DOCTOR OF PHILOSOPHY

Theoretical and computational aspects of organic chemistry = agweddau damcaniaethol a chyfrifiadurol yng nghemeg organeg

Davies, Richard Andrew

Award date:
1999

Awarding institution:
University of Wales, Bangor

THEORETICAL AND COMPUTATIONAL ASPECTS OF ORGANIC CHEMISTRY

AGWEDDAU DAMCANIAETHOL A CHYFRIFIADUROL YNG NGHEMEG ORGANEG

A thesis submitted to the University of Wales in candidature for the degree

Philosophiae Doctor

Thesis wedi ei gyflwyno i Brifysgol Cymru mewn ymgeisiaeth ar gyfer y radd

Doethur mewn Athroniaeth

by / gan

Richard Andrew Davies BSc CChem MRSC

December / Rhagfyr 1998

ABSTRACT

The first application of computational chemistry presented in this thesis is the prediction of electronic wavelength maxima, with particular emphasis on the cytological dyes used in the Papanicolaou cervical smear test. Initial studies of the staining mechanism of the alchemic Papanicolaou ingredients are also presented.

Wavelength maxima were initially calculated using semi-empirical HOMO-LUMO gaps for a training set of several hundred unsaturated organic compounds. Semi-empirical electronic wavelength maxima were substantially underestimated, with increasing deviations at longer wavelengths. The overestimation of the singlet excitation energy can be explained by the formation of a hole in the HOMO upon electronic excitation, causing the singlet excitation energy to be smaller than the HOMO-LUMO gap by an amount which approximates to $J_{H,L} - 2K_{H,L}$. Additionally, correlations were found between the calculated and experimental wavelength maxima. These correlations were interpreted in terms of Dähne's triad theory, allowing the electronic properties of any unsaturated organic compound to be characterised by varying degrees of polymethinic, aromatic and polyenic character. Wavelength maxima could be predicted to within 25 nm upon application of specific, character dependent correction factors.

More realistic excitation energies allowing the prediction of J and K, could be obtained by from post Hartree-Fock methods such as CIS configuration interaction approaches, and by explicit evaluation of $\Delta\Delta H_f$ (excited state *vs* ground state) values. Triad theory could again be used to predict wavelength maxima. Qualitative trends in absorption intensities could only be obtained from oscillator strength for certain unsaturated compounds.

First-principles density functional theory calculations employing the local density approximation were also performed for a training set of several hundred compounds. However, in complete contrast to the above semi-empirical results, singlet excitation energies obtained from DFT-LDA HOMO-LUMO gaps were underestimated, leading to larger values of wavelength maxima. Polar molecules, such as benzenesulphonic acid, greatly deviate from the correlations, especially for poor basis set quality, due to the failure of the LDA approximation.

The second application concerns the calculation of semi-empirical transition structures and activation energies for various pericyclic reactions. The studies focus on the diimide reductions, ene-dimerisations and Diels-Alder reactions of cyclopropenes. Due to a lack of experimental data, quantitative agreement with experiment cannot yet be determined for most compounds.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my family, especially Mam and Dad, who supported me through the emotional ups and down of research and thesis writing, enabling me to fulfil a lifelong dream.

I would also like to thank my supervisor, Professor Mark Baird, for his guidance, and input during my time at Bangor, and Florian Huber for translating the original Dähne paper on triad theory. The assistance of the technical staff and secretaries is also greatly appreciated.

I would like to thank Professor Nick Quirke for making me feel part of the team both at *Molecular Simulations Inc.* (formerly *Biosym Technologies SARL*), Orsay, France, and at the *Centre for Computational Chemistry* in Bangor. I am indebted to Dr. Erich Wimmer, who always had a free moment to talk about the results I had obtained.

I would like to thank Dr. Brian Jones, my Chemistry teacher, and *ex alumni* of Bangor, who showed me the wonders of chemistry through many interesting and explosive reactions.

The Royal Society of Chemistry William Briggs Scholarship, donated by the late Lady Lillian, widow of the eminent chemist Sir Jocelyn Field Thorpe, and daughter of William Briggs J.P, funded the research.

This thesis is dedicated to the memory of my late great aunt, Helena Winifred Hughes, the only other chemist in our family, who studied Chemistry at Bangor during the 1920s, and who passed away on the 29th April 1998, a week before my *viva voce*.

DON'T QUIT

When things go wrong, as they sometimes will,
When the road you're trudging seems all uphill,
When the funds are low, and the debts are high,
And you want to smile, but you have to sigh,
When care is pressing you down a bit,
Rest if you must, but don't you quit.

Life is queer with its twists and turns,
As everyone of us sometimes learns,
And many a failure turns about,
When he might have won had he stuck it out;
Don't give up though the pace seems slow,
You may succeed with another blow.

Success is failure turned inside out,
The silver tint of the clouds of doubt,
And you never can tell how close you are,
It may be near when it seems so far,
So stick to the fight when you're hardest hit,
It's when things seem worse, that you must not quit.

Anon

THE MOLECULAR CHALLENGE

Sir Ethylene, to scientists fair prey,
(Who dig and delve and peek and push and pry,
And prove their findings with equations sly)
Smoothed out his ruffled orbitals, to say:
" I stand in symmetry. Mine is a way
Of mystery and magic. Ancient, I
Am also deemed immortal. Should I die,
Pi would be in the sky, and Judgement Day
Would be upon us. For all things must fail,
That hold our universe together, when
Bonds such as bind me fail, and fall asunder.
Hence, stand I firm against the endless hail
Of scientific blows. I yield not." Men
And their computers stand and stare and wonder.

W.G. Lowe

GLOSSARY OF CYTOLOGICAL / HISTOLOGICAL TERMS

- **Acidophile**
  Easily stained with acid dyes.
- **Actinomycetes**
  Fungi found in the presence of IUDs - not implicated in cervical cancer.
- **Adenocarcinomas**
  A malignant tumour originating in glandular tissue.
- **Anulate squames**
  Large flat cells without a nucleus sometimes found when the cervix is prolapsed.
- **Atrophic smear**
  Thin, shrivelled cells common in post-menopausal women, reflecting oestrogen deficiency.
- **Bactericidal**
  A drug whose action destroys bacteria.
- **Bacteriostatic**
  A drug whose action arrests or inhibits bacterial growth and reproduction.
- **Basophil(e)**
  A cell, especially a white blood cell, having granules exhibiting an affinity for basic dyes.
- **Bathochromic shift (red shift)**
  A shift to longer wavelength.
- **Canada balsam**
  A viscous, yellowish, transparent resin obtained from the balsam fir and used as a mounting cement for microscope specimens.
- **Celloidin**
  A pure pyroxylin in which specimens being sectioned for examination are embedded.
- **Cervical intraepithelial neoplasia**
  The whole spectrum of abnormalities which occur on or in the skin of the cervix.
- **Chromatin**
  A complex of nucleic acids and proteins in the nucleus of a cell, characterised by intense staining with basic dyes.
- **Chromogen**
  A strongly pigmented or pigment-generating organ, organelle, or microorganism.
  A substance capable of chemical conversion into a pigment of dye.
- **Clue cells**
  Non-malignant change in the cells due to an infection by *Gardnerella*.
- **Colligator**
  A group that enables a dye to unite chemically with tissue elements or with certain other dyes and compounds.
- **Columnar endocervical epithelium**
  Membranous tissue of the endocervical canal, the lining of which is a single cell layer thick responsible for production of mucus.
- **Complementary secondary stain**
  A dye staining tissue elements that have not already been stained by the primary stain.
- **Compound dye**
  A dye produced by the reaction of two or more dyes.
- **Condylomata**
  Wart-like growths near the anus or external genitalia.
- **Connective tissue**
  Tissue arising chiefly from the embryonic mesoderm, consisting typically of a jelly-like matrix in which are embedded collagen and elastic fibres, fat cells, fibroblasts, mast cells *etc.* It supports and separates organs and occurs in tendons, ligaments, cartilage and bone.

- **Counterstain**
  A dye used to treat specimens that have already been treated with another dye.
- **Cytology**
  The branch of biology dealing with the study of cellular formation, structure and function.
- **Cytoplasm**
  The protoplasm outside a cell nucleus.
- **Denuded**
  Flesh with no epithelial skin covering.
- **Dichromatic**
  A compound dye or a mixture of several dyes capable of exhibiting only two colours.
- **Differentiation**
  To cause differences to develop in stained tissue by altering or modifying the stain by the addition of another dye / chemical.
- **Dyskaryosis**
  Alteration within the normal arrangement of the cell nucleus.
- **Dysplasia**
  Abnormal development of tissues, organs, or cells.
- **Ectropion / Erosion**
  The break down of the columnar epithelium.
- **Endocervical cells**
  Cells from the endocervical canal, the passage from the cervix to the uterus.
- **Endometrial cells**
  Cells from the endometrium (lining of the uterus) lost during menstruation.
- **Eosinophil(e)**
  A type of leucocyte with a lobed nucleus that stains with an eosin dye.
- **Epithelium**
  Membranous tissue, usually in a single layer, composed of closely arranged cells separated by very little intercellular substance and forming the covering of most internal surfaces and organs and the outer surface of the animal body.
- **Erythrocyte**
  A non-nucleated, disc-shaped blood cell containing haemoglobin - the "red blood cell".
- **Exfoliative**
  Skin which is removed / peeled in the form of flakes or scales.
- **Fast**
  Resistant.
- **Fix**
  To kill, harden and preserve a specimen for microscopic study using a fixative such as alcohol.
- **Fluorochrome**
  A molecular group or chromophore capable of fluorescence.
- **Fundus**
  The inner basal surface farthest away from the opening of the uterus.
- **Glandular endocervical cells**
  Identified by staining for the presence of mucin. Rarely reported but may indicate an adenocarcinoma of the cervix.
- **Histiocyte**
  A macrophage found in connective tissue.
- **Histology**
  The anatomical study of the microsopic structure of animal and plant tissues.
- **Hyperchromatic**
  Exhibiting a hyperchromic effect - intensely stained.
- **Hyperchromic effect**
  One that leads to increased intensity of absorption.

- **Hypochromic effect**
  One that leads to decreased intensity of absorption.
- **Hypsochromic shift (blue shift)**
  A shift to shorter wavelength.
- ***in tela***
  Designating biological processes occurring within tissues.
- ***in vitro***
  Designating biological processes made to occur in an artificial environment outside the living organism.
- ***in vivo***
  Designating biological processes occurring within the living organism.
- **Intracytoplasmic vacuolation**
  The process of forming any small cavity in the cytoplasm of a cell.
- **Koilocytes**
  Cells whose nucleus appears to be surrounded by a halo of cytoplasm.
- **Lake**
  A pigment consisting of organic colouring matter with an inorganic base or carrier.
- **Leucocyte**
  Any of the white or colourless nucleated cells occuring in blood - the "white blood cell".
- **Leucocytosis**
  A large increase in the number of leucocytes in the blood, generally in response to infection.
- **Macrophage**
  A large phagocytic cell present in connective tissue, bone marrow, lymph nodes etc.
- **Malignant crowded columnar epithelium**
  Cells from the endocervix are highly suspicious and suggest adenocarcinoma of the cervix.
- **Metachromatic**
  Changing of cells and tissues to a different colour from that of the dye used for staining.
  Designating a dye that is able to stain celols or tissues a different colour from its own.
- **Metaplasia**
  The change of cells from a normal to an abnormal state.
- **Metastasis**
  Transmission of a disease such as tuberculosis or cancer, from an original site to one or more sites elesewhere in the body.
- **Mordant**
  A reagent, such as alumina or tannic acid, used to fix colouring matter in textiles, leather or other materials.
- **Mount**
  To place a specimen on a microscope slide in preparation for microscopical examination.
- **Mucin**
  Any of a groups of glycoproteins produced by the mucous membranes.
- **Mucus**
  The viscous suspension of mucin, water, cells and inorganic salts secreted as a protective, lubricant coating by glands in the mucous membranes.
- **Neutrophil(e)**
  A phagocytic leucocyte of a type having a lobed nucleus and granular cytoplasm that stains with neutral dyes.
- **Nucleoli**
  Small, usually round bodies composed of protein and RNA in the nucleus of a cell.
- **Nucleus**
  A complex, usually spherical, protoplasmic body within a living cell that contains the cell's hereditary material and that controls its metabolism, growth and reproduction.

- **Pathognomonic**
  Distinctive or characteristic of a particular disease.
- **Phagocyte**
  A cell such as a leucocyte that engulfs and digests cells, microorganisms, or other foreign bodies in the bloodstream and tissues.
- **Plasmodia**
  Any protozoan of the genus *Plasmodium*, which includes the parasite that cause malaria.
- **Polychromatic**
  A compound dye or a mixture of several dyes capable of exhibiting many colours.
- **Pyelonephritis**
  Inflammation of both the kidney and the pelvis of the uterer.
- **Pyknotic nuclei**
  Short / powerful / endomorphic / thick / dense.
- **Serological**
  Pertaining to serum.
- **Squamocolumnar junction**
  The point at which the two different types of skin lining the vagina and cervix meet.
- **Squamous epithelium**
  Epithelium consisting of flat, scalelike cells.
- **Squamous intraepithelial lesions**
  Lesions which occur in the squamous epithelium.
- **Squamous metaplasia**
  The growth of squamous epithelium beneath the columnar epithelium, gradually replacing it.
- **Staining**
  To colour (specimens for the microscope) with a dye in order to heighten contrast between different structures, as of tissue.
- **Stratified epithelium**
  Epithelium layered into many strata.
- **Supplementary secondary stain**
  A dye staining only tissue elements that have already been stained by the primary stain.
- **Systemic**
  Of, pertaining to, or affecting the entire body.
- **Topical**
  Applied or pertaining to a local part of the body.
- **Transformation zone**
  The area on the post-adolsecent cervix covered by migrated columar epithelium from the cervical canal.
- **Trichomonad**
  Any of the various flagellate protozoans of the genus *Trichomonas*, occuring in the digestive and urogenital tracts of vertebrates.
- **Trichomoniasis**
  A vaginal infection caused by the protozoan, *Trichomonas vaginalis*, resulting in inflammation and discomfort.
- **Trypanosome**
  Any parasitic protozoans of the genus *Trypanosoma*, transmitted to the vertebrate bloodstream by certain insects, often causing diseases such as sleeping sickness.
- **Trypanosomiasis**
  Any disease caused by a trypanosome.
- **Uraemia**
  The presence of excess urea and other waster products in the blood.
- **Vesicular nuclei**
  Composed of or containing vesicles - a serum filled blister vacuole, cell or cavity.

CONTENTS

CHAPTER 1
COMPUTATIONAL QUANTUM CHEMISTRY

" ... in real life mistakes are likely to be irrevocable. Computer simulation, however, makes it economically practical to make mistakes on purpose. If you are astute, therefore, you can learn much more than they cost. Furthermore, it you are at all discreet, no one but you need ever know you made a mistake. "[1]

The ability of quantum mechanics to predict the total energy of a system of electrons and nuclei enables one to reap a tremendous benefit from a QM calculation. If total energies can be calculated, any physical property that can be related to a total energy or to a difference between total energies can be determined computationally.

This chapter gives a background to the exciting and developing field of computational chemistry. The fundamentals behind molecular mechanics and force fields will be discussed, followed by quantum mechanical (QM) methods, charting the development of *ab initio* basis sets and semi-empirical methodologies. Configuration interaction, correlation energy, calculation of orbital energies, and differences between restricted (RHF) and unrestricted Hartree-Fock (UHF) methods will also be addressed.

1.1 MOLECULAR MECHANICS

In the **molecular mechanics** model, a molecule is represented as an ensemble of spheres (possibly deformable) joined by springs, using simple potential functions and Newton's laws of classical physics to describe the atomic motions, allowing chemical systems of the order of thousands of atoms to be used. However, the method neglects explicit representation of electrons, and is restricted to the discussion of molecular ground states, disallowing the investigation of reactions. The results obtained are only as good as the potential functions and parameters used.

1.1.1 Energy calculations

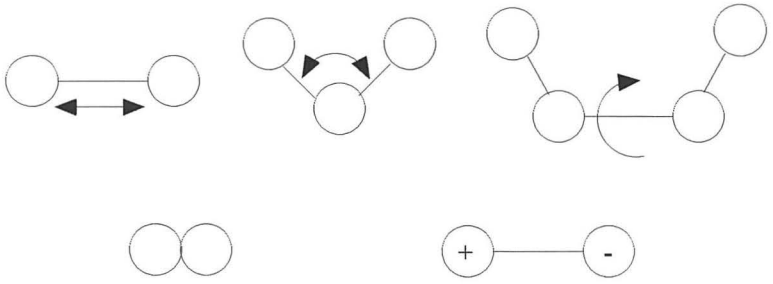
A total-energy function must be defined which is consistent with the above concept, yet allows accurate calculation of molecular properties. The energy of a molecule is calculated as a sum of the steric and non-bonded interactions present. Each bond

length, angle and dihedral is treated individually while non-bonded interactions represent the influence of non-covalent forces.

$$E_{tot} = E_l + E_\theta + E_\omega + E_{nb} \quad \text{... Equation 1.1}$$

where E_l, E_θ, E_ω and E_{nb} are the total bond, angle, dihedral and non-bonded energies respectively, as shown in *Figure 1.1*. Non-bonded energy usually comprises of van der Waals and electrostatic components.

Figure 1.1 **Pictorial representation of the terms included in a molecular mechanics force field - bond stretch, bond and dihedral angles, van der Waals and electrostatic interactions.**



a) *Bond stretch*

The typical vibrational behaviour of a bond is near harmonic close to its equilibrium value but shows dissociation at longer bond lengths. The most accurate description is the Morse function:

$$E_l = \sum D_e [1 - \exp\{-\alpha(l - l_0)\}]^2 \quad \text{... Equation 1.2}$$

where l_0 is the equilibrium bond length, D_e is the dissociation energy, and α is a force constant. Most force fields adopt a simple harmonic function:

$$E_l = \sum k_1(l - l_0)^2 \quad \text{... Equation 1.3}$$

k_1 being the stretching force constant describing the deformation. At extended bond lengths it is too steep and provides no representation of dissociation at very large

deformations. Other variations have been used to accommodate more accurate long distance behaviour, such as an additional cubic term, which suffers from inversion at long distances. A quartic term reverses the inversion.

b) Bond angles

Bond angles are treated in the same way as bond lengths, by a harmonic function:

$$E_\theta = \sum k_\theta(\theta - \theta_0)^2 \qquad \text{... Equation 1.4}$$

where k_θ is a force constant and θ_0 the equilibrium value for the bond angle. Higher order terms must again be added. In very strained ring systems, it is usually not possible to use the constants derived for unstrained and acyclic molecules, so separate three- and four-membered ring constants have been developed.

c) Dihedral angles

It was originally thought that this term could be omitted, as *gauche-trans* energy differences would then result from non-bonded interactions. This proved to be an impossible task. Dihedral angle terms in the form of a Fourier series were included:

$$E_\omega = \sum V_n\{1 + s\cos(n\omega)\} \qquad \text{... Equation 1.5}$$

where V_n is the rotational barrier, n is the periodicity of rotation ($n = 3$ for ethane, and $n = 2$ for ethene), $s = 1$ for staggered minima and s = -1 for eclipsed minima.

d) Non-bonded interactions

Non-bonded interactions are distance-dependent and are calculated as the sum over all atoms with a 1,4- or greater separation. It is usual to consider these interactions as having two components: van der Waals and electrostatic. The former can be considered as both a size parameter and representative of electron correlation (resulting from instantaneous dipole interactions), while the latter provides a quantitative measurement of polarity.

Many different functional forms have been used for van der Waals interactions, the most common being the Lennard-Jones (6-12) potential:

$$E_{vdw} = \sum \varepsilon \left[(r_m / r)^{12} - 2 (r_m / r)^6 \right] \quad \text{... Equation 1.6}$$

where ε is the well depth and r_m is the minimum energy interaction distance. Short-range repulsions are accounted for by the r^{-12} term whereas longer-range London dispersion-attraction forces are mediated by the r^{-6} component. Other forms have been proposed for the van der Waals interaction, as the r^{-12} term can be too steep at just less than optimal distances, the short contacts being important when investigating sterically crowded structures. In the Buckingham potential:

$$E_{vdw} = Ae^{-Br} - Cr^{-6} \quad \text{... Equation 1.7}$$

an exponential replaces the repulsive r^{-12} term. In most cases this function behaves similarly to the Lennard-Jones equation, but at very short interatomic distances the function inverts and goes to $-\infty$.

The second component of the non-bonded potential is the electrostatic term. This is usually calculated using partial charges, q, on the atom centres with the energy calculated using Coulomb's law:

$$E_{el} = \sum q_i q_j / D r_{ij} \quad \text{... Equation 1.8}$$

with the dielectric constant D taking a value appropriate to a given solvent or made proportional to the distance r_{ij} between the charges.

e) Other terms

For systems where hydrogen bonding is vital for stability, *e.g.* biological molecules, an additional explicit hydrogen bond energy function is often included:

$$E_{hb} = \sum (C_{ij} / r_{ij}^{12}) - (E_{ij} / r_{ij}^{10}) \quad \text{... Equation 1.9}$$

The four atoms of isolated unsaturated planar centres should be kept in the plane. The branch atom, *e.g.* the oxygen of a carbonyl group, can be distorted. If the distortion is measured as the height of the central atom above the plane formed by the other three atoms, then a simple restraining force can be used:

$$E_{opl} = \sum k_\chi \chi^2 \qquad \text{... Equation 1.10}$$

where k_χ is the force constant and χ is the height above the plane.

To fit vibrational frequencies, coupling must be included in the force field. For example from changing from *anti* to *syn* butane there is a change in the C–C bond length and an opening of the C–C–C bond angles, as shown in *Figure 1.2*. Commonly used cross terms include stretch-bend, bend-bend and torsion-bend terms.

Figure 1.2 Molecular geometries for *syn* and *anti* butane structures.

syn	*anti*
$l_1 = 1.56$ Å	$l_2 = 1.53$ Å
$\theta_1 = 116.4^\circ$	$\theta_2 = 112.9^\circ$

1.1.2 Energy minimisation

When a model is generated from standard molecular fragments or from a Z-matrix of internal co-ordinates, the energy obtained from the molecular mechanics calculation is too high. To obtain more reliable geometries and energies, the energy of the system must be minimised, either by varying the internal co-ordinates to find their optimum value, or more commonly, working in Cartesian space and optimising the atomic positions subject to restraining forces generated by the force field.

As most minimisation methods require first and second derivatives of the energy, the latter method is more convenient and the potential functions are easily differentiated. For a molecule of N atoms there are $3N - 6$ degrees of freedom to be minimised.

Many local energy minima may occur on the potential surface, which are higher in energy than the global minimum. Many minimisation programs are based on the Newton-Raphson method, requiring first and second derivative information about the energy surface. Potential functions must therefore be continuous and differentiable.

