to ovarian tumour cells. β -viridin was also later tested in Swiss 3T3s and found to be several times more active than viridin (see Fig. 11). After 5 days in a $1\mu\text{m}$ solution of inhibitor, β -viridin showed 7×10^{-4} cells remaining whereas viridin showed in the order of 10×10^{-4} . Thus β -viridin was a more potent inhibitor of cell growth.

Cell lines	Demethoxyviridin ID₅₀(μm)
Swiss 3T3	34.0 \pm 1.2
Mink type I	114.0 \pm 1.0
Mink type II	130.0 \pm 6.0
Mink type III	142.7 \pm 3.0
Mink type IV	113.7 \pm 2.0
Ovarian tumour type I	20.5 \pm 0.5
Ovarian tumour type II	32.0 \pm 0.6
Ovarian tumour type III	26.8 \pm 0.4
Breast tumour	123.3 \pm 9.0

TABLE 3- Cytotoxicity data for demethoxyviridin.

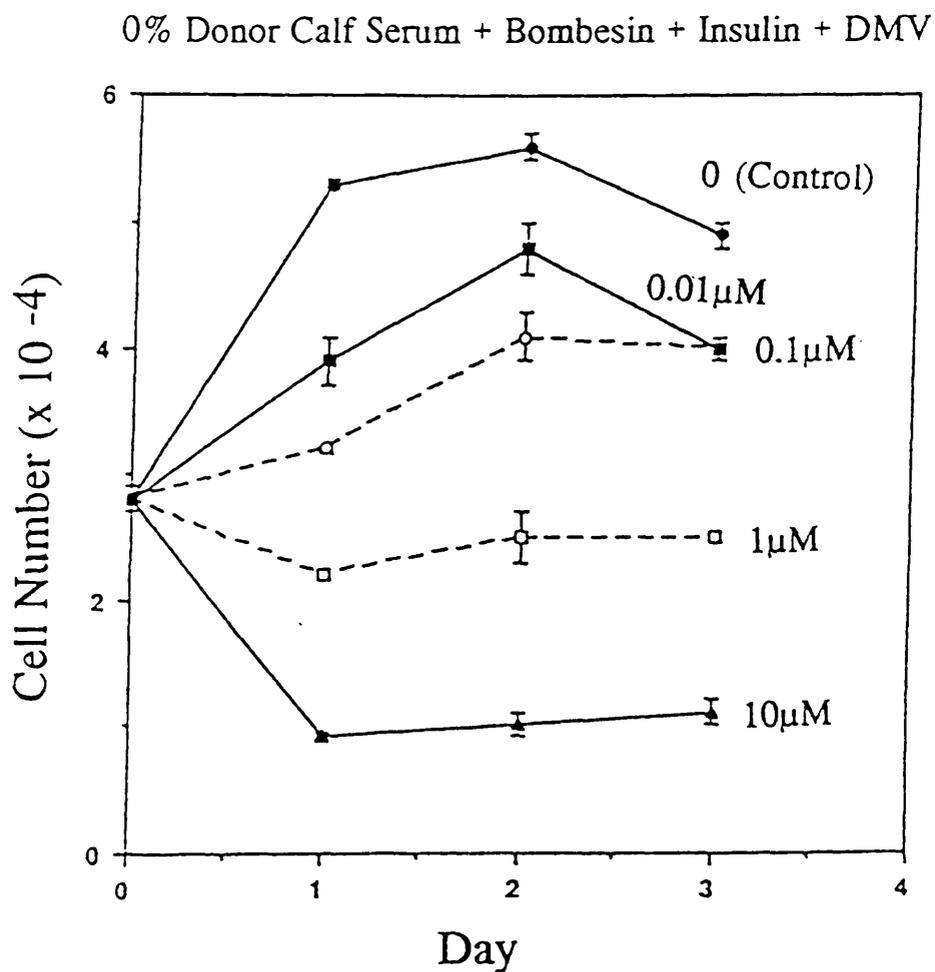


FIGURE 11 - Inhibition stimulated cell growth by β -viridin.

Bombesin = mitogen

FCS = foetal calf serum, growing medium for cells

Figure 12 shows results of the viridin light study where it was shown that growing the fungus in the dark increases the viridin yield and purity. The first peak is thought to be β -viridin.

It was seen that the viridin extracted from the fungus grown in the **light** had about **32%** impurity, thought to be β -viridin. From the **dark** sample, however, the β -viridin content was only **9%** (see Table 4). β -viridin was later tested and found to several times more active than viridin (see later). Although the percentage β -viridin was higher when the fungus was grown in the light, the combined sample extracted was much greater from the dark. It was therefore decided not to change the extraction conditions at this stage.

PEAK	LIGHT	DARK
1 (β -viridin.):		
AREA	0.555	0.127
%	32.18	9.05
2 (viridin.)		
AREA	1.060	1.275
%	61.48	90.62

TABLE 4 - HPLC analysis results of viridin light study.

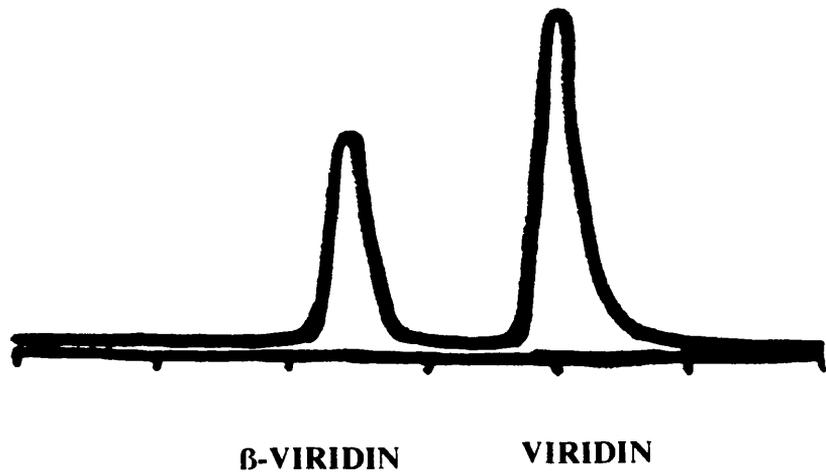
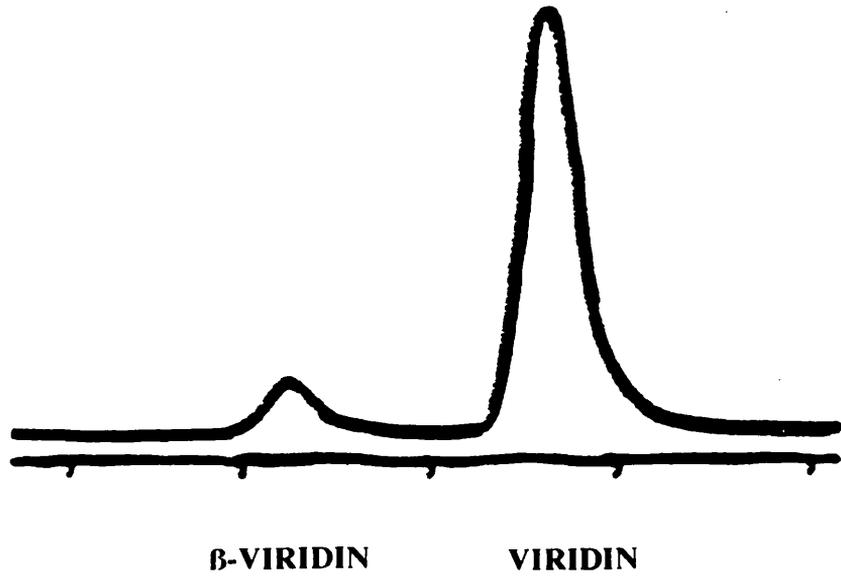


FIGURE 12 - HPLC analysis of viridin light study.

STRUCTURALLY RELATED COMPOUNDS

The Chemical Databank System (CDS) at Daresbury Laboratory provides a reliable service for the retrieval of chemical data, the graphical display of structures and the analysis of numerical data.

A retrieval system written at Daresbury provides a common method of access to a number of databanks and has a number of special features e.g. fast chemical structure retrieval facilities. The local software supports chemical connectivity searching facilities for five of the databases. The performance of this system depends on the pre-processing of data to produce structured indexed files. The service can be used from any terminal which can be connected to JANET, the Joint Academic Network. One of the databanks which can be accessed in this way is CSSR (Crystal Structure Search Retrieval), the Cambridge Structural Databank, which contains crystal structure data for about 90,000 organic and organo-metallic compounds. The program runs at Daresbury and provides fast retrieval of structural and bibliographic information on the VAX computer at Glasgow, allows logical manipulation of retrieved material, and display of results. Using an appropriate graphics terminal and the associated program VIEW, labelled stick diagrams or stereo-pairs of molecules can be displayed. These may be scaled or rotated, and bond length and angle calculations made.

An important feature of this program is its ability to use molecular fragments designed by the user in order to search for structures which contain these fragments or certain characteristics of them. A query structure is an

arrangement of atoms and bonds which is stored as a connection table which specifies how the atoms are connected. This table may be displayed as a conventional chemical diagram.

Using this program, the query structure in figure 13 was designed as a representation of the key chemical features of viridin and wortmannin. The crystal structures of viridin³⁴ and wortmannin³⁵ had been previously determined.

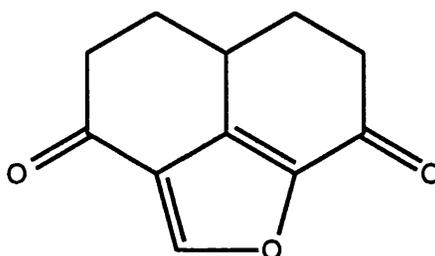


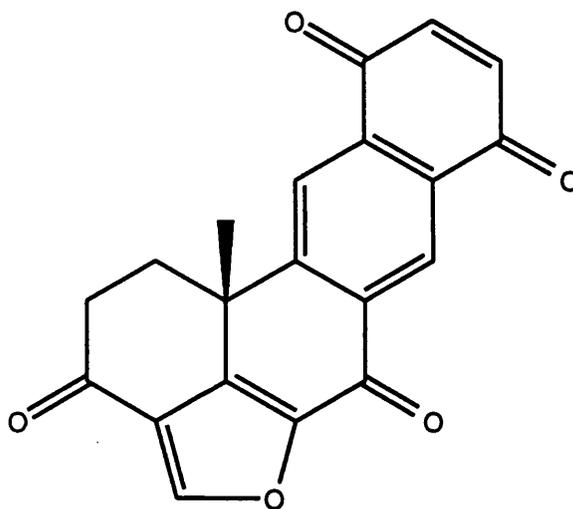
FIGURE 13- Query structure [5] used to search Cambridge database.

This structure was used to search all the compounds in the database. Similar compounds of interest are shown in figure 14 :

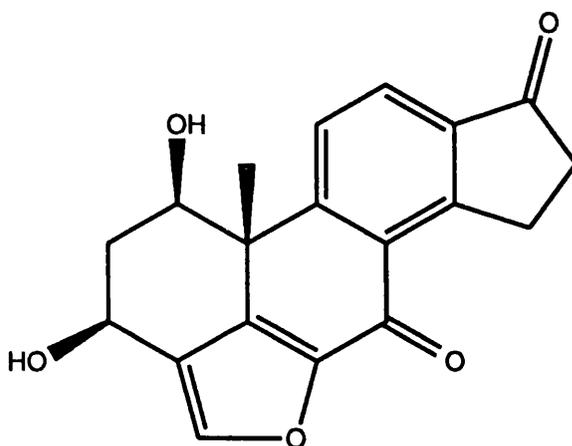
Halenaquinone - a pentacyclic polyketide isolated from a marine sponge *Xestospongia exigua*, first collected in Palau in the Western Caroline islands³⁶.

Demethoxyviridiol - a metabolite of the fungus *Nodulisporum hinnuleum*.³⁷

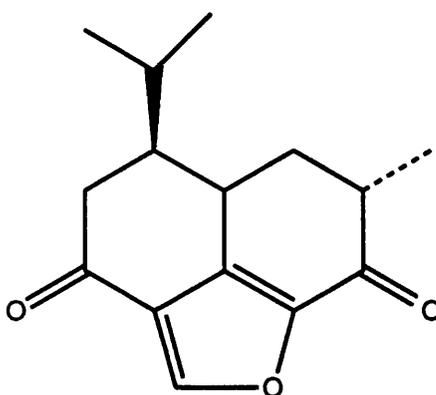
Hibiscone C - an extract from the Blue Mahoe, *Hibiscus elatus*, the national tree of Jamaica³⁸.



1. Halenaquinone [6]



2. Demethoxyviridiol [7]



3. Hibiscone C [8]

FIGURE 14 - Structural analogues found from CSSR search.

The co-ordinates of these compounds were stored in files and transferred to Glasgow via JANET. The co-ordinates of viridin had been published and those of wortmannin were requested and donated by Sandoz. The photosynthesis of hibiscone C had been documented and was carried out with a view to producing rigid structure analogues (see Chapter 5).

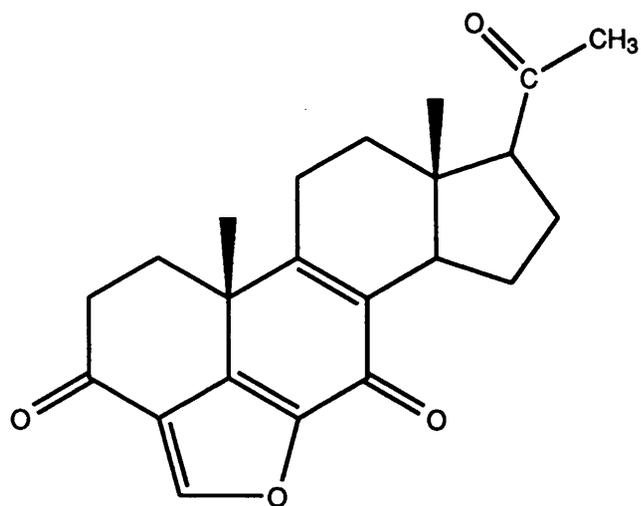
Further literature searching revealed several similar compounds for which the crystal co-ordinates were not known (see figure 15) :

Virone [9] - is produced by *Trichoderma viride* grown at a higher temperature than that which produces viridin³⁹.

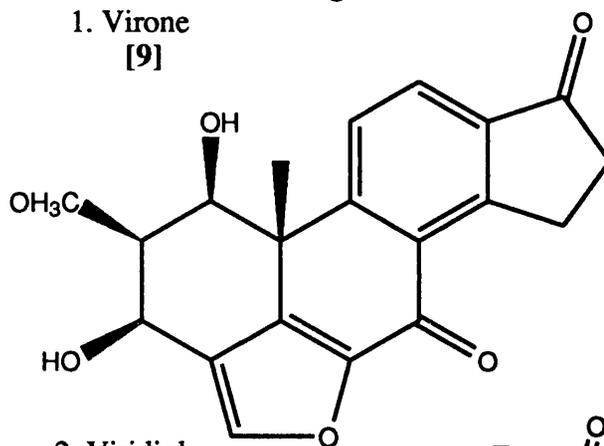
Viridiol [10] - a minor metabolite when the fungus is grown at the normal room temperature^{40,41}.

Wortmannolone [11] - is produced by *Penicillium dangeardii*³⁹.

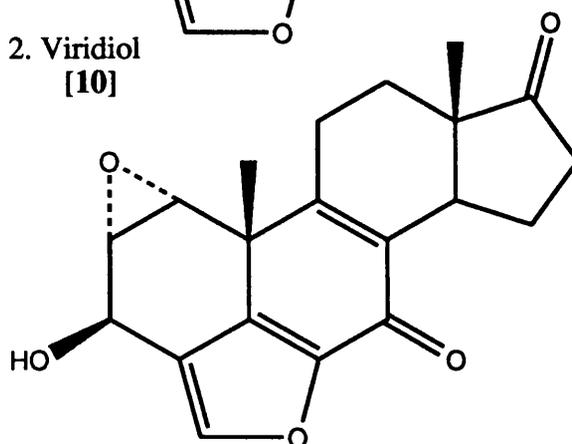
Demethoxyviridin [3] - produced by *Nodulisporum hinnuleum*³⁷.(see Fig.4)



1. Virone
[9]



2. Viridiol
[10]



3. Wortmannolone
[11]

FIGURE 15 - Virone, Viridiol and wortmannolone

Demethoxyviridin [3] was also donated by Wellcome and the crystal structure determined. (see Chapter 4).

CHAPTER 3
COMPUTER MODELLING

COMPUTER MODELLING

In studies of quantitative structure activity relationships (QSAR), the relative potencies of a series of drugs are subjected to analysis with the hope that biological potency will be described by a mathematical equation. QSAR is an actuarial or statistical method in which only objective data are used with no intrusion of models or mechanistic hypotheses. The equation that is obtained not only accounts for the relative potencies of the compounds, but from it are deduced predictions of the potencies of untested compounds. The method thus has the capability of yielding new (structurally related) drugs with desired potency, perhaps drugs with enhanced potency or fewer side effects.⁴⁵ Computers are also an invaluable aid to this method in terms of sheer magnitude of calculations and of visualisation of techniques and results on powerful graphics workstations.

In a series of structurally related compounds, some analogies seem to be traceable although there are always exceptions. For example, it is widely accepted that substitution of an aromatic ring by chlorine in a biologically active molecule 'deepens' activity.⁴⁴ This has been explained by invoking the Hammett substitution of chlorine and the push-pull electron density changes brought about by the substitution. It may be, however, that the chlorine atom blocks aromatic hydroxylation and thereby prevents metabolic removal of the active species. Although such speculation anticipates more contemporary ideas of drug action, the problems existed 85 years ago and seemed to resist solution. In 1893 the French physiologist Charles Richet tried to formulate SAR mathematically⁴⁴ but this procedure had to wait another six years when Hans Horst Meyer and E. Overton published studies in which narcotics were ranked according to their solubility in blood and lipids⁴⁴ (*i.e.* water and organic solvent). These theories of a relationship between the physical

properties of a compound and its biological properties were taken up by others and revived decisively by Corwin Hansch in 1964⁴⁴. The lipid theory was followed by ideas concerning other physical properties such as ionisation and a proposal by Linus Pauling that narcotics interfere with clathrate formation of ice crystals in the central nervous system and thereby prevent normal nervous transmission of pain impulses⁴⁴. These proposals, whether right or wrong at the time, were but rare attempts to rationalise an increasingly unruly accumulation of therapeutic data that defied explanation.

The programs used to study viridin, wortmannin and their analogues in an attempt to elucidate structural information from them which could be correlated with biological potency were MMS⁴⁵ (molecular modelling system), HYDRA⁴⁶, ORTEP⁴⁷, PYTHON⁴⁸ and SYBYL.⁴⁹

MMS, HYDRA AND ORTEP

Files with the crystal co-ordinates for the structural analogues available were converted from fractional to orthogonal co-ordinates using a locally written converting program GX-PDB⁵¹. These were transferred to the UNIX system for use on the Silicon graphics IRIS computer. Using the program MMS (Molecular Modelling System), a template molecule file was prepared and atom labels matched to those of the co-ordinate file. The molecule could then be displayed and manipulated in real time on the terminal.

More than one molecule can be displayed at a time hence qualitative comparisons can be drawn. For example, the A and B rings of hibiscone were slightly buckled in the opposite direction to both viridin and wortmannin. Ring C of halenaquinone and viridin were also angled away from each other although rings A and B were more or less identical. Figure 16 shows viridin (yellow) and wortmannin (blue) face on and translated slightly from their docking positions. What cannot be seen from this angle, however, is shown in Figure 17 where the molecules are shown side on and are seen to curve in slightly different directions. This may account for slight changes in potency of enzyme inhibition due to molecular recognition at the enzyme site.

Some basic overlap experiments were carried out to assess molecular similarity between the compounds for which crystal structures were known, namely viridin, hibiscone C, demethoxyviridin and wortmannin. The co-ordinates of the former two compounds were easily available in the literature, the crystal structure of demethoxyviridin was determined (see Chapter 4) and the co-ordinates of wortmannin were kindly donated by Sandoz of Switzerland on direct request. The atoms overlapped (Figure 18) in each compound were chosen because it has been shown that the unique ABE ring

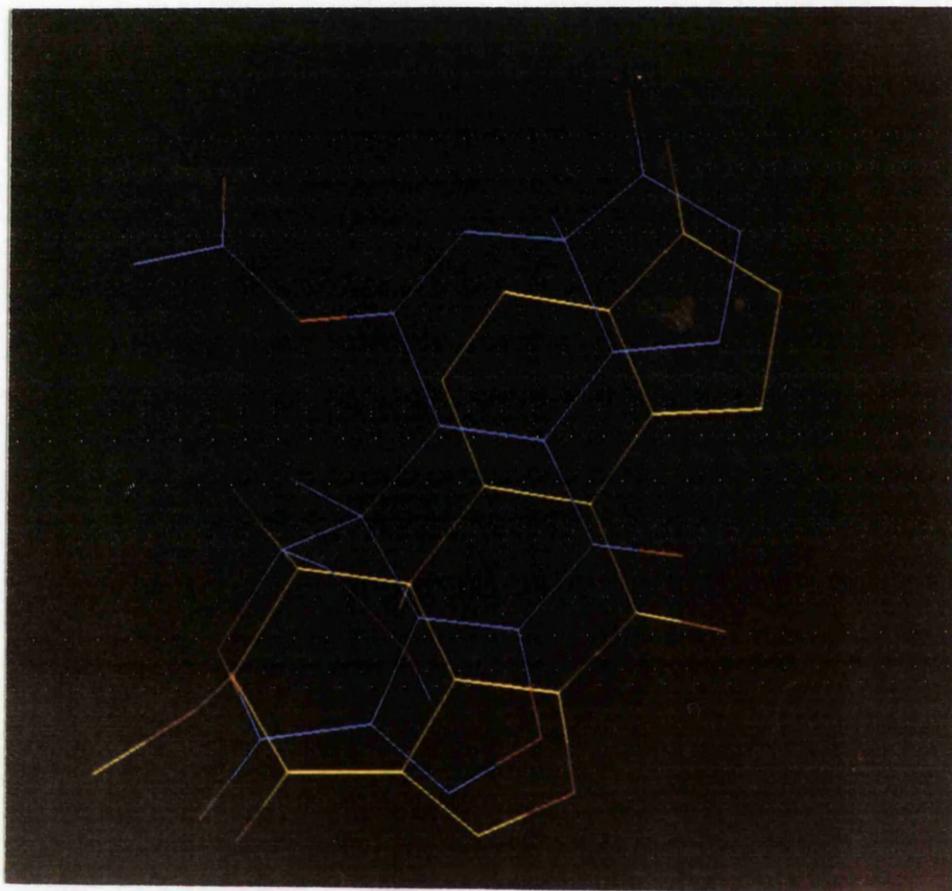


FIGURE 16 - Viridin (yellow) and wortmannin (blue) - face on and translated.

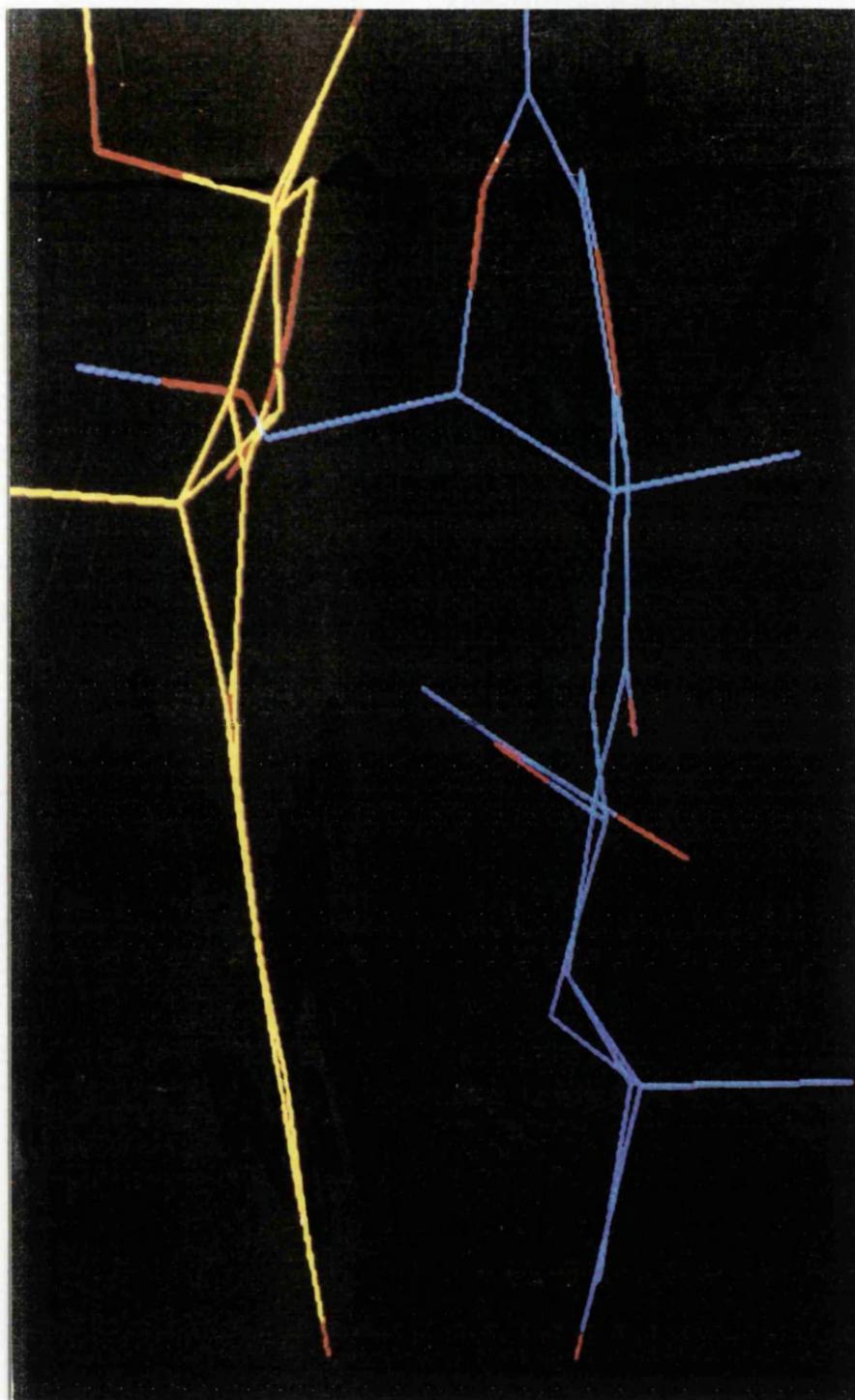


FIGURE 17 - Viridin (yellow) and wortmannin (blue) - Side view.

system is crucial for activity. The results of these experiments are shown in Table 5.

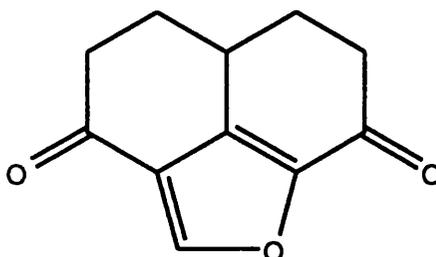


Figure 18 - Atoms overlapped and thought to be crucial for activity

Molecule 1	Molecule 2	Average displacement (Å)	Max. displacement (Å)
Viridin	Wortmannin	0.033	0.052
Viridin	Demethoxyviridin	0.013	0.018
Viridin	Hibiscone C	0.026	0.039
Wortmannin	Viridin	0.033	0.052
Wortmannin	Demethoxyviridin	0.024	0.044
Wortmannin	Hibiscone C	0.053	0.104
Hibiscone C	Viridin	0.026	0.039
Hibiscone C	Wortmannin	0.053	0.104
Hibiscone C	Demethoxyviridin	-	-
Demethoxyviridin	Viridin	0.013	0.018
Demethoxyviridin	Wortmannin	0.024	0.044
Demethoxyviridin	Hibiscone C	-	-

TABLE 5 - Results of overlap experiments

The average displacement is the average distance between corresponding atoms of molecule 1 and molecule 2 after overlapping according to the query structure in Figure 18. The maximum displacement is the greatest single

distance between these corresponding atoms. Thus these measurements are a crude measure of structural similarity.

Presuming demethoxyviridin is most potent, from test results (see later) and that the furan ring system is indeed necessary for activity, from the wortmannin analogue study, the results from Table 5 would predict a decreasing order of potency:

Demethoxyviridin

Viridin

Hibiscone C

Wortmannin

This assumes that the most structurally similar compound to the most active compound will be next in order of potency. Viridin is closest in structure to demethoxyviridin, as expected and seen from the results in Table 5, and hibiscone C is closer in structural distances than wortmannin to viridin.

Although activity data for the hibiscone C part structure analogue turned out to be disappointing, *i.e.* the compound was not at all active in the test systems used, demethoxyviridin was shown to be more active than viridin which in turn was more active than wortmannin.

Figure 19 shows an example of docked viridin and demethoxyviridin. It can be seen that although the two molecules are largely identical, when the ABE ring system is docked, rings C and D are slightly out of alignment. This difference in shape, in addition to the loss of a methoxy group, may account for the difference between the two molecules in activity at the receptor site.