For a one dimensional case, if a minimum on a curve is x^*, and the starting point is x, then using a Taylor series, it can be shown that:

$$x^* = x - f'(x) / f''(x) \qquad \text{... Equation 1.11}$$

For a molecule, each atom has three degrees of freedom - in the x, y and z directions. The term $f'(x)$ is replaced by a $3N \times 1$ matrix, **F**, containing $\delta V/\delta x_i$ terms, the derivatives of the potential energy, V, with respect to a change in the co-ordinate i. A corresponding second derivative matrix the Hessian, **H**, is constructed using all the cross-derivative terms ($\delta^2 V/\delta x_i \delta x_j$) involving each co-ordinate. The second term in *Equation 1.11* can be replaced by $\mathbf{H}^{-1}\mathbf{F}$ where $\mathbf{H}^{-1}$ is the inverse of the Hessian matrix.

By truncating the Taylor series, the assumption is made that the minimum is exactly quartic in behaviour. For a complex surface this will not hold true far from the minimum. The calculation must be carried out in a stepwise, iterative fashion, rather than reaching the minimum first time. For small molecules this approach is efficient and converges after relatively few steps. As the number of matrix elements in the Hessian rapidly increases as the number of atoms increases, more practical algorithms must be applied, usually through an approximation of the Hessian, including neglecting off-diagonal interactions - the **block diagonal Newton-Raphson** method.

A much more severe approximation to the Hessian is to consider it as a constant, producing the **steepest descent method**, which is driven purely by force gradients along the potential surface. The steepest descent method is very efficient when large forces are present, far from a minimum, and is ideal for tidying model geometries prior to further refinement. The new set of co-ordinates after step k is given by:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \lambda_k \mathbf{g}_k / |\mathbf{g}_k| \qquad \text{... Equation 1.12}$$

where λ_k is the step size, and $\mathbf{g}_k$ is the gradient vector.

In **pattern search** methods, the previous step is used to accelerate movement if the gradient is in the same direction as before. If the gradient changes, the pattern is abandoned and a new one set up. A more elegant improvement is that of **conjugate gradients** which, like pattern search, uses information from previous steps to modify the move in the next step, but does not abandon the history if the direction changes. In the first step, if the gradient vector is $\mathbf{g}_1$, the move is given by:

$$\mathbf{s}_1 = -\mathbf{g}_1 \qquad \text{... Equation 1.13}$$

The new direction from this step takes into account the previous gradient and follows the search direction:

$$\mathbf{s}_k = -\mathbf{g}_k + \gamma_k \mathbf{s}_{k-1} \qquad \text{... Equation 1.14}$$

where $\mathbf{s}_{k-1}$ is the search direction from the previous step, and γ_k is a scaling factor.

The scaling factor, γ_k, is given by the Fletcher-Reeves algorithm as:

$$\gamma_k = \mathbf{g}_k \bullet \mathbf{g}_k \,/\, \mathbf{g}_{k-1} \bullet \mathbf{g}_{k-1} \qquad \text{... Equation 1.15}$$

and by the Polak-Ribiere algorithm as:

$$\gamma_k = (\mathbf{g}_k - \mathbf{g}_{k-1}) \bullet \mathbf{g}_k \,/\, \mathbf{g}_{k-1} \bullet \mathbf{g}_{k-1} \qquad \text{... Equation 1.16}$$

The second term in *Equation 1.14* releases the constraint that the $k-1^{th}$ gradient should be orthogonal to the k^{th} gradient. As better search directions are usually obtained, conjugate gradients methods are unlikely to maximise rather than to minimise. This is a shortcoming of the Newton-Raphson family of methods, which use only derivative information to search for stationary points (maxima and minima).

A significant drawback to the "pure" second derivative methods such as Newton-Raphson is that the calculation of the inverse Hessian matrix is a time consuming operation. **Variable metric** (quasi-Newton) methods gradually build up the inverse Hessian matrix in successive iterations. At each iteration k, the new positions $\mathbf{x}_{k+1}$ are obtained from the current positions $\mathbf{x}_k$, the gradient $\mathbf{g}_k$, and the current approximation to the inverse Hessian matrix $\mathbf{H}_k$:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{H}_k \mathbf{g}_k \qquad \text{... Equation 1.17}$$

A line search is performed along the vector $\mathbf{x}_{k+1} - \mathbf{x}_k$. An effective compromise to the additional energy evaluations required by line searching, is to fit a function to the energy and gradient at the current point $\mathbf{x}_k$ and at the point $\mathbf{x}_{k+1}$, and determine the minimum in the fitted function. Having moved to the new position $\mathbf{x}_{k+1}$, $\mathbf{H}$ is updated from its value at the previous step according to a formula depending upon the specific method being used. The Broyden-Fletcher-Goldfarb-Shanno (BFGS)[2] method is routinely encountered, while mode following using the eigenvector following (EF) routine will used to calculate transition structures in *Chapter 10*. The performance of quasi-Newton algorithms can be improved by using a better estimate of the inverse Hessian (*e.g.* from molecular mechanics calculations) than the unit matrix, $\mathbf{I}$, which does not identify bonding, nor coupling between the various degrees of freedom.

1.1.3 Force field parameterisation

Examples of different force fields are MM2,[3] MM+,[4] AMBER (Assisted Model Building and Energy Refinement),[5] CHARMM (Chemistry at HARvard Macromolecular Mechanics),[6] BIO+,[7] OPLS (Optimised Potentials for Liquid Simulations),[8] CVFF (Consistent Valence Force Field),[9] and MM3.[10] The quality of a force field depends on the potential parameterisation, accomplished by:

1) Least-squares fitting to determine the set of parameters that gives the optimal fit to the data.[11]
2) Parameterisation by trial and error procedures in which the parameters are gradually refined to give better and better fits to the data, such as that used by Pranata and Jorgensen for the potent immunosuppressant FK506.[12]

1.2 QUANTUM MECHANICAL METHODS

1.2.1 Basics

The simplest form of the Schrödinger equation[13] is:

$$H\psi = E\psi \quad \text{... Equation 1.18}$$

where H, the Hamiltonian, operates on the wavefunction ψ, and E is the energy of the system.

The Schrödinger equation is a set of differential equations, with a function ψ_n corresponding to each allowed energy E_n. Spherical plane polar co-ordinates (r, θ and ϕ) are preferred over Cartesian x, y and z co-ordinates. Once the wave function is known for a system, any physical observable may be determined by:

$$\text{observable} = \int \psi^*(\text{operator})\psi \, d\tau \,/\, \int \psi^*\psi \, d\tau \quad \text{... Equation 1.19}$$

where ψ^* is the complex conjugate of ψ, and integration occurs over all space.

1.2.2 Spin -orbitals and antisymmetry

The electronic structure of helium, is written as $1s^{\alpha}1s^{\beta}$ where α and β represent the two opposite direction of electron spin. Writing $\psi He = \chi_{1s}{}^{\alpha}(1)\chi_{1s}{}^{\beta}(2)$ does not satisfy the Pauli principle,[14] which states that the wavefunction for the system must change sign if any pair of electrons are interchanged, as electrons are identical fermion particles. To overcome this the wavefunction for a helium atom is written as an **antisymmetrised** product of orbitals, in the form of a Slater determinant:[15]

$$\psi = 1/\sqrt{2}\,[1s^{\alpha}(1)1s^{\beta}(2) - 1s^{\alpha}(2)1s^{\beta}(1)] \quad \text{... Equation 1.20}$$

For beryllium, there are possible interchanges between four electrons giving 24 products:

$$\psi = N\,[1s^{\alpha}(1)1s^{\beta}(2)2s^{\alpha}(3)2s^{\beta}(4) - 1s^{\beta}(1)1s^{\alpha}(2)2s^{\alpha}(3)2s^{\beta}(4) + 2s^{\alpha}(1)1s^{\beta}(2)1s^{\beta}(3)2s^{\alpha}(4) - 2s^{\beta}(1)1s^{\alpha}(2)1s^{\beta}(3)2s^{\alpha}(4) + ...] \quad \text{... Equation 1.21}$$

Hartree produced some extremely accurate atomic functions, which were later fitted to analytic forms called Slater atomic wavefunctions,[16] having the form:

χ_i = Normalisation constant × (exponential function of r) × (spherical harmonic of θ and ϕ)

The spherical harmonic part is identical to the shape variation found for hydrogen atom wavefunctions. Differences from atom to atom are found only in the r-dependent, or radial, part of the orbital.

Atomic orbital functions are orthonormal, *i.e.* $\int \chi_i^2 \, d\tau = 1$, ... Equation 1.22

but $\int \chi_i \chi_j \, d\tau = 0$ if $i \neq j$ ... Equation 1.23

Normalised wavefunctions that are used are not orthogonal, *i.e.*:

$\int \chi_i \chi_j \, d\tau = S_{ij}$, the overlap integral. ... Equation 1.24

1.2.3 Molecular orbitals

The molecular wavefunction Ψ is given by:

$\Psi = \phi_1 \phi_2 \phi_3 \ldots \phi_n$ ... Equation 1.25

where each ϕ_i is a one-electron molecular orbital.

Each one-electron molecular orbital is written as a linear combination of known atomic orbitals (LCAO):

$\phi_i = \sum c_{ik} \chi_k$ ... Equation 1.26

The variation principle[17] states that the more flexible the wavefunction, the lower the energy and by implication the better the wavefunction. For Li_2 a 'better' energy is obtained by adding 2*p* orbitals.

1.2.4 Secular equations

A similar equation to the Schrödinger equation can be written for the MOs:

$$H\phi_i = \varepsilon_i\phi_i \quad \ldots \text{Equation 1.27}$$

H is now a one-electron Hamiltonian, and ε_i is the orbital energy.

ϕ can be expanded as a linear combination of known atomic orbitals:

$$H \sum c_{ik}\chi_k = \varepsilon_i \sum c_{ik}\chi_k \quad \ldots \text{Equation 1.28}$$

If $H_{lk} = \int \chi_l H \chi_k \, d\nu$ and $S_{lk} = \int \chi_l \chi_k \, d\nu$, the secular equations[18,19] can be written as:

$$c_{ik}\,(H_{lk} - \varepsilon_i S_{lk}) = 0 \quad \ldots \text{Equation 1.29}$$

which only has a non-trivial solution if the determinant, $|H_{lk} - \varepsilon S_{lk}| = 0$

The determinantal equation is the basis of all MO methods.

The matrix elements S_{lk} are called overlap integrals. The matrix elements H_{lk} include the one-electron operator H. Most MO methods differ in the definition of H. In more sophisticated MO methods H is precisely defined. In the less precise approximations it is never defined and all H_{lk} matrix elements are replaced by parameters.

1.2.5 Self-consistent molecular orbitals

The most clearly defined MO calculations are based on the Hartree-Fock method.[20] A Hamiltonian operator contains terms for the kinetic and potential energy of the system. The Schrödinger equation for the hydrogen atom with electronic charge, e and mass, m_e being unity, and the unit of length as the Bohr radius becomes:

$$(-\tfrac{1}{2}\nabla^2 - 1/r)\Psi = E\Psi \quad \ldots \text{Equation 1.30}$$

For the hydrogen molecule there are two electrons and the Schrödinger equation is:

$$\{-\tfrac{1}{2}\nabla_1^2 - \tfrac{1}{2}\nabla_2^2 - 1/r_{1A} - 1/r_{1B} - 1/r_{2A} - 1/r_{2B} + 1/r_{12}\}\Psi = E_{el}\Psi \qquad \ldots \text{Equation 1.31}$$

The above Hamiltonian can be separated into two H_2^+ Hamiltonians plus an additional $1/r_{12}$ interelectronic repulsion term. The nuclear-nuclear repulsion energy, $E_{nucl} = 1/R$, where R is the internuclear separation, must also be added to the electronic energy.

For one-electron equations, kinetic energy, nuclear attraction and electron-electron repulsion terms are included in the Hamiltonian. The kinetic and nuclear attraction terms are grouped together to give H_N. For a many-electron molecule, a major part of the potential energy is the electron-electron repulsion. For two electrons in orbitals ϕ_i and ϕ_j separated by a distance r_{12}, the repulsion between them is given by:

$$\int \phi_i^2(1)\ 1/r_{12}\ \phi_j^2(2)\ d\nu_1 d\nu_2 \qquad \ldots \text{Equation 1.32}$$

where $\int \phi_i^2(x)\ d\nu_x$ is the charge distribution of electron (x).

An expanded form of $H\phi = \varepsilon\phi$ might be:

$$\{-\tfrac{1}{2}\nabla^2 - \Sigma\ Z_v/r_{iv} + \Sigma\ \phi_j^2(2)1/r_{12}d\nu_2\}\phi_i(1) = \varepsilon_i\phi_i(1) \qquad \ldots \text{Equation 1.33}$$

This would be correct if the orbital wavefunction were a simple product and not a determinant. The full self-consistent field (SCF) or Hartree-Fock (HF) equations are:

$$H_{SCF}\ \phi i = \varepsilon_i\phi_i \qquad \ldots \text{Equation 1.34}$$

where $H_{SCF} = H_N + \Sigma J - \Sigma' K,$ ... Equation 1.35

$$J_j\ \phi_i(1) = (\int \phi_j^2(2)\ 1/r_{12}\ d\nu_2)\ \phi_i(1), \qquad \ldots \text{Equation 1.36}$$

and $K_j\ \phi_i(1) = \int \phi_j(1)\phi_i(2)\ 1/r_{12}\ d\nu_2.$ ... Equation 1.37

Hartree-Fock equations contain exchange terms, K, in addition to Coulombic interelectronic interactions, J. The exchange terms arise as a result of the Pauli principle, leading to determinantal wavefunctions rather than products, and exchange integrals between various cross-products of the expanded determinant.

The coefficients c_{ik} in $\phi_i = \Sigma\, c_{ik}\phi_k$ must be known at the start. Starting values of c_{ik} are given from educated guesses or from the results of simple calculations. The determinantal equation, det $|H_{SCF} - \varepsilon S_{lk}| = 0$, is solved after calculating all the integrals involved in H_{lk} and S_{lk}. Solution yields values of ε which are substituted in the secular equations to give new values of the various c_{ik}. The process is repeated until the results are **self-consistent**, *i.e.* the c_{ik} resulting from one cycle are identical within prescribed limits with those used in the previous cycle.

The various simplified methods of calculating MOs are essentially approximations of a greater or less drastic nature which result in a reduction of the number of integrals necessary to build the matrix element H_{lk} and S_{lk} in the determinantal equation.

1.3 *AB INITIO* METHODS

All molecular wavefunctions are approximate. The Schrödinger equation can be solved exactly for the hydrogen atom but not for H_2^+. For H_2^+, the adiabatic or Born-Oppenheimer approximation[21,22] is made, which separates electronic and nuclear motion, and calculates the electronic energy of the ion with a given fixed internuclear distance, and gives the total energy on adding the nuclear-nuclear repulsion term.

The term *ab initio* is used for calculations of orbital wavefunctions where the full Hartree-Fock SCF operator is used in:

$$\det |H_{lk} - \varepsilon S_{lk}| = 0$$

$$H = [H_N + \Sigma J - \Sigma' K] \qquad \ldots \text{ Equation 1.35}$$

and all the integrals implied in H_{lk} and S_{lk} are computed.

Each molecular orbital is a linear combination of atomic orbitals. If the expansion is infinite, the most flexible wavefunction within the constraints of the SCF Hamiltonian is achieved. The resulting energy is the Hartree-Fock limiting energy. Short expansions must be used for large molecules, to offset computational expense. The resulting energy will be far from the best possible limit.

When running an *ab initio* calculation, the starting point is a particular molecular geometry, the nature and co-ordinates of each atom being defined. Depending on which atoms are present, a basis set of atomic orbitals is decided. The program then computes all the integrals required in H_{lk} and S_{lk} using trial coefficients, builds and diagonalises the determinant, and produces a set of orbital energies and first-improved coefficients. This process is repeated until self-consistency is achieved.

1.3.1 Basis sets

The most obvious set of atomic orbitals, χ_k, to use in the expansion are Slater-type atomic orbitals, STOs, which were found by fitting analytical exponential functions to numerical atomic wavefunctions. Each of the atomic orbitals is of the form:

$$\chi_k = Ce^{-\zeta r}Y_{lm} \qquad \text{... Equation 1.38}$$

where C is a normalising constant, Y_{lm} is the angular part (spherical harmonic) and ζ is the orbital exponent. The Y_{lm} are known for 1*s*, 2*s*, 2*p* *etc.*, so a basis set can be specified by listing the exponents for each type of orbital used.

If one function is used for each atomic orbital, then the set of atomic functions or basis set is referred to as a **minimal** basis. If double the number is used, the basis set is of **double zeta** (DZ) quality. A **split-valence** (SV) basis set uses two STOs for each valence AO but only one STO for each inner-shell AO. As AOs are polarised upon molecule formation, basis-function STOs can be introduced whose l quantum numbers are greater than the maximum l in the corresponding free atom. **A double-zeta plus polarisation** set (DZ+P or DZP) adds a double-zeta set of five 3*d* functions for on each carbon atom and a set of three 2*p* functions on each hydrogen atom.

To overcome the problems in integrating Slater-type exponential wavefunctions, a gaussian form is used.[23,24] The gaussian radial dependence has the less appropriate $\exp(-\alpha r^2)$ form. Each basis function is taken as a linear combination of a smaller number of **primitive** gaussians, g, according to:

$$\chi_r = \sum d_r g \qquad \ldots \text{Equation 1.39}$$

where χ_r are contracted gaussian-type functions (CGTF). By using CGTFs instead of primitive gaussians, the number of variational coefficients are reduced.

An STO-3G basis indicates a minimal basis set of one STO per AO. Each STO is approximated as a linear combination of three gaussian functions. The STO-3G basis set for a compound of carbon and hydrogen is denoted as ($6s3p/3s$) contracted to [$2s1p/1s$] where parentheses indicate the primitive gaussians and brackets the contracted gaussians. An STO-4G calculation is likely to be closer to a true minimal basis calculation. Huzinaga[25] used a ($9s5p$) basis set of uncontracted gaussians to do SCF calculations on the atoms Li to Ne. For the ground state of the oxygen atom, the optimised orbital exponents of the nine s-type basis GTFs were found to be:

Table 1.1 Optimised orbital exponents of s-type basis GTFs.[25]

g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
7817	1176	273.2	81.2	27.2	9.53	3.41	0.940	0.285

The diffuse function g_9 has the smallest orbital exponent and so falls off most slowly as r increases. As the outer region of an AO changes the most upon molecule formation, the diffuse function can be taken as one of the basis functions. The $2s$ oxygen split-valence contracted basis functions are given as:

$$2s = N' (-0.196g_6 + 0.596g_8),\ 2s' = g_9 \qquad \ldots \text{Equation 1.40}$$

Dunning's DZ [$4s2p$] and Dunning and Hay's SV [$3s2p$] contractions of Huzinaga's ($9s5p$) first-row-atom AOs are used widely in molecular calculations.[26,27]

The 3-21G and 4-31G sets are commonly used split-valence basis sets of CGTFs. In the former, each inner-shell AO is represented by a single CGTF that is a linear combination of three primitive gaussians, while for each valence-shell AO there are two basis functions, one of which is a CGTF that is a linear combination of two gaussian primitives and one which is a single diffuse gaussian. The 6-31G* or 6-31G(*d*) is a split-valence basis set with some polarisation functions added and uses a linear combination of six *d*-type Cartesian gaussian polarisation functions for each non-hydrogen atom. The 6-31G** or 6-31G(*d*, *p*) basis set adds to the 6-31G* set a set of three *p*-type gaussian polarisation functions on each hydrogen atom. For second-row atoms, *d* orbitals contribute significantly to bonding.

The 3-21G(*) basis set is constructed by the addition of a set of six *d*-type gaussian functions on each second-row atom to the 3-21G set. For H-Ne, the 3-21G(*) is the same as the 3-21G set. Anions, compounds with lone pairs and hydrogen bonded dimers have significant electron density at large distances from the nuclei. 3-21+G and 6-31+G* basis sets are formed from the 3-21G and 6-31G* sets by the addition of four highly diffuse functions (s, p_x, p_y, p_z) on each non-hydrogen atom. The 3-21++G and 6-31++G* sets also include a highly diffuse *s* function on each hydrogen atom.

The STO-3G, 3-21G, 3-21G(*), 4-31G, 6-31G* and 6-31G** basis sets were developed by John Pople,[28] 1998 Nobel Chemistry Laureate, and widely used in *ab initio* programs such as GAMESS,[29] CADPAC,[30] and GAUSSIAN.[31]

1.4 SEMI-EMPIRICAL CALCULATIONS

Ab initio calculations are not always successful in reproducing experimental observations. Slater functions are not particularly amenable to implementation in MO calculations, as some of the integrals are difficult, if not impossible, to evaluate, particularly when the atomic orbitals are centred on different nuclei. Gaussian functions are used instead of STOs, but underestimate the long-range overlap between atoms and the charge and spin densities at the nucleus, due to the cusp nature of STOs at the origin. The number of integrals required in an *ab initio* Hartree-Fock approach increases roughly as the fourth power of the number of basis functions in the molecule, and many millions may be required for a small molecule.

Semi-empirical molecular orbital methods start with the determinantal equation but make a variety of approximations to reduce the amount of computer time required. In semi-empirical calculations, the basis set is restricted to valence orbitals only, which are built up using STOs rather than GTOs. The orthogonality of such orbitals enables further simplifications to be made to the Roothaan-Hall equations. The approximations result in a method that is about 100-10,000 times faster than *ab initio* HF procedures, meaning that the "throughput" of the method is potentially higher. Semi-empirical calculations usually used are for preliminary investigations, followed by a focus *using ab initio* methods on the most interesting systems to emerge.

All the commonly used techniques are valence electron calculations, neglecting 1*s* electrons as they play little part in chemical behaviour. The 1*s* electrons, together with *K* and *L* shell electrons for heavier atoms, are defined as part of the "core".

The SCF Hamiltonian thus becomes:

$$H = H_{\text{core}} + \Sigma J - \Sigma' K \qquad \ldots \text{ Equation 1.41}$$

H_{core} incorporates kinetic energy and attraction to a core rather than to a bare nucleus. Integrals or matrix elements involving H_{core} are replaced by empirical or calculated parameters. When molecules contain a heavy atom, such as a metal in an enzyme-binding site, an approximation is introduced by incorporating a pseudopotential or effective potential into the Hamiltonian. Savings in computer time of 50% or more are common. Parameters are chosen based either on experimental or *ab initio* results. Parameters based on experiment include correlation effects. Good agreement with experiment is obtained providing the experiment and source of parameterisation are related. Parameters based on *ab initio* calculations are, however, clearly defined.