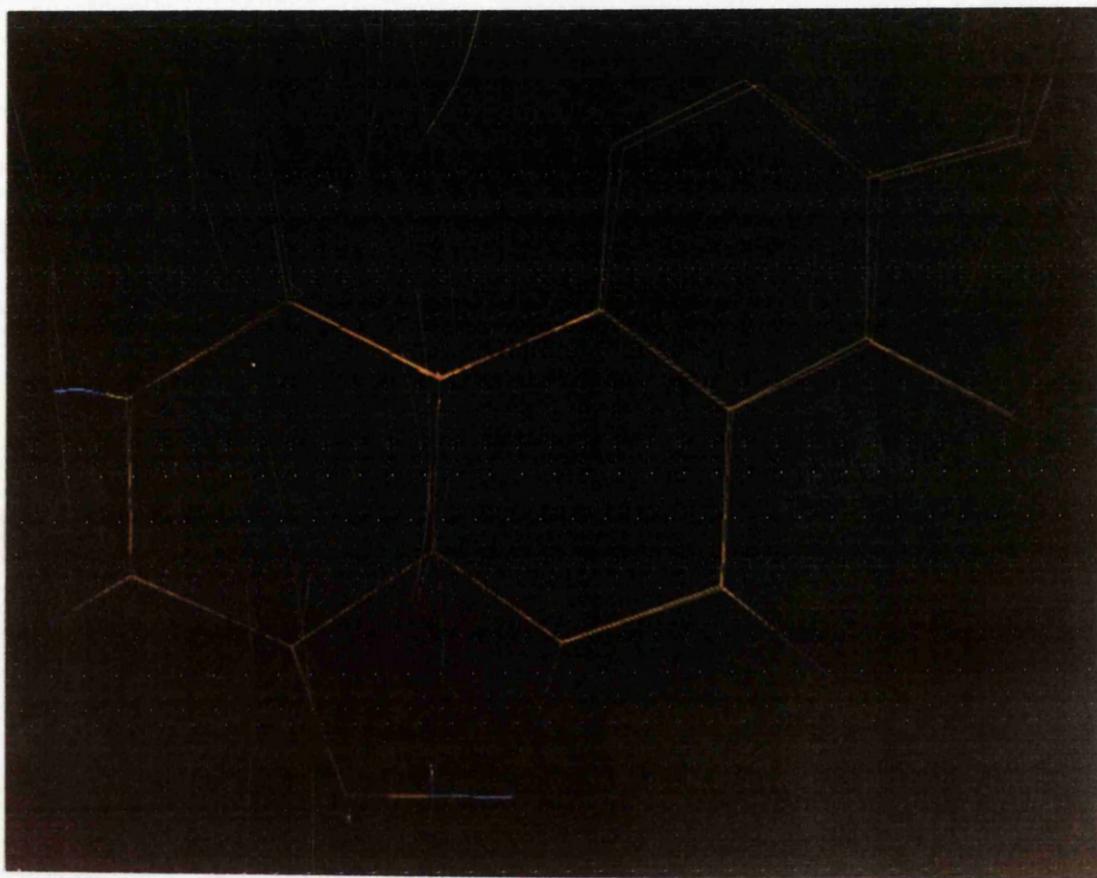


FIGURE 19 - Overlap of viridin and demethoxyviridin.

By loading an appropriate symmetry file, the molecules could be displayed as crystal packing diagrams (figure 20) where the regular zigzag arrangement of the structure can be seen, perhaps mimicking interactions taking place at the receptor site. The electron density of the molecule can also be generated by creating dot surfaces (figure 21) and space filling models (raster diagrams) (figure 22) can be depicted using the program HYDRA.

Using enlarged versions of crystal packing diagrams and distance calculations, possible hydrogen bonds could be identified between molecules in adjacent unit cells. These could be represented pictorially using a graphics terminal linked to the main VAX and the program ORTEP.

Figure 23 shows the unit cell of demethoxyviridin. The distances between atoms O(20') and O(24'), the hydroxyl and the carbonyl groups, was calculated to be 2.83Å, indicative of possible hydrogen bonding between these two atoms. A distance of 2.83Å was also calculated between O(18') and O(20) thus these molecules lie in a zigzag arrangement (seen in Figure 24) linked by two hydrogen bonds. This arrangement may be important as it may mimic the way in which the molecules approach the receptor site. Viridin, unit cell shown in Figure 25, is linked by a single hydrogen bond. The carbonyl groups of the AB rings do not seem to be involved in bonding in this molecule but a similar zigzag arrangement is seen (figure 26). There are no apparent hydrogen bonds in the unit cell of wortmannin (figure 27) and the combined unit cell arrangements show the molecules lining up in a pairwise manner (figure 28). Hibiscone, similarly, does not show any signs of obvious H-bonding (figure 29) but does show a slight zigzag pattern in the crystalline form (figure 30).

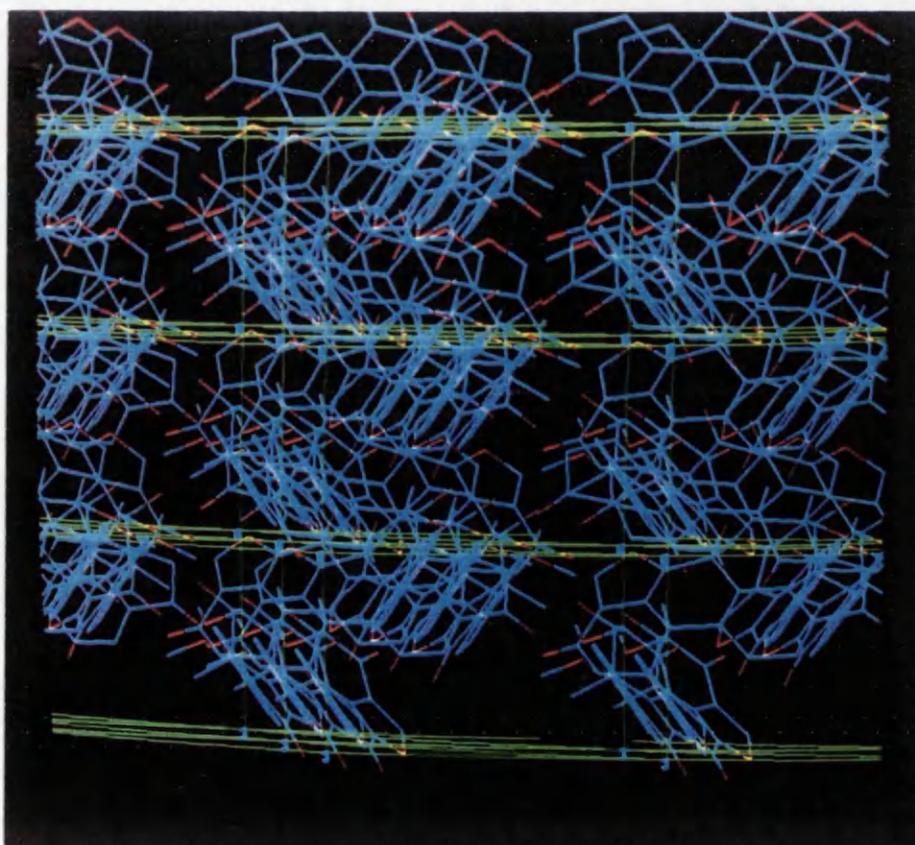
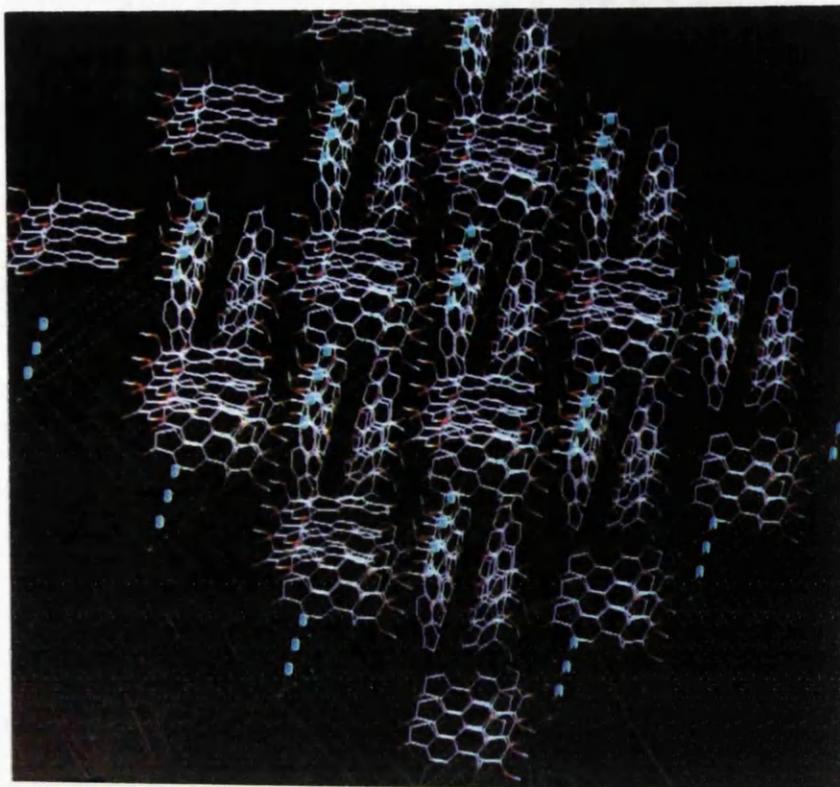


FIGURE 20 - Crystal packing diagrams of viridin (above) and wortmannin (below).

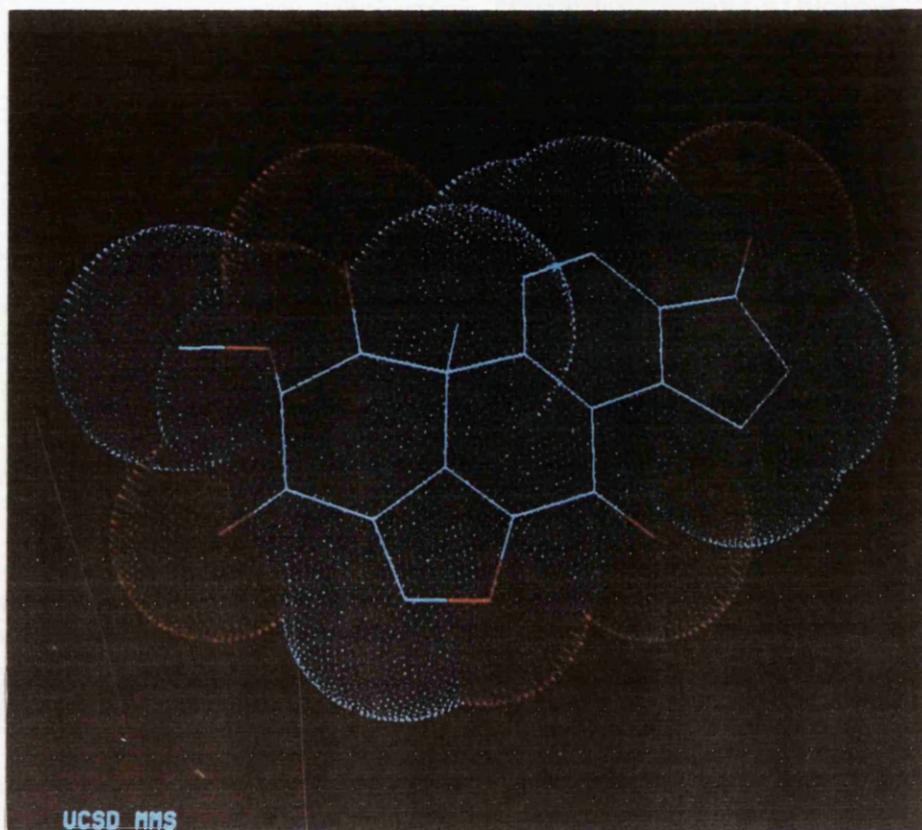


FIGURE 21 - HYDRA generated Van der Waals surface display of viridin.

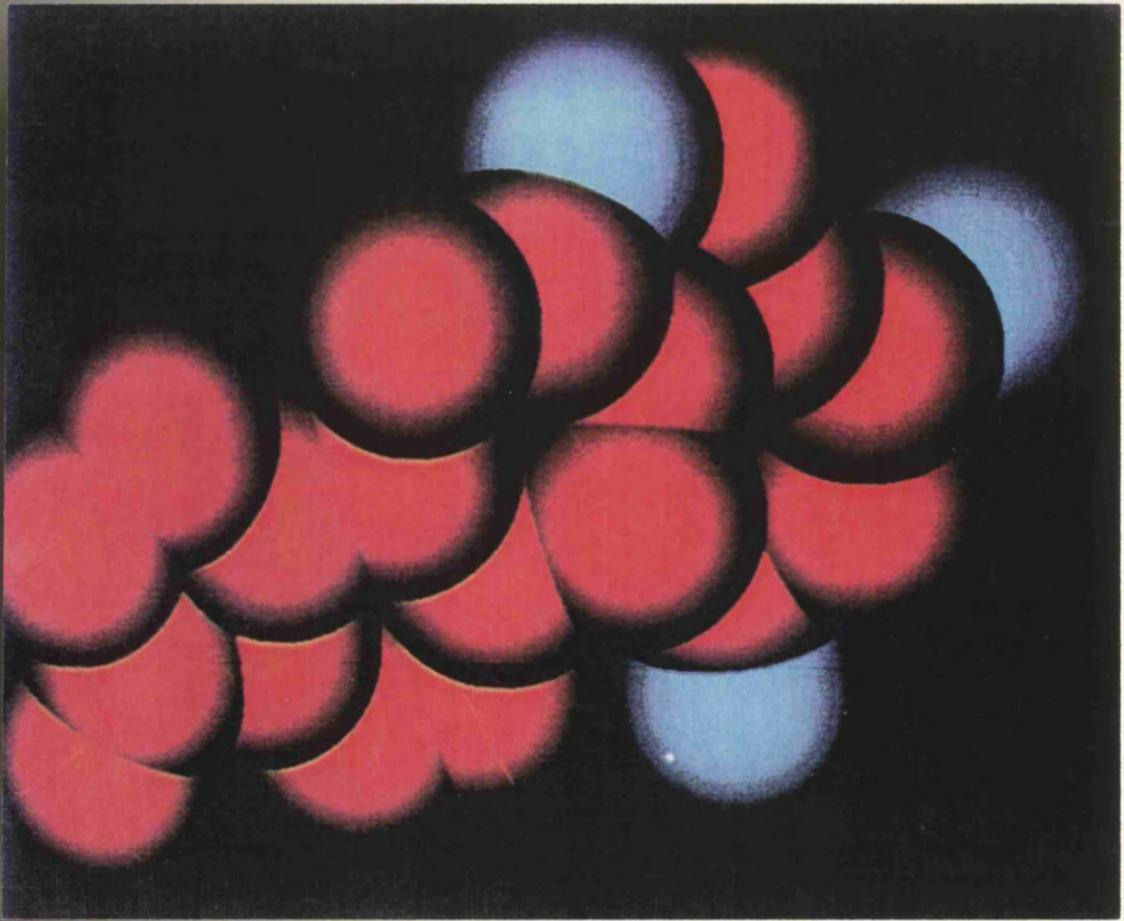


FIGURE 22 - HYDRA generated raster display of viridin.

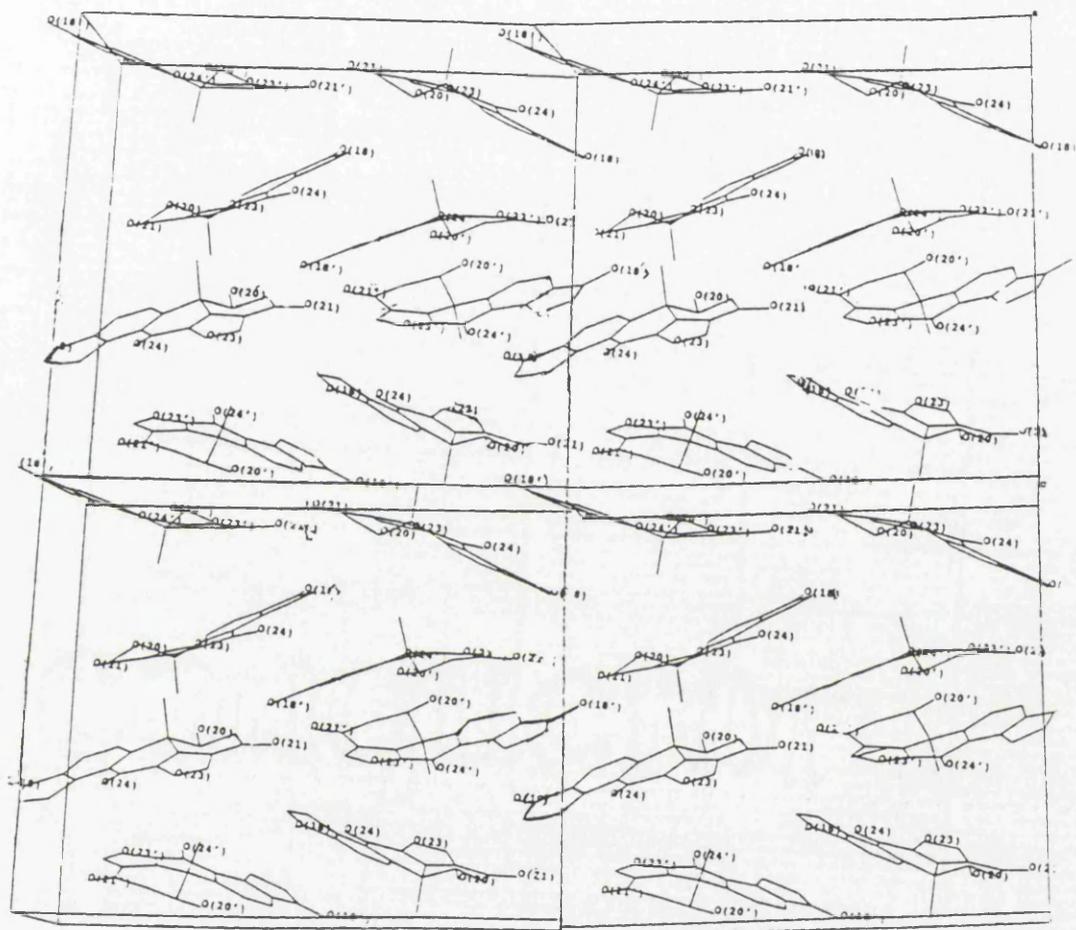


FIGURE 24 - Repeated unit cell of demethoxyviridin.

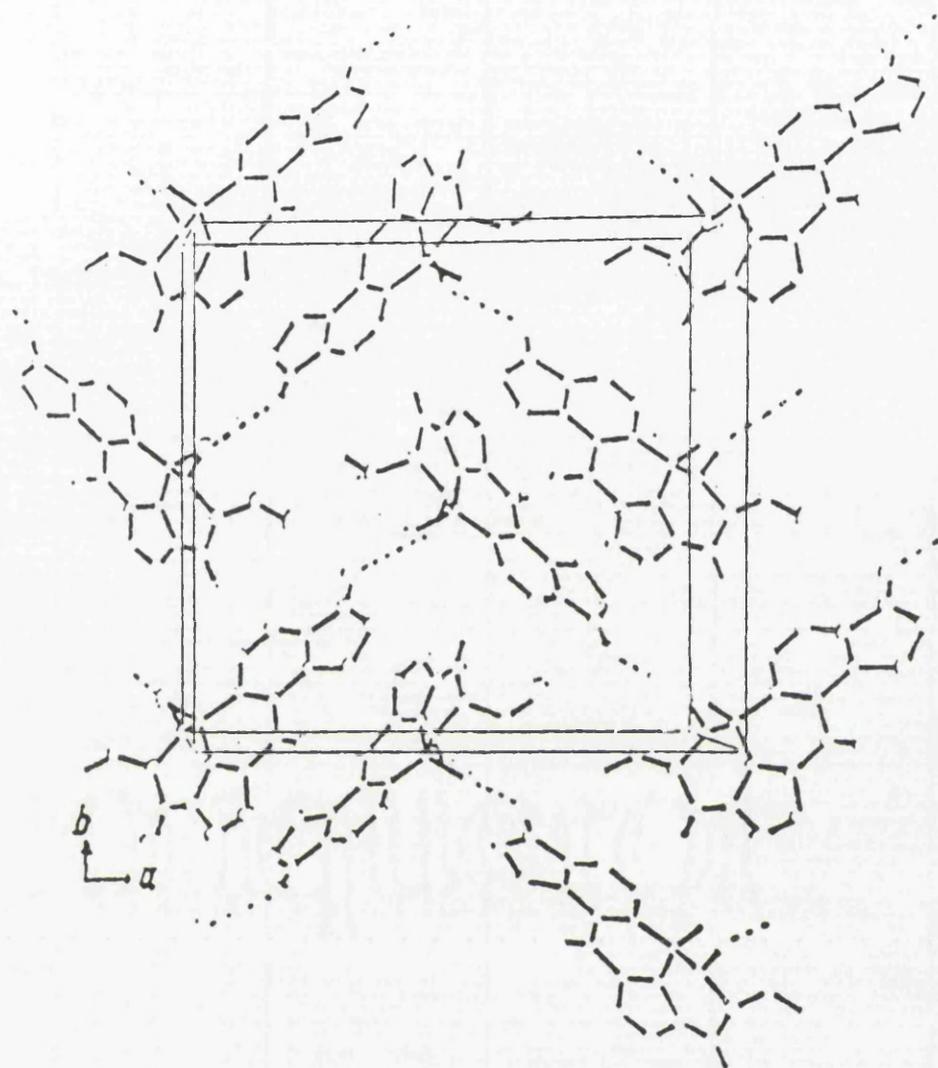


FIGURE 25 - Unit cell of viridin.

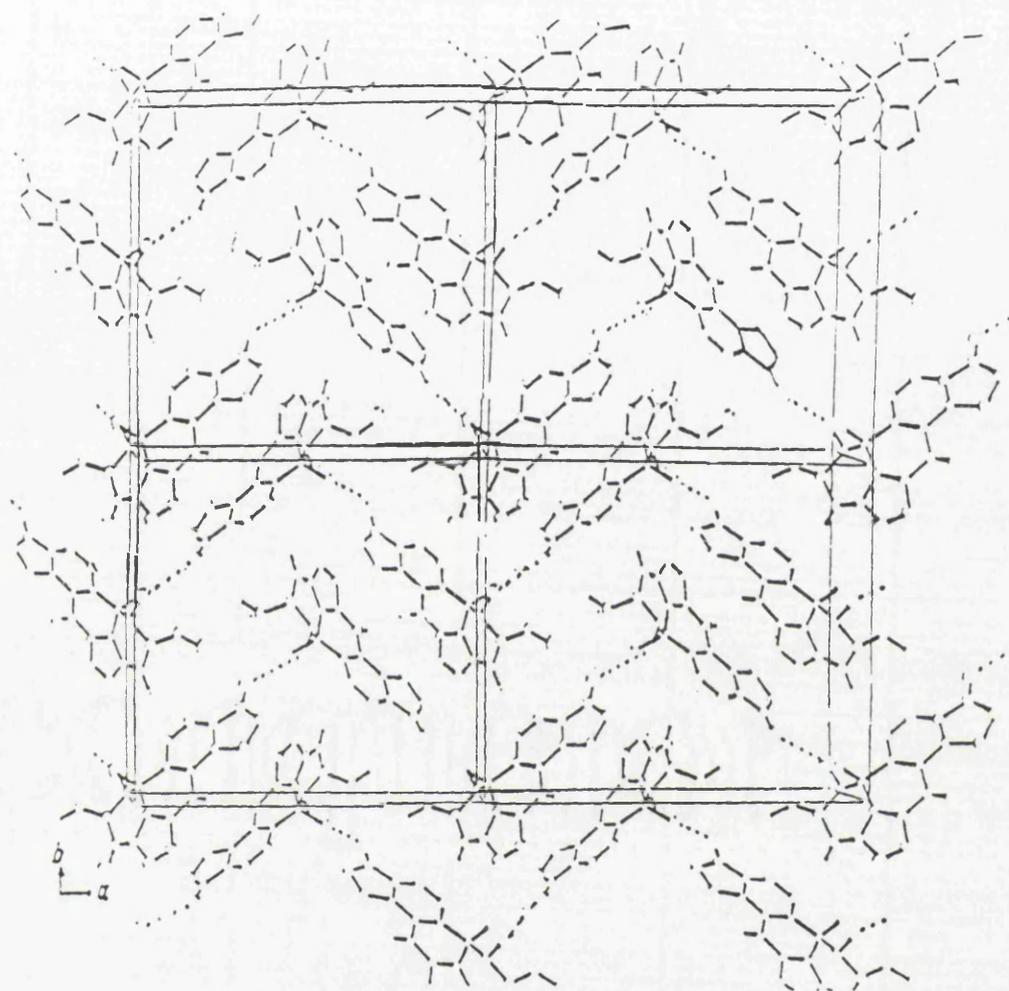


FIGURE 26 - Repeated unit cells of viridin.

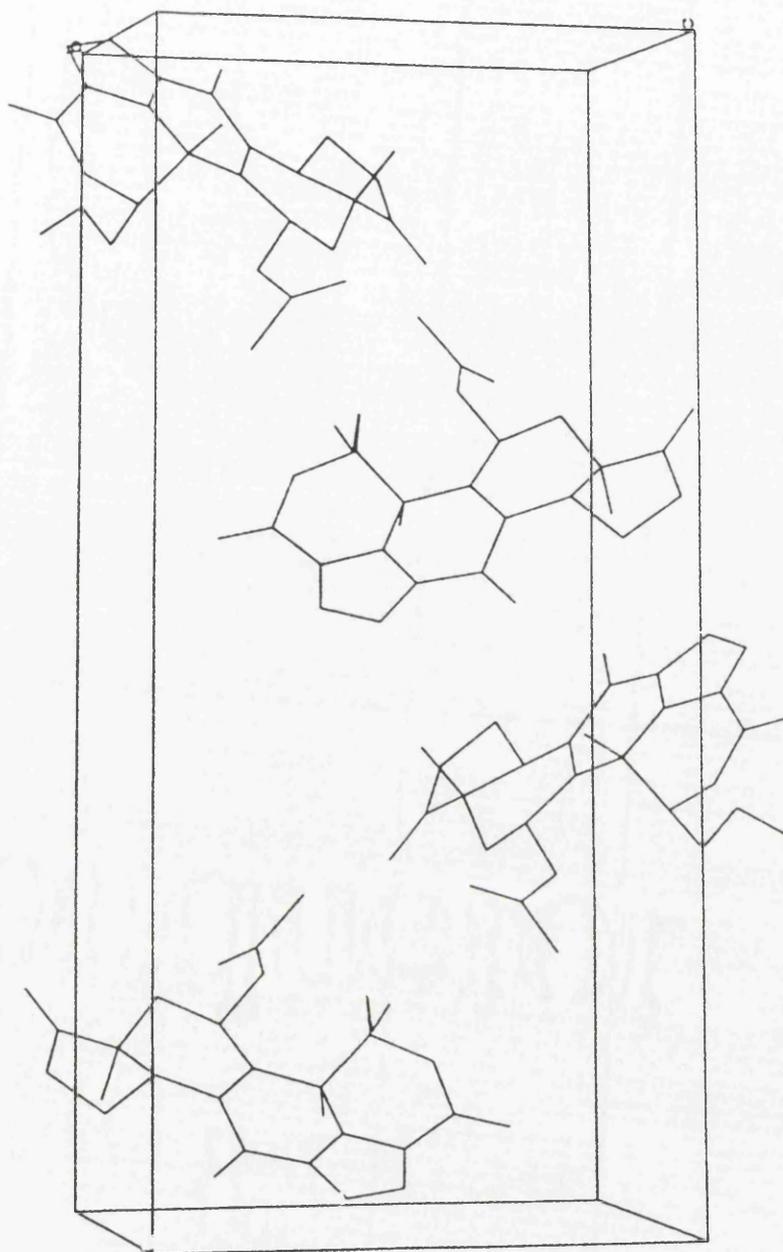


FIGURE 27 - Unit cell of wortmannin.

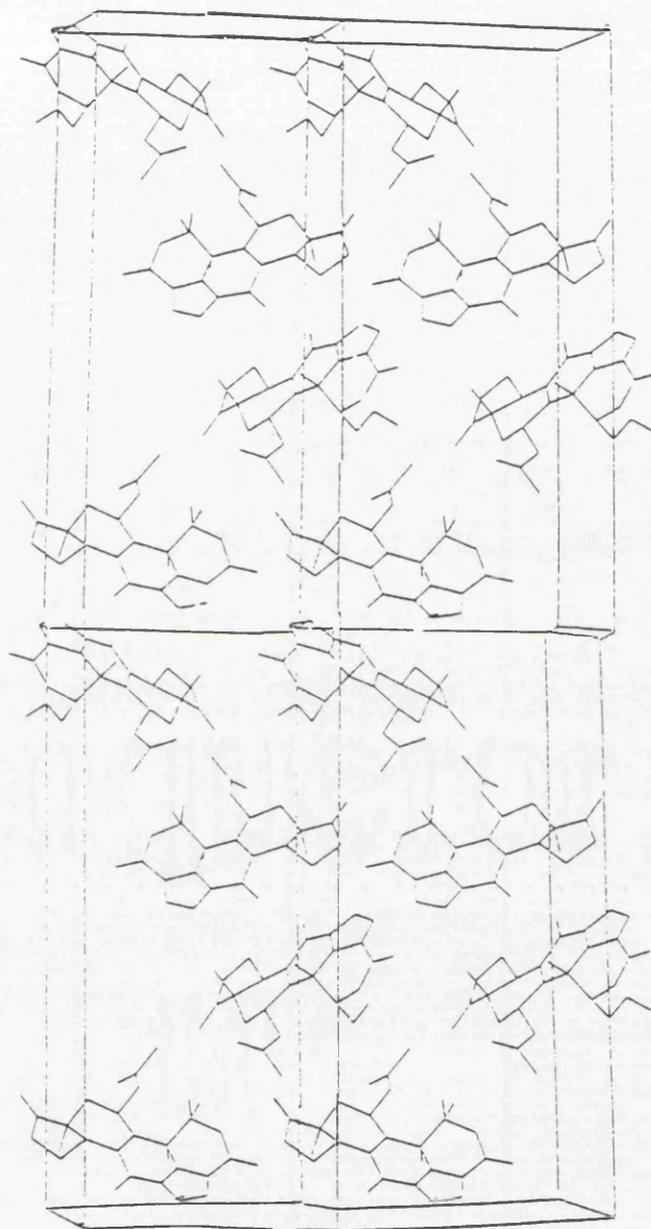


FIGURE 28 - Repeated unit cells of wortmannin.

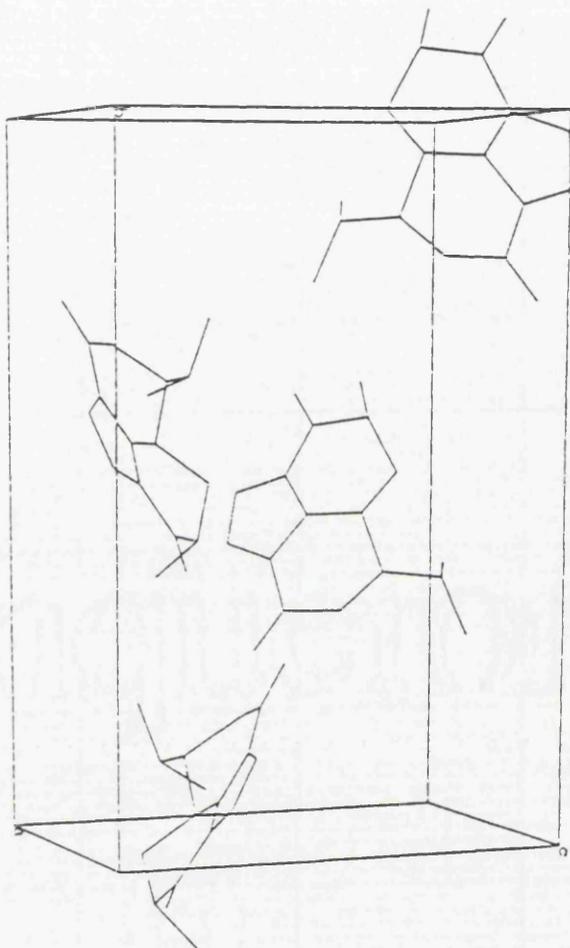


FIGURE 29 - Unit cell of hibiscone C.



FIGURE 30 - Repeated unit cells of hibiscone C.

STATISTICAL METHODS IN QSAR

One of the most important considerations in QSAR studies is the statistical analysis of the correlation of the observed biological activity with structural parameters. The coefficients of these parameters that establish the correlation with the biological activity can be obtained by a regression analysis. Since the models are constructed in terms of multiple additive contributions, the method of solution is also called multiple linear regression⁵¹. This method is based on three requirements:

- (i) the independent variable (structural parameters) are fixed variates and the dependent variable (biological activity) is randomly produced,
- (ii) the dependent variable is normally and independently distributed for any set of independent variables,
- (iii) the variance of the dependent variable must be the same for any set of independent variables.

The biological activity satisfies the second and third requirements, since each measurement is done independently. Upon repeated measurements the values will be normally distributed, and will have an equal variance. In the usual situation, m observed values are fitted to an equation with n variables and p parameters by a least squares procedure.

This procedure defines the best fit by requiring that the sum of differences between the observed and the calculated values can be minimal. Goodness of fit can then be defined by three quantities:

- (i) the multiple correlation coefficient, r ,
- (ii) the standard deviation of regression, s ,
- (iii) the overall test F value for the test of coefficient significance.

The correlation coefficient gives an indication of the correspondence of observed and calculated values. Its square, r^2 , can be interpreted as the fraction of variance in the observed data which is explained by the correlation. The standard deviation of the regression is a measure of the scatter of the experimental values from the mean. The overall F value indicates how much better is the correlation (within a certain degree of significance level) from the correlation in which all the coefficients are taken as zero.

PYTHON

Initial studies were done using the program PYTHON available from Oxford Molecular and data from a paper publishing results of anti-inflammatory data of a series of wortmannin analogues²⁹. This paper describes how the measure of inhibition of the respiratory burst in cells can be used as a direct measure of the anti-inflammatory potency of a compound. The program was used in York as part of a molecular graphics course. A spreadsheet was prepared using initially seven of the wortmannin analogues (see figure 31).

The columns entered included the log of the given IC₅₀ value (log A), indicator variables, Verloop values (i.e. a measure of the size and direction of each substituent), lipophilicity (Π) -related to log P, the latter two values being calculated automatically by the program.

Several graphs were plotted. No direct correlation was seen with activity and Verloop values or initially with lipophilicity so two rows were excluded from the table as they were structurally more different than the others. These five points were plotted and analysis done yielding the equation:

$$y = 0.077x_1 + 1.059x_2 + 0.913x_3 + 1.855$$

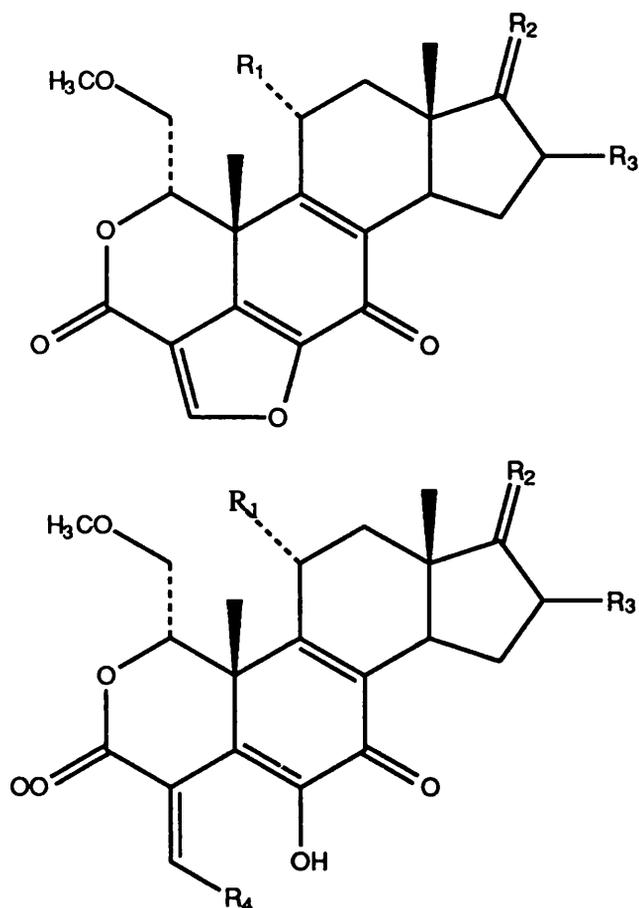
where $y = \log$ activity

$x_1 =$ indicator variable 1

$x_2 =$ indicator variable 2

$x_3 = \Pi$ (lipophilicity)

(Indicator variable 1 was the presence of an acetoxy group on R₁ and indicator variable 2 was the presence of a substituent other than H on R₃.)



Molecule 1	[1]	R ₁ =OAc	R ₂ =O	R ₃ =H	R ₄ =closed ring
Molecule 2	[12]	R ₁ =OAc	R ₂ =O	R ₃ =OAc	R ₄ =closed ring
Molecule 3	[13]	R ₁ =OAc	R ₂ =O	R ₃ =H	R ₄ =NHC ₆ H ₅
Molecule 4	[14]	R ₁ =H	R ₂ =O	R ₃ =H	R ₄ =closed ring
Molecule 5	[15]	R ₁ =H	R ₂ =O	R ₃ =H	R ₄ =OH
Molecule 6	[16]	R ₁ =OAc	R ₂ =OH	R ₃ =H	R ₄ =closed ring
Molecule 7	[17]	R ₁ =H	R ₂ =OH	R ₃ =H	R ₄ =closed ring

Figure 31 - Seven wortmannin analogues whose known anti-inflammatory data is used in a initial model using the program PYTHON. Molecules 3 and 5 were eventually discarded from the study due to significant structural differences i.e. the closed furan ring was seen to be critical for activity and a valid model.

This equation gave an R value of 0.989 *i.e.* the r^2 value was around 0.85 which meant that 85% variance was explained by this equation. The F value was 14.85 with 3 degrees of freedom.

From this PYTHON model activities could be predicted (see Table 6).

MOLECULE	ACTUAL IC₅₀	PREDICTED IC₅₀
1	1.447	1.377
2	2.436	2.436
4	1.892	1.962
6	1.230	1.300
7	1.954	1.885

TABLE 6 - Predictions of activity from PYTHON results.

These initial results were very exciting in that they showed that a valid statistical model could be formed for this set of compounds and could possibly be used to predict the activities of as yet unsynthesised and untested compounds. Both the order and value of predicted potency were correct within statistical limits.

SYBYL

The program PYTHON was only available during the course of a one week training seminar at York so the same theories were applied to the more sophisticated methodology and graphics equipment available at Glasgow. The program used was SYBYL by Tripos and the main methods used included CoMFA techniques and general QSAR methodology.

A major goal in chemical research is to predict the behaviour of new molecules, using relationships derived from analysis of the properties of previously tested molecules. Relationships derived primarily by empirical analysis of a data table, whose compounds are numerical property values and whose rows are compounds, usually taking the form of a linear equation, are called quantitative structure activity relationships (QSAR).

Especially in biological applications, it has long been agreed that the most relevant numerical property values would be shape dependent. Work on comparative molecular field analysis (CoMFA) began 12 years ago⁵² with two additional observations: (1) at the molecular level, the interactions which produce an observed biological effect are usually **non-covalent**; and (2) molecular mechanics force fields, most of which treat non-bonded covalent interactions only as steric and electrostatic forces, can account precisely for a great variety of observed molecular properties. Thus it seems reasonable that a suitable sampling of the steric and electrostatic fields surrounding a set of ligand or drug molecules might provide all the information necessary for understanding their biological properties. However, the emergence of a practical CoMFA methodology had to await a new method of data analysis, partial least squares (PLS), which can derive robust linear equations from

tables having many more columns than rows and a number of advances in the methodology of molecular graphics.

One of the studies using CoMFA was published in 1992 by a group of workers at Sterling Winthrop in New York⁵³. They showed that the CoMFA analysis of eight compounds related to disoxaril for which X-ray structures of the molecule bound to HRV-14 had been determined resulted in a strong positive correlation of activity with steric effects of the compounds, particularly toward the pore end of the compound binding site, and no correlation with electrostatic effects. These results confirm what had been previously found, that the activity of these compounds was highly dependent upon their hydrophobic nature expressed as log P or lipophilicity. The CoMFA study also confirmed results from the comparison of a series of active and inactive compounds using volume maps which showed that bulk at the pore end of the molecule was conducive to high levels of antiviral activity while excessive bulk around the ring led to poor activity.

Other 3D QSAR methodologies have been described. The molecular shape (MS) approaches, developed independently by Simon *et al*^{54,55} and by Hopfinger⁵⁶, compare net rather than location-dependent, differences in molecular connectivities, volumes and/or fields. A second approach, the distance geometry method of Crippen⁵⁷, provides validation of a site-point hypothesis, a list of binding set co-ordinates and properties that must be proposed by the investigator. In related work, for exploring binding modes of ligand to receptors, Goodford advocates the display of probe-interaction 'grids', similar to those used in CoMFA⁵⁸, while Hansch, Blaney, Langridge *et al*⁵⁹ have shown the complementarity of QSAR and molecular graphics in understanding enzyme inhibitory data.

CoMFA (Comparative Molecular Field Analysis) is a technique used only in the SYBYL program. It is a method which uses multivariate regression techniques to seek relationships between properties such as biological activity and various 3D structural features such as molecular shape and charge distribution.

Critical to a preparation of a CoMFA analysis is the building of a database of compounds similarly aligned and with their electron density and charge distribution calculated. Once minimised conformations of all the structures are in hand, the next step in the CoMFA process is to establish a mutual spatial alignment of the molecules within a series. This can be done by computing the fields associated with each of the molecules and then minimising the RMS differences between them by adjusting their positions (using the 'Field Fit' technique in SYBYL). With the wortmannin and viridin structures, however, the structural similarity was such that the molecules were aligned carefully by hand, using a process that seemed more intuitive and overlapping the ABE fused ring system of each compound.

The non-bonded energy fields associated with any structure are determined in CoMFA by first placing the structure in a 3-D lattice and then computing the steric (Lennard Jones) and electrostatic interactions within the molecule that are experienced by a probe atom⁶⁰ e.g. C⁺ at all of the lattice points around the molecule. A lattice grid spacing of 2Å in the x, y and z directions was used. These data are built into a CoMFA table which has 1 row per 'standard' compound and hundreds or thousands of columns to carry all the CoMFA energy field data. The statistical task at this stage is to establish by regression analysis over the whole set of compounds a relationship between these variables on the one hand, and the biological activity data on the other hand.

The biological activity data were entered in the CoMFA table as $1/\ln IC_{50}$ values.

Due to the very large number of independent variables in this system, the standard techniques of multiple regression cannot be applied directly and Partial Least Squares (PLS) analysis is employed to derive linear equations from the data. The quality of any model derived can be determined by a cross validation technique, sometimes referred to as the 'leave-one out' method. It measures the true predictive capability of a model by removing a subset ('cross-validation group') of one or more structures from the set of standard compounds ('the training set'), re-entering them as unknowns and predicting their activity. The sum of the squared differences between predicted and actual activities is compared to the standard deviation of the actual activities to yield a 'cross-validated r^2 ' as a measure of predictivity. This value can range from negative numbers to a maximum of +1.0. A cross-validated r^2 of above 0.4 is said to be statistically relevant and one of 1.0 corresponds to a perfect model, i.e. all predictions are 100% accurate. A value of 0 indicates that no statistical relationship exists between the CoMFA fields and the activities, and the predictions are no better than random numbers whose average is the mean of the measured activities. Negative values result from an inverse relationship between the actual and predicted activities and can be regarded as worse than no model at all.

In all cross-validation runs performed in the course of these studies, the number of cross-validation groups was chosen to equal the number of structures in the CoMFA table, *i.e.* the activity of each structure was predicted individually from all other structures. This guarantees reproducibility and comparability of subsequent cross-validation runs. The standard deviation threshold for exclusion of columns from the PLS analysis

(MINIMUM_SIGMA) was set to 0.5 for cross-validation runs, and to 0.0 for the final non-crossvalidated analyses.

CoMFA models can be built using a number of components in the linear equation derived by the PLS analyses. This number corresponds to the level of detail at which the PLS analysis has been conducted on the molecular fields, and is analogous to the number of coefficients in conventional multiple regression. Up to $n-1$ components can be used, with n the number of structures in the CoMFA dataset. In the crossvalidation runs, a crossvalidated r^2 value is calculated for each value π in the specified range of components ($\pi = 1..m$, $m < n-1$). CoMFA QSAR models which allow the prediction of the activity of novel compounds are derived from (final) non-crossvalidated runs. For these models, one specific number of components had to be chosen. A number is usually chosen which is lower than the one yielding the highest cross-validated r^2 , since the number of components at which the usually steep initial increase of the crossvalidated r^2 starts to level off has proven to yield better predictive CoMFA models. Higher numbers of components tend to 'overfit' the CoMFA model to irrelevant minutiae of the molecules in the dataset and may lead to less accurate predictions.

SYBYL - EXPERIMENTAL

While no actual cytotoxicity test results were yet available, a model was formed using the wortmannin analogues previously synthesised. This was based on the assumption that anti-inflammatory data would be related in some way to cytotoxicity data, and also comparable initial test result data showing that both wortmannin and viridin had potent cytotoxic potential.

The wortmannin analogues were built in SYBYL and energy minimised. Each molecule was added to the same database. Del-re charges⁶¹ were calculated for each molecule which was then aligned by hand using the virtual dial system so that a good 3D alignment was easily possible. Each molecule was held in this orientation by the option FREEZE_VIEW and re-added to the database. A new spreadsheet was made and each molecule to be in the CoMFA study was added. From the menu option in this spreadsheet, the alignments of the molecules were defined by the command sequence CoMFA ALIGNMENT DEFINE AS_IS MOLECULE_NAME MOLECULE_ALIGNMENT_NAME END. This defines each molecule to be critically aligned as the frozen conformation and allows the flexibility of different conformations if desired. The option AUTOCOMFA is then selected which automatically calculates the electrostatic and steric fields for each molecule, stores this information in the spreadsheet where it can be EXTRACTed and analyses the results by multivariate analysis. A good initial estimate of whether the analysis is successful or not is the critical value of r^2 which should be greater than 0.4.

An initial study was carried out using a selected group of compounds ([1], [12], [13], [14], [16], [17]) which were reasonably structurally similar. These

were placed in a table and an automatic CoMFA run was carried out. The resulting electrostatic and steric map can be seen in Figure 32.

The r^2 value for this analysis was 0.611. The relative contributions from each variable are shown in Table 7:

	Norm. Coefficient	Fraction (%)
1. Autocomfa (STERIC)	0.830	58.2
2. Autocomfa (ELECTROSTATIC)	0.597	41.8

TABLE 7 - Relative contributions from steric and electrostatic variables.

From the electrostatic map shown in blue and yellow (Fig. 32), the yellow regions represent areas where increased negative charge would lead to increased activity, and the blue regions represent regions where increased positive charge would lead to increased activity. The most strikingly apparent result from this study was the suggestion that a chemical group, e.g. containing a nitrogen or other atom which could easily protonate, substituted on the acetoxy group on ring C of wortmannin would lead to increased activity. The steric map is usually shown in red and green where the red regions represent areas where less bulk would lead to increased activity. In this case the green regions were more apparent, showing areas where increased bulk would lead to increased activity, in particular the acetoxy group mentioned above and the carbonyl substituent on ring D. Substitution at these two groups was therefore seen as a priority for the organic chemists working separately on this project, providing a clear lead to be assessed. Similar studies using molecules 1-10 with extremely careful alignment gave

an r^2 value of 0.624 (final non crossvalidated $r^2 = 0.957$) (see later). The results of each study can be easily validated by a number of different techniques such as **bootstrapping, crossvalidation, factor analysis and partial least squares.**

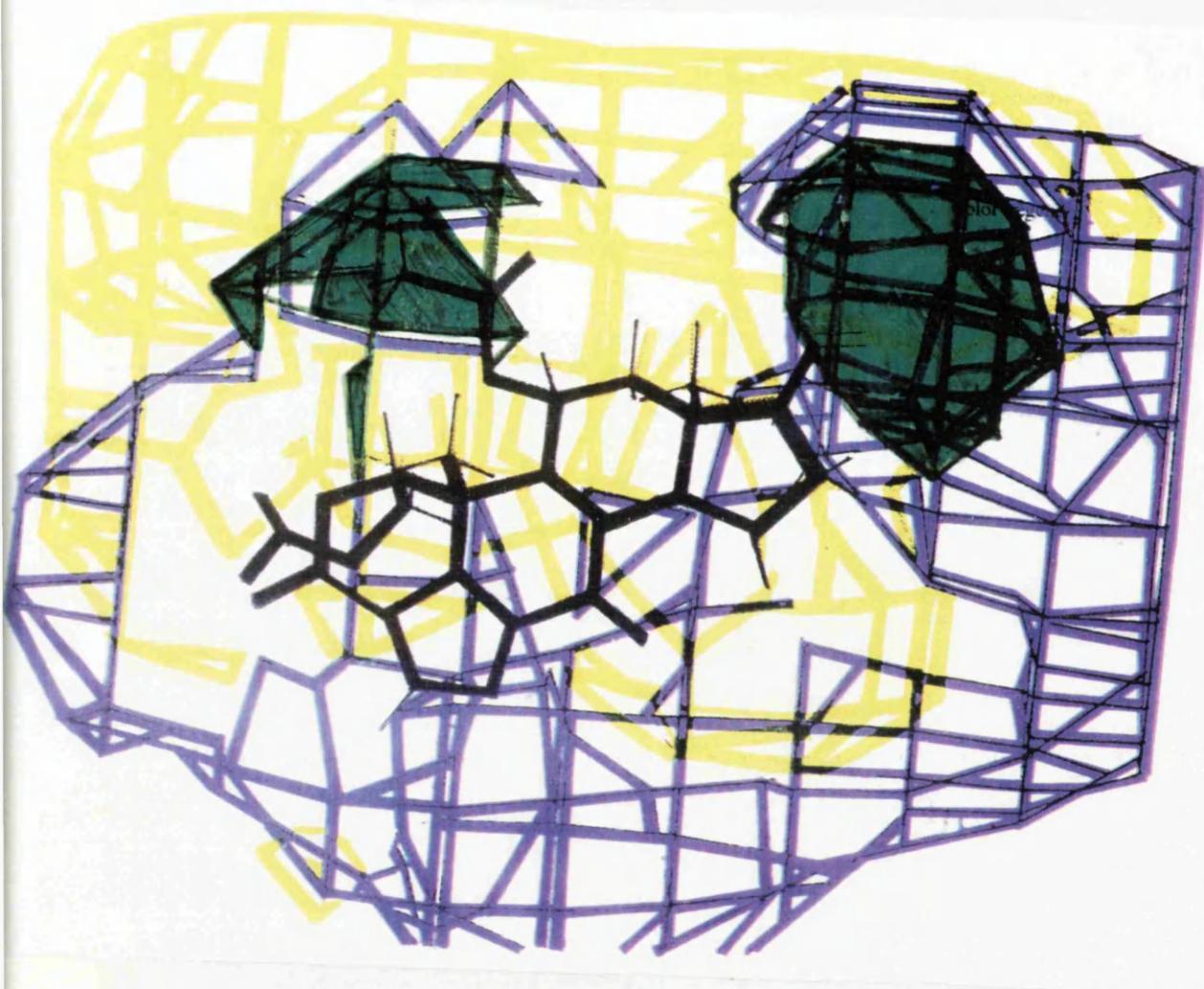


FIGURE 32 - Electrostatic (blue and yellow) and steric (green) maps from wortmannin CoMFA study.

Blue - increase positive charge

Yellow - decrease positive charge

Green - increase bulk

BOOTSTRAPPING

In many SYBYL/QSAR analyses, confidence intervals (mean and standard deviation) for the parameters to be estimated can be calculated by a modern validation method, the bootstrap. The name is derived from the old saying about pulling yourself up by your own bootstraps. The idea is to simulate a statistical sampling procedure by assuming that the original data set is the true population and generating many new data sets from it. These new data sets (called bootstrap samplings) are of the same size as the original data and are obtained by randomly choosing samples or rows from the original data, where repeated selection of the same row is allowed. The statistical calculation is performed on each of these bootstrap samplings, new values being calculated for each of the parameters to be estimated. The difference between the parameters calculated from the original data set and the average of the parameters calculated from the bootstrap samplings is a measure of **bias** of the original calculation. The calculated variance of the parameter estimates the **accuracy** with which any of the parameters can be estimated from the input data. This process is shown in figure 33⁶².

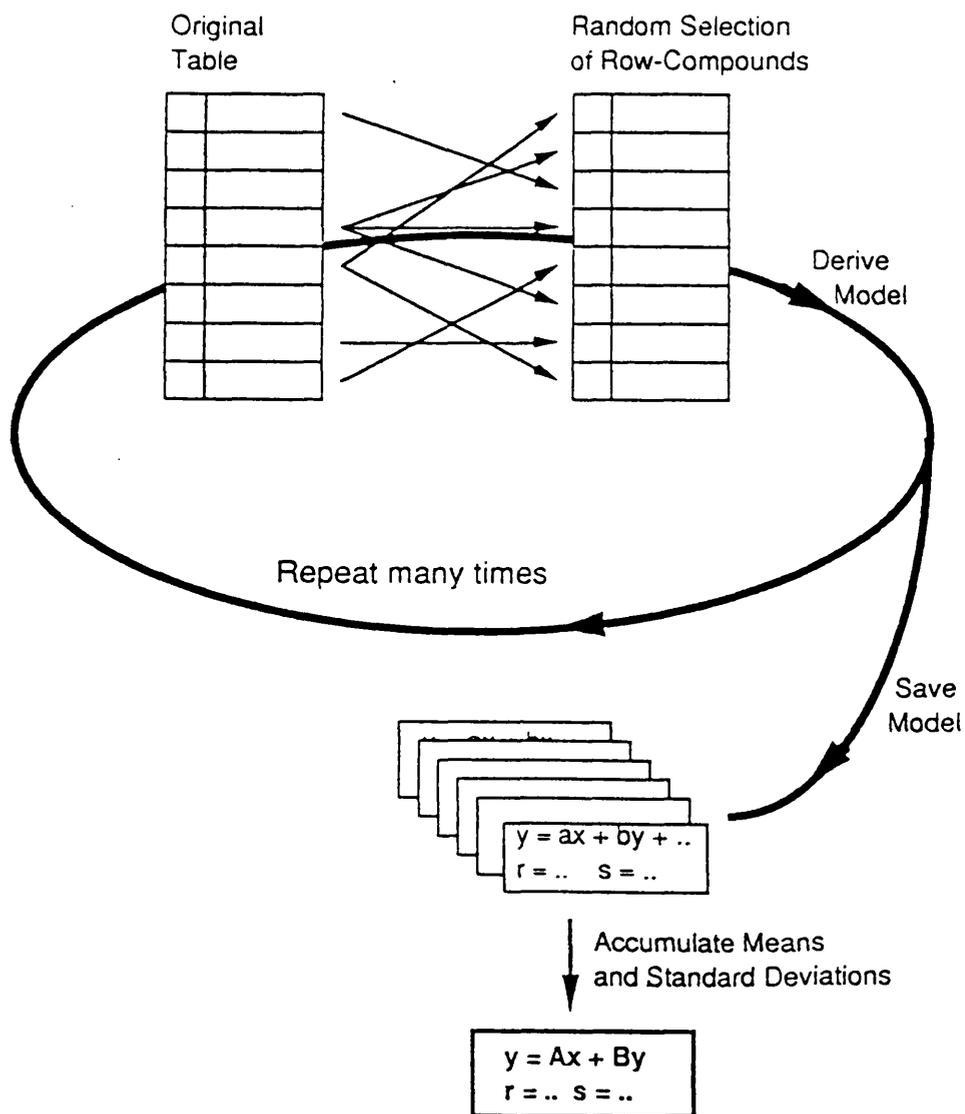


FIGURE 33 - The bootstrapping process

CROSSVALIDATION

Crossvalidation is an approach for selecting which model, among several with different levels of complexity, is most likely to have **predictive value**. It is particularly useful in PLS (Partial Least Squares), to establish the number of components which optimally distinguish signal from noise.

In contrast with the classical statistical techniques such as the F-test, crossvalidation substitutes modern computational power for the theoretical assumptions about data distributions. Crossvalidation assesses the probable predictive power of a particular model by the brute force method of attempting predictions of all the input target data values.

As shown in figure 34²³, crossvalidation proceeds by omitting one or more rows of input data, re-deriving the model, and predicting the target property values of the omitted rows. The rederivation-and-prediction cycle continues until all target property values have been predicted exactly once. The root mean square error of all target predictions, the **press (predictive sum of squares)**, is the basis for evaluating the current model.