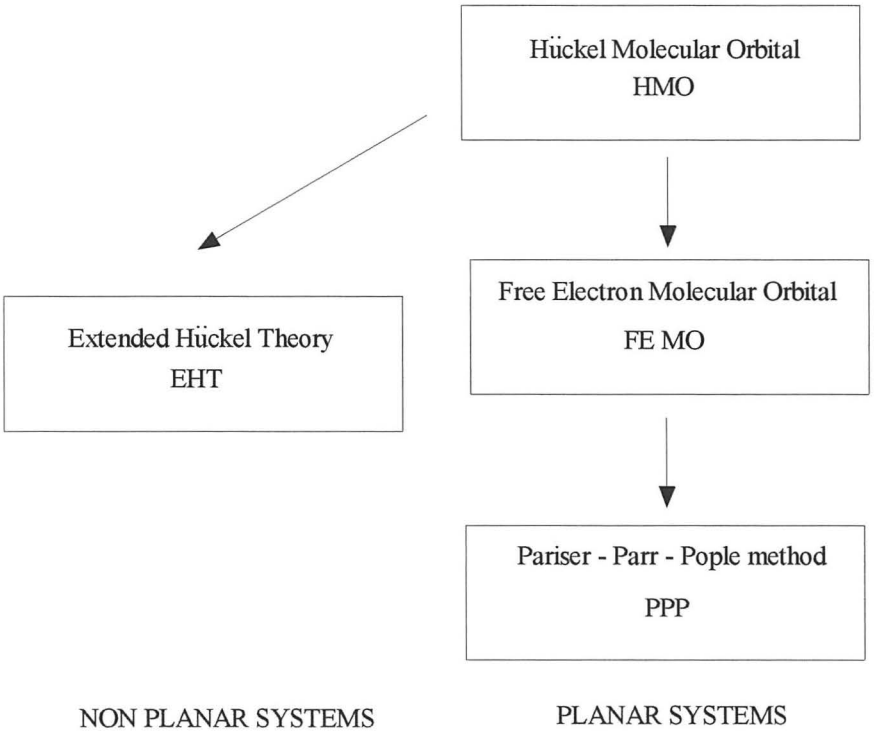
1.4.1 π-Electron methods

The development of π-electron methods is shown overleaf. In the π-electron approximation, π electrons are treated separately by incorporating the effects of σ electrons and nuclei into an effective Hamiltonian H_π:

$$H_\pi = \Sigma H_\pi^{core} + \Sigma 1/r_{ij} \quad \text{... Equation 1.42}$$

$$H_\pi^{core}(i) = -\tfrac{1}{2}\nabla_i^2 + V_i \quad \text{... Equation 1.43}$$

Figure 1.3 Development of π-electron methods.



The simplest semi-empirical π-electron method is the free-electron molecular-orbital (FE MO) method.[32-35] The interelectronic repulsions $1/r_{ij}$ are ignored, and the effects of the σ electrons are represented by a particle-in-a-box potential function. The most celebrated semi-empirical π-electron method is the Hückel molecular orbital (HMO) method, developed in the 1930s, the results of which have been tabulated for hundreds of compounds.[36-38] The π-MOs are approximated as linear combinations of $2p_x$ AOs. One-electron Hückel theory neglects interelectronic repulsions.

For butadiene we get:

$$\phi_1 = 0.372\chi_1 + 0.602\chi_2 + 0.602\chi_3 + 0.372\chi_4 \text{ , } E_1 = \alpha + 1.618\beta \quad \text{... Equation 1.44}$$

$$\phi_2 = 0.602\chi_1 + 0.372\chi_2 + 0.372\chi_3 - 0.602\chi_4 \text{ , } E_2 = \alpha + 0.618\beta \quad \text{... Equation 1.45}$$

$$\phi_3 = 0.602\chi_1 - 0.372\chi_2 - 0.372\chi_3 + 0.602\chi_4 \text{ , } E_3 = \alpha - 0.618\beta \quad \text{... Equation 1.46}$$

$$\phi_4 = 0.372\chi_1 - 0.602\chi_2 + 0.602\chi_3 - 0.372\chi_4 \text{ , } E_4 = \alpha - 1.618\beta \quad \text{... Equation 1.47}$$

where α is the Coulomb integral and β is the bond or resonance integral.

For planar conjugated molecules involving π-bonding to heteroatoms, the Coulomb and bond integrals for the heteroatoms must be modified. For heteroatoms X and Y:

$$\alpha_X = \alpha_C + h_X\beta_{CC} \quad \text{... Equation 1.48}$$

$$\beta_{XY} = k_{XY}\beta_{CC} \quad \text{... Equation 1.49}$$

A two-electron method taking electron repulsion into account is the Pariser-Parr-Pople (PPP) method.[39-41] The π-electron Hamiltonian including electron repulsions is used, and the π-electron wavefunction is written as an antisymmetrised product of π-electron spin orbitals. In addition to σ-π separability, the PPP method assumes as in Hückel theory that overlap is neglected. Consistent with the neglect of overlap integrals, the PPP method uses the zero differential overlap (ZDO) approximation, discussed in *Section 1.4.2*. The PPP method gives a good account of the electronic spectra of many, but not all, aromatic hydrocarbons.[42,43]

The most important one-electron semi-empirical MO method for nonplanar molecules is the Extended Hückel Theory (EHT).[44-47] In the EHT treatment of nonplanar hydrocarbons, each valence MO contains contributions from four ($2s$, $2p_x$, $2p_y$, $2p_z$) AOs on each carbon atom, and one $1s$ AO on each hydrogen atom. Overlap integrals are not neglected. Poor predictions of molecular properties are obtained.

1.4.2 Neglect of differential overlap

The matrix elements H_{lk} in the secular determinant involve a large number of integrals over atomic orbital functions χ of the type:

$$\int \chi_m(1)\chi_n(1)\ 1/r_{12}\ \chi_l(2)\chi_s(2)\ d\tau_1 d\tau_2 \quad \text{... Equation 1.50}$$

which are often written $(\mu\nu|\lambda\sigma)$.

The integrals $(\mu\nu|\lambda\sigma)$ are difficult to evaluate if the atomic functions, χ_m, are centred on different atoms. It is assumed that: $(\mu\nu|\lambda\sigma) = \delta_{mn}\delta_{ls}(\mu\mu|\lambda\lambda)$ where $\delta_{\mu\nu}$ and $\delta_{\lambda\sigma}$ are 'Kronecker deltas', equal to zero unless the subscripts are equal. By neglecting differential overlap of functions based on different centres, all three- and four-centre integrals, and most two- and one-centre integrals are eliminated where different AOs

are involved for either of the two electrons. The neglected integrals are significantly smaller than the remaining integrals, but are not negligible.

In the complete neglect of differential overlap (CNDO) method,[48,49] the first method to implement the zero-differential overlap approximation in a practical manner., electron repulsion integrals containing different valence AOs centred on the same atom are neglected. To overcome the problems of rotational invariance, the two-electron integrals $(\mu\mu|\lambda\lambda)$, where μ and λ are on two different atoms A and B, are set equal to a parameter γ_{AB}, the average electrostatic repulsion between an electron on atom A and an electron on atom B. Matrix elements of H_{core} are also parameterised, and in particular, atomic ionisation potentials are frequently used to replace integrals. There are a number of alternative parameterisations.

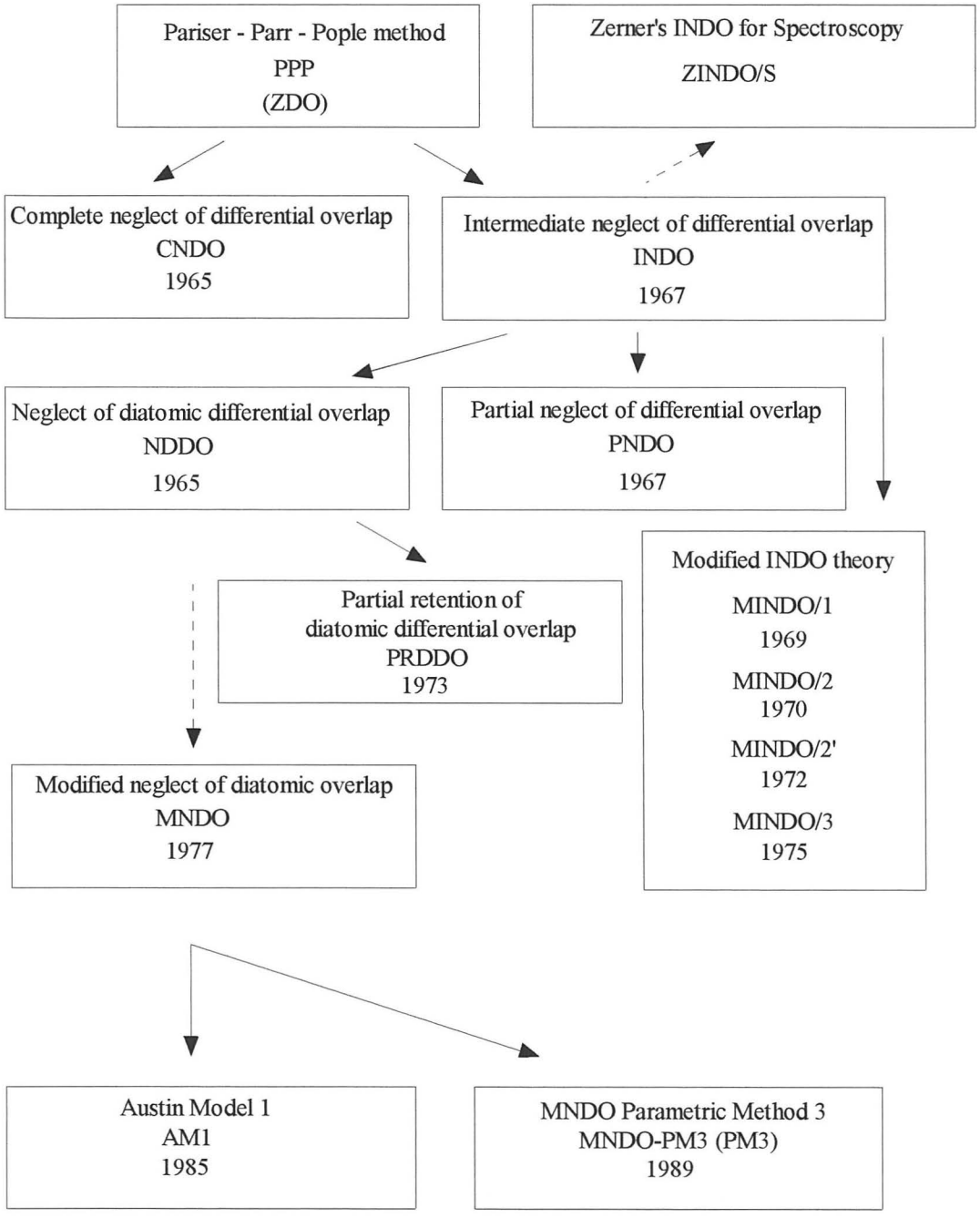
Intermediate neglect of differential overlap (INDO) schemes[49] are less dramatic in the application of neglect of differential overlap, and counter a defect in CNDO which results in there being no distinction between singlet and triplet electronic states. Their energy difference is related to the exchange integral, K. Integrals of the form $(\mu\sigma|\mu\sigma)$ are not neglected between different AOs centred on the same atom.

Versions of CNDO and INDO parameterised to predict electronic spectra are called CNDO/S and INDO/S.[50,51] Zerner's group modified INDO/S and produced ZINDO/1[52] for transition metals and ZINDO/S[53] for organic compounds, which will be used in conjunction with configuration interaction (CI) (*Section 1.5*) to calculate wavelength maxima of dyes in *Chapter 7*.

The Neglect of diatomic differential overlap (NDDO) method is an improvement on INDO in which differential overlap is neglected only between AOs centred on different atoms. Initial attempts at parameterising NDDO gave disappointing results.[54] NDDO was only used after Dewar and Thiel modified it to give the MNDO method. The partial retention of diatomic differential overlap (PRDDO) method was developed by Halgren and Lipscomb.[55,56] The PRDDO uses symmetrically orthogonalised or (Löwdin) AOs.[57,58] Fewer integrals are neglected than in NDDO. PRDDO calculations are close to those of an *ab initio* STO-3G calculation.[59] The

aims of the CNDO, INDO and PRDDO methods were to reproduce the results of minimal basis set *ab initio* calculations. The partial neglect of differential overlap (PNDO) method developed by Dewar and Klopman in 1967 was parameterised to reproduce molecular binding energies.

Figure 1.4 Development of NDO based semi-empirical methods.



In 1969, Baird and Dewar published the MINDO/1 (first version of modified INDO) theory.[60] Parameters were chosen to predict molecular heats of formation. MINDO/2

gave improved molecular geometries,[61] MINDO/2' gave better dipole moments and bond lengths to hydrogen,[62] culminating with MINDO/3 in 1975.[63-67] Average absolute errors in MINDO/3 calculated ionisation potentials are 0.8 eV.[68] As MINDO/3 did not meet Dewar's original aims, Dewar and co-workers developed the modified neglect of diatomic overlap (MNDO) method published in 1977,[67,69-73] based on the NDDO approximation. Average absolute MNDO errors are better than those of MINDO/3, *e.g.* 0.5eV in ionisation potentials.[68]

In 1985, Dewar and co-workers published an improved version of MNDO called AM1 (Austin Model 1, named after the University of Texas at Austin).[74-78] MNDO fails to reproduce intermolecular hydrogen bonds and AM1 partly corrects this. In 1989, Stewart reparameterised MNDO to give MNDO-PM3 or PM3 (MNDO parametric method 3; 1 and 2 being MNDO and AM1), reducing the errors of MNDO and AM1 in calculating heats of formation without loss of accuracy in molecular geometry and dipole moments.[77,79,80]

In 1980, Jug and Nanda published the SINDO1 (symmetrically orthogonalised INDO) method,[66,81,82] using symmetrically orthogonalised or (Löwdin) AOs[57,58] as basis functions, and pseudopotentials to eliminate inner-shell electrons. The accuracy of SINDO1 is similar to MNDO, being 0.8 eV in ionisation potentials.

The PCILO (pertubative configuration interaction using localised orbitals) method was developed in 1969.[83] A set of localised bonding, lone pair and antibonding orbitals is chosen for CI treatment. Pertubation theory is used to calculate the energy.

1.5 CONFIGURATION INTERACTION

The variation principle states that the more flexible the wavefunction, the better it will be in terms of the energy which results. One way of improving the wavefunctions is to allow the interaction of configurations, such as linear mixing:

$$\Psi_{improved} = a\Psi_0 + b\Psi_1 + c\Psi_2 + \ldots \qquad \ldots \text{Equation 1.51}$$

Ψ_0 is the wavefunction first found and Ψ_1 and Ψ_2 are wavefunctions appropriate for excited configurations of the same symmetry. The problem with CI is to know when to stop, as every extra configuration improves the energy. Although configurations closest in energy to the ground state may have an obvious influence, others, which are highly excited, may also be significant.

1.6 RHF AND UHF WAVEFUNCTIONS

For SCF MO calculations on closed-shell states, electrons paired with each other are almost always given precisely the same spatial orbital function. A Hartree-Fock wavefunction in which electrons whose spins are paired occupy the same spatial orbital is called a restricted Hartree-Fock (RHF) wavefunction.[84,85] For open-shell states there are two methods, the restricted open-shell Hartree-Fock (ROHF) method and the unrestricted Hartree-Fock (UHF) method.[86] In the ROHF method, electrons that are paired with each other are given the same spatial orbital. The ROHF wavefunction of the Li ground state is $|1s^{\alpha}1s^{\beta}2s^{\alpha}|$. In the UHF treatment, the spatial orbitals of spin-α electrons are allowed to differ from those of spin-β electrons. The UHF wavefunction for the Li ground state is $|1s^{\alpha}1s'^{\beta}2s^{\alpha}|$, where $1s \neq 1s'$. UHF wavefunctions give slightly lower energies than ROHF wavefunctions.

1.7 CORRELATION ENERGY

Any defect in the wavefunctions will result in the calculated energy being less than the true energy. Correlation energy is given by:[17]

$$E_{\text{correlation}} = E_{\text{true}} - E_{\text{Hartree-Fock}} - E_{\text{relativistic}} \quad \ldots \text{Equation 1.52}$$

and represents the remaining energy error between the limiting Hartree-Fock energy and true total energy taking into account the relativistic effect.

Correlation energy errors arise due to an unjustified approximation in the SCF Hamiltonian:

$$H = [H_{\text{N}} + \Sigma J - \Sigma' K] \quad \ldots \text{Equation 1.35}$$

Involved in J and K are terms of the type $\int \phi_j^2 . 1/r_{12}\, d\nu$, which are included to account for the interaction of one electron with another, the second electron being represented as a smoothed-out average electron density.

The correlation energy error results from electron pair effects and is reasonably constant as molecular geometry changes providing the electron pairings in the molecule do not change. It is not a constant as bonds are stretched to the point of dissociation. Errors are therefore serious in calculations of dissociation energies, ionisation potentials, pKa values or electronic transition energies.

1.8 CALCULATED ENERGY PROPERTIES

The ionisation potential and electron affinity of a molecule should, in principle, be simple to calculate - perform a calculation for the neutral molecule in its ground-state geometry and conformation and similar calculations for the positive and negative ions, then by difference the desired energies are obtained.

Such a procedure would only be expected to yield values in accord with experiment if the correlation energy error, the relativistic error and the discrepancy from the Hartree-Fock limiting energy were identical for the calculations on the molecule and on the ions. The correlation energy error, being dependent to a first approximation on the number of electron pairs in the molecule or ion will differ in the two situations even though the other two errors may cancel.

A more drastic but simply applied approximation is often employed, equating the ionisation potential of an electron with the orbital energy ε_i associated with that electron. This is known as Koopmans' theorem.[87] It implies that not only are relativistic and correlation energies the same in the molecule and in the ion but also there is no reorganisation of electronic structure or distribution on ionisation. Such approximations are not valid. However, when considering differences between ionisation potentials for a series of similar molecules many of the errors are constant and an acceptable indication of trends can be obtained. The molecular orbital whose energy approximates to the ionisation potential is the HOMO.

Similarly, the energy of the lowest unoccupied molecular orbital, LUMO, ε_i for the first virtual orbital, may give an indication of the ease with which the molecule may accept an electron. In addition to the above assumption, this approximation is made even less acceptable, ass there is no definite physical meaning for the orthogonal virtual orbitals. Electron affinities are not well calculated, as SCF calculations frequently do not converge for negative ions, although the simple equation of ionisation potential with orbital energy is satisfactory for comparative purposes.

1.9 COMPUTERS

The costs of computing have steadily fallen over the years, while the power in terms of hardware has grown, encouraging theoretical chemists, and later physical, organic and inorganic chemists to use computational packages. In some cases it is cheaper to model a system on a computer rather than experimentally. Software packages are usually written in FORTRAN although graphical programs are increasingly written in "C". Most programs have grown as generations of users add new functionality or options, perhaps cannibalising subroutines from other programs.

Hardware available to computational chemists extends from PCs, through workstations and mainframe machines to supercomputers and parallel devices. The bigger the problem (molecular size, number of integrals *etc.*), the more power or computer time is needed. PCs can be used to visualise molecules, although calculations beyond this level are somewhat restricted. Workstations can be used for almost the entire range of computational chemistry, and are favoured by industry, as they have a large amount of memory and are often joined *via* ethernet links to other computers. Supercomputers represent the top end of the scale of computational power, and are necessary for the biggest problems, particularly if many millions of integrals have to be performed.

1.10 APPLICATIONS OF COMPUTATIONAL QUANTUM CHEMISTRY

Applications of computational chemistry range from materials science to pharmaceuticals, such as:

atomistic simulations[88] of dopant substitution and oxygen vacancy migration in the $LaGaO_3$-based oxygen ion conductor using GULP;[89]

ab initio molecular dynamics simulations[90] of proton diffusion[91] in the protonic conductor Y-doped $SrCeO_3$ using ultrasoft pseudopotentials;

calculating energy differences between conformational isomers of calix[4]arenes;[92]

calculating the properties of reaction intermediates and transition states;[93]

grand canonical Monte-Carlo simulations of Lennard-Jones mixtures in slit micropores;[94]

studies of angiotensin-converting enzyme (ACE) inhibitors;[95]

studies of enzyme-substrate binding using DOCK;[96]

de novo ligand design using GRID,[97] LUDI[98] and CAVEAT;[99]

and the design of new pharmacophores and leads, such as HIV-1 protease inhibitors.[100]

1.11 SUMMARY

The underlying theories behind the various molecular mechanics, *ab initio* and semi-empirical quantum mechanical methods have been discussed. The roles that Koopmans' theorem, electron repulsion integrals, electron correlation, and configuration interaction, play in MO theory, will become apparent when in *Chapters 6*, *7* and *8*, the colour and the position of the electronic wavelength maximum, λ_{max}, of dyes are calculated using various semi-empirical and *ab initio* methods.

CHAPTER 2
DYES AND STAINING

Staining is the colouring of specimens with a dye, for examination under the microscope, in order to heighten contrast between different structures, such as tissue.

This chapter provides a brief review of the theory of colour, and the history of biological staining. The definition of colligators and a discussion of the interactions of acid and basic dyes will enable the reader to understand how certain dyes not only interact with tissues, but with other dyes. Some of the dyes used in the Papanicolaou staining technique, discussed in *Chapter 3*, form compound dyes, which may be colourless in the case of orange G and light green SF yellowish, or polychromatic in the case of acid fuchsin and light green SF yellowish, which form the falgic acids. The difference between complementary and supplementary secondary staining is also addressed. Gurr's reclassification of synthetic dyes based on colligator types, enabling histologists to develop staining using rational methods is also described. The effect of molecular morphology, acidification, and mordants on stains will be examined. A glossary of cytological / histological terminology is also contained in this thesis.

2.1 THEORY OF COLOUR

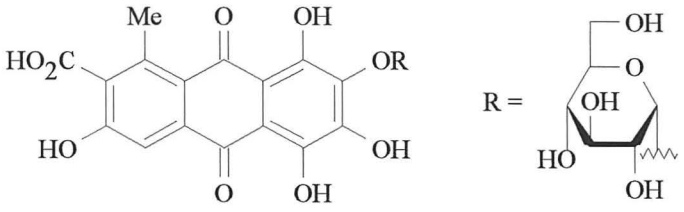
White light is composed of light of the visible spectrum. An object is coloured if it absorbs some, but not all, of the visible radiation. The unabsorbed complementary colour (*Table 2.1*) is reflected and is perceived by our eyes. As absorption often takes place in multiple regions of the visible spectrum, the observed colour also depends on the intensity of the absorption.

Table 2.1 The Visible Spectrum.