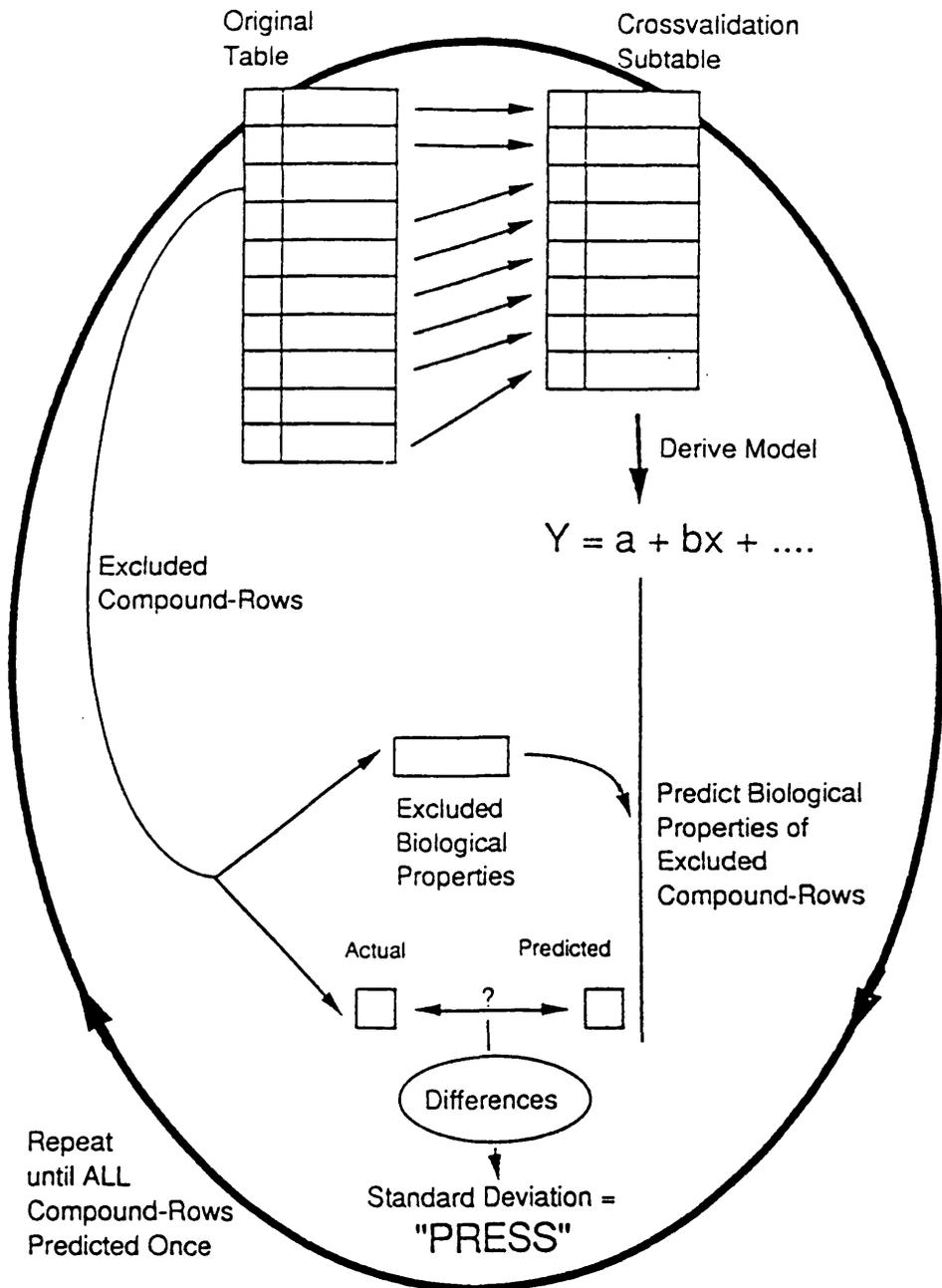


FIGURE 34 - The crossvalidation process

In SYBYL/QSAR, the intensity of the crossvalidation process is controlled by selecting the number of groups, or the number of times the crossvalidation process is to be carried out while predicting all rows (at each stage of model development). If the number of groups is set to zero, crossvalidation is not performed. With the minimum number of groups, two, compounds will be divided into two groups, as equally sized as possible, and a model derived from each half will be used to predict target property values for the other half. The number of groups typically used in these studies is five, meaning that each target property value will be predicted by a model based on about 80% of the available data.

Currently, crossvalidation group membership is random, and cannot be controlled or even observed by the SYBYL/QSAR user. This randomness means that values of press and hence crossvalidated r^2 will vary from run to run. Only by setting the number of crossvalidation groups equal to the number of compounds, so that every target property is predicted from all the other data, will these values be constant in consecutive runs.

The PRESS values are derived for models of increasing complexity, beginning with a 1-component model and computing up to a n-component model (n is set by the COMPONENTS option in PLS. An optimal number of components is reported as the number giving the smallest PRESS value. In practice, however, models with fewer components are more robust, and thus are to be preferred to models giving small decreases in PRESS. The final model was always obtained with the chosen number of components without crossvalidation, in order to make use of all available data.

FACTOR ANALYSIS

Factor analysis describes a family of techniques for detecting simple linear structures by performing a multidimensional rotation on an arbitrary table of data. In the chemical context of a molecular modelling, the process can be seen fairly easily. For example, if you look at an arbitrarily orientated benzene molecule shown in figure 35, and consider that since benzene is planar, there must be some 3D rotation/translation operation which will make all its z-coordinates zero. Factor analysis identifies this rotation, and the other panels of figure 35 show graphs of the three main types of data produced by such a factor analysis. The upper right panel plots the scores: the benzene atomic coordinates after simplification by factor analysis. The upper left panel plots the **loadings**: the elements of the rotation matrix which will convert the scores back into the original data. The lower left panel shows a bar graph of **eigenvalues** from this factor analysis, containing one line for each column of input data. Two of the lines are large and the third is miniscule, indicating that benzene is a flat molecule.

The aim of principal factor component analysis (FACTOR without the rotation of axes) is to find a new orthogonal co-ordinate system so that the sample variances with respect to these axes are maximised in decreasing order. In other words, the first new co-ordinate system describes the maximum variance among all possible directions, the second describes the next largest variance among all directions orthogonal to the first one, etc.

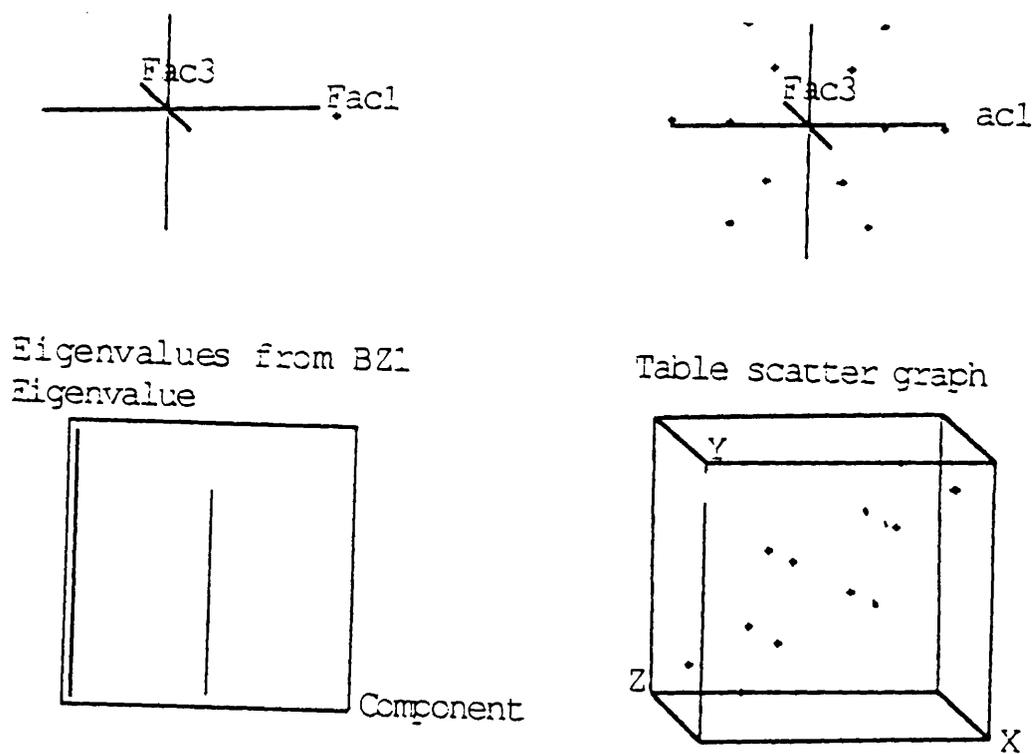


FIGURE 35 - Factor analysis of benzene co-ordinates.

The success or failure of this operation is reported as a series of eigenvalues, the fraction of total variance orthogonal to each of the successive new factors, conventionally scaled so that the total of the eigenvalues equals the number of eigenvalues. If a much simpler data structure exists, the sum of the first few eigenvalues will be nearly equal to the number of eigenvalues. The ratio of this sum to this number is conceptually identical to the conventional (non-crossvalidated) r^2 value in a PLS or regression analysis.

The other important results of this operation are different matrices known as the loadings or eigenvectors and the scores. The loadings matrix has the same number of columns as the input data and has p rows (where p is the number of principal components obtained), while the scores matrix has p columns and the same number of rows as the input matrix. Matrix multiplication of the scores by the loadings restores the input data matrix. The scores matrix is essentially the objective of the factor analysis, *i.e.* the input data in a form simplified by the new co-ordinate system.

PARTIAL LEAST SQUARES

Partial least squares is one of several techniques producing an equation, or QSAR, to describe or predict differences in the values in one or more table columns (the target properties) from differences in the values in other table columns (the explanatory properties). A more familiar example of such techniques is multiple regression (MR).

The outcome of PLS or MR is based on the knowledge about the explanatory properties to reduce uncertainty in the target properties. We may take the example of the case where it is desired to be able to predict the target property of some newly synthesised compound. In the absence of any other knowledge, the most reasonable guess for that value is would probably be the mean of all previous values, and the uncertainty of that guess would be the standard derivation of those values. The basic objective of any QSAR study is to find a predictive rule which performs better than this 'no rule at all' guess of using the average deviation of the average.

SOME QSAR TERMINOLOGY

Some of the basic terms used in QSAR terminology are:

s = a measure of the target property uncertainty still unexplained after the QSAR has been derived. Its magnitude depends also on the measurement scale for the target property;

r² = the proportion of the original uncertainty which is explained by using the QSAR. Being a ratio, its magnitude is unaffected by the measurement scale.

F-ratio = the ratio of r^2 to $1.0 - r^2$ (explained to unexplained), weighted so that the fewer the explanatory properties and the more the values of the target property, the higher the F-ratio.

residual = the difference between an actual target property and a calculated property. A large residual value indicates a compound that is not well modelled by the QSAR.

equation = the QSAR, a set of coefficients and an intercept or offset, used for prediction. An explanatory property value is supplies for each coefficient, and the prediction becomes the sum of the products of coefficient and explanatory property value, plus the intercept :

$$\text{prediction} = \text{intercept} + \text{explanatory}_1 * \text{coefficient}_1 + \\ \text{explanatory}_2 * \text{coefficient}_2 + \text{explanatory}_3 * \text{coefficient}_3 \text{ etc.}$$

Whether crossvalidation is used in conjunction with PLS, some of the above indices meanings change slightly and others are omitted as meaningless. The

key difference is in the definition of the s value. Usually s is the uncertainty remaining after the least-squares fit has been performed, but in crossvalidation, s becomes the uncertainty in prediction over all the compounds, often called PRESS. It is harder to predict values which are not used in deriving a model than it is to fit the same values while including them in a model, thus the 'predictive r^2 ' is always worse than the conventional r^2 for the same data. The crossvalidation indices, however, are reported to be much better indicators of how accurate the predictions are likely to be.

PLS uses the crossvalidation principle to control its central loop, the iterative derivation of successive components. With each new component, the press is recomputed. Typically the first few components contain signal, a consistent relationship between target and explanatory properties, and the press decreases while the crossvalidated r^2 increases. Eventually, however, additional components will only be fitting noise in the explanatory variables to noise in the target properties. this situation is signalled by an increase in press and a decrease in crossvalidated r^2 ; the additional components are improving the fit between target and explanatory properties, but only by distorting coefficient values in a way which steadily worsens the predictive value of the QSAR. The optimal number of components correspond to the highest crossvalidated r^2 . If there are more than one dependent columns, the optimal number of components refers to the one with the highest sum of crossvalidated r^2 derived for each component.

The **PRESS** value can be given by:

$$\sum^Y (Y_{PRED} - Y_{ACTUAL})^2$$

and the formula for the predictive or crossvalidated r^2 is :

$$1 - \frac{\sum^Y (Y_{PRED} - Y_{ACTUAL})^2}{\sum^Y (Y_{PRED} - Y_{MEAN})^2}$$

where **YPRED** is a predicted value;

YACTUAL is an actual or experimental value

YMEAN is the best estimate of the mean of all values that might be predicted;

the summations are over the same set of **Y**.

The crossvalidated standard error of estimate used is :

$$\sqrt{\frac{PRESS}{n - c - 1}}$$

where **n** = number of rows

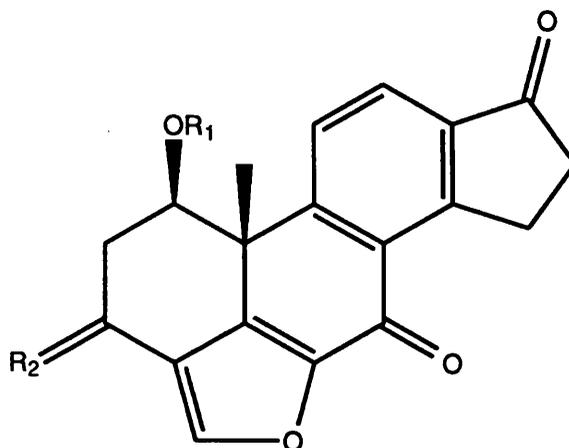
c = number of components.

Once the model has been formed and validated, predictions are relatively straightforward to carry out. Each new molecule is built, added to the database in the same manner as the model compounds, with its orientation frozen and Del-re charges calculated. QSAR CoMFA ALIGNMENT DEFINE is then carried out to restrict and name the conformation of each new molecule. This ALIGNMENT is then VIEWED and QSAR ANALYSIS PREDICT is carried

out under this named alignment. CoMFA fields are calculated automatically for the compound unless they have been calculated separately earlier, in which case when prompted 'USE PREVIOUS FIELDS (default no)' , one can reply yes. The field values are entered into the model equation and a value of biological activity based upon these is predicted.

CoMFA RESULTS

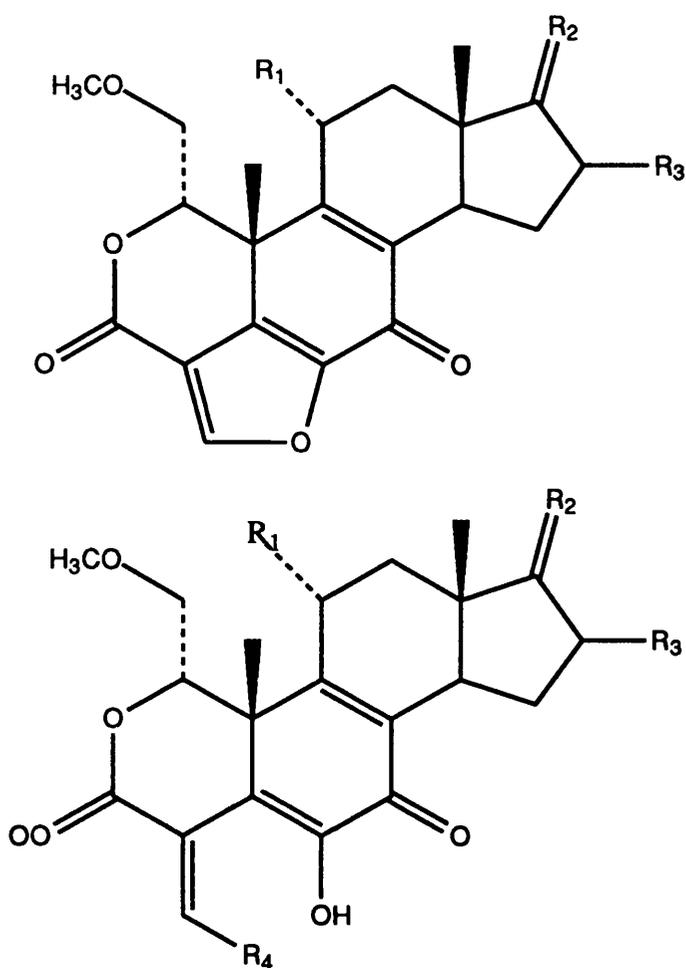
The first four analogues of viridin synthesised were based on initial modelling studies and feasibility of synthesis (see Chapter 5) and are shown below.



$R_1=H$	$R_2=O$	Demethoxyviridin	[3]
$R_1=H$	$R_2=N-OH$	Analogue 1	[18]
$R_1=OCOCH_3$	$R_2=O$	Analogue 2	[19]
$R_1=OCOCH_2CH_3$	$R_2=O$	Analogue 3	[20]
$R_1=OCOCH_2CH_3$	$R_2=O$	Analogue 4	[21]

FIGURE 36 - The first four demethoxyviridin analogues synthesised.

The first model built was from wortmannin analogues 1-10:



[1]	R ₁ =OAc	R ₂ =O	R ₃ =H	R ₄ = CLOSED RING
[12]	R ₁ =OAc	R ₂ =O	R ₃ =OCOCH ₃	R ₄ = CLOSED RING
[13]	R ₁ =OAc	R ₂ =O	R ₃ =H	R ₄ = NHC ₆ H ₅
[14]	R ₁ =H	R ₂ =O	R ₃ =H	R ₄ = CLOSED RING
[16]	R ₁ =OAc	R ₂ =OH	R ₃ =H	R ₄ = CLOSED RING
[17]	R ₁ =H	R ₂ =OH	R ₃ =H	R ₄ = CLOSED RING
[22]	R ₁ =OAc	R ₂ =CH=CH ₂	R ₃ =H	R ₄ = CLOSED RING
[23]	R ₁ =OAc	R ₂ =OH	R ₃ =CH ₂	R ₄ = CLOSED RING
[24]	R ₁ =OAc	R ₂ =ONO(OH)	R ₃ =H	R ₄ = CLOSED RING
[25]	R ₁ =H	R ₂ =OCH ₂ OCH ₃	R ₃ =H	R ₄ = CLOSED RING

FIGURE 37 - The ten wortmannin analogues used to build a CoMFA model.

The crossvalidated r^2 for this study was 0.624 with a non-crossvalidated r^2 of 0.957. The results of the study are shown in the table below.

MOLECULE	ACTUAL 1/lnIC ₅₀	PREDICTED 1/lnIC ₅₀	RESIDUAL
[1]	0.690	0.747	0.057
[12]	0.410	0.392	0.018
[13]	0.333	0.333	0.004
[14]	0.529	0.539	0.010
[16]	0.813	0.751	0.062
[17]	0.513	0.499	0.014
[22]	0.513	0.523	0.010
[23]	0.476	0.491	0.015
[24]	0.444	0.436	0.008
[25]	0.357	0.373	0.016

TABLE 8 - Results from the CoMFA model study of wortmannin analogues.

As can be seen from the table above the predictions of activity are very accurate, with extremely small residual values. The standard error of estimate for these results is 0.046, with an F value of 17.783. The relative contributions from each variable (%) were 61.2% steric and 38.8% electrostatic. This suggests that both factors are important, although steric factors carry twice as much weight in this case. The electrostatic value is fairly high in comparison to other studies ⁶⁴, where electrostatic contributions vary from only 8.0 to 12.3%.

From this model, predictions of synthesised analogues were carried out.

The results are shown in table 9 below:

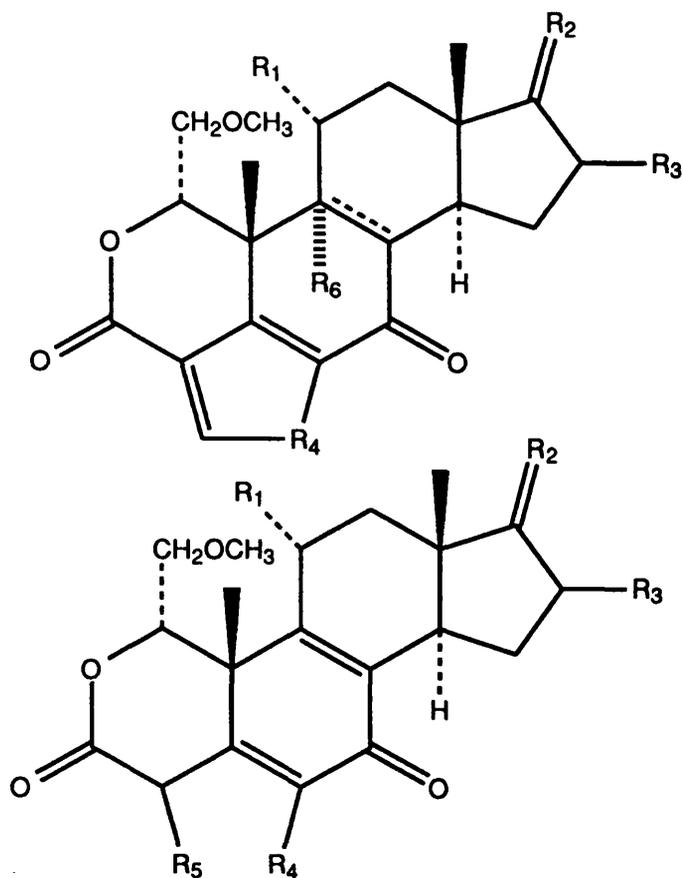
MOLECULE	PREDICTED 1/lnIC₅₀	ACTUAL 1/lnIC₅₀
VIRIDIN [2]	0.671	0.817
MOL1 [18]	0.561	0.353
MOL2 [19]	0.671	0.593
MOL3 [20]	0.707	0.940
MOL4 [21]	0.669	0.817

TABLE 9 - Predicted vs. actual results for the first analogues synthesised.

The resulting IC₅₀s are based on different assays so the actual values are different. The correlation in order of potency, however, is good and regression analysis by partial least squares gives an r^2 of 0.828. These results were very exciting because they showed not only that activity could actually be predicted, at least in terms of order of potency, but also suggested that there was perhaps a mechanistic link in the cytotoxic and anti-inflammatory action of these compounds, perhaps an increase in activity was activated by means of the same key enzyme.

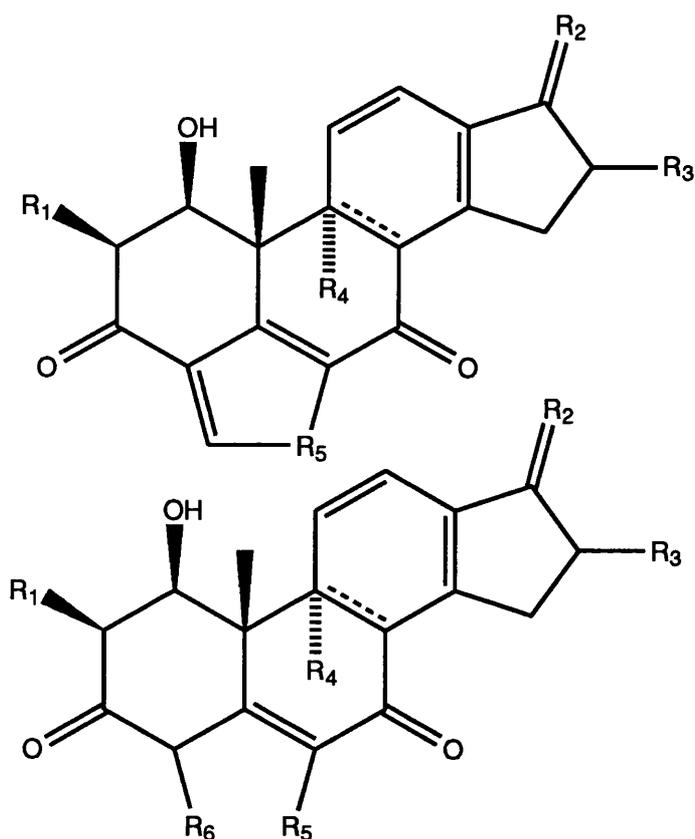
From the reported result of a steroid binding study⁶⁴ it was decided to build some wortmannin analogues based on the changes in the steroid structures which led to the highest increases in potency of binding to the receptor. This was based simply on the fact that wortmannin and viridin both have a very rigid steroid-like skeleton. Changes made to the steroids included the addition of a bulky group to ring D to increase lipophilicity, and therefore ease transport across membranes, addition of fluorine as a very electronegative

influence to pull electrons away from the carbonyl group on ring A thereby making the oxygen more delta negative and more likely to bind to a delta positive hydrogen. Binding through these atoms was thought to be critical in the action of steroids. Other ideas such as introducing a corticoid side chain onto ring D or changing the furan oxygen to a nitrogen were also tried. The additional compounds overleaf were built and energy minimised.



Molecule	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆
[26]	OAc	O	H	N	ring closed	-
[27]	OAc	O	H	S	ring closed	-
[28]	OAc	COCH ₂ OAc	H	O	ring closed	-
[29]	H	COCH ₂ OH	H	O	ring closed	-
[30]	OAc	OCOCH ₂ OCOC1	H	O	ring closed	-
[31]	OAc	OCOCH ₂ OCOB _r	H	O	ring closed	-
[32]	OAc	O	H	O	ring closed	F
[33]	OAc	O	H	CH ₃	-	-
[34]	OAc	O	CH ₃	O	ring closed	-
[35]	OAc	O	H	OH	CHN(Et) ₂	-
[36]	OAc	O	H	CH ₂	ring closed	-
[37]	OAc	O	H	F	CH ₃	-
[38]	OAc	CO(CH ₂) ₂ COCH ₃	CH ₂ Pr ^t	O	ring closed	-

FIGURE 38 - Theoretical wortmannin analogues.



Molecule	R ₁	R ₂	R ₃	R ₄	R ₅	R ₆
[39]	H	O	H	-	N	closed ring
[40]	H	O	H	-	S	closed ring
[41]	H	OCOCH ₂ OAc	H	-	O	closed ring
[42]	H	OCOCH ₂ OH	H	-	O	closed ring
[43]	H	OCOCH ₂ OCOC _l	H	-	O	closed ring
[44]	H	OCOCH ₂ OCOB _r	H	-	O	closed ring
[45]	H	O	H	F	O	closed ring
[46]	H	O	H	-	OH	CH ₃
[47]	H	O	CH ₃	-	O	closed ring
[48]	H	OH	H	-	OH	CHN(Et) ₂
[49]	H	O	H	-	F	CH ₃
[50]	H	CO(CH ₂) ₂ COCH ₃	CH ₂ Pr ^t	-	O	closed ring

FIGURE 39 - Theoretical viridin analogues.

Hydrocortisone (cortisol) was also tested:

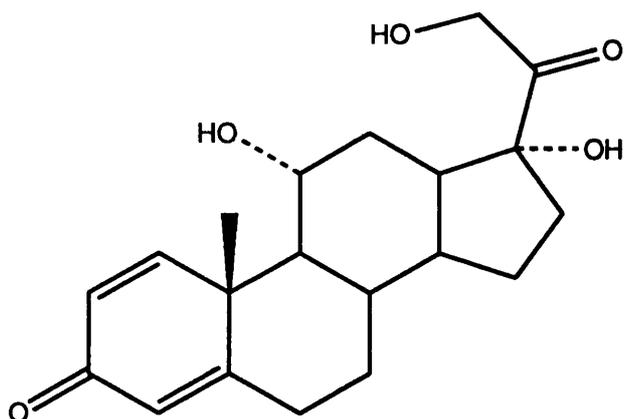


FIGURE 40 - Hydrocortisone (cortisol) [51]

The predictions of biological activity of these molecules, based on the CoMFA model, is shown in the table overleaf.

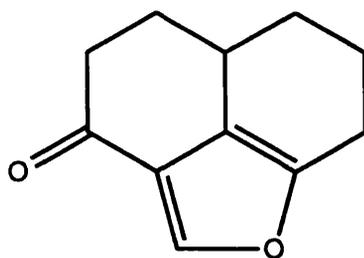
Molecule	Actual IC ₅₀	Actual 1/log IC ₅₀	Predicted IC ₅₀	Predicted 1/log IC ₅₀
[1]	28.1	0.690	22.6	0.738
[2]	3.4	1.881	29.8	0.678
[3]	1.5	5.679	28.8	0.685
[4]	-	-	28.8	0.685
[5]	-	-	58.9	0.565
[6]	-	-	32.6	0.661
[7]	-	-	29.6	0.680
[8]	-	-	3.7	0.760
[9]	-	-	28.7	0.686
[10]	-	-	32.9	0.659
[11]	-	-	36.5	0.640
[12]	274.8	0.410	256.8	0.415
[13]	1000	0.333	44.1	0.608
[14]	77.7	0.529	81.0	0.524
[15]	1000	0.333	29.6	0.680
[16]	17.0	0.813	19.9	0.770
[17]	89.0	0.513	84.5	0.519
[18]	17.0	0.813	93.8	0.507
[19]	5.4	1.365	27.9	0.692
[20]	2.9	2.163	28.1	0.690
[21]	3.4	1.882	30.2	0.676
[22]	89.0	0.513	92.2	0.509
[23]	126.1	0.476	54.1	0.577
[24]	178.8	0.444	196.6	0.436
[25]	632.6	0.357	134.2	0.470
[26]	-	-	21.9	0.746
[27]	-	-	23.7	0.727
[28]	-	-	256.8	0.415
[29]	-	-	599.5	0.360
[30]	-	-	55.6	0.573

TABLE 10 - CoMFA predictions based on wortmannin model.

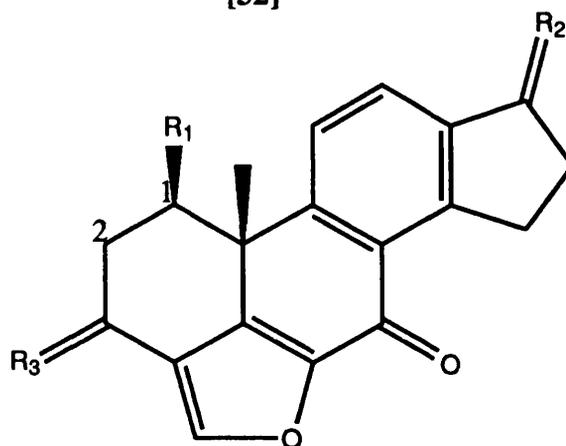
Molecule (cont.)	Actual IC₅₀ (cont.)	Actual 1/logIC₅₀ (cont.)	Predicted IC₅₀ (cont.)	Predicted 1/log IC₅₀ (cont.)
[31]	-	-	330.3	0.397
[32]	-	-	68.4	0.545
[33]	-	-	54.5	0.576
[34]	-	-	52.5	0.582
[35]	-	-	48.3	0.594
[36]	-	-	63.8	0.554
[37]	-	-	54.5	0.576
[38]	-	-	345.2	0.394
[39]	-	-	28.8	0.685
[40]	-	-	28.8	0.685
[41]	-	--	36.1	0.642
[42]	-	-	35.9	0.643
[43]	-	-	37.6	0.635
[44]	-	-	37.1	0.637
[45]	-	-	27.2	0.697
[46]	-	-	28.8	0.685
[47]	-	-	28.0	0.691
[48]	-	-	27.6	0.694
[49]	-	-	27.2	0.697
[50]	-	-	69.4	0.543
[51]	-	-	206.4	0.432

TABLE 10 (cont.)- CoMFA predictions based on wortmannin model.

The following compounds were also synthesised and predictions tested.



[52]



Molecule	R ₁	R ₂	R ₃	Position 1,2
[53]	OCOPh	O	O	saturated
[54]	OH	O	N(CH ₃)OH	saturated
[55]	OCOCH ₃	O	N-OH	saturated
[56]	OCOCH ₃	OH	OH	saturated
[57]	-	O	O	unsaturated

FIGURE 41 - Synthesised hibiscone C and demethoxyviridin analogues.

The results from CoMFA predictions of these molecules are shown in the table below.

Molecule	Actual IC₅₀	Actual 1/logIC₅₀	Predicted IC₅₀	Predicted 1/logIC₅₀
[52]	111.7	0.488	58.4	0.566
[53]	1.62	4.773	28.5	0.687
[54]	-	-	95.5	0.505
[55]	-	-	82.4	0.522
[56]	-	-	128.7	0.474
[57]	0.033	-0.675	28.0	0.691

TABLE 11 - CoMFA predictions of hisbicone C and demethoxyviridin analogues based on viridin model.

There are several points to note from all these predictions. The most active (actual) compounds were [3] **demethoxyviridin** with an IC_{50} of 1.5, [53] the synthesised viridin analogue with the **phenyl** group substituent and IC_{50} of 1.62 and [57] the synthesised dehydro-viridin analogue with an IC_{50} of 0.033 the most active compound to date. All three of these compounds were predicted to have an IC_{50} of around 28, which although is higher, is the lowest (and in the most active) range of predictions. The lowest predicted value was for [8] **hibiscone C** as a result of which, a synthesis of this molecule was attempted with a view to producing analogues for testing. The synthesis proved difficult, however, with a very low yield and the part structure synthesised [52] was almost totally inactive with an IC_{50} of 111.7. The part structure itself was predicted to be fairly active with an IC_{50} of 58.4. These results were disappointing, although a fairly large structural difference between hibiscone C and viridin analogues may account for the difference in activity. From these results it was seen, in particular, that a shortcut to active analogues would not be possible. It appears that the rest of the structure is required to stabilise what are thought to be the active centres.

Other compounds with predicted IC_{50} s of less than 28 were [1] **wortmannin**, which was the second most potent compound in the anti-inflammatory test, [16] the wortmannin analogue with ring D carbonyl **reduced**, the most active in the test, so the correlation here was very good. Theoretical analogues with potent IC_{50} s (less than 28) were [26] and [27], wortmannin analogues where the oxygen in the furan ring E is substituted with nitrogen or sulfur respectively. The two similarly modified viridin analogues [39] and [40] also had IC_{50} s of about 28. Although this activity is identical to viridin, although slightly better than demethoxyviridin, this small change in structure may have been worth following up if time had allowed as the analogues may be more stable.

Other theoretical analogues predicted to be particularly active included [48] and [49] with IC_{50} s of 27.6 and 27.2 respectively. These compounds were 'broken ring' analogues, however, so it was thought that in practice activity would be lost, as is the case for the anti-inflammatory analogues of wortmannin. This theory, however, has not yet been tested. Introducing a fluorine atom into ring B produced a theoretically active compound [45] with an IC_{50} of 27.2. This step was taken by workers in steroid binding studies where the electronegative fluorine was thought to draw electrons away from the carbonyl oxygen, thus making it more susceptible to hydrogen bonding at the receptor. In fact, the most potent orally active steroid to date is the fluorine containing compound betamethasone. This suggests perhaps this carbonyl is critically involved in binding to a receptor and that the electronegativity of the oxygen would be of fundamental importance.

Changes to the molecules which rendered them particularly inactive were, in general, the addition of bulky groups onto ring D, *e.g.* [12] the addition of $OCOCH_3$ onto the position adjacent to the carbonyl of ring D of wortmannin increased the IC_{50} to over 250, as did the conversion of this carbonyl to a corticoid side chain. Other bulky substitutions as in molecules [31] and [38] increased the IC_{50} to over 330 and an extended corticoid group without the acetyl group in [29] led to an IC_{50} of almost 600. These results indicate that either shape or lipophilicity is critical at this position on ring D, perhaps even the receptor is not completely internalised by membrane so a very high lipophilicity is not necessary for binding.

Following on from these results, the next logical step appeared to be to build a model from the viridin analogues themselves. This would mean the predicted activities of any new compounds would be directly relevant to cytotoxic activity. Nine molecules had the available cytotoxicity test data at this point, [2], [3], [18], [19], [20], [21], [52], [53], and [57]. An initial attempt to use all nine compounds was unsuccessful with a negative r^2 resulting. Deletions of [53], shown to have a very high CoMFA value in the CoMFA column, [3] and [57] for reasons of having particularly high activities led to a crossvalidated r^2 of 0.601 with a final r^2 of 0.874. The F value was 27.7. and the standard error of estimate was 0.025. The relative contributions from each variable were as follows:

	Norm. Coefficient	Fraction (%)
1. Autocomfa (STERIC)	0.153	15.3
2. Autocomfa (ELECTROSTATIC)	0.847	84.7

TABLE 12 - Relative contributions from steric and electrostatic variables
(viridin model).

From the electrostatic (blue and yellow) and steric (red) map shown (Figure 42) a strongly yellow region was seen around the top ring D suggesting that increased activity could be achieved by decreasing bulk around this region. A blue region was also apparent at this point, suggesting increased positive charge would be beneficial. A distinct red area around the carbonyl group in ring A suggests an increased negative charge at this point would be conducive to favourable activity *e.g.* addition of an electronegative influence such as fluorine.

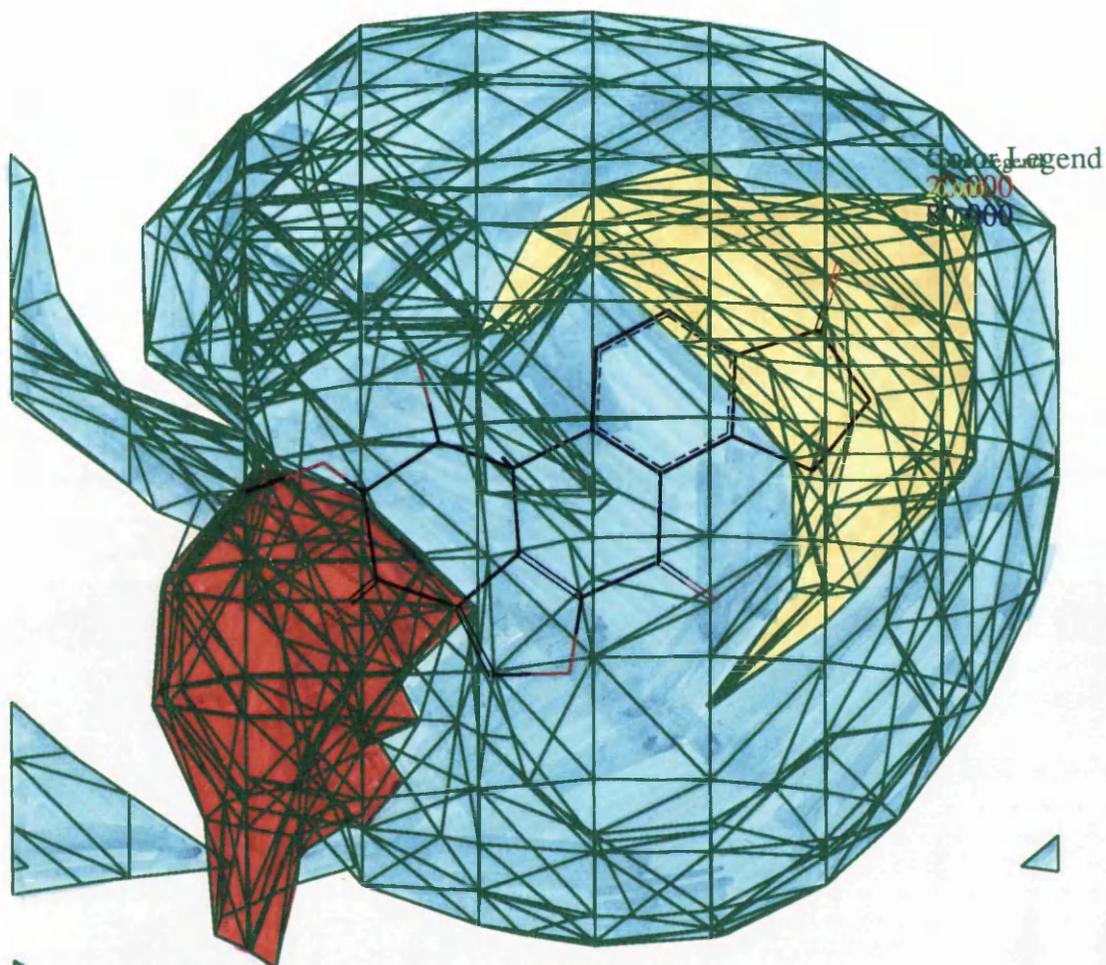


FIGURE 42 - Electrostatic (blue and yellow) and steric (red) map from viridin CoMFA study.

Blue - increased positive charge

Yellow - increased negative charge

Red - decreased bulk

A table of the activity and residuals of the viridin model is seen in table 13.

Molecule	Actual 1/log(IC₅₀x100)	Predicted 1/log(IC₅₀x100)	Residual
[2]	0.40	0.39	0.01
[18]	0.31	0.28	0.03
[19]	0.37	0.39	-0.03
[20]	0.41	0.39	0.01
[21]	0.40	0.39	0.0018
[52]	0.25	0.27	-0.03

TABLE 13 - Actual vs Predicted activity results from viridin CoMFA model.

From this model, predictions were re-calculated as before, and are shown in Table 14 overleaf.

Molecule	Actual IC₅₀	Actual 1/log(IC₅₀x 100)	Predicted IC₅₀	Predicted 1/(logIC₅₀x100)
[1]	28.1	0.290	4.1	0.382
[2]	3.4	0.395	3.8	0.387
[3]	1.5	0.500	6.8	0.353
[4]	-	-	3.9	0.286
[5]	-	-	49.9	0.270
[6]	-	-	4.1	0.382
[7]	-	-	3.9	0.385
[8]	-	-	4.4	0.378
[9]	-	-	4.1	0.382
[10]	-	-	3.9	0.386
[11]	-	-	3.9	0.384
[12]	274.8	0.225	21.7	0.300
[13]	1000	0.200	16.8	0.310
[14]	77.7	0.257	25.5	0.294
[15]	1000	0.200	4.2	0.382
[16]	17.0	0.310	4.1	0.382
[17]	89.0	0.253	22.2	0.299
[18]	17.0	0.310	36.0	0.281
[19]	5.4	0.366	4.3	0.380
[20]	2.9	0.406	3.6	0.392
[21]	3.4	0.395	3.5	0.393
[22]	89.0	0.253	4.1	0.383
[23]	126.1	0.244	3.5	0.393
[24]	178.1	0.235	17.0	0.310
[25]	632.6	0.208	22.5	0.298
[26]	-	-	4.1	0.382
[27]	-	-	4.1	0.382
[28]	-	-	27.7	0.290

TABLE 14 - Predicted results from viridin model (continued overleaf).

Molecule	Actual IC ₅₀	Actual 1/log(IC ₅₀ x100)	Predicted IC ₅₀	Predicted 1/log(IC ₅₀ x100)
[29]	-	-	29.7	0.288
[30]	-	-	4.1	0.383
[31]	-	-	27.4	0.291
[32]	-	-	23.2	0.297
[33]	-	-	22.8	0.298
[34]	-	-	23.1	0.297
[35]	-	-	17.5	0.308
[36]	-	-	24.5	0.294
[37]	-	-	21.8	0.299
[38]	-	-	28.1	0.290
[39]	-	-	3.9	0.386
[40]	-	-	3.9	0.386
[41]	-	-	3.9	0.386
[42]	-	-	3.9	0.386
[43]	-	-	3.9	0.386
[44]	-	-	3.9	0.386
[45]	-	-	3.9	0.386
[46]	-	-	3.9	0.386
[47]	-	-	3.9	0.386
[48]	-	-	3.9	0.386
[49]	-	-	3.9	0.386
[50]	-	-	4.0	0.383
[51]	-	-	18.7	0.305
[52]	111.7	0.274	44.3	0.274
[53]	1.62	0.453	3.6	0.391
[54]	-	-	35.6	0.282
[55]	-	-	32.7	0.285
[56]	-	-	38.6	0.279
[57]	0.033	1.929	4.1	0.383

TABLE 14 - CoMFA predictions from viridin model (cont.).

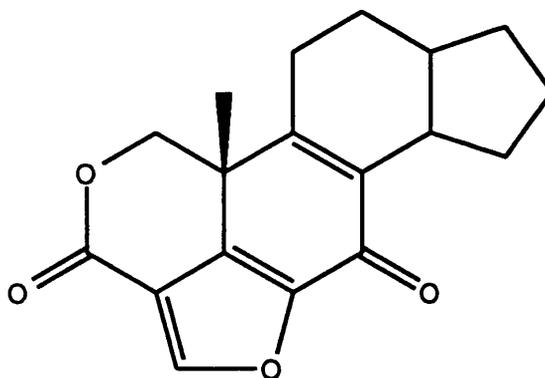
From this model there are also several points to note about the resulting predictions. The most active compounds [3], [53] and [57] were predicted to have IC_{50} s of 6.8, 3.6 and 4.1 respectively (actual IC_{50} s were 1.5, 1.62 and 0.033) so again although activity for these extremely active compounds could not be predicted in terms of the actual value, the order of potency was apparently high. The lowest predicted value in this model was for [5], the 'query' hibiscone-like structure with predicted IC_{50} of almost 50. This low activity was seen in practice. One of the main differences in this model was that hibiscone C [8] was predicted to be fairly active ($IC_{50} = 4.4$). The actual part structure [52] synthesised, however, was predicted, and in fact was, very inactive (predicted $IC_{50} = 44.3$, actual = 111.7). This suggests that completion of the synthesis of hibiscone C, if time had allowed, may have been fruitful.

Wortmannin [1], and its hydroxy derivative [6] both had predicted IC_{50} s of 4.1, *i.e.* very active. Among the theoretical analogues, the highest activities were seen from the wortmannin [26] and [27] and viridin [39] and [40] analogues with substituted furan ring oxygen to a nitrogen or sulfur atom. Their predicted activities were comparable with wortmannin and only marginally worse than viridin.

Changes in the theoretical molecules which rendered them particularly inactive were as before, *i.e.* [29] $IC_{50} = 29.7$ when the furan ring is broken and a fluorine atom is introduced. Also seen in both models is a large decrease in activity associated with an increase in lipophilicity seen by substitution in ring D as in [38].

A recent publication explained the synthesis of a simplified analogue of wortmannin (shown below). Broka *et al*⁷⁴ found tested the analogue for its

ability to block *f*-Met-Leu-Phe-stimulated generation of superoxide by neutrophils and found it to be inactive under conditions where wortmannin itself possessed an IC_{50} of $0.1\mu\text{m}$. It was therefore decided to build this molecule using SYBYL and predict its activity. In fact, SYBYL predicted an IC_{50} of $442\mu\text{m}$ for this compound, *i.e.* extremely inactive.



CHAPTER 4
CRYSTALLOGRAPHY

BASIC THEORY

The primary aim of crystal structure analysis by X-ray diffraction is to obtain a detailed picture of the contents of the crystal at the atomic level. Once this information is available, and the positions of the individual atoms known precisely, one can calculate the inter-atomic distances, bond angles and other features of the molecular geometry that are of interest, such as polarity of a particular group of atoms, the angles between planes, and conformational angles around bonds. For the purposes of this project this information is used mainly as a tool for drug design, *i.e.* an attempt to relate structural features to chemical or biological properties.

X-rays are scattered by the electrons in atoms but in contrast to the situation with visible light, such radiation cannot normally be focused by presently known experimental techniques⁶⁵. This is because no electrical or magnetic field or material exists that can refract X-rays sufficiently to form an X-ray lens. An alternative technique used in practice is to intercept the diffracted or scattered radiation by a detecting system, although the relationships between the phases of scattered waves are lost and only the intensities can be measured, therefore information is lost.

It is possible, however, to simulate the recombination of scattered X-rays by a mathematical calculation called a **Fourier synthesis**. This is a fundamental step in crystal structure determination. The major complication is referred to as the phase problem, *i.e.* discovering the relative phases of the diffracted beams that are needed to compute a recognisable image of the molecule. When these relative phases are known, the molecule will be revealed as a result of such a Fourier synthesis.

CRYSTAL STRUCTURE OF DEMETHOXYVIRIDIN

Demethoxyviridin, $C_{19}H_{14}O_5$, (figure 44), has been previously isolated from extracts of the mycelium of the unidentified fungus ACC3199. It also comes from *Apiospora camptospora* and *Nodulisporum hinnuluem* (ATCC24911)⁶⁶. It is structurally related to several known metabolites including viridin, for which the crystal structure has been solved³⁴. The crystal structure of demethoxyviridin therefore provides useful data for detailed comparison of these compounds.

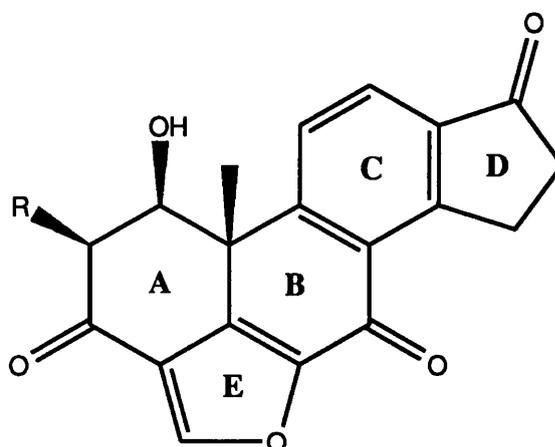


FIGURE 44 - Viridin R=OMe [2]
Demethoxyviridin R=H [3]

EXPERIMENTAL

The crystal sample donated by Wellcome consisted of pale yellow crystals of varying size, depth of colour and opacity. The chosen crystal was pale yellow and acicular, having dimensions 3.2 x 1.5 x 0.5 mm. Diffraction data were collected on an Enraf-Nonius CAD-4 diffractometer.

CRYSTAL DATA

$C_{19}H_{14}O_5$, $M=321.31$, orthorhombic space group $P2_12_12_1$, $a=8.47$, $b=19.33$, $c=17.91$, $V=2935 \text{ \AA}^3$, $Z=8$, $D_x=1.454 \text{ gcm}^{-3}$, $(\text{CuK } \alpha)=1.5418 \text{ \AA}$, $\mu=4.19 \text{ cm}^{-1}$, $F(000)=668$, final $R=0.062$ for 3158 unique observed reflections.

RESULTS

Structure solution was achieved by direct phasing methods using MITHRIL⁶⁸. Least squares were used to refine lattice parameters. From an initial 25 reflections, two standard intensities were chosen to monitor variations in intensity, < 3% variation was observed. The correct phasing was established only after a tricyclic ring system was incorporated into NORMAL with random position and orientation. 150 negative quartets were also used for both phase expansion and figures of merit. The resulting solution clearly indicated the positions of all the main atoms for one molecule. It was therefore decided not to recycle MITHRIL at this stage. Initial stages of least squares refinement resulted in a rise in the R factor from approximately 0.76 to around 2.2 although it is not unusual for R to rise before dropping to a more realistic value. Block matrix refinement was repeated for a further 2 cycles until R was 0.49. A difference Fourier map was then calculated to determine the positions of the H atoms. It became apparent at this stage that there were 2 molecules in the asymmetric unit. The main atoms for the second molecule were appended to the model file and refinement continued. Block matrix refinement was then rerun for a further 4 cycles without refining the temperature factors. This brought the R value down to 0.13 with several H atoms missing. Final refinement revealed the Hs and achieved a final R value of 0.062 for 3158

unique observed reflections. All calculations were performed using the Glasgow GX package⁶⁸.

DISCUSSION

Final positional and equivalent isotropic thermal parameters are given in Appendix 1. Bond lengths, bond angles and torsion angles are listed in Appendices 2, 3 and 4 respectively. An ORTEP⁴⁷ diagram, Figure 45 illustrates the numbering scheme for the molecule. Figure 46 shows a stereoview of the cell contents.

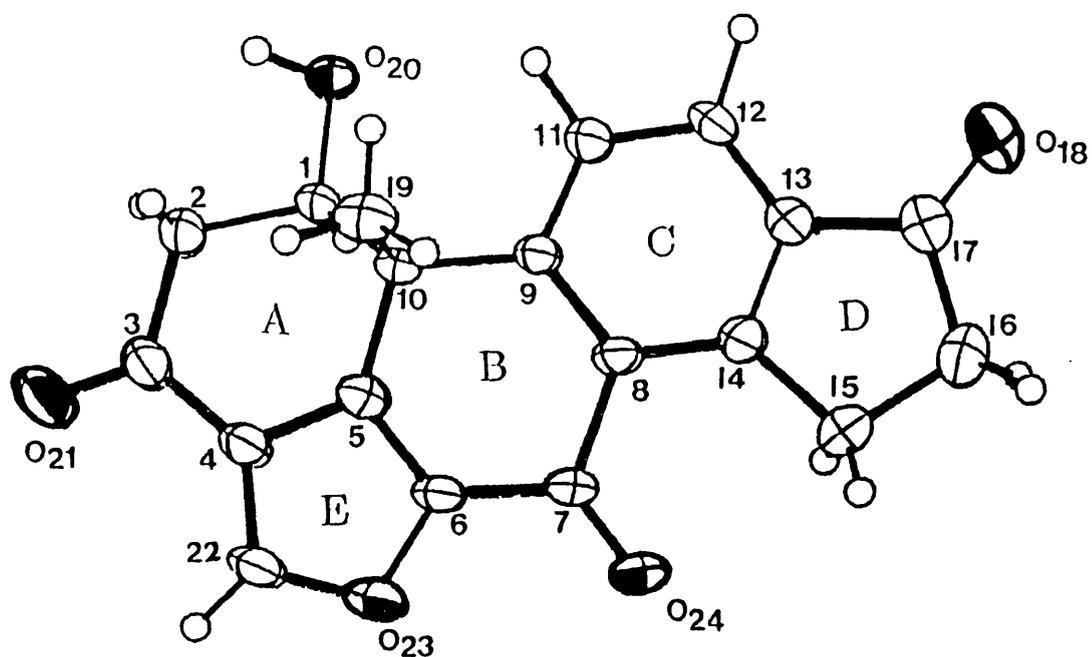


FIGURE 44 - ORTEP⁴⁸ diagram of demethoxyviridin showing the numbering scheme with vibrational ellipsoids at 50% probability level.

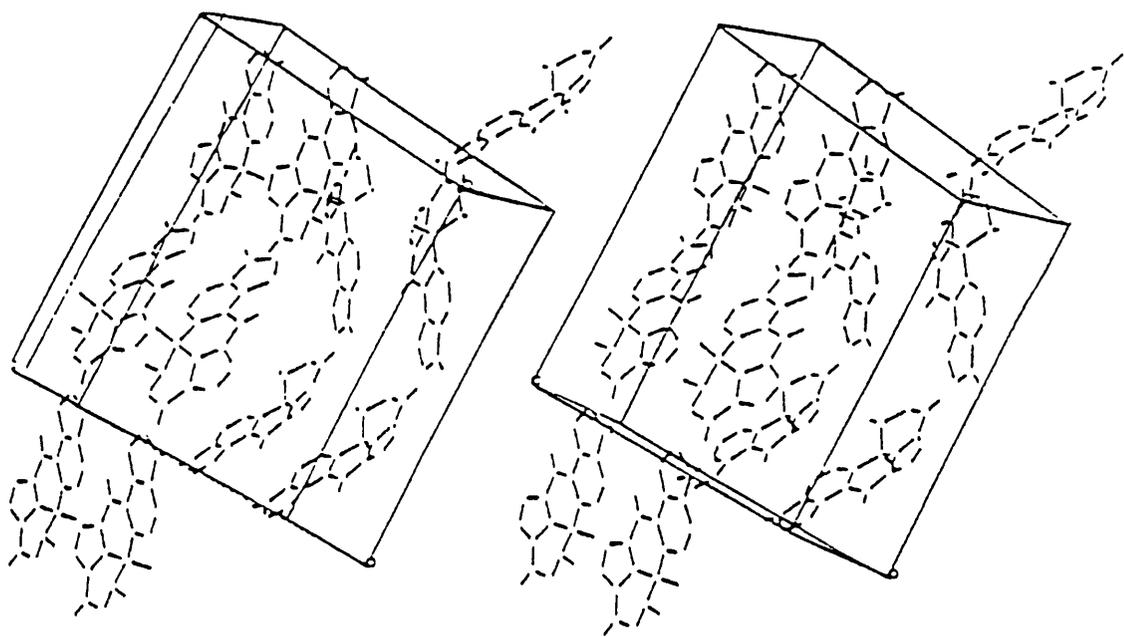


FIGURE 45 - Stereoview of demethoxyviridin unit cell contents.

DESCRIPTION OF MOLECULE

There are two molecules in the asymmetric unit which are identical within experimental error. Analysis using molecular graphics revealed the molecule to be largely planar. Each molecule was identical around the two chiral centres present.

The atoms best describing the planes for each of the 5 rings are listed in Table 15. Deviations from these same planes were examined by Neidle, Rogers *et al* for viridin³⁴ These deviations are compared to viridin and listed in Table 16. The largest of these is seen in ring D. This observation was confirmed in molecular graphics experiments docking the two molecules together using HYDRA. It was found that the basic structures are almost identical but ring D and to a lesser extent ring C had changed conformation slightly as a result of the removal of the methoxy group on C(2). There was no apparent difference in ring A.

RING	ATOMS
A	C(2) C(3) C(4) O(21)
B	C(6) C(7) C(8) O(24)
C	C(8) C(9) C(11) C(12) C(13) C(14)
D	C(13) C(14) C(17) O(18)
E	C(4) C(5) C(6) C(22) O(24)

TABLE 15 - Atoms best describing the planes of demethoxyviridin.

RING	ATOM	DEMETHOXYVIRIDIN	VIRIDIN
A	C(1)	0.001	0.000
	C(3)	-0.004	0.002
	C(4)	0.001	0.000
	O(21)	0.001	-0.001
B	C(6)	-0.005	-0.003
	C(7)	0.016	0.011
	C(8)	-0.005	-0.003
	O(24)	-0.005	-0.004
C	C(8)	-0.006	-0.004
	C(9)	0.004	0.000
	C(11)	0.001	0.004
	C(12)	-0.004	-0.004
	C(13)	0.002	0.000
	C(14)	0.004	0.004
D	C(13)	-0.008	0.001
	C(14)	0.011	-0.001
	C(17)	-0.028	0.003
	C(18)	0.017	-0.002
E	C(4)	-0.011	-0.011
	C(5)	0.012	0.019
	C(6)	-0.014	-0.020
	C(22)	0.003	-0.001
	C(24)	0.004	0.013

TABLE 16 - Deviations (Å) from planes compared to viridin.

Many of the conformational features of demethoxyviridin are due to the strain induced by the unique unsaturated ring E, on the A and B rings. Ring A is buckled and ring B is a slightly flattened boat with both bow and stern deflected β to the main plane of the molecule. Ring C is benzenoid and ring D is flat and essentially co-planar with C. Ring E is also flat. This basic molecular shape can also be applied to the viridin molecule.

Ring A is perhaps best described as a distorted chair cyclohexenone, with C(1) and C(4) deflected respectively above and below the plane. Ring B shows considerable deviations from the usual geometry of a cyclohexadienone. C(5), C(6), C(8) and C(9) are closely co-planar, with C(7) and C(10) displaced 0.116Å and 0.219Å respectively from the plane. These values are comparable for viridin, *i.e.* 0.108Å and 0.304Å respectively. This ring can be best described as a slightly flattened boat. Steric considerations probably play a major role in the case of demethoxyviridin, with angular strain on the ABE ring junctions producing a situation in which ring B is forced to pucker and adopt a boat shape.

Ring C is a typical benzene ring although substituents C(7) and C(10) are displaced slightly out of plane, probably due to ring fusion strain mentioned above.

Ring D is almost planar with C(15) being displaced 0.0321Å above the plane of the ring. This shift may be due to the proximity of O(24).

Ring E is a furan ring and is planar within experimental error although its substituents C(3) and C(7) lie below the plane. The angles between the planes are essentially indicative of a largely planar structure (Table 17 overleaf).