Wavelength / nm	Absorbed colour	Complementary colour
400-430	Violet	Greenish-yellow
430-490	Blue	Yellowish-orange
490-510	Blue-green	Red
510-530	Green	Purple
530-560	Yellowish-green	Violet
560-590	Yellow	Blue
590-610	Orange	Greenish-blue
610-750	Red	Blue-green

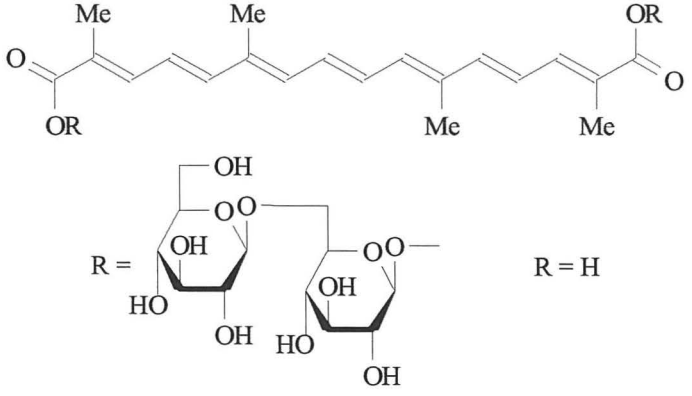
2.2 HISTORY OF BIOLOGICAL STAINING

The first uses of stains in microscopy are accredited by Gurr,[101] Rooseboom[102] and others to Grew[103] in 1682 (cochineal extract on potatoes), Leeuwenhoek[104] in 1719 (saffron on bovine muscle), Sarrabat[105] in 1733, and Hill[106] in 1770. Cochineal **(1)** comes from the dried bodies of the female of the species, *Coccus cacti coccinellifera*, a tropical insect, which feeds on the cactus *Nopalea cochinellifera*. The crude product, carminic acid, is extracted, although carmine, the alumina **lake** of cochineal is preferred. The first stain was of animal origin and was used on plant tissue.



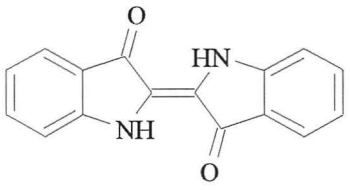
(1)

The first plant stain used on animal tissues was an extract of saffron, the dried stigmata of *Crocus sativus L.* (fam. *Iridaceae*) used to stain bovine muscle.[104] The main constituent of saffron is crocin **(2)**, the di-gentiobiose ester of crocein **(3)**.



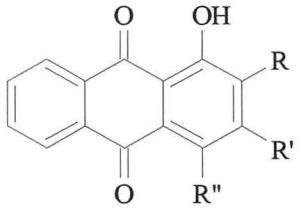
(2) **(3)**

Before Perkin founded the synthetic dye industry with mauve in 1856, the stains used by microscopists were natural colouring matters, such as: cochineal **(1)**,[107,108] indigo **(4)** from *Indigofera tinctoria* and *Isatis tinctoria*, madder from plants of the order *Rubiaceae*, seaweed extract, and inorganic substances such as copper sulphate.



(4)

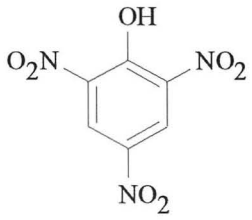
Madder is mainly alizarin **(5)**[109,110] and a mixture of other dyes, such as purpurin **(6)**,[111,112] purpuroxanthin **(7)**[113,114] pseudopurpurin **(8)**,[115,116] rubiadin[117,118] and munjistin.[114,119]



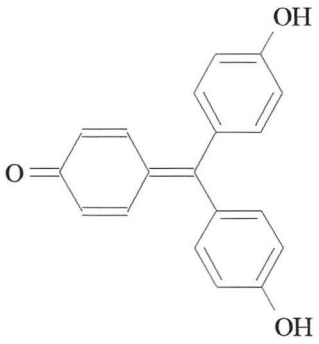
(5) R = OH, $R^1 = R^2 = H$ **(6)** R = H, $R^1 = R^2 = OH$

(7) $R = R^2 = H$, $R^1 = OH$ **(8)** R = COOH, $R^1 = R^2 = OH$

Two early examples of synthetic dyes are picric acid **(9)**,[120] and rosolic acid or aurin **(10)** prepared by oxidising phenol.[121,122]

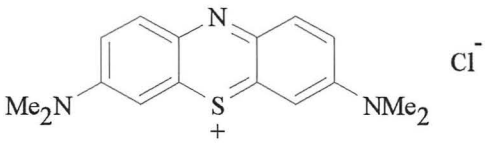


(9)



(10)

Beneke[123] used the first synthetic dye, "lilac aniline", in microtechnique. Multiple staining techniques became popular with increasing availability of synthetic dyes. For example, although Schwartz stained tissues with picric acid after ammonia carmine primary staining,[124] Ranvier combined these two stains into one solution, starting the modern picro-carmine technique.[125]



(11)

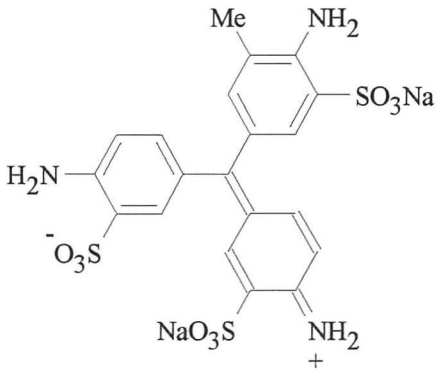
Up to the time of Ehrlich[126,127] all staining procedures were empirical, with the use[128] of red cabbage,[129] bilberry juice,[101] blackcurrant juice,[101] walnut juice[101] and litmus[101] as biological stains. Single staining has progressed to multiple staining, and from empirical to rational methods. As more becomes known of the chemical reactions of synthetic dyes with each other and with tissues, and the physio-chemical conditions that produce the best results, empirical methods are being transformed into rational methods. However, conglomerations of alchemic ingredients of certain staining solutions may never be rationalised.

2.3 INTERACTIONS OF DYE MOLECULES

Tissues are built of complex organic substances, whereas dyes are composed of homogeneous molecules. The most complex dyes are simple in comparison with tissue components such as proteins, nucleic acids, lipids, and polysaccharides.

Gurr defined a **colligator** as a group that enables a dye to unite chemically with tissue elements or with certain other dyes and compounds.[101] Colligators can be divided into two groups, acidic and basic. Acidic colligators such $-SO_3^-$, $-SO_3H$, $-SO_3Na$, $-CO_2H$, $-CO_2Na$, -OH, -ONa or -OK enable acid dyes to bind to tissue elements which are basic in reaction, to basic dyes or other basic substances. Basic colligators such as $-NH_2$, -NHR, $=N^+R_2$, or $-N^+R_3$ enable basic dyes to bind to acidic tissue elements and acid dyes. In amphoteric anionic dyes, these groups enable the dyes to bind to more acidic dyes. Reactive groups common to both dyes and tissues are amino, imino, quaternary nitrogen, carboxyl and hydroxyl.

Ehrlich was the first to realise that organic dyes were divisible into acid, neutral and basic dyes.[126,127] He found that **neutrophile** granules in **leukocytes** and **phagocytes** were stained by the neutral methylene blue and methyl green compound dyes of acid fuchsin **(12)**, but not when the ionic parent dyes were used singly, suggesting that unionised neutral compound dye molecules are responsible for the staining.[126]



(12)

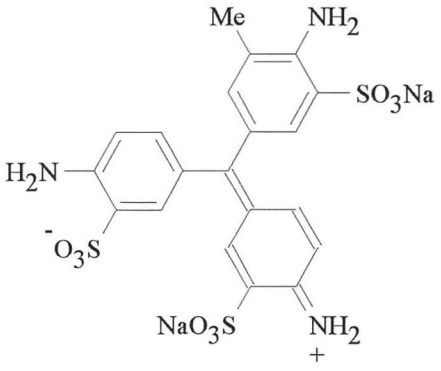
Similarly, the Jenner stain,[130,131] (see *Section 2.6*) an eosinate of methylene blue, is ionised when the alcoholic solution is diluted with water during staining. Acidic eosinate anions stain certain basic tissue elements, basic methylene blue cations stain certain acidic tissue elements, while undissociated molecules stain neutrophile tissues.

2.4 INTERACTIONS BETWEEN ACID DYES

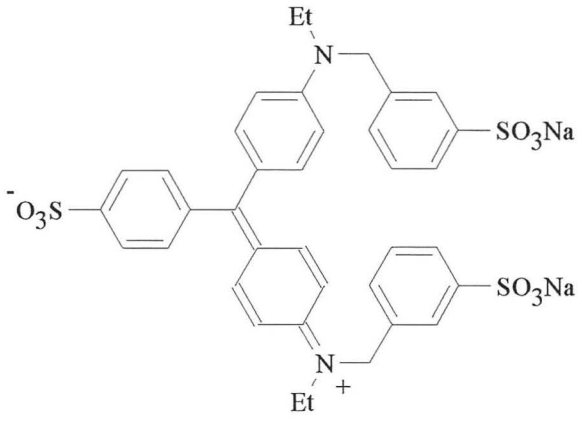
An acid dye will:

(i) stain tissue elements that are basic in reaction, and

(ii) interact with basic dyes to form compound dyes, which may be either **dichromatic** or **polychromatic**, provided the basic dyes are different in colour from the acid component.



(12)

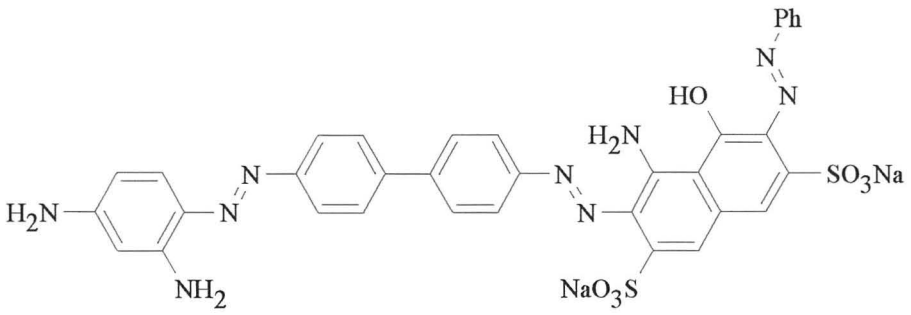


(13)

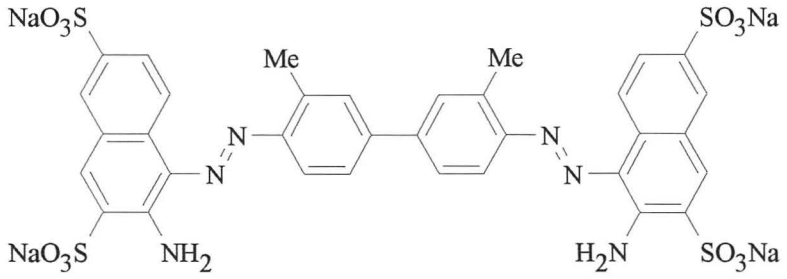
Sections of spinal cord stained with the red anionic dye acid fuchsin **(12)**, washed with water, and then **counterstained** with the green anionic dye light green SF yellowish **(13)**, the Falg technique, gave a polychrome picture in shades of red,

mauve, violet and blue. Although the mauve and violet secondary colours can be produced by mixing red and green, the blue primary colour, also seen in aqueous solution, was caused by a **compound** dye - trifalgic acid,[132-134] formed by the reaction of acid fuchsin with light green SF yellowish. The same polychrome mixture was found with other plant and animal tissues. Both acid fuchsin **(12)** and light green SF yellowish **(13)** are components of Shorr's cervical smear stain (*Section 3.6*).

Many anionic acid dyes are amphoteric, *i.e.* in addition to their acidic groups, *e.g.* hydroxyl or sulphonic acid groups, the dye ions also have basic groups. A few cationic basic dyes also have carboxyl groups and therefore amphoteric properties, but are weakly acidic. The anionic dye acid fuchsin **(12)** has unique properties, being the only common dye with the same number of acidic and basic groups. Most dyes have an inequality, such as chlorazol black E **(14)** with three amino and two sulphonic acid groups, and trypan red[131,135] **(15)** with four sulphonic acid and two amino groups.



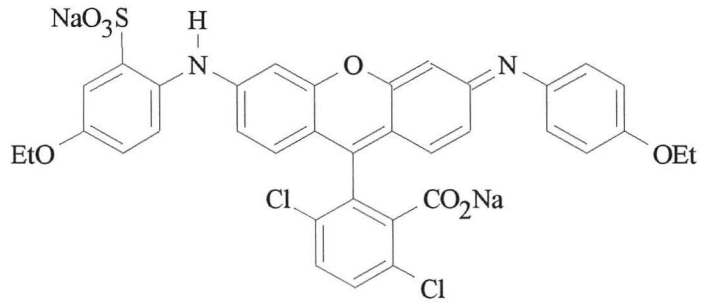
(14)



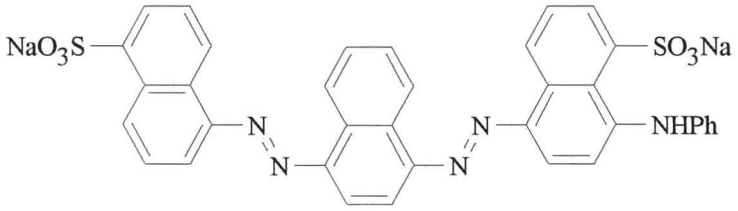
(15)

Acid fuchsin **(12)** reacts with other acid dyes through its basic amino groups to form water-soluble polychrome compound dyes. It also reacts through its acidic sulphonic groups with basic dyes to form polychrome compound dyes which are insoluble in

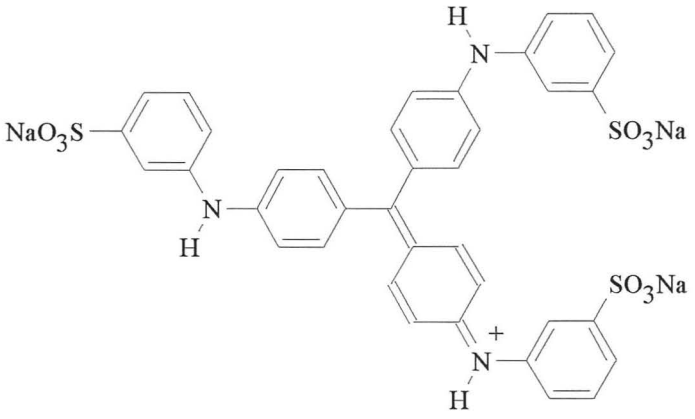
water, but soluble in alcohols and other organic solvents. When applied without a **mordant**, acid fuchsin stains only acidic tissue elements. Acid fuchsin **(12)** acts as a mordant between certain basic tissue elements and other anionic dyes, including light green SF yellowish **(13)**, violamine 3B **(16)**, naphthalene violet **(17)** and aniline blue aqueous **(18)**.[132,136] Using a dilute solutions of acid fuchsin and light green as the primary and counterstain respectively, MacConaill found that the counterstain stained only those tissue elements that had already been stained by acid fuchsin. Under suitable conditions any anionic dye having one or more basic colligators is capable of interaction, as a base, towards certain other more strongly acid anionic dyes.



(16)

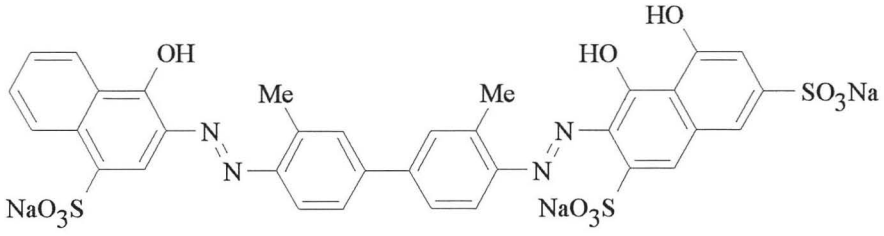


(17)

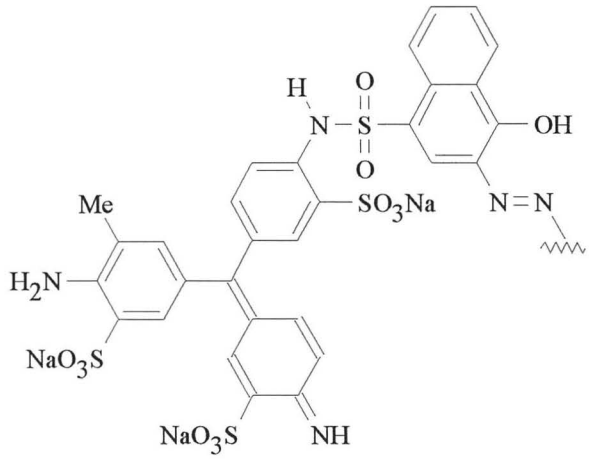


(18)

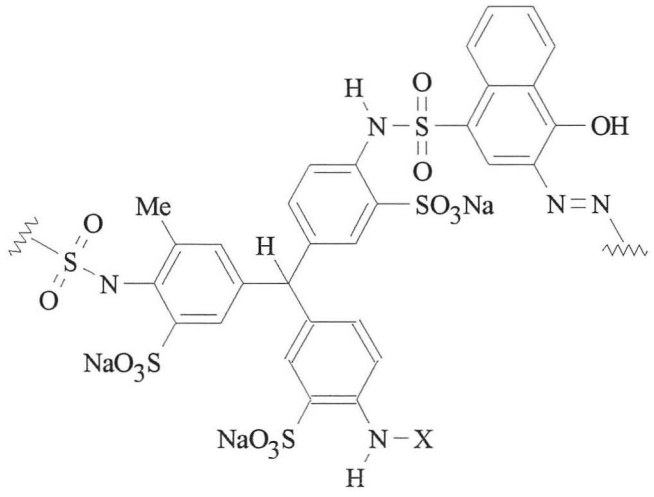
Over one hundred compound anionic dyes, such as mono-, di- and tridianofuchsinic acids **(20)**, **(21)** and **(22)**, formed *in vitro* by the interaction of one acid fuchsin molecule with one, two or three molecules of dianil blue 2R **(19)** respectively,[138] were synthesised *in vitro* and isolated in the solid state by Gurr and MacConaill.[134,137-139]



(19)



(20)

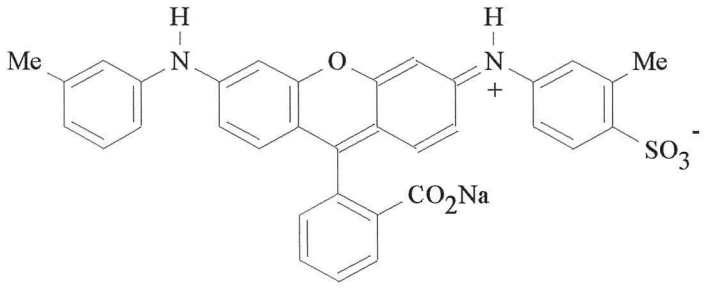


(21) X = H **(22)** X = SO_2-WW

where WW is the remainder of the dianyl blue 2R moiety.

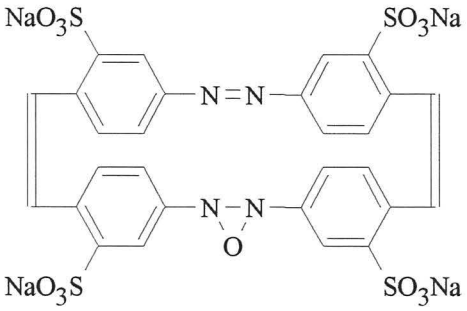
If drops containing aqueous solutions of the three compound dyes **(20)**, **(21)** and **(22)**, are allowed to spread out separately on a piece of filter paper, three spots identical in colour are produced, each blue spot ringed by a moist, colourless zone containing *leuco* acid fuchsin. On heating, the colourless ring becomes red, the same colour as acid fuchsin. The red circular band is narrowest in the case of tridianofuchsinic acid as there is only one acid fuchsin moiety to three dianil blue moieties in the latter.

Other compound dyes of interest are the falgic acids, the xalgic acids, and the anilofuchsinic acids. The anilofuchsinic acids are formed by the reaction of acid fuchsin with aniline blue aqueous. Mono-, di- and trifalgic acids, FaLg, $FaLg_2$ and $FaLg_3$, are formed both *in tela* and *in vitro* by interactions between acid fuchsin (Fa) **(12)** and one, two or three light green SF yellowish (Lg) **(13)** molecules respectively. Mono-, di- and tri-xalgic acids, XaLg, $XaLg_2$ and $XaLg_3$ are formed both *in tela* and *in vitro* by interactions between xanthene red (Xa) and one, two or three light green SF yellowish (Lg) molecules respectively. In the xalgic acids, light green SF yellowish has a **bathochromic** effect on xanthene red **(23)**, whereas, in the case of the falgic acids, acid fuchsin has a bathochromic effect on light green SF yellowish.



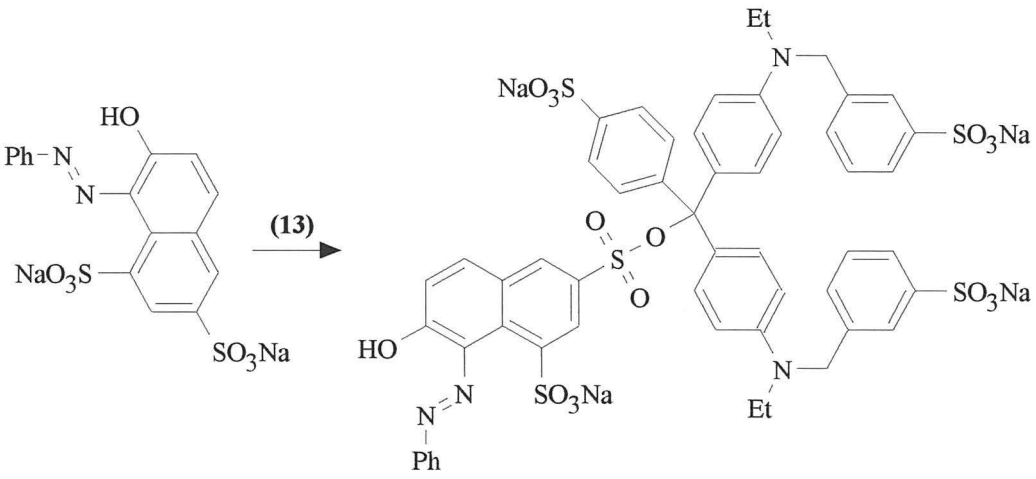
(23)

Certain strongly acidic anionic dyes unite with certain other dyes and decolourise the latter. The decolourised dye is always a quinonoid dye, while the decolourising dye may or may not be a quinonoid dye. The decolourised quinonoid dye has a bathochromic effect on the decolourising dye, *e.g.* when acid fuchsin and light green SF yellowish interact to form trifalgic acid. Sun yellow G **(24)**, a strong decolourising dye, acts on violamine 3B **(16)**, acid fuchsin **(12)**, light green SF yellowish **(13)** and other dyes, bringing about a reduction or complete loss of colour.



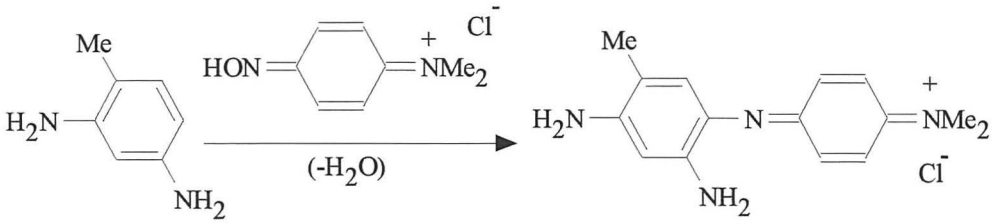
(24)

Orange G **(25)** is another decolourising dye, but its action is not as strong as sun yellow G **(24)**. It also forms compounds with acid fuchsin, acting as a decolourising agent towards the latter as well as towards certain other quinonoid dyes such as light green SF yellowish **(13)**, producing the colourless compound dye **(26)**.



(25) **(26)**

Some of the compounds formed by the interactions of pair and trios of anionic dyes can be regarded as salts. Most prepared *in vitro* can be described as condensation products. The compound dyes synthesised *in vitro* from pairs and trios of anionic dyes, although produced after prolonged heating, were prepared without the aid of a condensing agent and in the presence of a large amount of water. This is unusual, but other examples are known in dyestuff chemistry, *e.g.* the preparation of toluylene blue **(27)**,[140,141] the condensation product of *m*-toluylenediamine and *p*-nitrosodimethylaniline hydrochloride.



(27)

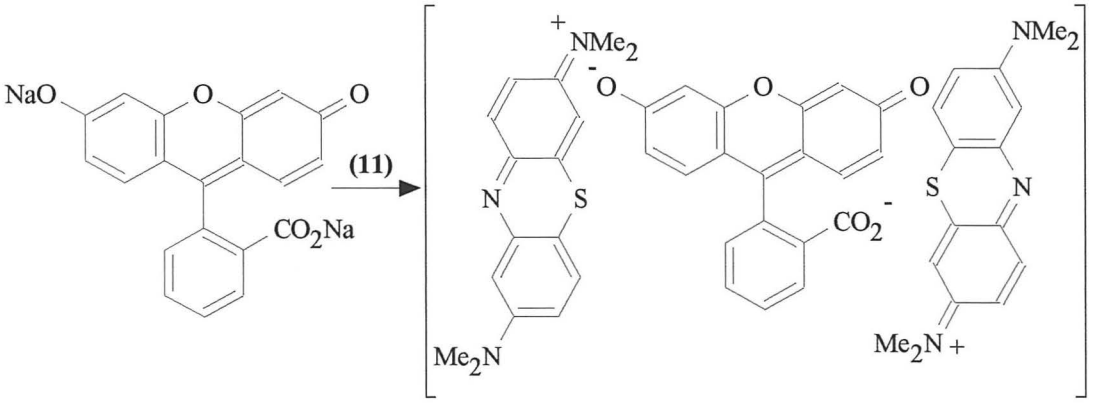
Anionic dyes containing basic groups, may be regarded as synthetic amino acids. Compound dyes formed by interactions of these dyes may therefore be considered to be simple synthetic polychromatic peptides.[137] While natural proteins / peptides are composed primarily of about 20 specific "standard amino acids", compound dyes can be prepared from over a thousand amphoteric anionic dyes.

2.5 COMPLEMENTARY AND SUPPLEMENTARY STAINING

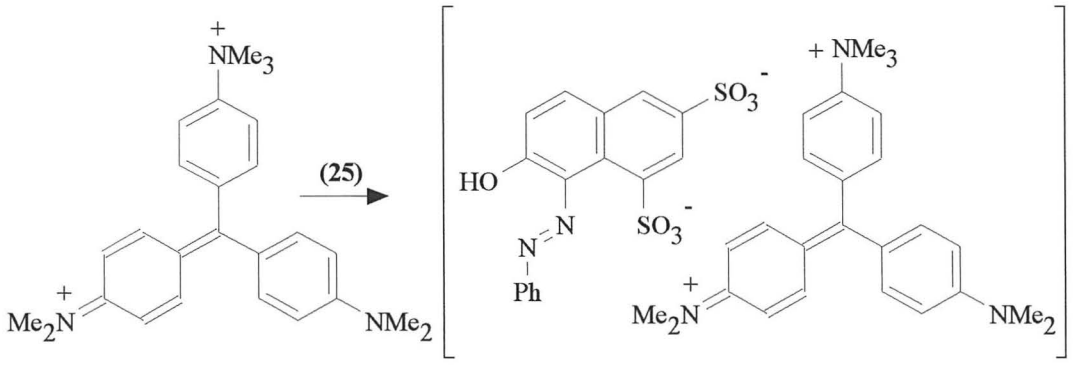
In the hæmatoxylin-eosin technique,[101] eosin may be regarded as a **complementary** secondary stain, attaching itself to a different moiety of tissue element from that to which the hæmatoxylin primary stain has already attached itself. In the Falg[131,132,136] and similar techniques[131,138,142,143] in which acid fuchsin is the primary stain, and light green or some other anionic dye is the secondary stain, the latter attaches itself only to those tissue elements that have already been stained by the primary stain. In **supplementary** staining, the primary stain acts as a mordant for the secondary stain.

2.6 INTERACTIONS BETWEEN ACID AND BASIC DYES

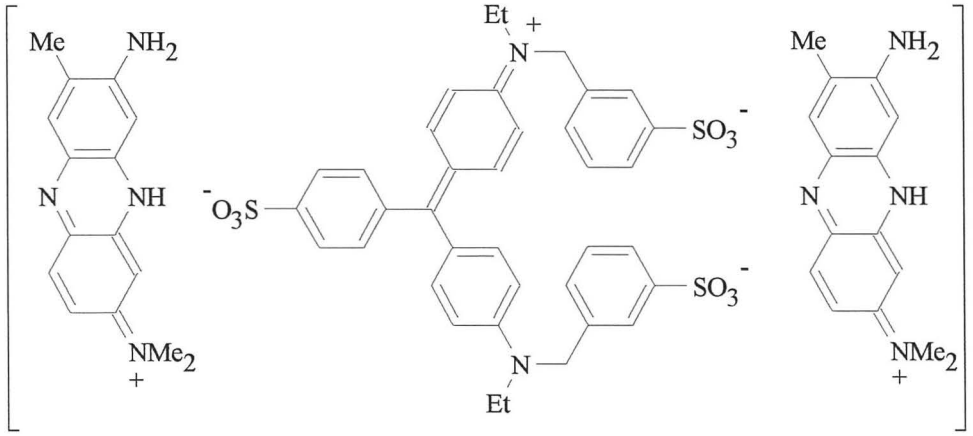
Interactions take place between acid and basic dyes with the production of polychrome dyes, which are definite chemical compounds. The Jenner stain[130,131] **(28)**, an eosinate of methylene blue, is prepared by complete precipitation from an aqueous solution of methylene blue **(11)** with eosin Y **(29)**. A similar stain, the Leishman stain[144] is an eosinate of polychrome methylene blue. The Kardos stain[145] **(30)** is formed by the reaction of orange G **(25)** with methyl green **(31)**, while the Twort stain[146] **(32)** is formed by the reaction of neutral red with light green. The Jenner, Kardos and Twort stains are related to the important Romanowsky[147] and Giemsa[148] stains, used to detect malarial parasites and **trypanosomes** in blood.



(29) (28)



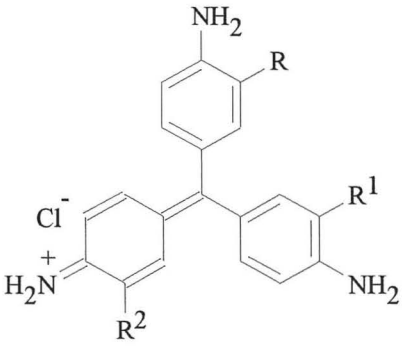
(31) (30)



(32)

As discussed in *Section 2.4*, the Falg and several other methods give a polychrome picture to basic elements of tissues, the Falg method staining the least basic tissue blue and the most basic red. The colours of the Falg method produced by the union of one molecule of acid fuchsin with one, two or three molecules of light green.

A natural sequence to this polychromatic **differentiation** of basicity would be to devise a method by which a polychromatic differentiation of acidity of tissue elements would be achieved. The Parafalg method, discussed below, stains the least basic tissue red and the most basic blue. The same colours may also be formed by the interaction of light green with pararosaniline chloride **(33)**, rosaniline chloride **(34)**, new magenta **(35)**, or basic fuchsin {a hybrid mixture of **(33)**, **(34)** and **(35)**}, since each of the four dyes have three unsubstituted amino groups. Pararosaniline chloride **(33)** was found to be the best primary stain, giving rise to the Parafalg technique for the polychromatic differentiation of acidic tissue elements.[143] Basic tissue elements which would react with an excess light green are blocked by treating with benzenesulphonic acid.



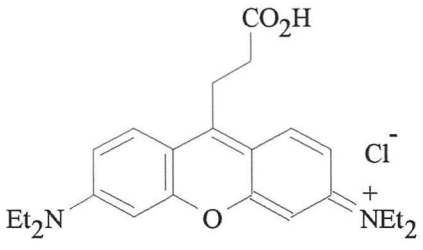
(33) $R = R^1 = R^2 = H$

(34) $R = R^1 = H$, $R^2 = Me$

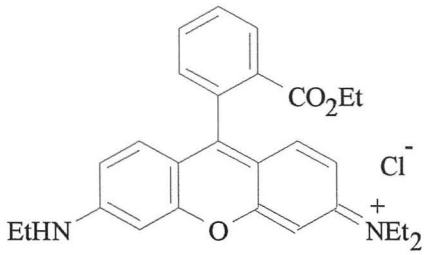
(35) $R = R^1 = R^2 = Me$

2.7 INTERACTIONS BETWEEN BASIC DYES

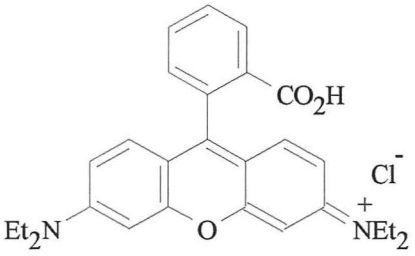
For a chemical interaction to take place between a pair of cationic dyes, one of them must be amphoteric, *i.e.* having one or more acidic side-chains capable of forming salts. Only a few of the many cationic dyes, such as rhodamine S **(36)**, rhodamine G **(37)**, rhodamine B **(38)**, and gallocyanin chloride **(39)** are definitely amphoteric.



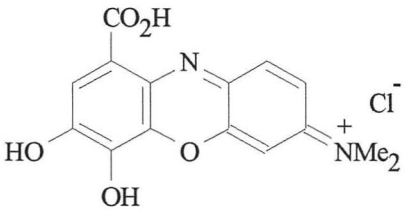
(36)



(37)



(38)

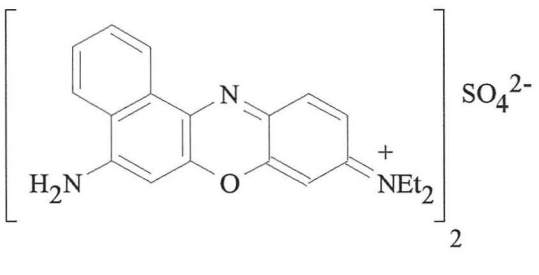


(39)

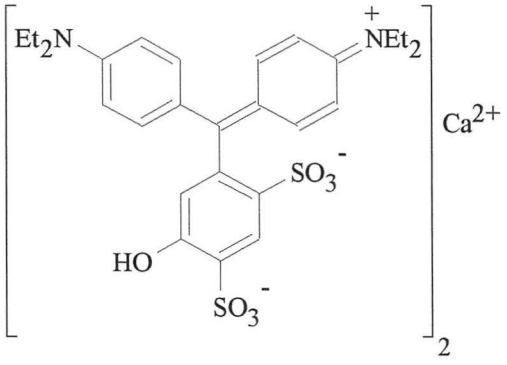
2.8 CLASSIFICATION OF SYNTHETIC DYES USED IN BIOLOGY

Dyes have been classified according to their chromophores. Many dyes however, having the same chromophores, but different types of colligators, undergo widely different staining reactions from each other. Other dyes, with the same types of colligators, but different chromophores, undergo the same staining reactions. The reclassification of organic dyes according to the types of colligators will be described, which enabled rational staining methods to be developed. Although molecular weights[131,135,149,150] of dyes such as Nile blue sulphate **(40)** and patent blue V **(41)**, which form univalent dye ions, and morphology[133,151] may be important factors, the number and kinds of colligators were, however, found to be more important.

Dyes having both acidic and basic colligators are capable of combining *in vitro* with other substances, including tissue elements, which could be either basic or acidic in reaction. Dyes containing wholly basic colligators, act *in vitro* as bases towards acidic substances, while those containing wholly acidic colligators, act *in vitro* as acids towards basic substances.

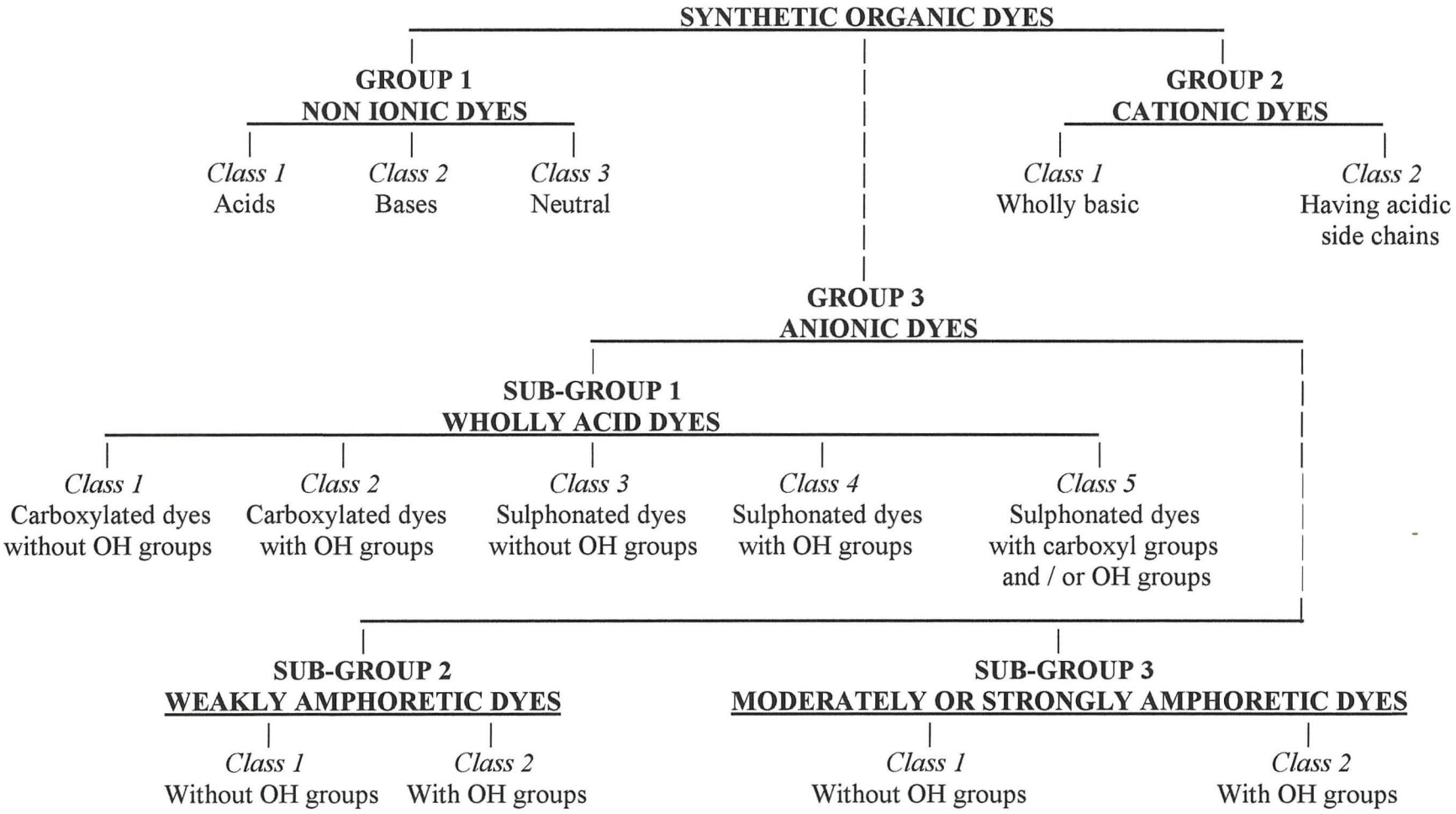


(40)



(41)

Figure 2.1 Classification of dyes according to Gurr.[101]



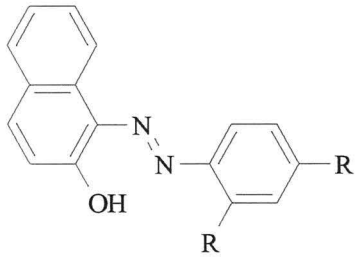
Some wholly acid dyes stain acidic, as well as basic elements of tissues, without the use of mordants. These are called pseudo-amphoteric anionic dyes, and have hydroxyl groups which enable the dye to attach to both acidic and basic tissues. Anionic dyes of this type are called acid-mordant dyes in the textile industry, and are used with metallic (*e.g.* Cr^{3+}) salts, for dyeing fabric with chrome colours.

Gurr reclassified synthetic organic dyes according to the types of colligators, as shown above in *Figure 2.1*.[101] Each dye in a class should have the same staining property, although the chromophore and colour may be different. The classes will be examined in detail in the following sections, using suitable examples.

2.9 GROUP 1 : NON-IONIC DYES

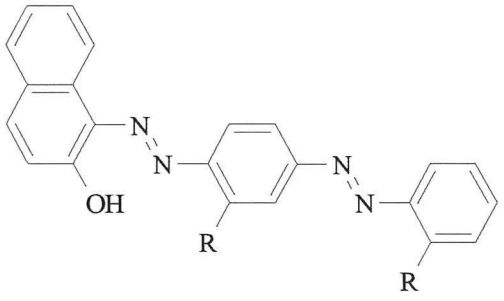
2.9.1 CLASS 1 : Non-ionic dyes with acidic side chains

Dyes of the (1:1) group, such as the mono-azo dyes, Sudan I **(42)** and Sudan II **(43)**, and the bis-azo dyes, Sudan III **(44)** and Sudan IV **(45)**, have one hydroxyl group attached to an aromatic carbon, are insoluble in water, but soluble in organic solvents.



(42) R = H

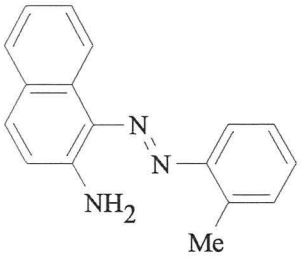
(43) R = Me



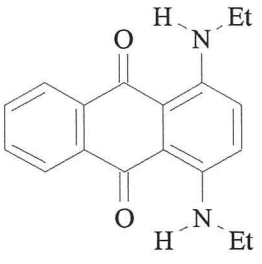
(44) R = H

(45) R = Me

2.9.2 CLASS 2 : Non-ionic dyes with basic side chains



(46)

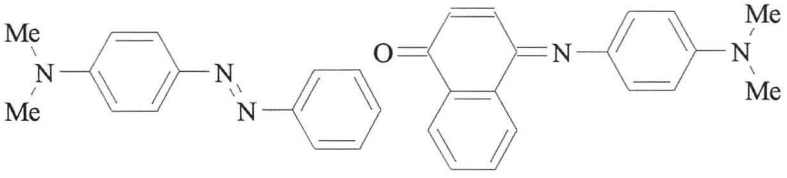


(47)

Dyes of the (1:2) group, such as oil yellow OB **(46)** and Sudan blue **(47)**, have one or more imino or amino groups attached to an aromatic carbon, are insoluble in water, but soluble in organic solvents.

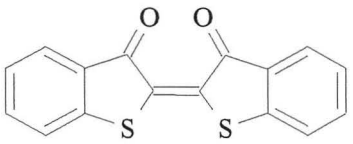
2.9.3 CLASS 3 : Non-ionic neutral dyes

Dyes of the (1:3) group, such as oil yellow D **(48)** and indophenol **(49)**, are devoid of either acid or basic groups, and are insoluble in water but soluble in organic solvents. Thioindigo red **(50)** is, however, slightly soluble in water.



(48)

(49)



(50)

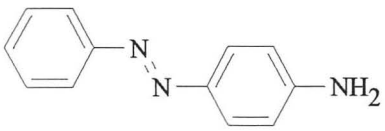
Group 1 dyes may be used for the differential staining of lipids.

2.10 GROUP 2 : CATIONIC DYES

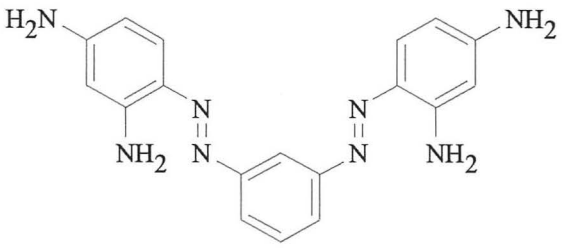
2.10.1 CLASS 1 : Wholly basic cationic dyes

All (2:1) dyes are soluble in water and a number of polar organic solvents such as glycols. Most are less soluble in water than organic solvents, and insoluble in non-polar solvents. Apart from hæmatoxylin, most nuclear stains are (2:1) dyes. There are approximately 10 anionic dyes for every known cationic dye.

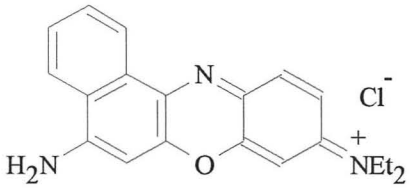
The hydrochloride salt of the mono-azo dye oil yellow B **(51)**, and the oxazine dye Nile blue chloride **(52)** are two of the few fat stains that are soluble in water. Used together they act as a differential stain for lipids and lipoproteins. The oxazine dye Meldola blue **(53)** has been used for staining lipids in bacteria.[152] The bis-hydrochloride salt of Bismarck brown Y **(54)** is a component of the Papanicolaou EA stain (*Section 3.6.1*), while the triarylmethane dye, crystal violet **(55)**, and the azine dye, safranin O **(56)**, are used in the Gram staining of bacteria.[153] A large number of (2:1) dyes contain the triarylmethane chromophore. All (2:1) dyes interact with the anionic group 3 dyes. Many (2:1) dyes will also interact with certain (2:2) dyes which have acidic side chains.