PLANE	ANGLE (°)
A-B	170.45
A-E	171.95
B-E	12.79
B-C	171.40
C-D	177.26

TABLE 17 - Angles between the planes of the rings.

The entire crystal structure consists of infinite chains of molecules linked by a single hydrogen bond between O(20)-H and O(18), spiralling around the a axis, packed in a zigzag arrangement of chains (Fig. 47).

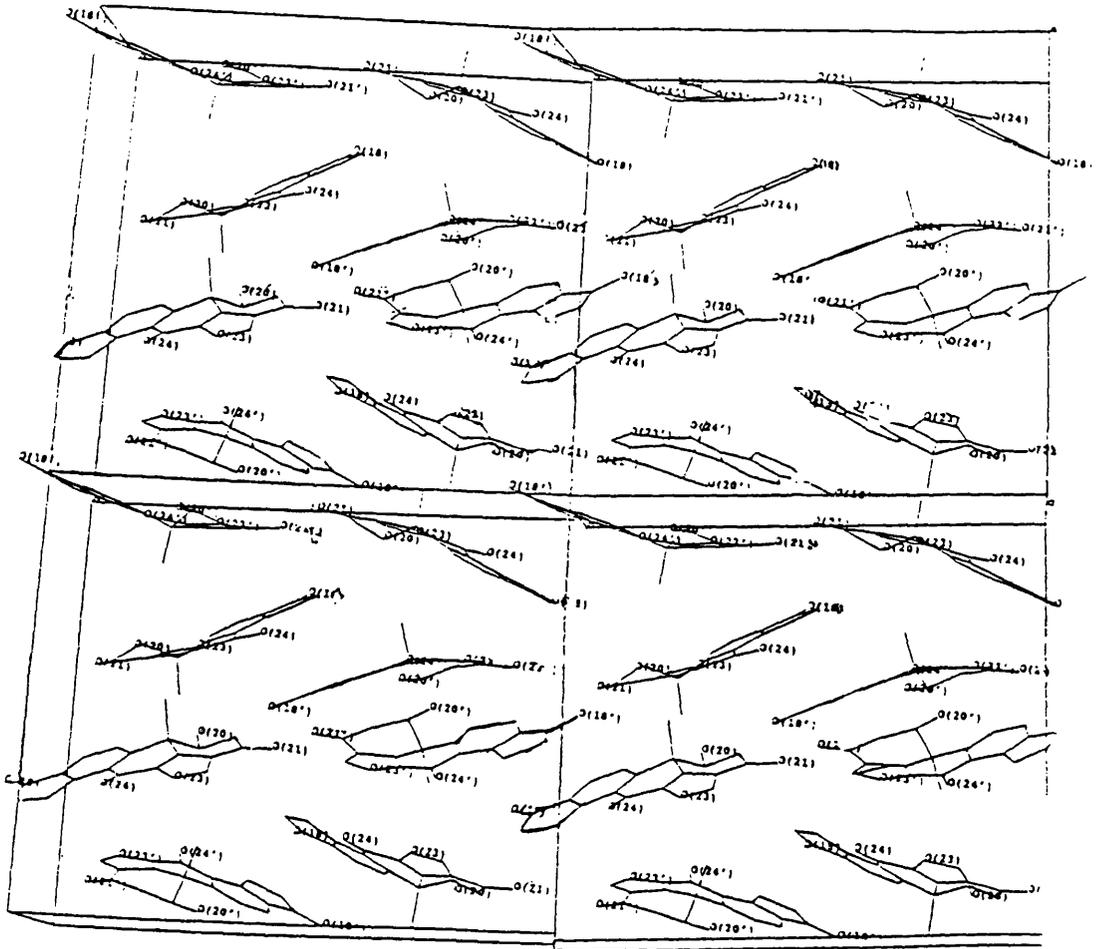


FIGURE 47 - Repeated unit cells of demethoxyviridin.

CHAPTER 5
ORGANIC SYNTHESIS

1. SYNTHESIS OF A HIBISCONE C ANALOGUE

From a computer search of the Cambridge crystallographic database (Chapter 2), one of the compounds retrieved by use of the query structure was hibiscone C [8], a furanosesquiterpene. This compound was isolated independently by Pelter⁶⁹ (as **gmelofuran** from the roots of *Gmelina aborea*) and Thomson⁷¹ (as **hibiscone C** from the heartwood of *Hibiscus elatus*, or the Blue Mahoe, national tree of Jamaica) in the late 1970s. The compounds were later shown to be identical. The structure (shown below) represents an analogue of the tricyclic furan part of demethoxyviridin, viridin and wortmannin.

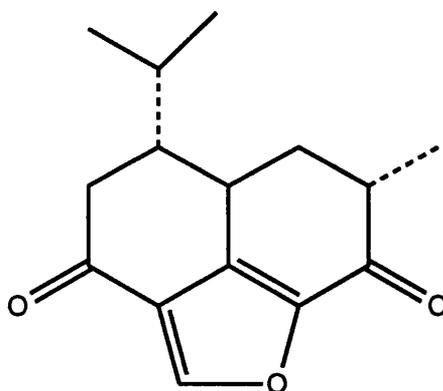
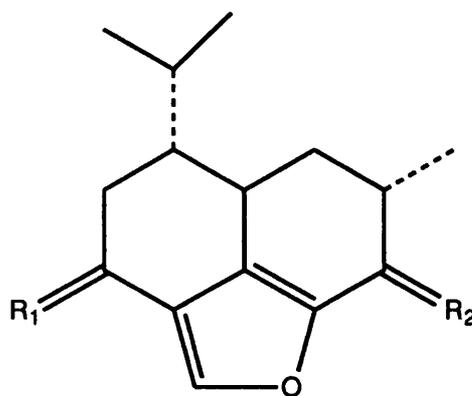


FIGURE 48 - Hibiscone C [8] (Gmelofuran)

This compound was synthesised by A.B. Smith III *et al* in 1984⁷¹ and is a member of a small group of tricyclic furanosesquiterpenes that now includes hibiscone A [59], B [60], D [61] and agarol [62], the latter being independently isolated in the late 1970s⁷⁰.



Molecule	R ₁	R ₂
Hibiscone A [59]	H ₂	O
Hibiscone B [60]	α-OH, β-H	O
Hibiscone D [61]	O	α-OH, β-H
Agarol [62]	O	OH

FIGURE 49 - Other tricyclic furanosesquiterpenes.

The key architectural feature of this family, the tricyclicofuran ring skeleton, is thought to be critical for the cytotoxic activity of the more complex fungal antibiotics wortmannin [1] and viridin [2]. Attempts to obtain a sample of hibiscone C [8] for testing this hypothesis were unsuccessful therefore it was decided to synthesise the molecule according to the published method with a view to producing several structural analogues.

Interest in hibiscone C [8] as a synthetic target stemmed from the realisation that the tricyclic nucleus might be assembled from intermediate [65] via a novel [2+2] photocycloaddition of the acetylenic moiety to the enone functionality (Figure 50). The resultant cyclobutene [64] was then

exploited as a latent furan by a sequence of oxidative cleavage to a 1,4-dicarbonyl compound and acid catalysed cyclisation/dehydration to the furan [63]. Further elaboration, *i.e.* introduction of a carbonyl group at C(2) and then the axial methyl at C(3) led to racemic hibiscone C [8]. The synthesis proceeded in 10 steps with an overall yield of 2.5%.

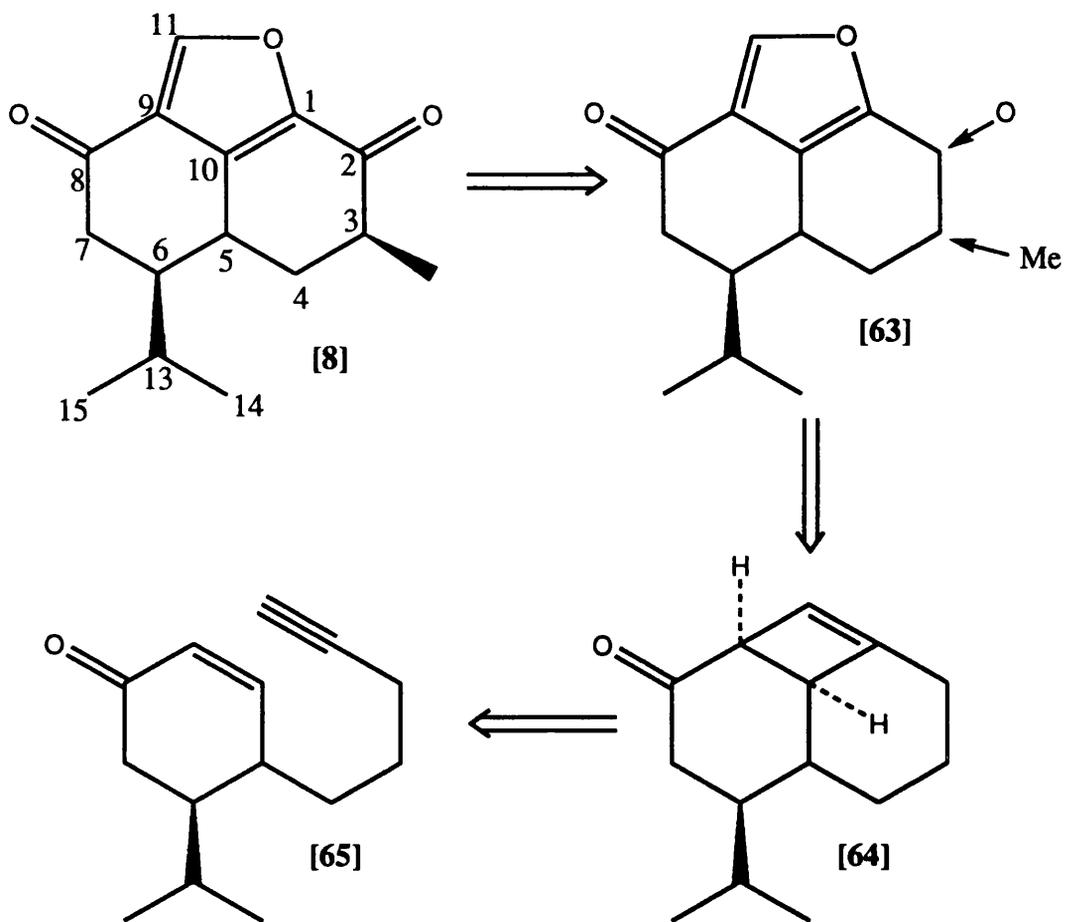


FIGURE 50 - Retrosynthesis of hibiscone C [8].

It was decided to attempt the synthesis of the simplified hibiscone C analogue [66], as this retained what were thought to be the main structural features for biological activity.

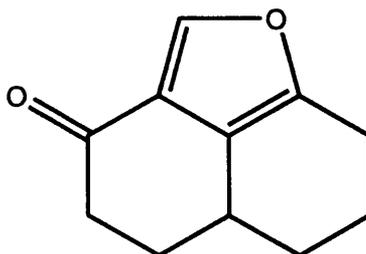


FIGURE 51 - Simplified hibiscone C analogue [66].

The reaction scheme for this synthesis is outlined in Figure 52.

The enol ether of dihydroresorcinol is formed using ethanol and *p*-toluene sulphonic acid. This compound then undergoes alkylation with 5-iodo-1-pentyne, and is further reduced then hydrolysed in one step to form the cyclohexenone. Photolysis of this compound produces an unstable cyclobutene ring which is easily opened during subsequent ozonolysis to produce the tricyclic hibiscone C analogue [72] (for experimental details see Appendix 6).

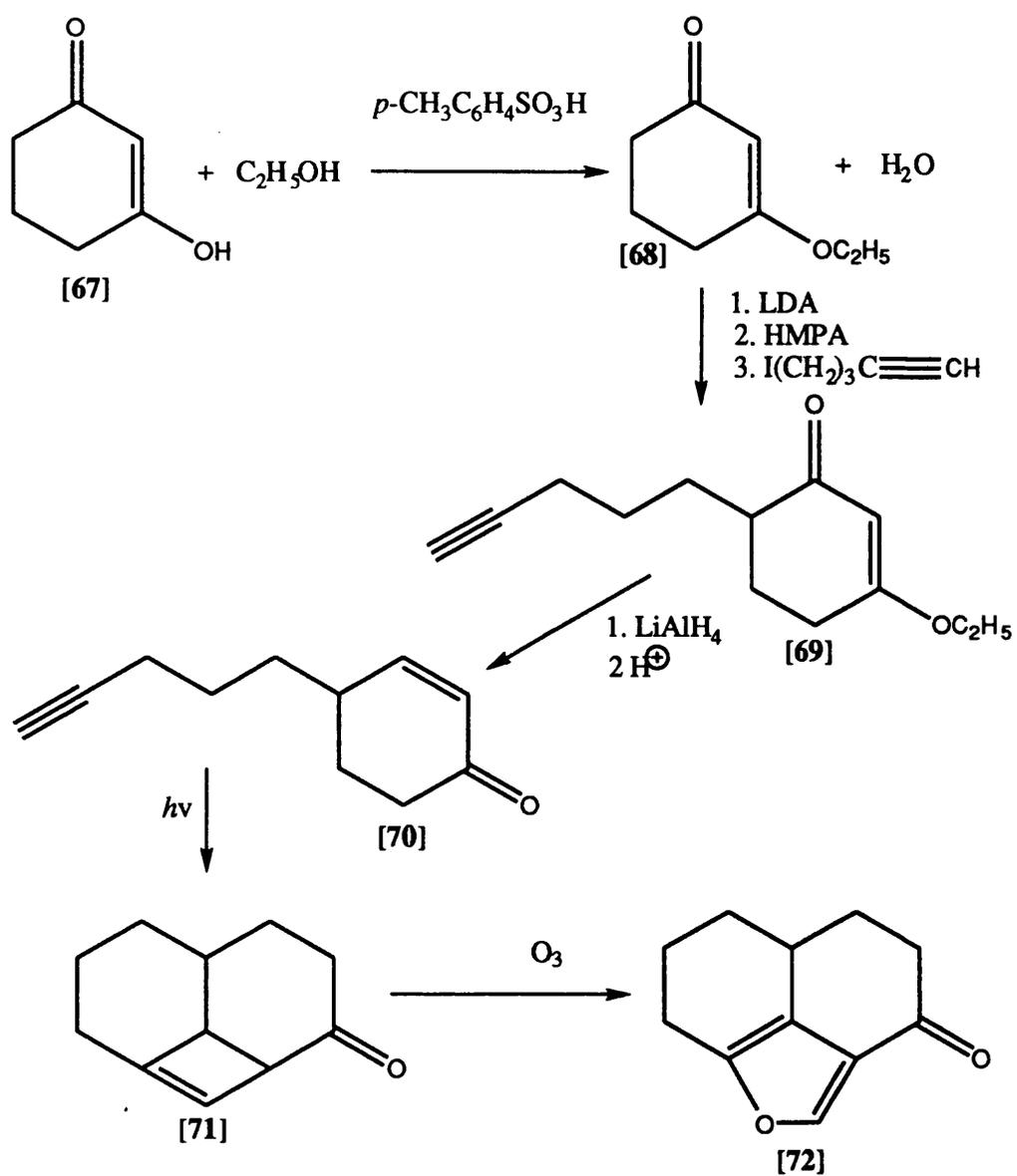


FIGURE 52 - Reaction scheme for hibiscone C analogue [72].

This compound, however, showed little or no cytotoxic activity and was not seen to inhibit the specified enzymes in biochemical tests. A second dicarbonyl analogue,[73] synthesised by a fellow postdoctoral worker⁷⁴, was also shown to be inactive.

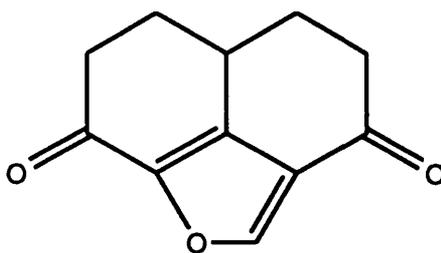


FIGURE 53 - Dicarbonyl hibiscone C analogue [73].

2. SYNTHETIC STUDIES TOWARDS WORTMANNIN

A report by workers in California⁷⁴ describes the first synthesis of the highly reactive furanocyclohexadienone lactone sub-unit of wortmannin (Fig. 54).

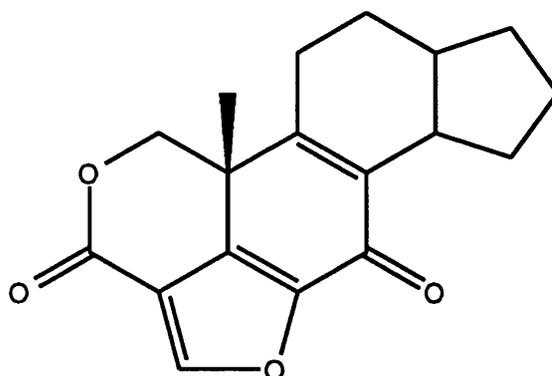


FIGURE 54 - Simplified wortmannin structure [74].

These workers acknowledged that wortmannin had been shown to inhibit *f*-Met-Leu-Phe stimulated superoxide production⁷⁵ and phospholipase D activation⁷⁶ in human neutrophils, perhaps by disrupting the function of the GTP-binding protein involved in the transduction of the *f*-Met-Leu-Phe signal. In spite of their interesting biological activities and challenging structures, these compounds have received little synthetic attention. This fact may be due, in part, to their chemical instability which promises to complicate efforts towards their synthesis. The route published circumvents many of these difficulties and should prove versatile enough to allow for the preparation of wortmannin itself.

The retrosynthetic analysis is outlined overleaf in Figures 55, 56 and 57.

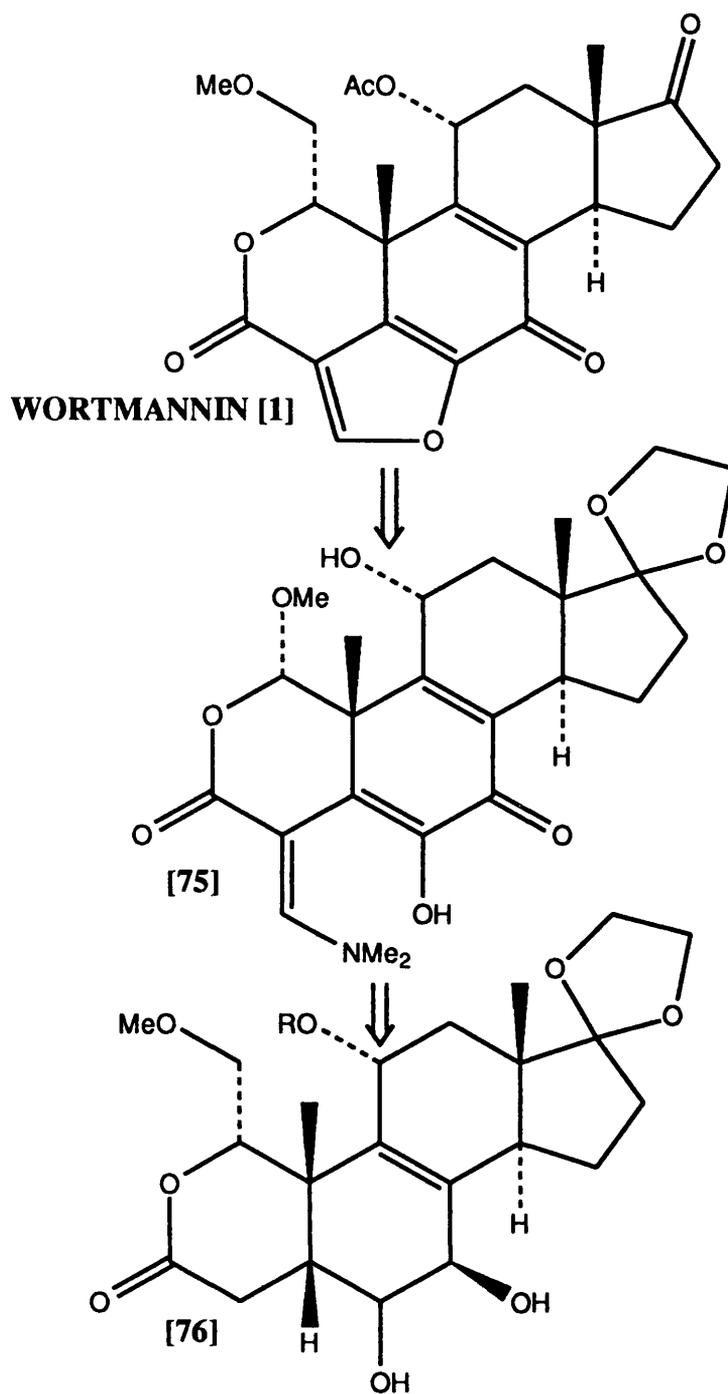


FIGURE 55 - Retrosynthetic analysis of wortmannin.
(continued overleaf)

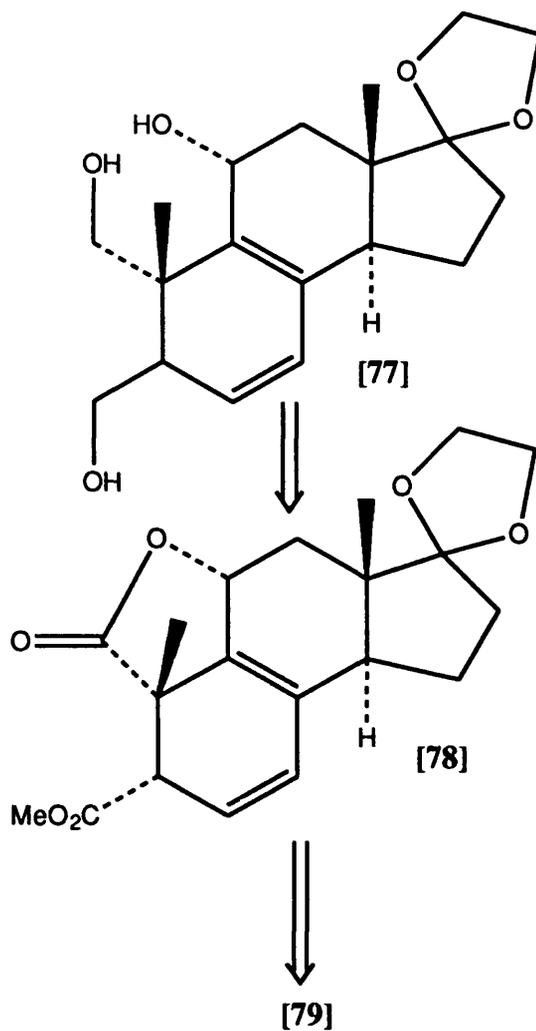


FIGURE 56 - Retrosynthetic analysis of wortmannin.
(cont. overleaf)

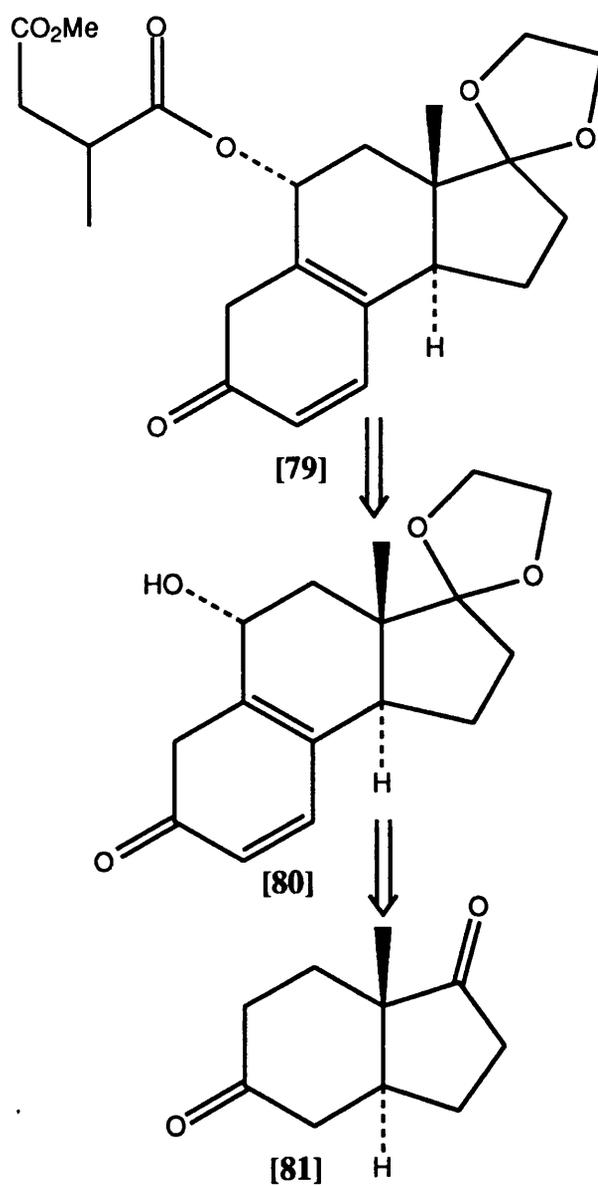
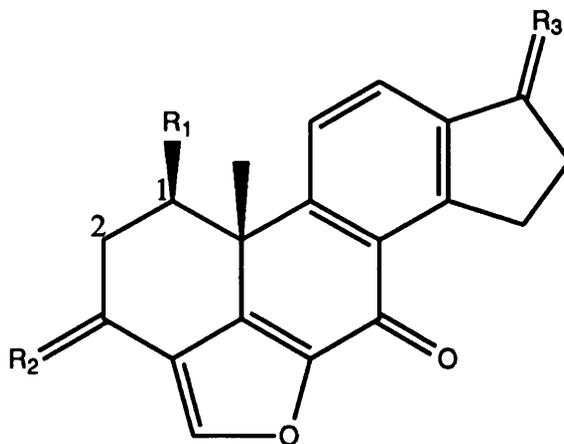


FIGURE 57 - Retrosynthesis of wortmannin (cont.)

The idea to use [75] as a precursor for the synthesis of wortmannin came from investigation by workers at Sandoz.⁷⁷ It was shown that wortmannin is an active alkylating agent, reacting with ammonia, aniline and diethylamine to give derivatives exactly analogous to [75]. In contrast to wortmannin, these aminomethylene lactones are stable to a range of nucleophilic conditions. Furthermore, treatment of these materials with acid regenerates wortmannin. Thus these lactones may be thought of as protected forms of wortmannin, a fact which may be of major importance if these compounds are to be delivered to the body as drugs (see Chapter 6, Discussion). In fact, the preparation of these compounds was considered by Broka *et al* to be the equivalent to the preparation of wortmannin itself. It seemed reasonable that [75] could be obtained from the diol lactone [76] and that this compound could in turn be obtained from the Diels-Alder derived intermediate [77]. The workers did not, however, wish to undertake such a lengthy synthesis without the prior assurance that the key transformations outlined would take place. Accordingly, the simplified wortmannin analogue [74] (Fig. 54), which still contains the reactive functionality present in natural systems, was selected as an initial target. This compound was synthesised in 13.5% overall yield.

3. SYNTHESIS OF DEMETHOXYVIRIDIN ANALOGUES

Following from the recently published synthesis of wortmannin, this approach, rather than that of Smith III *et al* , was adopted in the synthesis of a number of demethoxyviridin analogues by a postdoctoral student also involved in this project. The analogues synthesised are shown below⁷³.



Molecule	R ₁	R ₂	R ₃	Position 1,2
[18]	OH	N-OH	O	saturated
[19]	OCOCH ₃	O	O	saturated
[20]	OCOCH ₂ CH ₃	O	O	saturated
[21]	OCO(CH ₂) ₂ CH ₃	O	O	saturated
[53]	OCOPh	O	O	saturated
[54]	OH	N(CH ₃)OH	O	saturated
[55]	OCOCH ₃	N-OH	O	saturated
[56]	OCOCH ₃	N-OH	OH	saturated
[57]	-	O	O	unsaturated

FIGURE 58 - Demethoxyviridin analogues synthesised.

4. BIOSYNTHESIS

Although neither viridin or wortmannin have been synthesised there has been extensive study into the biosynthesis of these compounds. In 1969⁷⁸ Grove *et al* realised that the structure of viridin strongly suggested a steroidal biogenetic pathway

The biosynthesis of steroids may be divided into two main phases⁷⁹. The first phase involves the assembly of the prenyl pyrophosphate units, from mevalonic acid, into two head to head condensed farnesyl pyrophosphate units and subsequent formation of squalene oxide and cyclisation to lanosterol (see Figs. 59 and 60 overleaf). The second phase involves the oxidative degradation and rearrangements of C₃₀ lanosterol to give the individual families of steroids.

FIGURE 59 (overleaf) - Formation of squalene from farnesyl PP.