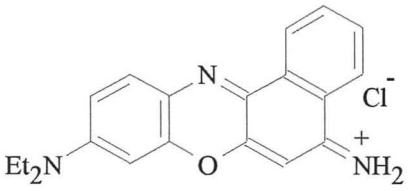
(51)



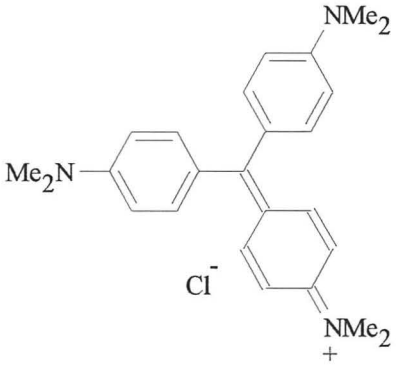
(54)



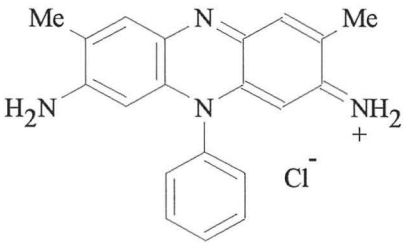
(52)



(53)

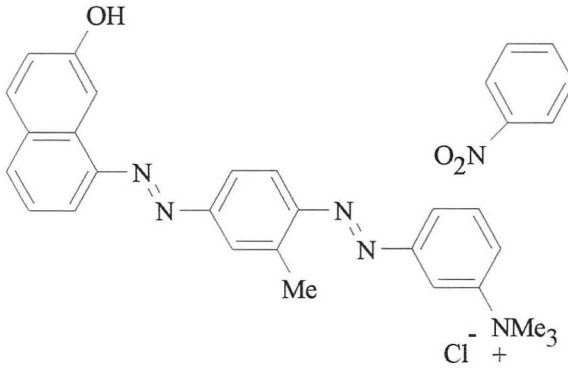


(55)

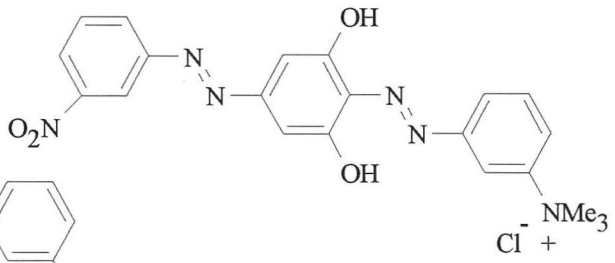


(56)

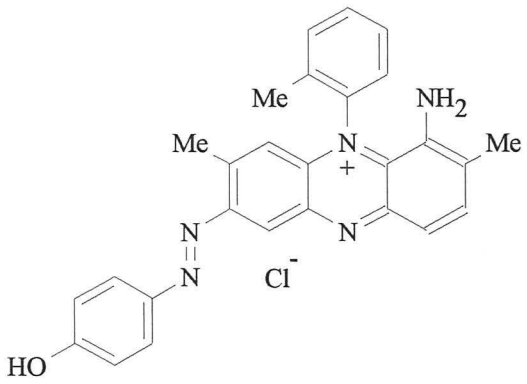
2.10.2 CLASS 2 : Basic cationic dyes with acidic side chains



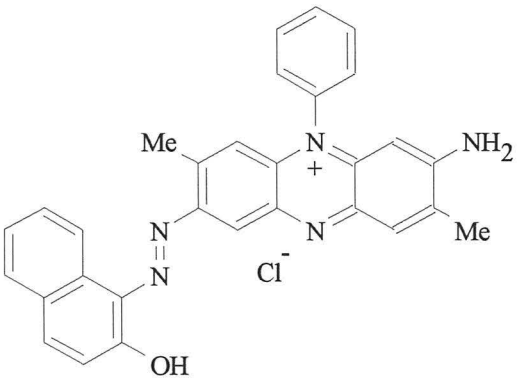
(57)



(58)



(59)



(60)

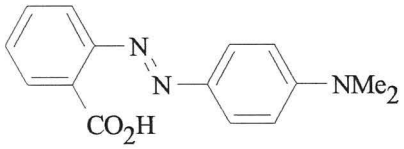
These are soluble in water and polar organic solvents such as glycols and cellosolve, the most important being the xanthene dye rhodamine S **(36)**. Other examples are Janus red **(57)**, Janus yellow **(58)**, Janus black B **(59)**, and Janus blue **(60)**.

2.11 GROUP 3 : ANIONIC DYES

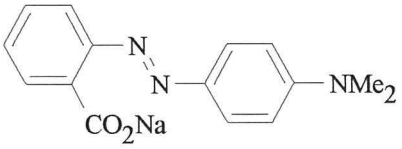
2.11.1 SUB-GROUP 1 : WHOLLY ACIDIC ANIONIC DYES

a) CLASS 1 : Carboxylated dyes without hydroxyl groups

The (3:1:1) group is very small. The mono-azo dye methyl red, alcohol soluble **(61)** and its water soluble sodium salt **(62)** are used as hydrogen-ion indicators and microscopic stains.[135]



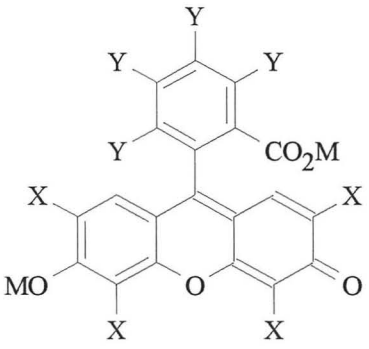
(61)

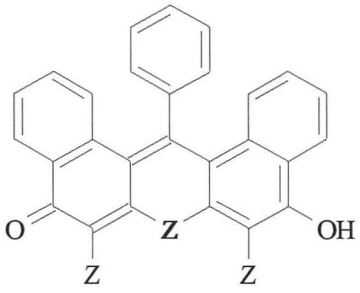


(62)

b) CLASS 2 : Carboxylated dyes with hydroxyl groups

The dyes of the (3:1:2) group are soluble in alcohol and other polar organic solvents. A few are soluble in water or non-polar solvents. When applied to animal tissues, as aqueous or alcoholic solutions without mordants, many of these dyes can be completely extracted by methanol or ethanol. Colouration by anionic dyes is probably due to physical and not chemical attachment. The majority of (3:1:2) dyes such as eosin Y **(29)**, rose Bengal 2B **(63)**, erythrosin B **(64)** and phloxine B **(65)**, are prepared by the condensation of phthalic anhydride (or acid) with phenols.





(29) M = Na, X = Br, Y = H

(63) M = K, X = I, Y = Cl

(64) M = Na, X = I, Y = H

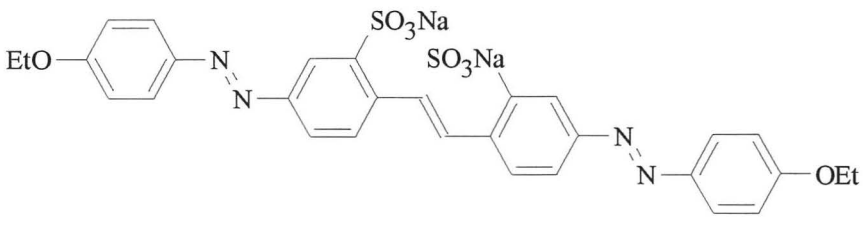
(65) M = Na, X = Br, Y = Cl

(66) Z = CO_2Na

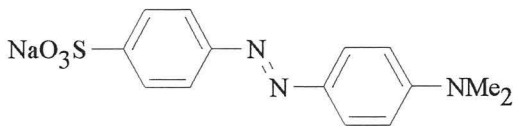
Under suitable conditions, the hydroxy-triarylmethane dyes of this groups, classed as mordant dyes, will stain both acidic and basic tissue elements polychromatically. They can also be used for demonstrating certain metals in tissues, the metals acting as mordants, *e.g.* naphthochrome green **(66)** is used for the detection of beryllium.[131,154]

c) CLASS 3 : Sulphonated dyes without hydroxyl groups

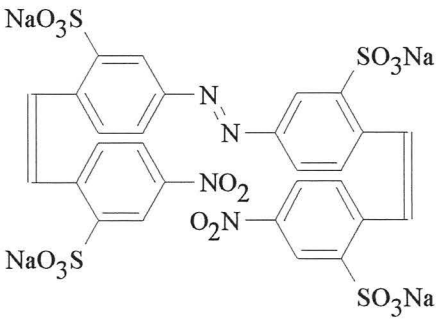
Chrysophenine **(67)** is the dye of this group readily soluble in alcohol. Methyl orange **(68)**, a counterstain and pH indicator of cell contents,[135] and paramine fast yellow **(69)** are slightly soluble, while the rest are insoluble in alcohol. The vinylene groups in chrysophenine, paramine fast yellow and sun yellow G **(24)**, cause fluorescence in UV light, making them useful as fluorochromes in fluorescence microscopy.



(67)



(68)

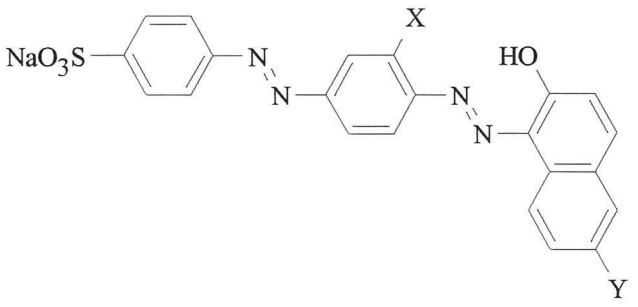


(69)

All dyes of this class interact chemically with dyes of Group 2 and with amphoteric dyes of Group 3, forming polychromatic compounds. Some members, notably sun yellow G, combine with and decolourise certain amphoteric dyes of Group 3.

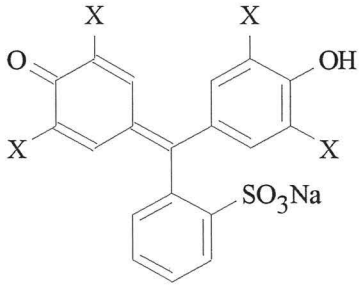
d) CLASS 4 : Sulphonated dyes with hydroxyl groups

The most useful dyes of this class are orange G **(25)**, Biebrich scarlet **(70)**, and ponceau de xylidiene **(71)**.[135] Under suitable conditions, most, if not all, (3:1:4) dyes interact *in vitro* with cationic dyes, forming water insoluble polychrome dyes. Most (3:1:4) dyes are azo dyes, the only triarylmethane (3:1:4) dyes being hydrogen-ion indicators, the sulphonaphthaleins. Phenol red **(72)** has been used as a stain for the study of carcinogenesis in white mice treated with benzopyrene and other carcinogens.[155] Bromophenol blue **(73)** has been used as a stain for basic proteins.[135]



(70) X = SO_3Na, Y = H

(71) X = H, Y = SO_3Na



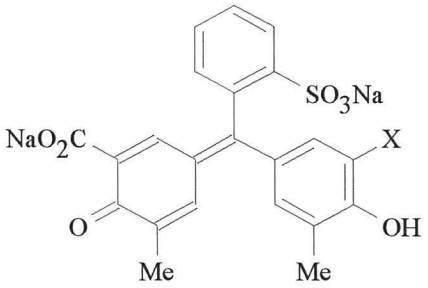
(72) X = H

(73) X = Br

Certain (3:1:4) dyes, notably orange G, erase certain other dyes such as acid fuchsin and light green SF yellowish from stained tissues, as described earlier in *Section 2.4*.

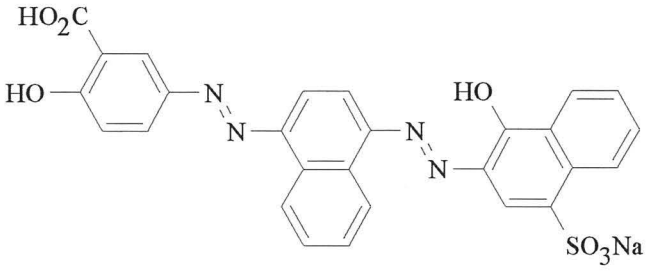
e) *CLASS 5 : Sulphonated dyes with hydroxyl and carboxyl groups*

The very small (3:1:5) group contains mordant dyes, such as eriochrome cyanin BA **(74)**, and solochrome black F **(75)**, readily soluble in water and polar organic solvents, and used are used for the demonstration of certain metal ions, *e.g.* Ca^{2+} and Al^{3+}, which act as mordants between dye and tissues. The roles of mordants are examined in *Section 2.14*. Pearse found solochrome cyanin R **(76)** to be a highly specific histochemical stain for aluminium deposits, which were coloured a deep rose-red.[156]



(74) X = SO_3Na

(76) X = CO_2Na

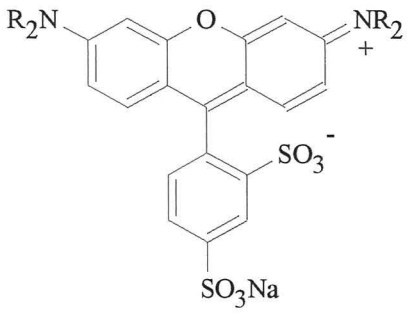


(75)

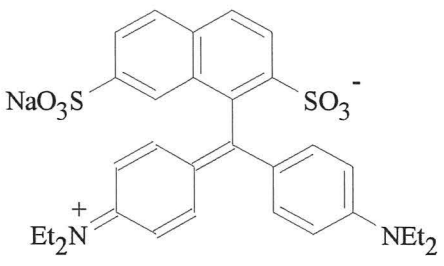
2.11.2 SUB-GROUP 2 : WEAKLY AMPHOTERIC ANIONIC DYES

a) *CLASS 1 : Sulphonated dyes without hydroxyl groups*

(3:2:1) Dyes are generally highly soluble in water and a number of polar organic solvents, but insoluble in non-polar solvents such as xylene, benzene and toluene. With their basic groups, they are able to unite chemically with certain more acidic dyes. As described in *Section 2.4*, light green SF yellowish **(13)** acts as a base towards orange G or sun yellow G, and is reduced and decolourised in the process. Other examples of (3:2:1) dyes are the xanthene dyes sulpho rhodamine **(77)** and xylene red B **(78)**, and the triarylmethane dye alkali green **(79)**.



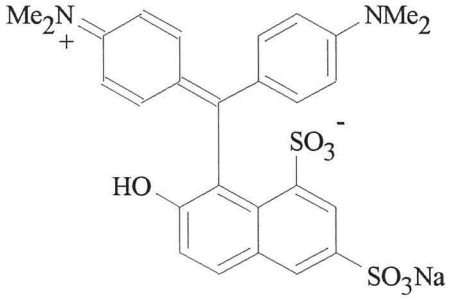
(77) R = Me **(78)** R = Et



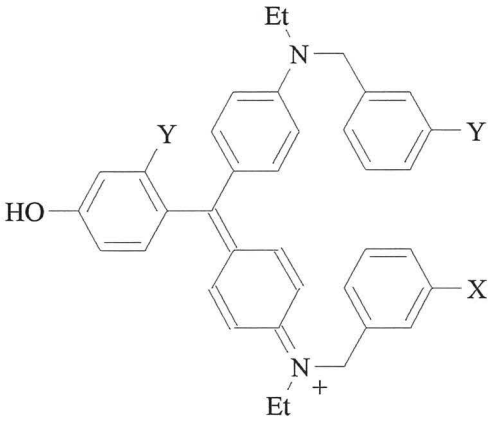
(79)

b) *CLASS 2 : Sulphonated dyes with hydroxyl groups*

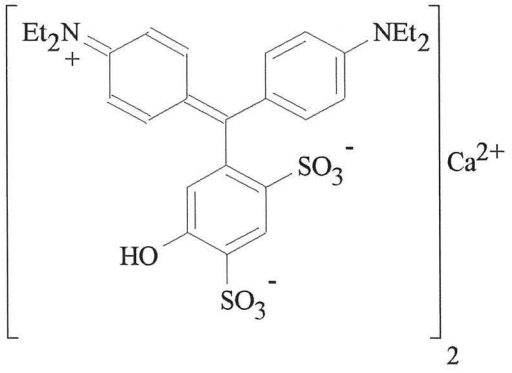
Wool green **(80)**, fast green FCF **(81)**, patent blue A **(82)** and patent blue V **(41)** are the only members of this group. Fast green FCF **(81)** which can be used in place of light green SF yellowish, is bluer in shade and less subject to fading. Gurr and MacConaill, showed that patent blue A **(82)** stained mammalian neurological tissue elements, but patent blue V **(41)** did not. Patent blue V is used as an redox indicator,[135] and is probably reduced and decolourised by neurological tissues.



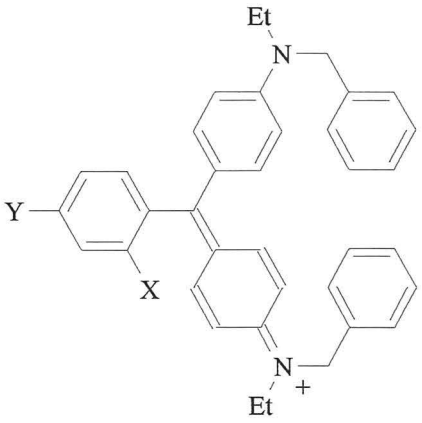
(80)



(81) X = SO_3^-, Y = SO_3Na



(41)



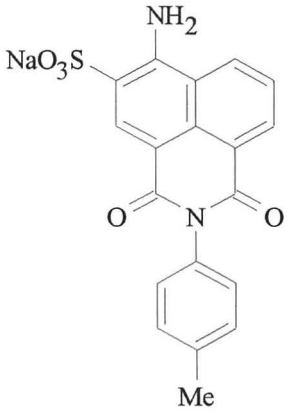
(82) X = SO_3^-, Y = SO_3Na

(3:2:2) Dyes can interact with many cationic and other anionic dyes to form compounds, which for the former are water-insoluble, while with the latter are water-soluble. Fast green FCF and light green SF yellowish, combine with safranin causing the latter dye to decolourise. Dioxane also has a decolourising action on light green SF yellowish.

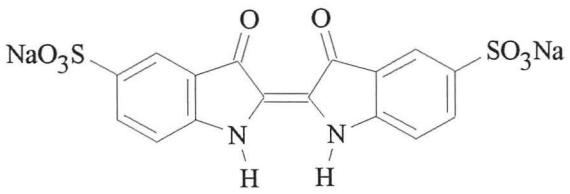
2.11.3 SUB-GROUP 3 : MODERATELY / STRONGLY AMPHOTERIC

a) CLASS 1 : Sulphonated and / or carboxylated acid dyes without hydroxyl groups but with one or more imino and / or amino groups

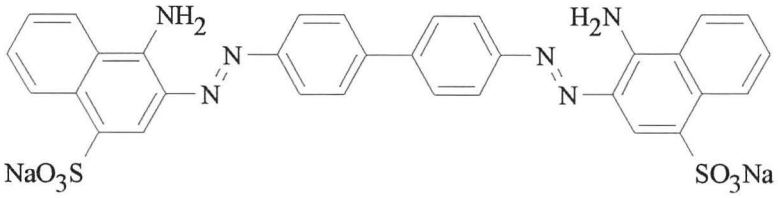
Other (3:3:1) dyes of importance, apart from acid fuchsin **(12)** and xanthene red **(23)** are brilliant sulphaflavine **(83)**, indigo carmine **(84)**, and Congo red **(85)**.[119]



(83)



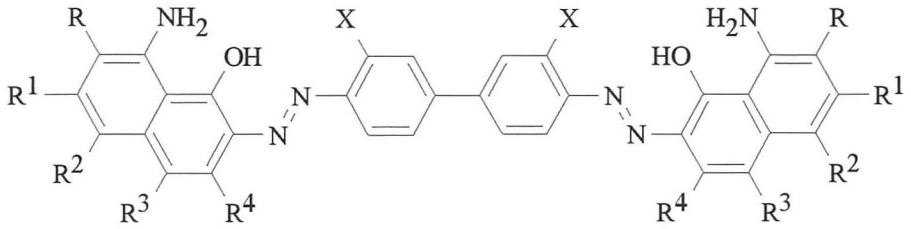
(84)



(85)

b) CLASS 2 : Sulphonated and / or carboxylated acid dyes with hydroxyl groups and with one or more imino and / or amino groups

Most (3:3:2) dyes are azo dyes, the remainder being mostly anthraquinone or triarylmethane dyes. Trypan blue **(86)**, Evans blue **(87)**, chlorazol sky blue FF **(88)**, and pontamine sky blue 5BX **(89)** are widely used as vital stains. Trypan blue **(86)** was used for the treatment of trypanosomiasis. Many of these dyes are amphoteric, not only towards other dyes, but also towards certain tissue elements. When used as simple aqueous or alcoholic solutions, chlorazol black E **(14)** and other dyes act as single dye polychromatic differential stains. Due to their hydroxyl groups, (3:3:2) dyes are less predictable in their staining reactions than (3:3:1) amphoteric dyes.

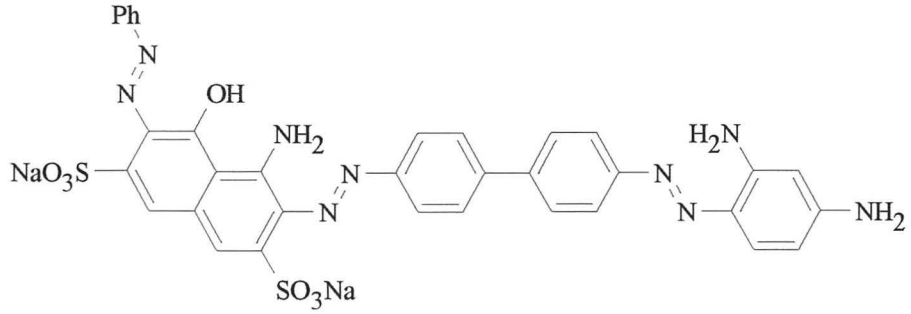


(86) X = Me, R = R^2 = R^3 = H, R^1 = R^4 = SO_3Na

(87) X = Me, R = R^2 = SO_3Na, R^1 = R^3 = R^4 = H

(88) X = OMe, R = R^2 = SO_3Na, R^1 = R^3 = R^4 = H

(89) X = OMe, R = R^2 = R^3 = H, R^1 = R^4 = SO_3Na



(14)

2.12 MOLECULAR MORPHOLOGY

Large molecules are three-dimensional, existing in many forms such as α- and β-helices, chains and sheets. The morphology of a stain is therefore important.[131,133,135,149,150]

Certain dyes are known to deeply stain collagen, but will stain keratin weakly; other dyes are known to deeply stain keratin, but will stain collagen weakly; while others stain both keratin and collagen weakly. No correlation has been shown to exist between acidity or basicity of CK dyes, and the diffusiveness of staining, below an ionised molecular weight of 350 Daltons. Between 350 and 590 Daltons, keratin is stained preferentially, and above 700 Daltons, collagen is the only one of the two types of fibre stained. The intermediate group stains both with equal intensity. In the cervical smear test, the best stains for cytoplasm are anionic dyes with a molecular weight between 350 and 590 Daltons, such as eosin Y **(29)** and orange G **(25)**. Similarly, the best stains for collagen are anionic dyes with a molecular weight above 700 Daltons, such as light green SF yellowish **(13)** or fast green FCF **(81)**.

2.13 EFFECT OF ACETIC ACID

When aqueous solutions of certain dyes are applied they impart little colour. In many cases, acidification using acetic, hydrochloric or sulphuric acid intensifies the colour. In the Falg, Falgog and Faviol techniques, acid fuchsin is used as the primary stain, applied as an aqueous solution to which 4% acetic acid has been added. After staining and washing, the sections are counterstained with an aqueous solution of light green SF yellowish. The sulphonic colligators attach the amphoteric acid fuchsin to basic tissue groups. Light green attaches itself onto tissue elements (histofuchsinates) already stained with acid fuchsin, whose amino groups allow it to act as a mordant.

In the Falgog technique, after application of acid fuchsin followed by light green SF yellowish, the preparations are washed in distilled water for five minutes. They are then immersed in 2% aqueous acetic acid solution for five minutes before being immersed in a dilute aqueous solution of orange G. The latter dye erases all the Falg colours except the red in the most strongly erythrophile elements.[157] If the protective 2% acetic acid bath is omitted, orange G erases the red also, with the strongly erythrophile tissue elements taking up orange G.

There is evidence that acetic acid added to solutions of anionic dyes has a direct effect on the dyes themselves apart from its effect on tissue elements.[133] Acid fuchsin and light green are disodium salts of trisulphonic acids. If acetic acid is added to solutions of these dyes, sodium acetate and the free dye acid are formed. The free dye acid reacts more strongly with basic tissue elements than it does as a sodium salt.

2.14 MORDANTS

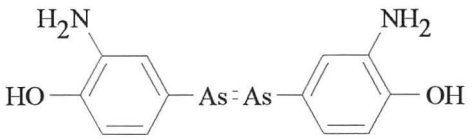
Salts of certain metals radically alter the behaviour of some dyes. These metallic salts are called **mordants** because they were thought to bite into certain textile fibres, giving attachment to dyes which would not work satisfactorily by themselves.[158,159] The great advantage of the use of mordants in microtechnique, is that once the tissue-mordant-dye complex has been formed, it is insoluble in all the neutral fluids normally used in histology, so that subsequent colouring with other dyes is easy, and there is no need for dehydration.[158]

A mordant forms an insoluble compound called a **lake**, with the dye, although some dyes form soluble lakes with certain mordants. In microtechnique, the lake is formed *in tela* as in the case of hæmatoxylin. A lake of hæmatoxylin can be differentiated by the very mordant which was responsible for its deposition in the tissues,[158] as changes in the metal cation not only affects the electronic spectrum and therefore colour of the lake, but more importantly binding strengths to organic tissues.[158]

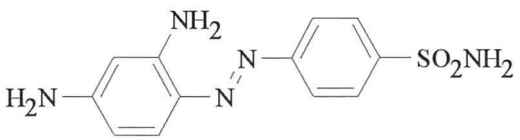
Lakes are basic in reaction, and **fast** to alcohol. The mordants used in microtechnique are salts, especially sulphates, of iron, aluminium and chromium. Lead has also been used.[157,160] The metal of a mordant will eventually be held to tissues by bonds similar to those that hold it to the dye.[158] The primary linkage is with acidic groups in tissues, *e.g.* phosphoric and carboxylate groups of nucleic acids and proteins respectively.

2.15 CHEMOTHERAPEUTIC USES OF DYES

New antibiotics, antimetabolites and enzyme inhibitors are usually screened with natural products, but can be screened with synthetic chemicals, such as dyes. Paul Ehlrich, founder and prophet of chemotherapy, discovered that synthetic dyes were absorbed differentially into tissues, and could kill parasites and bacteria without affecting mammalian cells. This led to Salvarsan **(90)** in 1910, an arsenic compound for treating syphilis.[161-164] Several chemical companies followed up this discovery by establishing their own research programmes seeking drugs to treat venereal diseases.

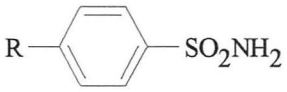


(90)



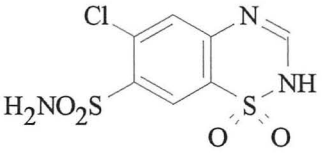
(91)

In 1931, Gerhard Domagk began screening sulphonamide derivatives of azo dyes, leading to Prontosil® Rubrum (red), sulphamidochrysoidine, **(91)**,[165-170] the first effective chemotherapeutic agent for any generalised bacterial infection. Prontosil is converted *in vivo* to sulphanilamide **(92)** and mafenide **(93)**. The subsequent development of the sulphonamide class of antibacterials, such as chlorothiazide **(94)**, led to a new class of diuretics.[171]



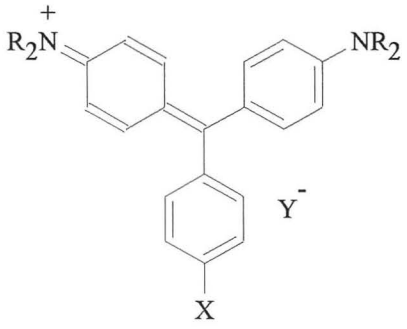
(92) R = NH_2

(93) R = CH_2NH_2



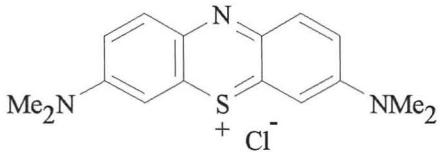
(94)

Dyes that exhibit an antibacterial action are, as a rule, **bacteriostatic** rather than **bactericidal**. Although basic dyes with their electropositive character are specific for Gram-positive micro-organisms, and electronegative acid dyes, for Gram-negative bacteria,[172] the quantitative determination of the uptake of crystal violet by bacterial cells has failed to show any correlation with their Gram-character.[173] The inhibitory action of basic dyes such as triphenylmethane and acridine derivatives, is attributed to their tendency of forming nonionising complexes with bacterial cellular constituents, probably nucleoproteins in the cytoskeleton.[174] Similarly, for acid dyes such as acid fuchsin, the anion is believed to react with a basic bacterial cell receptor, forming a nondissociating complex.[175] Bacteria become increasingly sensitive to basic dyes as the pH is increased, and to acidic dyes as it is decreased. Proceeding from the weakest members (*e.g.* 1,9-diaminoacridine) to the strongest (*e.g.* 5-amino-1-methyl-acridine) there is a progressive increase in antibacterial potency.[176] The relationship between ionisation and antibacterial activity also holds for other dyes.[177]



(95) R = Et, X = H, Y^- = HSO_4^-

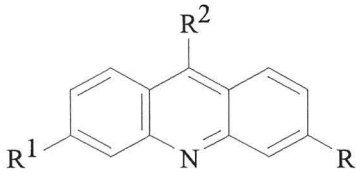
(55) R = Me, X = NMe_2, Y^- = Cl^-

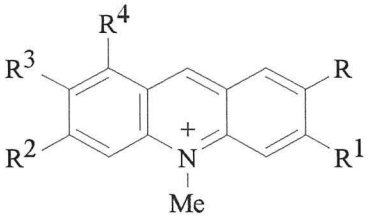


(11)

Brilliant green **(95)** shows a bactericidal capacity for *Staphylococcus aureus* in a dilution of 1:50,000. Using 10% blood broth as a reversing agent, **(95)** is not germicidal at 1:2000. Crystal violet **(55)** suppresses the growth of *Lactobacillus*

acidophilus and *Micrococcus citreus* in a dilution of $1:10^6$ or higher,[178] and is recommended as a **topical** antiseptic for the prevention of secondary infection of war wounds and burns.[179] Methylene blue **(11)** stains **plasmodia** and inhibits several pathogens including *Mycobacterium tuberculosis*.[180] Acriflavine, a mixture of **(96)** and **(97)**, originally considered for treatment of trypanosomiasis, gained importance as a **topical** antiseptic with bacteriostatic action, but has limited use as a **systemic** drug.[181] Rivanol, the lactate of 2-ethoxy-6,9-diaminoacridine **(98)**, exhibits antibacterial action *in vitro* comparing favourably with acriflavine.[182] Flavicid, **(99)**, has been used as a germicide for wound dressing, and topically for gonorrhea.[183]





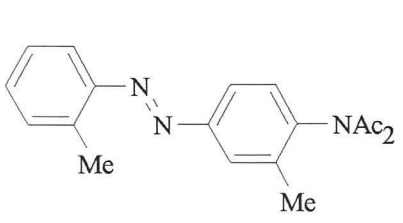
(97) $R = R^1 = NH_2$, $R^2 = H$

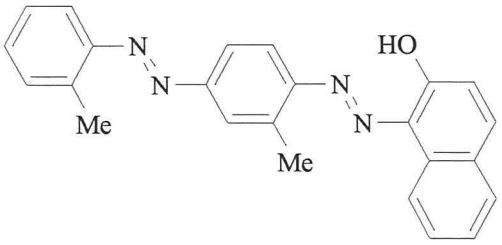
(96) $R^1 = R^2 = NH_2$, $R = R^3 = R^4 = H$

(98) $R = OEt$, $R^1 = R^2 = NH_2$

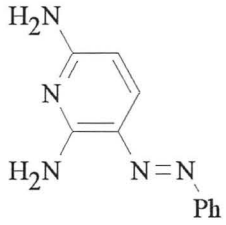
(99) $R = NMe_2$, $R^1 = R^3 = Me$, $R^2 = H$, $R^4 = NH_2$

Dimazon, diacetamidoazotoluene **(100)**, is chemically related to Sudan IV **(45)**. It is applied in powder or ointment form to **denuded** areas as a means of controlling bacterial multiplication and of stimulating epithelisation. Phenazopyridine hydrochloride **(101)**, is used primarily as a urinary antiseptic because of its ability to exert bacteriostatic action in both acidic and alkaline urine. It is inhibitory for *staphylococci*, *streptococci*, *gonococci*, and *coli bacilli*.[184] Salicylazosulphapyridine, **(102)**, is a dye of the sulphonamide series.[185]

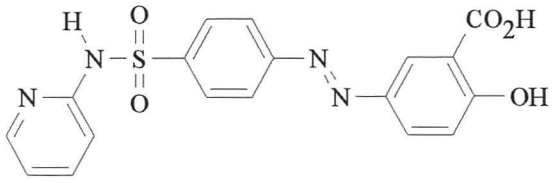




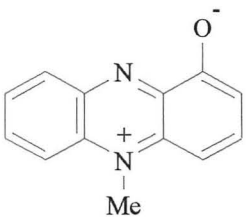
(100) **(45)**



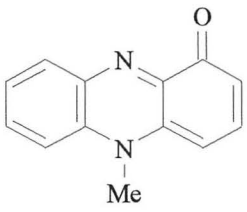
(101)



(102)



(103a)



(103b)

The pigments produced by some **chromogenic** bacteria are inhibitory for other species. The antibiotic effect upon *staphylococci* and *streptococci* of pyocyanin **(103)**, the pigment of *Pseudomonas aeruginosa*, has stimulated a systematic investigation of the series of phenazines to which pyocyanin belongs.[186] Several of these dyes exhibit bacteriostatic effectiveness in the dilution range of $1:10^3$ to $1:10^4$ against *Staphylococcus aureus*. Much higher dye concentrations (~5%) were needed for *Escherichia coli* or *Proteus vulgaris*. The success of these discoveries in chemotherapy dominated research thinking in the phamaceutical industry for many years. Screening procedures were also established for non-infectious diseases, such as anticonvulsants for the treatment of epilepsy, non-addictive analgesics, antihypertensives, anti-inflammatory, and antiulcer drugs.

2.16 SUMMARY

The development of stains and staining techniques, from natural pigments to synthetic dyes, and the interactions of dyes with each other to form polychromatic compound dyes have been examined. Gurr's reclassification of dyes enabled histologists to develop staining methods in a rational manner. The effects of molecular weight, acidification, and mordanting on staining reactions have also been discussed. The next chapter will concentrate on the use of staining methods in the detection and prevention of cervical cancer, without which, thousands of British women would needlessly die each year from the advance stages of cervical cancer.

CHAPTER 3
CERVICAL CANCER SCREENING

The general reactions of stains have been discussed in *Chapter 2*. The principal aim of this thesis is to calculate the wavelength maxima and therefore colour, of the cytological stains used in the detection of cervical cancer. The empirical cytochemical methods used in the detection of cervical cancer have not changed since the 1940s when Papanicolaou and Shorr published their staining methods. As the stains are conglomerations of alchemic ingredients, it should be possible to design more selective stains by analysing stain-cell interactions.

The cellular evolution, epidemiology, causes, screening and treatment of cervical cancer will be examined. The advantages and disadvantages of cervical cancer screening using smears, colposcopy and biopsies will be discussed. Finally, the Papanicolaou, Shorr and other cervical screening methods will be reviewed.

3.1 THE CERVIX

Cancer occurs when cells have advanced from a pre-cancerous state into one where they grow uncontrollably, penetrating and invading surrounding tissues, forming malignant tumours. The cells can break off and **metastasise** *via* the lymphatic channels to other organs. Although cervical cancer can be detected at an early stage, it is the second most frequent cancer in women world-wide, accounting for ~2% of all cancer deaths in the West, and is the most frequent cancer in developing countries.[187]

There are three types of skin in the genital tract, each designed to fulfil a specific function. The vagina and the part of the cervix in the vagina are lined with skin called **squamous epithelium**. The passage from the cervix to the uterus, the **endocervical canal**, is lined with a completely different kind of skin called **endocervical epithelium**, of columnar cells. This lining is a single cell layer thick and is responsible for producing important mucus secretions.

Being at the junction of the vagina and uterus, the arrangement of the cells in the surface epithelium of the cervix changes.[188] Externally, as in the rest of the vagina, the

cells form a many-layered **stratified epithelium**, in which the cells, as they are pushed to the surface, become flattened, lose their nuclei, and are shed like scales. At the immediate entrance to the uterus, the arrangement changes to one in which the cells are elongated and form columns lying at right angles to the plane of the surface. This **columnar epithelium** is usually only one or two cells deep. The cells at the surface do not undergo the rapid, programmed cell death occurring in the neighbouring stratified epithelium.

The point at which the two different types of skin lining the vagina and cervix meet, is called the **squamocolumnar junction**. It is here that the initial pre-cancerous changes begin. The area on the post-adolescent cervix covered by migrated columnar epithelium from the cervical canal is called the **transformation zone**. The acidic vaginal fluid breaks down the migrated columnar epithelium, a process called **erosion**. The squamous epithelium grows beneath the columnar epithelium, gradually replaces it – the replacement known as **squamous metaplasia**.

The squamocolumnar junction and the transformation zone are the most common sites at which pre-cancerous and cancerous changes originate. These areas are the most vulnerable to infection, inflammation and possible cancer induction, especially before metaplasia is complete. Cancer of the cervix originates in the stratified epithelium, usually very close to its junction with the columnar epithelium of the cervical canal.[189] It is not known why it should be confined to that particular region of the epithelium, but it makes it easier to study the changes that precede the appearance of cancer. Cells can be scraped from the epithelium and observed microscopically - the Papanicolaou "Pap" smear test. Quite large samples of the epithelium can also be excised and studied microscopically or cultured *in vitro* so that the program of cell renewal can be observed. The intact surface can also be observed directly.

Many abnormalities have been described as a result of mass routine screening, ranging from minor variations in the developmental pattern of the epithelium, through more severe changes, carcinoma *in situ*, to invasive metastasising cancer.[190] Minor changes, such as abnormal thickening of the epithelium, or some drifting in the boundary between columnar and stratified epithelium are common. **Dysplasia** (*lit.* false structuring), occurs when the cells vary in size and arrangement in a rather

haphazard fashion from one area to another, and dividing cells are no longer confined to the deepest layer, but are scattered throughout most layers of the epithelium,[191,192] as if they are no longer receiving precise positional information to multiply and differentiate in a completely ordered manner.

The next level of abnormality, **carcinoma *in situ***, occurs where the signs of differentiation may be minimal, and the only microscopic feature distinguishing the situation from a fully developed cancer is that the dividing line between epithelium and underlying connective tissue is still precisely maintained. When the abnormal epithelium is seen to be invading the underlying tissues, the condition is classified as a cancer, irrespective of its size or of the detailed appearance of its cells. It is these locally invasive cancer cell populations that are capable of entering lymph and blood vessels and thereby undergoing metastasis to distant sites.

Both dysplasia and carcinoma *in situ* are sometimes followed by the development of a fully invasive cancer. Judging from surveys of women of all ages, the timing is as follows:[193-195] Dysplasia tends to arise at about the age of 30 and persists for 10 to 20 years. Carcinoma *in situ* arises rather later and persists for five to 10 years. Finally invasive cancer does not reach maximum incidence until after the age of 60. It is usually present and detectable by screening for one to four years before it starts to produce symptoms. By studying the age distribution of the various abnormal states and from observing what happens to women who are not treated,[196,197] dysplasia and carcinoma *in situ* are common and usually undergo spontaneous regression. Fully invasive cancers can sometimes emerge after little or no warning.[198]

3.1.1 Cervical pre-cancer (CIN)

The term "pre-cancer" covers a range of abnormalities which occur on or in the skin of the cervix, but which are still confined in the skin. The general term used to cover the whole spectrum of change is **cervical intraepithelial neoplasia** (CIN). CIN is a histological description of the disturbance within a cell, the cell arrangement and the distribution of cells within the cervix tissue. Pre-cancerous cells are chemically, structurally, and functionally different from other normal cells in the cervical skin. As long as they remain confined to the cervical skin and do not spread, the cells are pre-cancerous. Changes in nuclear appearance are described as mild (CIN 1), moderate

(CIN 2), or severe **dyskaryosis** (CIN 3). Mild dyskaryotic (pre-cancerous) changes may be reversible. Removal of all dyskaryotic cells should halt the progress of cancer. While the pathological diagnosis of invasive carcinoma is straightforward, the cytological and histological diagnosis of CIN can be difficult, *e.g.* distinguishing CIN from sub-clinical human papilloma virus (HPV) infection by morphology only. The Bethesda System of grouping **squamous intraepithelial lesions** (SIL) in two categories has been proposed - low-grade SIL (CIN-I), and high-grade SIL (carcinoma *in situ* or CIN-II and CIN-III).

3.1.2 Cervical adenocarcinoma

Up to 10% of cervical cancers are **adenocarcinomas**. **Glandular neoplasia**, occurs in the glands of the columnar epithelium, deep in the endocervical canal. It is harder to detect, being beyond the reach of the cervical smear, and exists in virginal women.

3.2 EPIDEMIOLOGY

The epidemiological evidence for spread of cancer from one human being to another - "horizontal" transmission, as opposed to "vertical" transmission of viruses from parent to child, is rather scant. There is some evidence of horizontal transmission in cancer of the cervix. Cervical cancer is virtually unknown in nuns and is most common in women who have had several sexual partners.[199] Its incidence is higher in the wives of previously married men if a former wife had cancer of the cervix.[200] There is evidence therefore that it is a venereally transmitted disease, perhaps due to infection by one of the *herpes*-viruses[201] such as *the Herpes simplex type II* virus.[202]

3.2.1 Incidence

Very high incidence rates are found in Third World countries. Relatively high rates are found in eastern, western and northern Europe, and in Black and Hispanic groups in the USA. Low rates are observed in Causasian populations in North America, Southern Europe, the Middle East and Australia. The lowest rates are found in Finland, Israel, Kuwait, Ireland and Spain.[203]

3.2.2 Time trends

Incidence and mortality rates steadily declined in most populations during the last half century partly due to increased use of cervical and cytology screening.[204]

3.2.3 Age distribution

Invasive cancer is very rare before 20 years of age. Rates increase progressively with age until the fifth decade, after which they either decrease or level off.[203]

3.2.4 Etiological inferences

The very low rates in North America and in European Jewish women, and the lower rates in Moslem as compared to Hindu women are consistent with circumcision. Low rates in Finland are probably the result of population-wide screening over the last 30 years. Rates in Israel, Kuwait, Ireland and Spain are consistent with low exposures to a putative sexually transmitted agent, due to conservative sexual behavioural habits.

3.3 CAUSES

3.3.1 Sexual behaviour

Only female sexual behaviour has been associated with cervical cancer, but epidemiological studies indicate that male sexual behaviour also increases the risk of cervical cancer.[205] Geographic clusters of high rates of both cancers of the cervix and penis have been described,[206] while increased risk of cervical cancer has been reported both among wives of men with penile cancer[207] and among second wives of men previously married to women with cervical cancer.[200] A small case-control study of monogamous women with CIN lesions showed an eight-fold increased risk for women whose husbands had 15 or more sexual partners outside marriage.[205] Recent results from studies in Latin America suggest that the contribution of male sexual behaviour to the risk of cervical cancer in their wives is less than suggested in earlier studies.[208] A joint study by the Imperial Cancer Research Fund, the University of Oxford, and the John Radcliffe Infirmary in Oxford showed that the relative risk of pre-cancer was higher the younger the age of first intercourse.[209]

3.3.2 Reproductive factors

Early epidemiological studies revealed associations of cervical cancer with age at first pregnancy and number of pregnancies. Studies have shown a significant risk for cervical cancer with the number of pregnancies, even after adjusting for possible confounders such as age at first sexual intercourse or number of sexual partners.[210,211] It has been postulated that the transient state of immunosuppression which occurs during pregnancy facilitates the oncogenic role of an infectious agent.

3.3.3 Sexually transmitted infectious agents

Various sexually transmitted infections[204,211] have long been described as associated with cervical cancer, including *syphilis, trichomonas vaginalis, candida albicans, chlamydia trachomatis, herpes simplex virus type 2* (HSV-2), *cytomegalovirus* (CMV) *and human papillomavirus* (HPV).[212] HSV-2 has been the most extensively studied. Infection by *chlamydia trachomatis* was found in 27% of the same group.[213]

a) HSV-2

Neutralising antibodies to HSV-2 are more frequent among cervical cancer cases than among controls. Sensitive **serological** tests for antibodies to type-specific glycoproteins of HSV-1 and HSV-2 have been developed.[214] The prevalence of HSV-2 infection has been associated with the recent increase in CIN in Australia.[215] The prevalence of HSV-2 antibodies is higher in Greenland than in Denmark, as is cervical cancer.[216] Laboratory studies do not provide clear-cut molecular, immunological or experimental evidence supporting an oncogenic effect of HSV-2.[217] Hit and run effects[218] or synergistic effects with HPV[219-221] have been postulated.

b) HPV

Although certain types of HPV (16,18,31,33,35,39,41-45,51-56), in particular, HPV 16 and 18 are associated with an increased risk of cervical cancer, most studies do not completely satisfy basic epidemiological criteria for causality. While the experimental data suggest an oncogenic potential for HPV, the human data is limited.[212] Allerding[199] showed that **koilocytes**, cells whose nucleus appears to be surrounded by a halo of cytoplasm, indicating the presence of HPV, were found in 25% of women who had carcinoma *in situ* or invasive cancer.