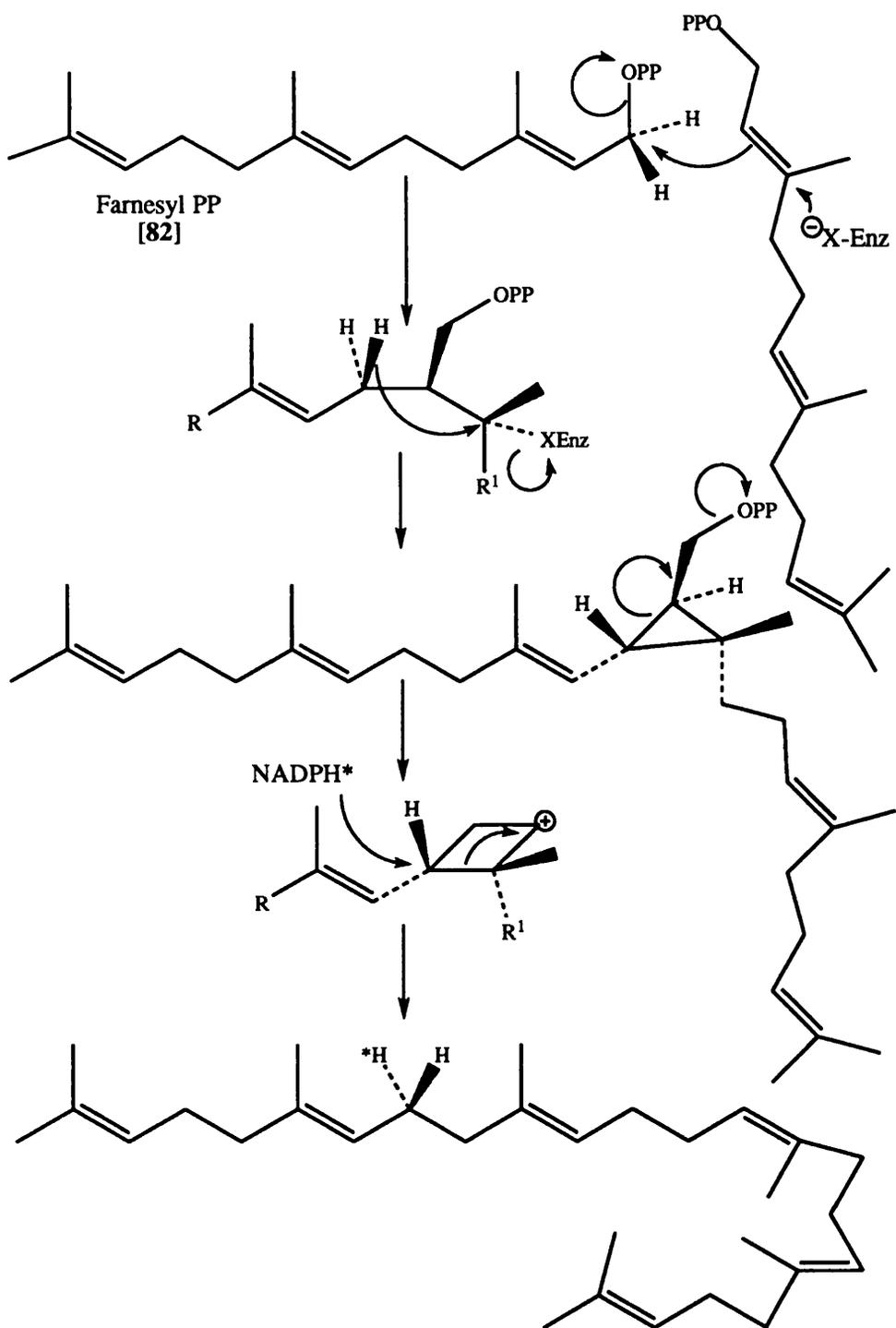


FIGURE 59 - Formation of squalene [83].

The cyclisation of squalene is initiated by the acid catalysed ring-opening of squalene mono-epoxide, and may be a completely concerted process, or perhaps interrupted at some intermediate stage to allow rearrangement to occur.

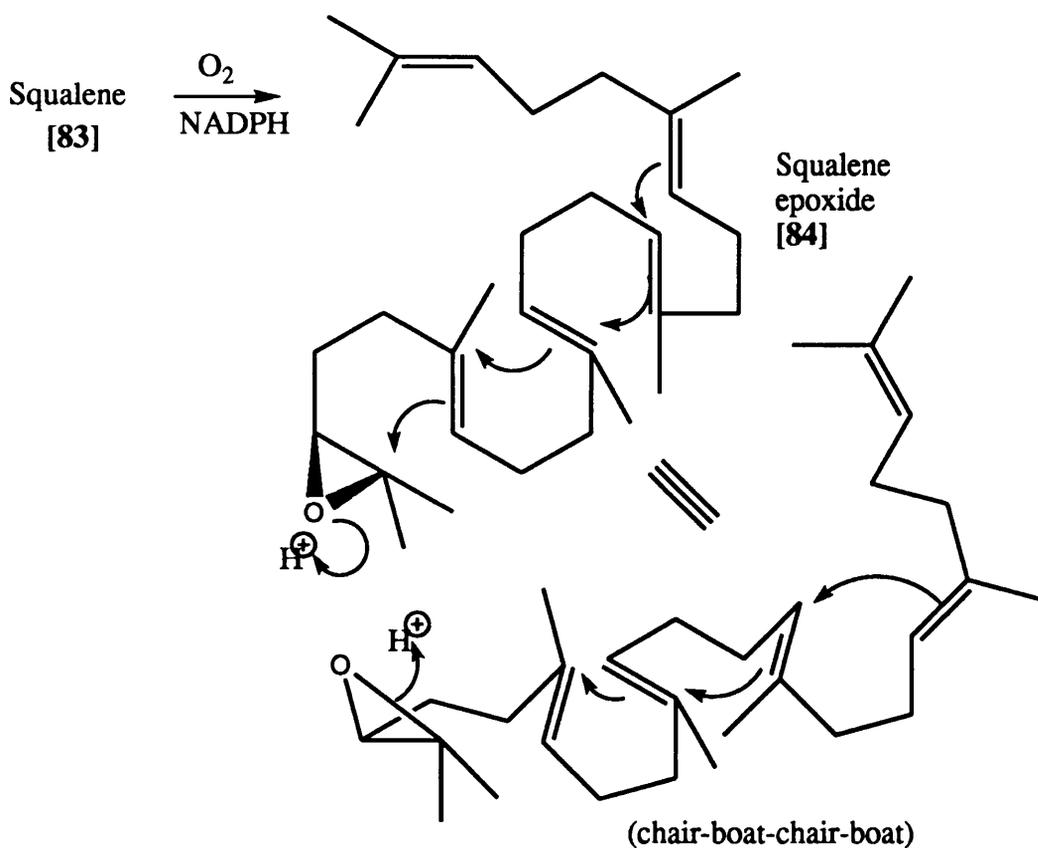


FIGURE 60 - Epoxidation of squalene.

The cyclised molecule then undergoes rearrangement to form lanosterol and degradation to form various other steroids such as cholesterol, sex hormones, bile acids and vitamin D.

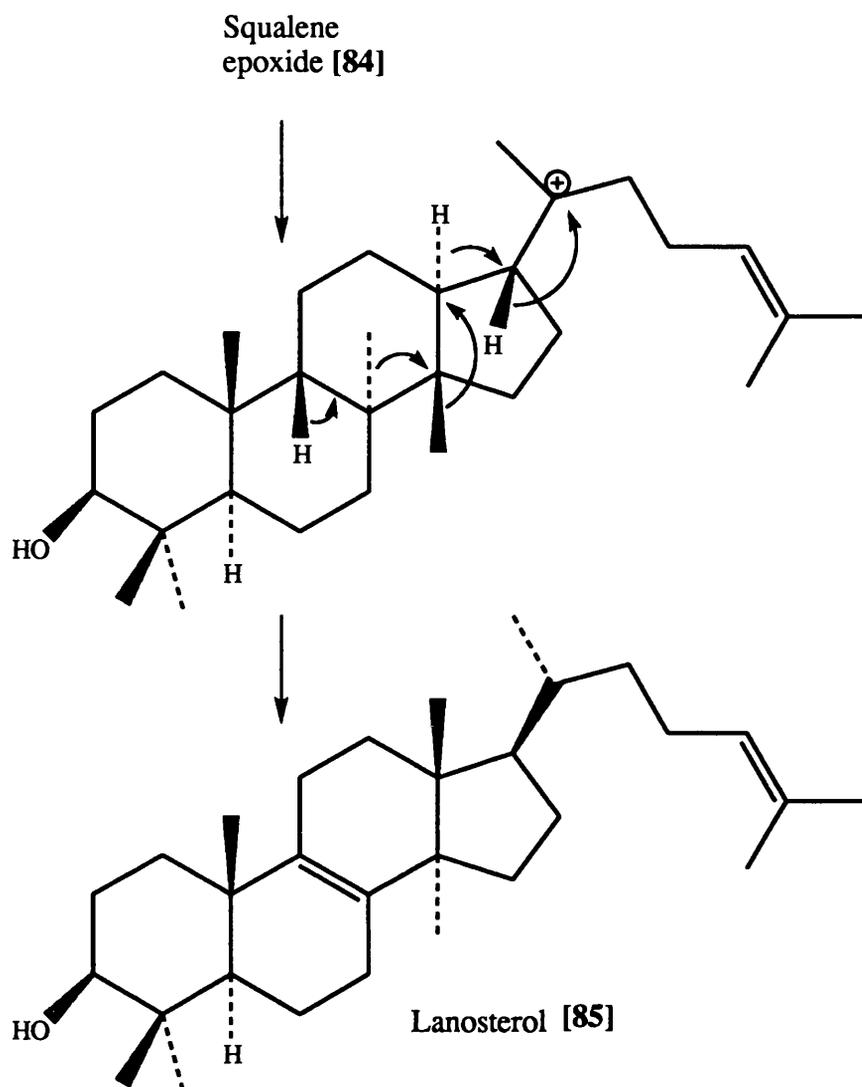


FIGURE 61 - Generation of lanosterol [85].

Biosynthetic studies of viridin have been fairly extensive. It was known that [^{14}C]viridin, derived from [$2\text{-}^{14}\text{C}$]mevalonate, had a labelling pattern consistent with its biogenesis by tail-to-tail condensation of two farnesyl residues.⁷⁸ The steroidal biogenetic pathway of viridin was shown to be from (R)-mevalonic acid [86] by way of farnesyl pyrophosphate, squalene epoxide [84] and lanosterol [85].

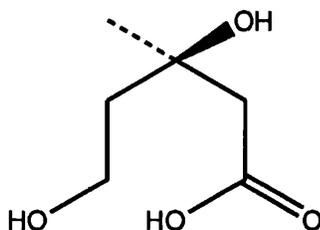


FIGURE 62 - (R)-Mevalonic acid [86].

Biosynthetic studies of wortmannin were undertaken by a Simpson *et al*⁸⁰ in 1979. By enrichment of *Penicillium dangeardii* cultures with sodium [$1,2\text{-}^{13}\text{C}_2$]acetate, it was shown that the biogenesis of this metabolite was consistent with the anticipated pathway *via* triterpenoid intermediates. Like viridin, the acetate was incorporated *via* mevalonic acid, squalene, and lanosterol.

CONCLUSION

A variety of techniques has been employed to provide valuable detailed structural information about this series of fungal metabolite analogues with potential novel cytotoxic mechanisms of action .

The extraction of viridin from *Trichoderma viride* was optimised and carefully monitored until the yield was increased four fold. This was fairly low, however, so synthesis of part structure analogues was attempted. Also, attention was turned to the fungus *Nodulisporum hinnuleum* which generated high yields of demethoxyviridin. Although the synthesis of part structure analogues was in itself successful, the compounds produced were shown to be inactive. A more rational approach was then adopted by means of computer modelling and drug design. A QSAR model was built using SYBYL which enabled prediction of activity of new molecules. This study enabled guiding in the synthesis of new molecules and was eventually a contributing factor in the synthesis of a molecule several hundred times more active than the lead compound.

Much encouraged by this result, a patent is being applied for this analogue while consideration is being made of the stability required to enable formulation of the drug to a form acceptable for administration to animals and for phase I-III clinical trials in humans.

CHAPTER 6
DISCUSSION

DISCUSSION

1. EXTRACTION

If time had allowed, a greater degree of optimisation of the extraction conditions may have afforded a higher yield of viridin, *i.e.* the growing medium may have been changed to determine the optimum nutrients required for maximum growth and production of viridin by the fungus. More detailed pH or temperature studies may have been beneficial. Growth may also have been simplified by growing in one fermentation vessel if the necessary equipment was available.

Growth and extraction of wortmannin from *Penicillium dangeardii* was delayed due to several contaminated batches of fungus being sent from Kew initially. The extraction is now in process.

The extraction of demethoxyviridin from *Nodulisporum hinnuleum* proved to be much more viable. Reported literature yields⁸¹ were approaching 1g compound per litre of broth. The production of this compound is now being undertaken on a large scale at SmithKline Beechams' production plant.

2. CRYSTALLOGRAPHY

The crystal structure determination of several more analogues is now in progress. The structures assume that the conformation adopted by the molecules in the unit cell is similar to that in solution as it is the most energetically favourable. To investigate how realistic this assumption was, a model of viridin was built and energy minimised using SYBYL, and

compared to a structure derived from the imported crystal co-ordinates from the Cambridge Structural Database. The structures were found to be exactly the same when overlapped. This may be due to the rigidity of these types of molecules. Their planar and pentacyclic ring structure does not lend itself to conformational changes. For other, less rigid molecules, differences may well be expected between crystal structure and energy minimised model.

3. MODELLING

Initial modelling studies were based on the assumption that the anti-inflammatory structure-activity relationship of the wortmannin analogues would be at least in some way directly related to the cytotoxicity test results of these compounds. This assumption was taken in the absence of any other test data, and because wortmannin had a roughly similar order of potency in both the anti-inflammatory test system and in cytotoxicity tests. This assumption, in fact, turned out to be justified and it was shown that molecules which had a predicted activity according to the wortmannin anti-inflammatory model had the same order, if not magnitude, of potency when predicted by the cytotoxic viridin model.

The building and validation of this model could then predict activities of new compounds, before they were synthesised, and was a central part of this project allowing the synthetic efforts to be guided in a more rational way.

A more helpful situation would obviously be where the identity and structure of the target protein is known, *i.e.* where and how exactly the molecules are binding. The solubilisation and purification of phospholipase D, with a view to crystallisation and determining the crystal

structure, is in process in the Biochemistry department. Rat-tissue phospholipase D has previously been purified⁸². It is becoming increasingly apparent with recent results, however, that the target proteins may be tyrosine kinase. The planar structure of the molecules themselves suggest they are binding in a cleft or groove of a particular protein. In the future, an unambiguous way to determine where they are binding would obviously be to identify, purify and crystallise the target protein, determine the crystal structure with and without the compound bound to its active site, and compare the two.

4. SYNTHESIS

From the viridin, demethoxyviridin and wortmannin structures, and the knowledge that all anti-inflammatory activity of wortmannin is lost when the furan ring is broken, this furan flanked by two carbonyl groups is thought to be central to the activity of these compounds. Modification of existing lead compounds to establish a structure-activity relationship was established by computer modelling studies. Results to date suggest that the more planar the molecule, the higher its biological activity. This, again, may be due to its binding into a narrow cleft in a protein.

Synthesis of an analogue of hibiscone C, was successful although the molecule itself was not active. A second approach to the problem is to synthesise a number of furan derivatives containing suitable functional groups in order to attempt to build up the minimum structure for activity. Derivatives of demethoxyviridin such as the 1-epimer, derivatisation at C1, functionalisation of ring D and modification of the furan ring are being synthesised. Crystal structure determination of these analogues will be performed in order to assist in modelling studies.

Thus a combination of different techniques has led to the design, synthesis and testing of analogues several hundred times more active than the original lead compound. The synthetic efforts are now being directed towards the problems of stability.

PROBLEMS OF STABILITY

These compounds are very unstable in an aqueous environment at pH greater than 5, with a half life in the order of minutes, and so would rapid degradation at physiological pH would be expected. Presuming their stability, *e.g.* expressed as half-life in solution at certain pH, is structure dependent, the possibility exists to model these compounds using the CoMFA technology, in order to predict stability.

To deliver these compounds as drugs would present numerous problems. The entire formulation process would have to be strictly pH controlled with the only obvious stable form of drug as a freeze dried powder for reconstitution and injection. Appropriate time stability studies would be necessary to indicate stability of the compound in the solid state, although data so far suggests little degradation takes place. A light stability study may also be appropriate since increased production of the compound by the fungus is seen in the dark, possibly due to photo-degradation in the light. This may necessitate storage of the compound in amber or light resistant vials.

The very short half life of the compound may necessitate protection of the drug until it reaches the target site. Here, a variety of drug delivery methods may come into play, *e.g.* pro-drugs or carrier vehicles.

The compound could be delivered as an inactive pro-drug, *e.g.* with a broken ring (see Chapter 5), to reduce the chances of hydrolysis in the bloodstream. This derivative would have to be designed so that it is released at the active site as an active drug. This method would need more detailed studies of precise mechanism of action a particular enzymes could

be used as a means to cleave the functional group from the pro-drug and allow recyclisation of the furan. Some of the broken ring adducts, *e.g.*, those produced on reaction of wortmannin or viridin with diethylamine, are stable to many forms of nucleophilic attack and easily recyclise in the presence of acid. This would suggest the stomach as an ideal site of release and a possible way to treat stomach cancer.

The drug could also be delivered by means of a carrier, protected from the external aqueous environment by a hydrophilic matrix, *e.g.* liposomes or cyclodextrins. This vehicle would still need to be delivered to the active site. This may happen via normal bloodstream flow, although problems of poor tissue perfusion and decreased blood flow are often seen at the site of a solid tumour. one method of targeting therefore may involve attachment of a semi-synthetic antibody for a particular protein to several molecules of drug. A problem, again, with this method is the difficulty in achieving sufficient loading of the antibody with enough drug to produce a therapeutic effect. The affinity of the drug to the antibody must be maximised. Problems of rejection of this foreign body by the body's immune system must also be overcome. Although site specific delivery is ideal in theory, in practice it can be fraught with difficulties.

Additional formulation problems which may be foreseen are non-homogenous particle size. This may necessitate salt formation or experimenting with different salts and conditions of crystallisation to achieve the desired result. Compatibility with excipients, *e.g.*, diluent, glidant, coating agent of a tablet formulation, would be assessed as well as possible interactions with other drugs.

Cancer patients are often on a complex multiple dosage system comprising many different drugs which are administered concurrently or at optimum intervals to provide least depression of the white blood cell count. If a new drug is to be introduced into this complex regimen, all possible interactions or cumulative side effects must be carefully assessed to maximise benefit and minimise risk to a patient.

In conclusion, the studies reported in this thesis have guided and led to the synthesis of a compound not only with high cytotoxic activity but also a possible novel mechanism of anti-cancer action which would provide selectivity of action in future clinical trials.

APPENDICES

APPENDIX 1

TABLE 1 - Fractional atomic co-ordinates and isotropic temperature factors (angstrom squared), with standard deviations in the least significant digits in parentheses. For anisotropic atoms, the equivalent isotropic temperature factors are shown (continued overleaf).

ATOM	X/A	Y/B	Z/C	U
C(1)	0.1452(4)	0.1651(2)	0.8683(2)	0.030
C(2)	0.2305(4)	0.1022(2)	0.9012(2)	0.039
C(3)	0.4079(4)	0.0961(2)	0.8895(2)	0.039
C(4)	0.4833(4)	0.1614(2)	0.8717(2)	0.033
C(5)	0.3953(3)	0.2237(2)	0.8691(2)	0.027
C(6)	0.4911(4)	0.2730(2)	0.8416(2)	0.032
C(7)	0.4437(4)	0.3424(2)	0.8204(2)	0.030
C(8)	0.2664(4)	0.3510(2)	0.8220(2)	0.028
C(9)	0.1632(4)	0.3007(1)	0.8535(2)	0.027
C(10)	0.2265(3)	0.2343(2)	0.8910(2)	0.026
C(11)	0.0021(4)	0.3142(2)	0.8551(2)	0.032
C(12)	-0.0613(4)	0.3746(2)	0.8268(2)	0.036
C(13)	0.0382(4)	0.4228(2)	0.7955(2)	0.034
C(14)	0.2032(4)	0.4120(2)	0.7924(2)	0.032
C(15)	0.2824(5)	0.4709(2)	0.7526(2)	0.042
C(16)	0.1532(5)	0.5234(2)	0.7375(3)	0.048
C(17)	-0.0011(5)	0.4913(2)	0.7645(3)	0.044
C(19)	0.2169(4)	0.2457(2)	0.9767(2)	0.037
C(22)	0.6307(4)	0.1778(2)	0.8479(2)	0.041
O(18)	-0.1309(4)	0.5172(2)	0.7625(2)	0.067
O(20)	-0.0167(3)	0.1619(2)	0.8890(2)	0.040

TABLE 1 - Fractional atomic co-ordinates and isotropic temperature factors (angstrom squared), with standard deviations in the least significant digits in parentheses. For anisotropic atoms, the equivalent isotropic temperature factors are shown (continued).

ATOM	X/A	Y/B	Z/C	U
O(21)	0.4767(4)	0.0411(2)	0.8941(2)	0.063
O(23)	0.6394(3)	0.2462(1)	0.8291(2)	0.040
O(24)	0.5343(3)	0.3876(1)	0.8005(2)	0.046
C(1')	1.0163(4)	0.2549(2)	0.5627(2)	0.033
C(2')	1.0829(4)	0.3246(2)	0.5891(2)	0.039
C(3')	0.9623(5)	0.3823(2)	0.6063(2)	0.039
C(4')	0.8076(5)	0.3549(2)	0.6267(2)	0.039
C(5')	0.7782(4)	0.2822(2)	0.6296(2)	0.032
C(6')	0.6213(4)	0.2744(2)	0.6409(2)	0.050
C(7')	0.5378(4)	0.2118(2)	0.6375(2)	0.050
C(8')	0.6375(4)	0.1531(2)	0.6098(2)	0.050
C(9')	0.8031(4)	0.1602(2)	0.5950(2)	0.050
C(10')	0.8962(4)	0.2256(2)	0.6191(2)	0.050
C(11')	0.8856(4)	0.1047(2)	0.5635(2)	0.050
C(12')	0.3119(5)	0.4564(2)	0.4557(2)	0.050
C(13')	0.6522(5)	0.0365(2)	0.5589(2)	0.050
C(14')	0.5632(4)	0.0898(2)	0.5915(2)	0.050
C(15')	0.3940(5)	0.0687(2)	0.6008(3)	0.050
C(16')	0.3829(5)	-0.0033(2)	0.5665(3)	0.050
C(17')	0.5462(5)	-0.0224(2)	0.5424(2)	0.050
C(19')	0.9750(4)	0.2080(2)	0.6942(2)	0.050
C(22')	0.1656(5)	0.1137(2)	0.3628(3)	0.050
O(18')	0.5886(4)	-0.0773(1)	0.5134(2)	0.050
O(20')	1.1462(3)	0.2108(1)	0.5488(2)	0.050
O(21')	0.9939(4)	0.4430(1)	0.5998(2)	0.050
O(23')	0.5488(3)	0.3389(1)	0.6459(2)	0.050
O(24')	0.3975(3)	0.2056(1)	0.6526(2)	0.050

APPENDIX 2

TABLE 2 - Bond lengths (Å) of both molecules of demethoxyviridin compared to viridin.

BOND	MOLECULE 1	MOLECULE 2	VIRIDIN
	Demethoxyviridin	Demethoxyviridin	
C(1) - C(2)	1.534(5)	1.536(5)	1.537
C(1) - C(10)	1.559(5)	1.563(5)	1.554
C(1) - O(20)	1.422(4)	1.414(5)	1.417
C(2) - C(3)	1.522(6)	1.544(6)	1.518
C(3) - C(4)	1.451(5)	1.461(6)	1.441
C(3) - O(21)	1.216(5)	1.210(5)	1.213
C(4) - C(5)	1.416(5)	1.427(5)	1.427
C(4) - C(22)	1.357(5)	1.361(6)	1.376
C(5) - C(6)	1.346(5)	1.354(5)	1.348
C(5) - C(10)	1.498(5)	1.474(5)	1.490
C(6) - C(7)	1.541(5)	1.403(5)	1.444
C(6) - O(23)	1.377(5)	1.393(5)	1.398
C(7) - C(8)	1.512(5)	1.499(5)	1.491
C(7) - O(24)	1.216(5)	1.225(5)	1.229
C(8) - C(9)	1.425(5)	1.435(5)	1.402
C(8) - C(14)	1.398(5)	1.415(5)	1.424
C(9) - C(10)	1.545(5)	1.536(5)	1.538
C(9) - C(11)	1.390(5)	1.399(5)	1.409
C(10) - C(19)	1.553(5)	1.553(5)	1.538
C(11) - C(12)	1.381(5)	1.380(6)	1.374
C(12) - C(13)	1.377(5)	1.385(6)	1.387
C(13) - C(14)	1.415(6)	1.404(5)	1.393
C(13) - C(17)	1.474(5)	1.480(6)	1.478
C(14) - C(15)	1.502(5)	1.500(6)	1.500
C(15) - C(16)	1.518(6)	1.524(6)	1.533
C(16) - C(17)	1.526(7)	1.496(6)	1.510
C(17) - O(18)	1.209(6)	1.235(5)	1.218
C(22) - O(23)	1.367(5)	1.358(5)	1.368

APPENDIX 3

TABLE 3 - Bond angles ($^{\circ}$) of both molecules of demethoxyviridin compared to viridin (continued overleaf).

BOND	MOLECULE 1	MOLECULE 2	VIRIDIN
	Demethoxyviridin	Demethoxyviridin	
C2-C1-C10	112.3(4)	111.7(3)	114.8
C10-C1-O20	112.9(4)	114.3(3)	112.6
C2-C3-C4	113.7(5)	113.3(3)	113.4
C4-C3-O21	125.0(5)	124.9(4)	124.8
C3-C4-C22	133.3(5)	132.2(4)	132.3
C4-C5-C6	107.2(5)	106.4(3)	107.7
C6-C5-C10	125.7(5)	125.4(3)	124.2
C5-C6-O23	110.2(5)	110.6(3)	110.4
C6-C7-C8	112.1(4)	112.4(3)	111.4
C8-C7-O24	123.5(5)	123.8(4)	125.4
C7-C8-C14	118.8(5)	118.7(3)	117.9
C8-C9-C10	122.2(4)	120.7(3)	120.9
C10-C9-C11	119.3(4)	119.0(3)	118.3
C5-C10-C9	108.9(4)	108.9(3)	108.6
C9-C10-C19	106.7(4)	106.2(3)	106.6
C11-C12-C13	119.0(5)	119.5(4)	118.0
C12-C13-C17	128.6(6)	128.9(4)	127.5
C8-C14-C13	120.8(5)	120.2(4)	118.6
C13-C14-C15	109.5(5)	110.6(4)	111.7
C15-C16-C17	106.8(5)	106.2(4)	106.2
C13-C17-O18	126.1(6)	125.5(4)	125.3
C4-C22-O23	110.9(5)	111.3(4)	111.7
C2-C1-O20	108.9(4)	107.9(3)	108.9

TABLE 3 - Bond angles ($^{\circ}$) of both molecules of demethoxyviridin compared to viridin (continued).

BOND	MOLECULE 1 Demethoxyviridin	MOLECULE 2 Demethoxyviridin	VIRIDIN
C1-C2-C3	118.4(5)	116.6(3)	117.1
C2-C3-O21	121.2(5)	121.8(4)	121.4
C3-C4-C5	120.8(5)	121.0(4)	121.3
C5-C4-C22	105.6(5)	106.4(4)	105.1
C4-C5-C10	127.1(5)	128.1(4)	128.1
C5-C6-C7	125.7(5)	125.5(4)	125.6
C7-C6-O23	123.9(5)	123.4(4)	123.8
C6-C7-O24	124.3(5)	123.6(4)	123.1
C7-C8-C9	122.1(5)	123.1(3)	123.7
C9-C8-C14	119.1(5)	118.0(3)	118.3
C8-C9-C11	118.4(5)	120.2(3)	120.6
C1-C10-C5	104.2(4)	104.6(3)	105.3
C1-C10-C19	111.4(4)	110.0(3)	112.1
C5-C10-C19	109.2(4)	110.4(3)	110.2
C9-C11-C12	122.3(5)	120.7(4)	121.3
C12-C13-C14	120.4(5)	121.4(4)	123.3
C14-C13-C17	111.0(5)	109.6(4)	109.2
C8-C14-C15	129.6(5)	129.3(4)	129.7
C14-C15-C16	105.3(5)	105.1(4)	104.9
C13-C17-C16	107.2(5)	108.4(4)	107.9
C16-C17-O18	126.6(6)	126.1(4)	126.9
C6-O23-C22	106.0(4)	105.4(3)	105.1

APPENDIX 4

TABLE 4 - Torsion angles of both molecules of demethoxyviridin (continued overleaf).

TORSION ANGLE	MOLECULE 1 Demethoxyviridin	MOLECULE 2 Demethoxyviridin
C10-1-2-3	-51.3(4)	-54.9(4)
C2-1-10-5	52.7(3)	53.0(3)
C2-1-10-9	172.8(4)	173.3(4)
C2-1-10-19	-64.3(4)	-65.5(4)
O20-C1-2-3	-177.2(5)	-178.8(5)
O20-C1-10-5	176.0(4)	175.0(4)
O20-C1-10-9	-63.9(3)	-64.7(4)
O20-C1-10-19	59.0(3)	56.6(3)
C1-2-3-4	20.6(4)	25.4(4)
C1-2-3-O21	-158.7(6)	-151.3(6)
C2-3-4-5	2.3(4)	0.2(4)
C2-3-4-22	-169.9(6)	-171.0(6)
O21-C3-4-5	-178.5(6)	-176.8(6)
O21-C3-4-22	9.3(5)	5.6(5)
C3-4-5-6	-171.7(5)	-172.7(5)
C3-4-5-10	7.2(4)	4.8(4)
C3-4-22-O23	171.7(6)	171.7(7)
C22-4-5-6	2.3(4)	0.5(4)
C22-4-5-10	-178.7(5)	-178.0(5)
C5-4-22-O23	-1.3(3)	-0.5(4)
C4-5-6-7	171.5(5)	171.0(5)
C4-5-6-O23	-2.5(3)	-0.4(3)
C4-5-10-1	-34.8(4)	-31.5(4)
C4-5-10-9	-159.6(5)	-156.8(5)
C4-5-10-19	83.4(4)	86.6(4)
C6-5-10-1	144.0(4)	145.5(5)
C10-5-6-7	-7.5(4)	-6.6(4)

TABLE 4 - Torsion angles of demethoxyviridin (continued overleaf).

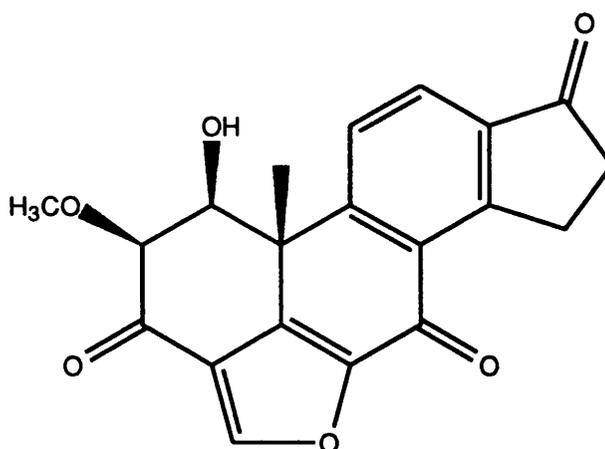
TORSION ANGLE	MOLECULE 1 Demethoxyviridin	MOLECULE 2 Demethoxyviridin
C6-5-10-9	19.1(4)	20.2(4)
C6-5-10-19	-97.8(4)	-96.4(5)
C10-5-6-O23	178.5(5)	177.9(5)
C5-6-7-8	-8.4(4)	-7.4(4)
C5-6-7-O24	174.6(6)	175.2(7)
C5-6-O23-22	1.7(3)	0.1(4)
O23-C6-7-8	164.8(5)	162.9(5)
C7-6-O23-22	-172.5(5)	-171.5(5)
O23-C6-7-24	-12.1(4)	-14.5(4)
C6-7-8-9	10.6(3)	5.0(4)
C6-7-8-14	-170.8(5)	-169.5(5)
O24-C7-8-9	-172.4(6)	-177.6(6)
O24-C7-8-14	6.2(4)	7.9(4)
C7-8-9-10	2.2(3)	10.1(3)
C7-8-9-11	177.5(5)	174.8(5)
C7-8-14-13	-177.7(5)	-174.0(5)
C7-8-14-15	5.2(4)	5.2(4)
C14-8-9-10	-176.4(5)	-175.3(5)
C14-8-9-11	-1.1(4)	-0.3(4)
C9-8-14-13	1.0(4)	0.7(4)
C9-8-14-15	-176.1(6)	-179.9(6)
C8-9-10-1	-133.0(4)	-139.2(5)
C8-9-10-5	-16.0(3)	-21.1(3)
C8-9-10-19	102.2(4)	97.9(4)
C8-9-11-12	0.4(4)	1.7(4)
C11-9-10-1	51.7(4)	45.8(4)
C11-9-10-5	168.6(4)	163.9(5)
C10-9-11-12	175.9(5)	176.8(6)
C11-9-10-19	-73.1(4)	-77.1(4)
C9-11-12-13	0.4(4)	-2.0(4)

TABLE 4 - Torsion angles of both molecules of demethoxyviridin
(continued).

TORSION ANGLE	MOLECULE 1 Demethoxyviridin	MOLECULE 2 Demethoxyviridin
C11-12-13-14	-0.5(4)	-0.9(4)
C11-12-13-17	-177.7(6)	178.5(7)
C12-13-14-8	-0.2(4)	0.4(4)
C12-13-14-15	177.4(5)	179.8(6)
C12-13-17-16	179.8(6)	177.7(6)
C12-13-17-18	1.2(5)	2.1(4)
C17-13-14-8	177.4(5)	177.6(5)
C17-13-14-15	-4.9(4)	1.8(4)
C14-13-17-16	2.4(4)	0.1(4)
C14-13-17-18	-176.3(6)	-179.9(6)
C8-14-15-16	-177.3(6)	-176.4(6)
C13-14-15-16	5.4(4)	-2.9(4)
C14-15-16-17	-3.7(4)	-2.8(4)
C15-16-17-13	1.0(4)	1.8(4)
C15-16-17-18	179.6(6)	178.4(6)
C4-22-O23-6	-0.2(4)	-0.2(4)

APPENDIX 5

VIRIDIN [2]

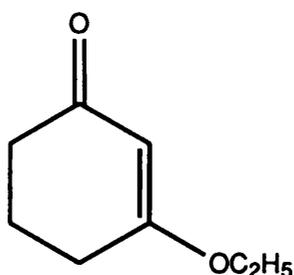


($R_f = 0.47$) M. Pt. 222-224°C, NMR: H1, 4.45 (J=5Hz); H2, 3.95; H11, 8.10 (J=8Hz); H12, 8.80, H15, 2.90; H16, 3.80; H18, 1.73; H19, 8.45; H20, 3.75. Calculated molecular weight: 352; found 352.

APPENDIX 6

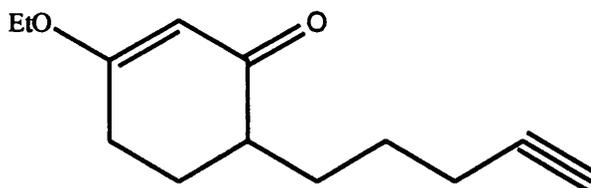
ORGANIC SYNTHESIS - EXPERIMENTAL

1. 3-ETHOXY-2-CYCLOHEXENONE [68]



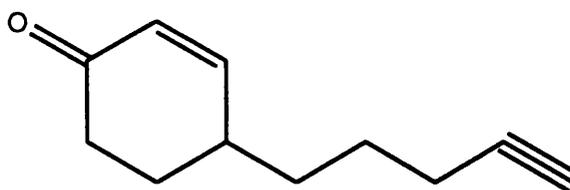
Into a flask fitted with a Dean Stark apparatus was placed a solution of dihydroresorcinol (5.3g), *p*-toluene sulphonic acid (0.23g) and absolute ethanol (25ml) in benzene (90ml)⁸⁴. The mixture was heated to boiling and the azeotrope composed of toluene, alcohol and water was removed at the rate of 10ml per hour. When the temperature of the distilling vapour reached 78°C after 6-8 hours, the distillation was stopped and the residual solution washed with portions of 10% aqueous sodium hydroxide (4 x 10ml) saturated with sodium chloride. The resulting organic solution was washed with successive portions of water (5ml) until the aqueous washings were neutral and then concentrated under reduced pressure to produce dihydroresorcinol monoethyl ether (3-ethoxy-2-cyclohexenone) [68]. (4.06g, 76.6%): IR (CCl₄) 3010 (s), 2940 (s), 1740 (s), 1710 (s), 1230 (s), 1190 (s) cm⁻¹; NMR (CDCl₃, 90MHz) 5.2 (s, 1H), 3.75 (q, 2H), 2.2 (t, 4H), 1.85 (m, 2H), 1.2 (t, 3H); mass calculated for C₈H₁₂O₂: 140; found 140.

2. 3-ETHOXY-6-(4-PENTYNYL)-2-CYCLOHEXANE [69]



The next step was an alkylation using 5-iodo-1-pentyne as the alkylating agent⁸⁵. All glassware was flame dried under vacuum before starting the reaction. To a solution of diisopropylamine in 5ml dry THF (0.91ml) at 0°C under vacuum was added *n*-butyl lithium (4.91; 1.6M). After 15 minutes at 0°C, the solution was cooled to -78°C and 3-ethoxy-2-cyclohexenone [68] (0.92g) in 2.5ml THF was added dropwise over 45 minutes via syringe. The reaction mixture was then allowed to warm up to -40°C and HMPA (1.6g) in THF (2ml) was added followed by 5-iodo-1-pentyne (1.91g). This was allowed to warm to 25°C then left at this temperature for a further hour. Saturated ammonium chloride (about 10ml) was then added and the organic phase diluted with hexane (approximately 20ml), washed with water, dried using magnesium sulphate and concentrated under reduced pressure to give 3-ethoxy-6-(4-pentynyl)-2-cyclohexenone [69]. (1.36g 71.3%): IR (CCl₄) 3320 (s), 2940 (s), 2100 (w), 1660 (s), 1610, 1380 (s), 1200cm⁻¹; NMR (CDCl₃, 250MHz) 1.34 (t, J = 6.8 Hz, 3H), 1.4-2.24 (m, 10H), 2.4 (t, J = 5.8 Hz, 2H), 3.9 (q, J = 6.8 Hz, 2H), 5.26 (s, 1H); mass calculated for C₁₃H₁₈O₂:206; found 206.

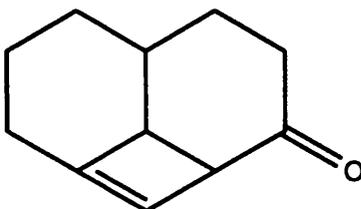
3. 4-(4-PENTYNYL)-2-CYCLOHEXENONE [70]



The next stage was a 2 step 1,3-carbonyl transposition by first reducing the carbonyl in [69] with lithium aluminium hydride, followed by hydrolysis of the enol ether to the ketone then elimination to form [70]. 3-Ethoxy-6-pentynyl-2-cyclohexenone [69] (0.92g) in THF (2.5ml) was added dropwise to a stirred suspension of LiAlH_4 (0.13g) in THF (10ml) under nitrogen. the reaction mixture was stirred at 25°C for two hours. Excess hydride was destroyed by addition of solid sodium sulfate decahydrate. Water (2.5ml) was then added followed by sufficient 2N HCl to adjust the pH of the aqueous phase to 2 (approximately 3ml). The reaction mixture was again left to stir for 30 minutes, the organic phase removed and washed with saturated sodium bicarbonate and brine and dried to give [70]. (0.199g; 21.5%):IR (CCl_4) 3320, 3020 (w), 2940 (s), 1685 (s), 1250, 900 (s), 620 cm^{-1} ; NMR (CDCl_3 , 250 MHz) 1.98 (t, $J = 2.5\text{ Hz}$, 1H), 1.5-2.56 (m, 11H), 6.0 (dd, $J = 10, 2.5\text{ Hz}$, 1H), 6.86 (d, $J = 10\text{ Hz}$, 1H); mass calculated for $\text{C}_{11}\text{H}_{14}\text{O}$: 162; found 162.

4. 1a, 3, 4, 4a, 5, 6, 7, 7b-OCTAHYDRO-2H-
CYCLOBUT[de]NAPHTHALEN-2-ONE

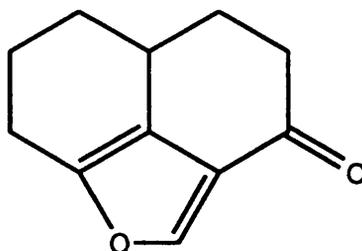
[71]



Photolysis of [70] produces an intramolecular [2+2] cycloaddition of the enone and the acetylene to form a cyclobutene [71] which acts as a latent furan. The ketone [70] (200mg) was irradiated for 6 hours in acetonitrile (400ml) under nitrogen. A Pyrex glass filter was used to minimise the incidence of unwanted side reactions. Prior to the irradiation, nitrogen was bubbled through the acetonitrile solution for 15 minutes. The solution was evaporated to dryness and the product purified by column chromatography to give [71]. (480mg; 32%): IR (CCl₄) 3050 (w), 2940 (s), 1710 (s), 900cm⁻¹; NMR (CDCl₃, 250 MHz) 1.4-1.8 (m, 4H), 2.0-2.3 (m, 6H), 2.6 (d, J = 14 Hz, 1H), 3.0 (s, 1H), 3.35 (m, 1H), 5.54 (s, 1H); mass calculated for C₁₁H₁₄O: 162; found 162..

5. 4,5,5a,6,7,8-HEXAHYDRO-3H-NAPTHO[1,8-*bc*]FURAN-3-ONE

[72]



Ozonolysis then introduces the oxygen into the furan to form [72]. Ozone was bubbled into a solution of [71] (500mg) in dry dichloromethane (20ml) at -78°C for about 2 hours until the solution turned blue. After bubbling ozone through for an additional 5 minutes after the colour change, excess ozone was removed by bubbling nitrogen into the solution for 20 minutes until the blue colour had disappeared. Dimethyl sulphide (2ml) was added dropwise and the solution allowed to warm slowly to room temperature with stirring. The solution remained stirring at room temperature for a further 4 hours then the solvent removed by evaporation. A solution of the crude product in benzene (50ml) containing tosic acid (100mg) was then refluxed for 5 hours using a Dean Stark apparatus. The solution was allowed to cool and washed with sodium bicarbonate solution and brine, dried and evaporated to give the crude product [72]. (10%): IR (CCl_4) 3150 (w), 2940 (s), 2850, 1690 (s), 1530, 1190, 1120cm^{-1} ; NMR (CDCl_3 , 250 MHz) 1.18 (m, 1H), 1.48 (m, 1H), 1.9 (m 1H), 2.06-2.24 (m, 3H), 2.44-2.8 (m, 5H), 7.82 (s, 1H); mass calculated for $\text{C}_{11}\text{H}_{12}\text{O}_2$:176; found: 176.

APPENDIX 7

Melting points were determined on a Kofler hot-stage apparatus and are uncorrected.

Infra-red spectra were recorded for potassium bromide discs on a Perkin-Elmer 983 spectrophotometer. The following abbreviations are used: s - strong, w - weak, br - broad.

Proton nuclear magnetic resonance spectra were measured at 200MHz on a Bruker AM200 SY or a Bruker NP 200 SY spectrometer and at 90Hz on a Perkin-Elmer R32 or a Varian EM390 spectrometer. Carbon nuclear magnetic resonance spectra were determined at 50MHz on a Bruker AM200 SY or a Bruker WP200 SY spectrometer in the pulsed FT mode. DEPT editing was used to assist peak assignment. Spectra were recorded in deuteriochloroform with tetramethylsilane as internal standard. The following abbreviations are used: s - singlet, d - doublet, t - triplet, q - quartet, m - multiplet and b - broad.

Mass spectra were obtained using a MS12 or MS 902 spectrometer.

Flash chromatography was carried out on silica gel HF₂₅₄ and TLC on 0.2mm silica gel 60F₂₅₄.

Tetrahydrofuran was freshly distilled from sodium/benzophenone. Organic solutions were evaporated on a rotary evaporator under reduced pressure; solutions in organic solvents were dried over anhydrous magnesium sulfate.

REFERENCES

- 1 Henderson, B. E., Ross, R. K. and Pike, M. C. 'Towards the primary prevention of cancer', *Science*, 1991, **254**, 1131-1138.
- 2 *Monthly Index of Medical Specialities*, July 1993, 266-270.
- 3 *British National Formulary*, 'Drugs used in the treatment of malignant Disease and for immunosuppression', March 1993, **25**, 317-332.
- 4 *Pharmaceutical Journal*, April 24th, 1993, 570.
- 5 *Pharmaceutical Journal*, March 1993, 301.
- 6 Sneader, W. 'The history of drug design', *Chemotherapy*.
- 7 Hickman, J.A. 'Membrane targets in cancer chemotherapy', *Eur. J. Cancer*, 1988, **24**, 8,1385-1389.
- 8 Gescher, A. and Dale, I. L. 'Mini-Review: Protein Kinase C - a novel target for rational anti-cancer drug design?' , *Anti-Cancer Drug Design* , 1989, **4**, 93-105.
- 9 Milligan, G. and Wakelam, M. J. O., *Prog. Growth. Factor. Res.*, 1989, **1**, 161-177.
- 10 Stryer, 'Biochemistry', 985-988.
- 11 Berridge, M. J. 'Inositol triphosphate and diacylglycerol: two interacting second messengers.', *Ann. Rev. Biochem.* , **56**, 159-193.
- 12 Billah, M. M., Anthes, J. C. 'Review article: The regulation and cellular functions of phosphatidylcholine hydrolysis.' *Biochem. J.* , 1990, **269**, 281-291.
- 13 Bell, R. M. 'Protein kinase C activation by diacylglycerol second messengers' , *Cell*, 1986, **45**, 631-632.
- 14 Nishizuka, Y., 'The role of protein kinase C in cell surface signal transduction and tumour promotion' , 1984, *Nature*, **308**, 693-698.
- 15 Blumberg, P. M. 'Protein kinase C as the receptor for the phorbol ester tumor promoters: Sixth Rhoads memorial award lecture.' , *Cancer Research*, 1988, **48**, 1-8.

- 16 Rozengurt, E. 'Signal transduction pathways in mitogenesis', *British Medical Bulletin*, 1989, **45**, 515-528.
- 17 Rozengurt, E. 'Early signals in the mitogenic response', *Science*, 1986, **234**, 161-166.
- 18 Pelech, S. L. and Vance, D. E. 'Signal transduction via phosphatidylcholine cycles.', *Trends Biochem Sci.* 1989, **14**, 28-30.
- 19 Nishizuka, Y. 'The molecular heterogeneity of protein kinase C and its implications for cellular regulation.', *Nature*, 1988, **334**, 661-665.
- 20 Cantley, L., Fleischman, L. and Whitman M. 'The role of lipid-derived second messengers in cell growth and transformation'., *Anti-Cancer Drug Design* , 1987, **2**, 129- 138.
- 21 Jeffrey, A. M., and Liskamp, R. M. 'Computer assisted molecular modelling of tumour promoters: Rationale for the activity of phorbol esters, teleocidin B and aplysiatoxin.', *Proceedings of the National Academy of Sciences USA*, 1986, **83**, 241.
- 22 Wender, P. A., Cribbs, C. M., Koehler, K. F., 'Modelling of the bryostatins to the phorbol ester pharmacophore on protein kinase C'., *Proceedings of the national academy of sciences USA*, 1988, **85**, 7197.
- 23 Wender, P. A., Koehler, K. F., Sharkey, N. A., Dell 'Aquila, Blumberg, P. M. 'Analysis of the phorbol ester pharmacophore on protein kinase C as a guide to the rational design of new classes of analogs.', *Proceedings of the national academy of sciences USA*, 1986, **83**, 4214.
- 24 Cook, S. J. and Wakelam, M. J. O. 'Hydrolysis of phosphatidylcholine by phospholipase D is a common response to mitogens which stimulate inositol lipid hydrolysis in Swiss 3T3 fibroblasts'., *Biochem . et Biophys. Acta*, 1991, **1092**, 265-272.
- 25 Cook, S. J. and Wakelam, M. J. O. *Biochem J*, 1989, **263**, 581-537.
- 0 Bonser, R. W., Thompson, N. T., Randall, R. W., Tateson, J. E., Spacey, G. D. Hodson, H. F. and Garland, L. G. 'Demethoxyviridin and wortmannin

- block phospholipase C and D activation in the human neutrophil', *Br. J. Pharmacol.*, 1991, **103**, 1237-1241.
- 26 Pai, J. K. and Billah, M. M. 'Phospholipase D catalyses phospholipid metabolism in chemotactic peptide-stimulated HL-60 granulocytes', *J. Biol. Chem.*, **263**, 12472-12477.
- 27 Grove, J. F. 'Viridin. Part 1. Isolation and characterisation', *J. Chem. Soc.*, 1965, **696**, 3803-3811.
- 28 Bonser, R. W., Thompson, N. T., Randall, R. W., Garland, L. G. 'Phospholipase D activation is functionally linked to superoxide generation in the human neutrophil.', *Biochem J.*, 1989, **264**, 617-620.
- 29 Baggiolini, M., Dewald, B., Scmyder, J., Ruch, W., Cooper, P. H. and Payne, T. G. 'Inhibition of the phagocytosis-induced respiratory burst by the fungal metabolite wortmannin and some analogues.', 1987, *Exp. Cell Research*, **169**, 408-418.
- 30 Brian, P. W., Curtis, P. J., Hemming, H. G. and McGowan, J. C. 'The production of viridin by pigment-forming strains of *Trichoderma viride*.', 1946, *Ann. Appl. Biol.* **33**, 190-200.
- 31 Curtis, B., Hemming. 'Raulin-Thom medium.' *Transactions of Brit. Myc. Soc.*, 1946, 173.
- 32 Cole, 'Handbook of toxic fungal metabolites', **17**, Viridiol group, 747.
- 33 Plumb, J., CRC Department of medical Oncology, Beatson Institute, Garscube, Glasgow- unpublished work.
- 34 Neidle, S., Rogers, D. and Hursthouse, M. B. 'Crystal and molecular structure of viridin.', 1972, *J.C.S. Perkin II*, 760-766.
- 35 Petcher, T. 'Crystal and absolute configuration of wortmannin and wortmannin *p*-bromobenzoate.', 1972, *J.C.S. Chem. Comm.*, 1061-1072.
- 36 Roll, D. M. 'Halenaquinone, a pentacyclic polyketide from a marine sponge.', 1983, *J.A.C.S.*, **105**, 6177-6178.

- 37 Cole, R. J. 'Desmethoxyviridiol, a new toxin from *Nodulisporum hinnuleum*.' 1975, **14**, 1429-1432.
- 38 Ferreira, M. 'Naturally occurring quinones. Part 27. Sesquiterpenoid quinones and related compounds from *Hibiscus elatus* : Crystal structure of Hibiscone C (Gmelofuran).', 1980, *J.C.S.*, 249-256.
- 39 Blight, M. M. and Grove, J. F. 'Viridin, Part 8. Structures of the analogues virone and wortmannolone.', 1986, *J.C.S.*, 1317-1322.
- 40 Moffat, J. S. 'Viridiol, a steroid-like product from *Trichoderma viride*', 1969, *J.C.S. C.C.* , 839.
- 41 Jones, R. W. and Hancock, J. G. 'Conversion of viridin to viridiol by viridin-producing fungi', 1987, *Can. J. Microbiol*, **33**, 963-966.
- 42 Aldridge, D. C., Geddes, A. J., Sheldrick, B., and Turner, W. B., 1975, *J. Chem. Soc. Perkin Trans. I*, 943.
- 43 Olson and Christofferson 'Computer-Assisted Drug Design.', 1979, *A.C.S. Symposium Series* , **112**, 21.
- 44 Sneader, W. 'A History of drug design' , Chapter 1.
- 45 M.M.S. - Dempsey, S. 'The U.C.S.D. Molecular Modelling System Reference Manual', Chemistry Computer Facility, Dept., of Chemistry, B-014, University of California, San Diego, La Jolla, CA 92093.
- 46 HYDRA- Hubbard, R., 'Harvard York Drawing Package', Chemistry Dept., York University.
- 47 ORTEP Johnson, C. K., 'ORTEP II. A fortran thermal-ellipsoid plot program for crystal structure illustrations. Report ORNL-3794 (2nd revision) and supplementary instructions.', Oak Ridge National Laboratory, Tennessee, U.S.A.
- 48 PYTHON - available from Oxford Molecular Systems.
- 49 SYBYL - Molecular Modelling Software, Tripos Associates, Inc. Subsidiary of Evans and Sutherland. Version 6.0.

- 50 Gilmore, C. J., Mallinson, P. R., and Muir, K. W. 'The GX package - A collection of compatible programs for crystal structure determination.', 1985, Glasgow University .
- 51 Osman, R., Weinstein, H., and Green, J. P. 'Parameters and methods in quantitative structure-activity relationships.' , 1979, *A.C.S. Review* , 21-77.
- 52 Burkett, U., Allinger, N. L., 'Molecular mechanics', 1982, *American Chemical Society, Washington D.C.*.
- 53 Diana, G.D., Kowalczyck, P., Treasurywala, A.D., Oglesby, R.C., Pevear, D.C., Dutko, F.J. 'CoMFA analysis of the interactions of antipicornavirus compounds in the binding pocket of human rhinovirus-14.', 1992, *J. Med. Chem.*, **35**, 1002-1008.
- 54 Simon, Z., Badileuscu, I. and Racovitan, T.J. *Theor. Biol.*, 1977, **66**, 485.
- 55 Simon, Z, Badileuscu, I., Plauchithiu, M. G., Holban, S., Glatt, S., Kerek, F. 1980, *Eur. J. Med. Chem.I*, **15**, 521.
- 56 Hopfinger, A. J., 1985, *J. Med. Chem.* , **28**, 133.
- 57 Crippen, G. M. and Ghose, A. K. 'General distance-geometry three-dimensional receptor model for diverse dihydrofolate reductase inhibitors', 1984,*J. Med. Chem.*, **27**, 901-14.
- 58 Goodford, P. J. 'Identifying targets for bioreductive agents: using GRID to predict selective binding regions of proteins', 1989, *J. Mol. Graphics*, **7**, 103.
- 59 Hansch, C., Klein, T. E., Langridge. 'Molecular graphics and QSAR in the study of enzyme-ligand interactions. On the definition of bioreceptors', 1986, *Acc. Chem. Res.* , **19**, 392-400.
- 60 Sybyl Tutorial Manual. Version 6.0 Tripos Inc.
- 61 Hinze, J., Jaffe, H.H., 1962, *J.A.C.S.*, **84**, 540-546.
- 62 SYBYL Theory manual. Version 6.0. Tripos Inc.
- 63 Nicklaus, M. C., Milne, G. W. A., Burke Jr., T. R. 'QSAR of conformationally flexible molecules: Comparative molecular field analysis of

- protein tyrosine kinase inhibitors.', 1992, *J. Comp-Aided. Mol. Design*, **6**, 487-504.
- 64 Cramer III, R. D., Patterson, D. E. and Bunce, J. D. 'Comparative Molecular Field Analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins', 1988, *J.A.C.S.*, **110**, 5959-5967.
- 65 Glusker, J. P. and Trueblood, K. N. 'Crystal structure analysis - a primer'.
- 66 Cole, 'Handbook of toxic fungal metabolites', **17**, Viridiol group, 747
- 67 Gilmore, C. J. 'MITHRIL. A computer program for the automatic solution of crystal structures from x-ray data', 1983, University of Glasgow.
- 68 Gilmore, C. J., Mallinson, P. R., Muir, K. W. 'The GX package - A collection of compatible programs for crystal structure determination', 1985, University of Glasgow.
- 69 Joshi, K. C., Singh, P., Pardusani, R. T., Pelter, A., Ward, R. S., Reinhardt, R., 1978, *Tetrahedron. Lett.*, 4917.
- 70 Thomson, R. H., Ali, S, King, T. J., Ferreira, M. A., 1980, *J. Chem. Soc. Perkin Trans.*, **1**, 249.
- 71 Koft, E. R., Smith III, A. B. 'Intramolecular [2+2] Photocyclisations. 2. Total synthesis of (\pm)-Hibiscone C (Gmelofuran).', 1984, *J. Am. Soc.*, **106**, 2115-2121.
- 72 Thomson, R. H., Ali, S., King, T. J., Ferreira, M. A. *J. Chem. Soc., Perkin Trans.* 1980, **19**, 1969.
- 73 Macaulay, G. S., Chemistry Department, University of Glasgow, unpublished work.
- 74 Broka, C. A., Ruhland, B. 'Synthetic studies on wortmannin and 11-desacetoxywortmannin', 1992, *J. Org. Chem.*, **57**, 4888-4894.
- 75 Baggiolini, M., Dewald, B., Schnyder, J., Ruch, W., Cooper, P., H., Payne, *Exp. Cell Res.*, 1987, **169**, 408.
- 76 Bonser, R. W., Thompson, N. T., Randall, R. W., Garland, L. W., *Biochem J.*, 1989, **264**, 617.

- 77 Haefliger, W. and Hauser, D. ,*Helv. Chim. Acta.*, 1973, **56**, 2901.
- 78 Grove, J. F. 'Viridin. Part VI. Evidence for a steroidal pathway in the biogenesis of viridin from mevalonic acid', 1969, *J. Chem. Soc.*, 549-551.
- 79 Mann, J. 'Secondary Metabolism', Oxford Chemistry series, Chapter 3.
- 80 Simpson, T.J., Lunnon, M.W. and MacMillan, J. 'Fungal products. Part 21. Biosynthesis of the fungal metabolite wortmannin, from [1,2-¹³C₂]-Acetate', 1979, *J. Chem. Soc.*, 931-934.
- 81 Aldridge, D. C., Geddes, A. J., Sheldrick, B., and Turner, W. B., 1975, *J. Chem. Soc. Perkin Trans. I*, 943.
- 82 Kobayashi, M. and Kanfer, J. N. 'Solubilisation and purification of rat-tissue phospholipase D', 1991, *Methods in Enzymology*, **1**, 63-66.
- 83 Gannon, W. F. and House, H. O. '3-ethoxy-2-cyclohexenone (dihydroresorcinol monoethyl ether)', *Org. Synth. Coll.*, **V**, 539-541.
- 84 3-ethoxy-6-methyl-2-cyclohexenone', 1985, *J.A.C.S.*, **106**, **7**, 2118.