Molecular cloning of HPV-DNA has led to the identification of more than 60 types of HPV which are being studied in cancerous and normal cervical tissue.[222] HPV 16 and to a lesser extent, HPV 18 are associated with advanced CIN or invasive carcinoma, while HPV 6 and 11 are more often found in lower grade CIN and **condylomata**.[223] Progression to more advanced CIN was greater with HPV types 16 or 18 infection than in types 6 or 11.[224,225] HPV 6 and 11, were found to be equally present in lesions that progressed to carcinoma *in situ*.[226] The prevalence of HPV 16/18 was 1.5 times higher in Denmark, although the incidence of cervical cancer was six-fold lower in

Greenland, in contrast to HSV-2 infection.[227] Different results were obtained from Recife and Sao Paulo where cervical cancer rates correlated with HPV prevalence.[228]

Little information exists on the prevalence and role of other HPV types, which have also been observed occasionally with cervical neoplasia. A few serological case-control studies and hybridisation experiments searching for CMV-DNA in cervical cancer cells have yielded negative or inconclusive results.[229,230]

3.3.4 Oral contraceptives

The relationship between oral contraceptives and CIN has been extensively studied with conflicting results. Oral contraceptive use is highly correlated with other key risk factors, such as the number of partners, age at first intercourse, and screening history. Some case-control studies show a moderate increase in risk for squamous cell and adenocarcinoma of the cervix[231] but others do not.[232-234]

3.3.5 Barrier contraceptives

A protective effect of barrier contraceptives has been noted in some studies.[232]

3.3.6 Genital hygiene

Although poor genital hygiene has been associated with an increased risk of cervical cancer in Indian women,[235] no consistent relationship has been found in the USA, nor with vaginal douching.[231,232]

3.3.7 Male circumcision

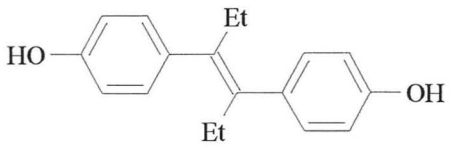
Low rates in Jewish women have been ascribed to early circumcision in males. In populations where males are circumcised around puberty, *e.g.* Moslem and African, cervical cancer rates tend to be lower. Although circumcision did not affect the risk of cervical cancer in monogamous Latin American women, poor genital hygiene did.[208]

3.3.8 Tobacco smoking

Epidemiological studies investigating the role of tobacco smoking in cervical carcinogenesis have yielded conflicting results.[232,236-238]

3.3.9 Vaginal adenocarcinoma

Tumours of the vagina are exceedingly rare. In the 1970s a small epidemic of clear cell adenocarcinomas was reported in young females aged 17 to 19 years. These arose in the offspring of women treated during pregnancy with large doses of diethylstilbestrol (DES) **(104)** to prevent spontaneous abortion.[239,240] The treatment proved to be ineffective and was discontinued. The tumours have essentially disappeared. The association between reproductive factors and adenocarcinoma of the cervix is not clear.[231,233] Minor developmental abnormalities of the vagina, possibly relevant to pathogenesis, have also been found in offspring without cancer.[241]



(104)

3.3.10 Other factors

An increased risk associated with primary and acquired immunodeficiencies has been suggested, as has a protective effect of β-carotene or other unidentified components of a vegetable rich diet, but the evidence is inconclusive. Women most willing to be screened tend to be better educated, and have lower incidences of cervical cancer.[242]

3.4 SCREENING

Cervical cancer can exist in a woman long before it gives rise to any symptoms. The cervical smear is the chief method of detecting abnormalities such as pre-cancerous and cancerous change and is 70-90% accurate; 60-90% of women who die of cervical cancer never have had a smear. Pelvic examination at the time of screening can also detect other abnormalities such as ovarian and uterine cancers.

3.4.1 The Cervical Smear

The modern detection of cervical cancer was developed by the father of cytology, George Nicholaus Papanicolaou. Through his study of vaginal and cervical cells, he showed that there were cyclical changes in nuclear shape and size, and also changes in the way the cells were stained by various dyes. The first paper was published in

1925, and presented at a conference in Battlecreek, MI in 1928. In the early 1940s, Papanicolaou and Traut published a series of papers, discussed in *Section 3.6*, on detecting cervical carcinomas. In 1943, Papanicolaou showed that cervical cancer could be detected in women exhibiting symptoms. Professor Sir Dugald Baird pioneered cervical smear testing in Britain in the 1950s.

3.4.2 Histology of the cervical smear

Once the cells have been collected, they are stained with various dyes and observed under a microscope by a cytologist. Pre-cancerous cells are **hyperchromatic**, staining darker, and show a different texture from normal cells.[243]

Two kinds of cytological investigations are available. In the **static procedure**, the stained cells are examined under a microscope. In **flow cytometry**, the cells are stained with special dyes and allowed to flow through a glass chamber, causing cancer cells to fluoresce under the presence of laser light. The automated measurement of fluorescent light indicates the presence of cancer protein. New staining techniques to stain viral and cancer proteins in cells selectively are being developed.

Problems can occur with inadequate samples with no endocervical cells, or poor fixing, which causes distortion of cell shape, making it impossible to analyse correctly. The cells may also be obscured by discharges containing blood, semen, bacteria or fungi. Infection and inflammation can cause the appearance of abnormal cells, which are difficult to distinguish from possible pre-cancerous ones. Human and administrative errors also occur, giving rise to false positive or negative smears. False positive smears indicating abnormalities, which do not exist on rechecking, are less common than false negative smears where existing abnormalities are not detected. The misdiagnosis of smears at the Kent and Canterbury Hospital Trust from 1990 to 1995 resulted in a review of 91,000 smears.

3.4.3 Colposcopy

Colposcopy has only recently replaced cervical cone biopsy as the routine diagnostic procedure for all women with severely abnormal smears. The colposcope, invented by Hinselman in 1925, is an illuminated binocular microscope, giving a magnified view of the cervix when a smear or case history suggests the presence of an abnormality.

During a colposcopic examination the vulvovaginal area is first inspected for abnormalities such as genital warts. A closed speculum is then inserted into the vagina, and the inside walls inspected. Cervical smears and high vaginal swabs can then be taken. A saline solution is applied which improves inspection of blood vessels on the cervix, giving information on the possible presence of disease. The cervix is then painted with a 3 to 5% acetic acid solution, which brings out the abnormal **aceto-white** areas as white patches. Lugol's iodine[244-246] can also be applied to the cervix, and stains all the normal squamous cervical skin and vaginal epithelium, which contain glycogen, deeply brown. This procedure is known as Schiller's test. Abnormal areas on the cervix, normal endocervical epithelium, immature metaplasia and old and ageing areas do not contain glycogen and appear pale yellow or white.

The most recent technique for visualisation of the cervix is cervicography, where photographs of the cervix are taken using a specialised camera. Again, the cervix is painted with acetic acid. Time is allowed for abnormal areas to become visible.

3.4.4 Biopsy

a) Punch biopsy

Punch biopsies are very small amounts of tissue taken from abnormal areas, sent for histological examination to identify the presence, distribution and depth of the abnormality observed by colposcopy.

b) Cone biopsy

For cases such as adenocarcinoma of the cervix, where the cancer occurs in the folds of the columnar epithelium, colposcopy must be followed by a cone biopsy. Cone biopsies are also carried out when:

i) a positive smear result occurs, and no abnormality is seen on colposcopy,
ii) the abnormality disappears into the endocervical canal and the top of it cannot be seen by colposcopy,
iii) the biopsy has shown the presence of a possible (micro)invasive cancer.

A cone biopsy involves the removal of a cone-shaped piece of tissue from the cervix to determine the source and depth of the abnormality, using either a cold knife or laser

treatment, which is preferred as it causes cauterisation. The area to be excised is identified by colposcopic visualisation and staining of the cervix with iodine.

c) Loop excision

Diathermy loop excision (DLE) or large loop excision of the transformation zone (LLETZ) has grown in popularity in the last few years and may replace punch and cone biopsies, as a way of taking samples for histological examination. The technique involves an electrically heated wire loop to remove a section of tissue from the surface of the cervix, it taking a larger sample of tissue than a punch biopsy.

3.5 TREATMENT

3.5.1 Treatment of cervical pre-cancer (CIN)

CIN can be treated locally. Local ablative treatment (LAT) is very successful in eradicating the disease whilst preserving the structure and function of the cervix. Treatment involves removing the abnormal tissue by one of the following methods:

a) Cryocautery

Cryocautery, recommended for small CIN1 or 2 lesions, is performed on an outpatient basis and takes as little as fifteen minutes to complete, with first treatment success rates of 80-90 %. A colposcope determines the extent of the diseased area, which is then frozen for a few minutes using CO_2 or N_2O.

b) Electrodiathermy

Electrodiathermy involves destruction of the abnormal area under general anaesthetic, using electrically generated heat (1000 °C) applied through needle and ball electrodes. Diathermy is highly effective in eradicating all varieties of CIN, with up to 97% of patients needing only one treatment. Neighbouring normal tissue can be damaged.

c) Cold coagulation

A probe through which heat (100 °C) is generated is applied to the cervix and the abnormal area destroyed. General anaesthesia is not required and a 95% cure rate for CIN 3 lesions using cold coagulation is reported.

d) Laser evaporation

Lasers were first used for the treatment of CIN in Britain in August 1977 at the Birmingham and Midland Hospital for Women. The entire affected area and the whole of the transformation zone on the cervix is vaporised and is not available for future histological reference. As the laser beam is thin, minimal damage is done to neighbouring tissues. The laser also seals the local tissues, reducing bleeding.

e) Diathermy loop excision (LLETZ)

Loop excision of CIN from the cervix is cheap and can be performed at the same time as colposcopy. Preserved tissue is removed for histological examination.

3.5.2 Treatment of microinvasive and invasive cancer

If CIN 3 or carcinoma *in situ* pass undetected, 20-30% of cases progress into invasive cancer, occurring when cancerous cells have not only penetrated and broken through the deepest layer of the skin, but may also have entered the lymphatic system. The penetration, extent, and metastasis of invasive cancer in the body, are described by the International Federation of Obstetrics and Gynæcology:

Table 3.1 Staging classification of cervical pre-cancer and cancer.

Stage 0	CIN3 (carcinoma *in situ*) – cancer is still confined to cervical skin.
Stage Ia	Microinvasive carcinoma – cancerous cells have broken through basement membrane but not entered lymphatic or blood channels.
Stage Ib	Occult invasive carcinoma – cancerous cells have broken beyond 3 mm from the basement membrane – lymphatic channels may be involved.
Stage IIa	Invasion into upper part of the vagina.
Stage IIb	Invasion into tissue around the cervix - the parametrium.
Stage III	Invasions into lower vagina and pelvic wall.
Stage IV	Invasion beyond genital area into nearby organs.

Stages 0, 1a and 1b can be determined using cytology, punch and cone biopsies. Once the cancer has spread beyond the basement membrane, 90% of invasive carcinoma cases are of a clinically recognisable type. Bleeding and ulceration are also diagnostic tools. Lower abdominal pain can result if the tumour extends into the uterosacral

ligaments, or from lymph nodes that adhere to the sacral plexus. Death is most commonly due to uraemia, pyelonephritis, haemorrhage or effects of metastases.

Examination under anaesthetic (EUA) involves a dilatation and curettage (D and C) together with cystoscopy or proctoscopy if necessary. If the cancer has spread into the lymphatic channels then several techniques can be used. An intravenous urogram involves injecting a radio-opaque dye into a vein in the arm, which concentrates in the urinary system where X-rays can detect tumours. A lymphogram involves injecting radio-opaque dye *via* lymph channels in the feet, and X-raying the lower torso to highlight tumours. Computerised Axial Tomography (CAT) using a radio-opaque dye, involves taking X-rays at various positions, building a 3-D image of the body. Magnetic Resonance Imaging (MRI), an offshoot of NMR techniques is similar to CAT, but is very expensive and requires trained personnel. NMR and CAT scans cannot identify tumours less than 2 cm in diameter. Ultrasound scans are not as accurate and reliable as CAT and NMR, but are widely available, safe and quick.

At the beginning of the twentieth century, cervical cancer was considered incurable, until Enrst Wertheim designed a surgical procedure, which removed the uterus and surrounding lymph glands. Microinvasive carcinoma is treated either by conisation or hysterectomy. Conisation is performed when the cancer does not extend beyond 3 mm beneath the basement membrane. In the early stages of invasive cancer when the disease is confined to the womb (Stage Ib), or has encroached into the top of the vagina (Stage IIa), a radical hysterectomy is recommended. As the ovaries are also removed, to diminish the risk of potential ovarian cancer, long-term hormone replacement therapy (HRT) may be required to counteract side effects such as bone loss. Radiotherapy involves either internal treatment with small tubes or packets containing radioactive caesium or cobalt, or external radiotherapy where a beam of high-energy X-rays or gamma rays is directed at tumours. Treatment of Stage Ib by radiation or surgery has roughly identical results, with ~85% of patients surviving for five years. If the lymph nodes are clear from any cancerous cells, there is only a 10% risk of recurrence. In the 12% of women who have nodal invasion by the malignant cells, their five year survival drops drastically to about 20% if five or more nodes are involved, or 50% for four nodes or less. For women with Stage IIb, III and IV, radiotherapy is usually the treatment of choice. For women with Stage II, the five-

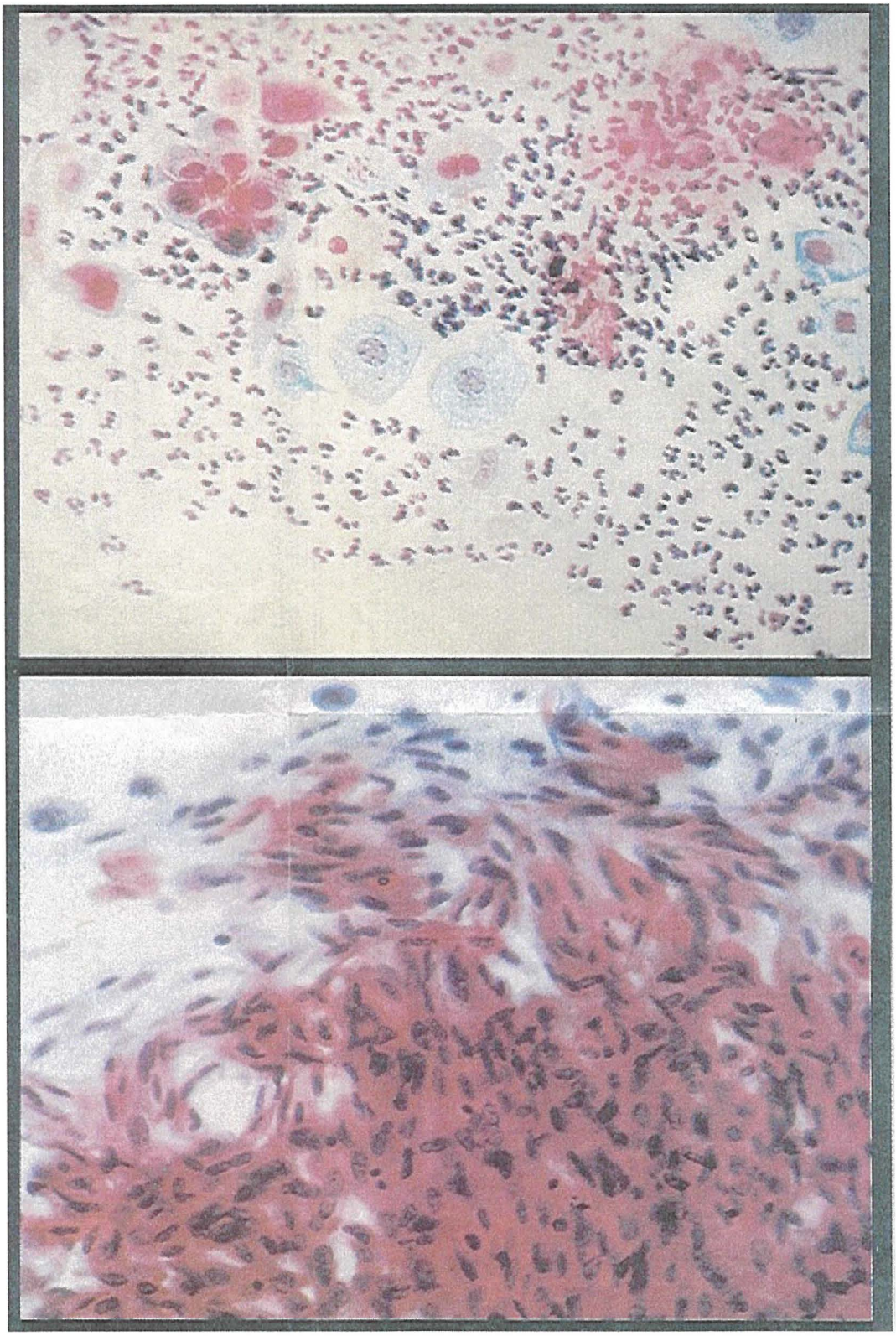
year survival outlook is between 45-65%, 25-35% with Stage III, and a maximum of 15% in Stage IV. If the disease reoccurs or radiotherapy has failed to destroy the disease completely, exenteration of the pelvic organs is necessary. The five-year survival following exenteration is reasonable (30-60%) in women with advanced but localised disease. Chemotherapy with anti-cancer drugs does not have a great effect in improving the outlook of victims of advanced cervical cancer.

3.6 STAINING TECHNIQUES

Papanicolaou studied the normal and abnormal variation in the vaginal smear in women and in animals.[247,248] He became aware of the fact that carcinoma of the **fundus** and carcinoma of the cervix are **exfoliative** lesions, in which cells at the free surface of the growth tend to be dislodged into the vagina. He also perfected the collecting and empirical staining technique.[249,250]

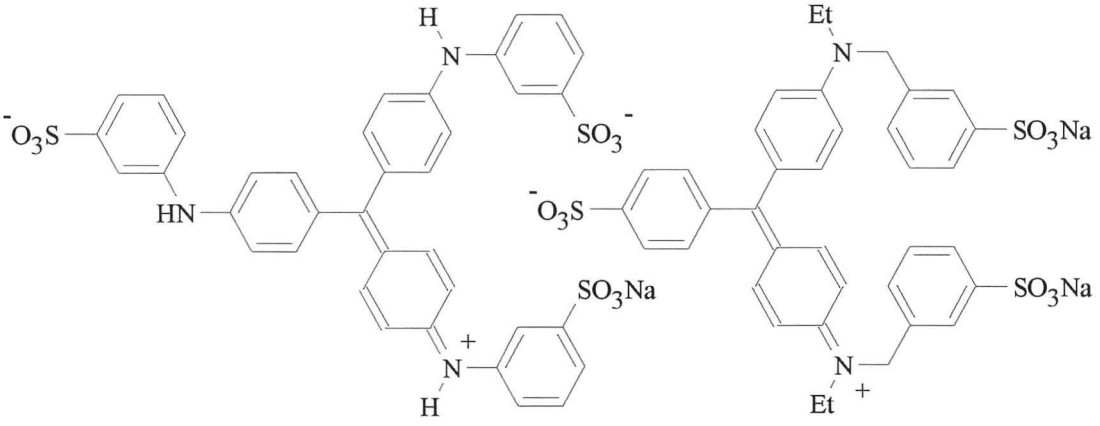
The difference between normal and abnormal cervical smears is shown below in *Figure 3.1*, reproduced from *The Guardian*. The most characteristic feature of abnormal cells is the atypical form and structure of the nuclei, which are often very large. The **chromatin** frequently shows a characteristic distribution in the form of conspicuous granules and of one or more small nucleoli. The nuclei are often **hyperchromatic** and stain intensely. The cytoplasm also shows abnormal changes, often being dense and hyperchromatic, particularly in the cells of the basal type. In some cases there is an excess of cornified cells in the early stages of malignancy due to a hyperestrin condition. **Cornification** is sometimes due to other pathologic conditions such as vaginitis, chronic cervicitis or due to the presence of *trichromonas*. Blood elements are quite conspicuous in the malignant cervical smears. **Erythrocytes** and **leucocytes** are found in large numbers. Many erythrocytes may have lost their hæmoglobin. **Eosinophiles** are rarely found in normal smears and are more frequent in malignant smears. *Trichomonas*, *cocci*, *diplococci*, and other chronic infections can also cause **leucocytosis**. **Histiocytes** are present in carcinoma smears and display a high **phagocytic** activity, phagocyting dead leucocytes, erythrocytes and other cellular debris.

Figure 3.1 **Front-page extract from *G2* of *The Guardian* (21st May 1997) "Spot the difference – These are slides of cervical smear tests. One is cancerous, the other normal. Clear enough, you'd think. So why are women dying despite having routine check-ups?"**

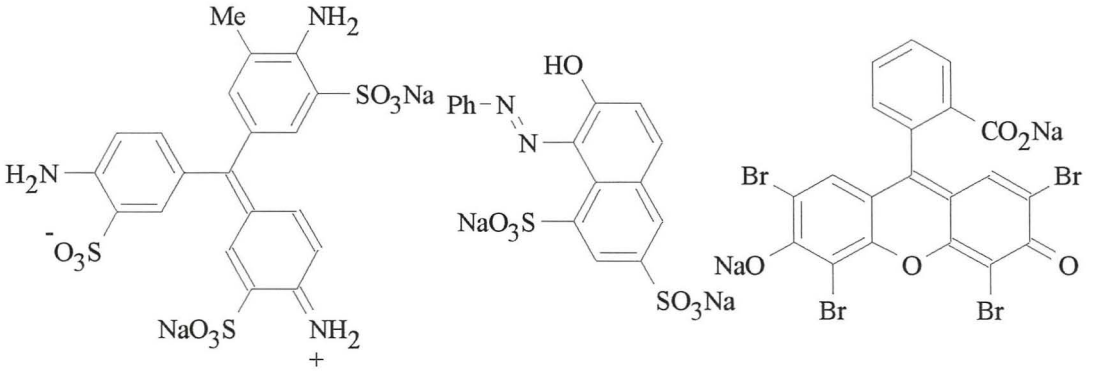


3.6.1 Papanicolaou staining techniques

The eosin-water blue staining method suggested by Papanicolaou in 1933,[248] permits a good differentiation between the blue **basal** cells, the purple mucous cells, and the red **cornified** cells. The method consists in treating vaginal smears, previously stained with Ehrlich's hæmatoxylin,[251] as good nuclear staining is essential for recognition of malignant changes, with a 0.5% aqueous solution of eosin Y **(29)** for 4 minutes, and then after rinsing in water, in a 0.5% aqueous solution of water blue, an obsolete dye, for about 1 minute. The staining method was simplified by a combination of eosin and water blue.[252] An increase in the amount of eosin results in sharper staining of the cornified cells.[253] An increase in the colour range of the cornified cells was attempted by the addition of orange G **(25)**, acid fuchsin **(12)** and other stains. Various shades of orange, pink, and red were obtained by modifying the amount of stain and acid. Aniline blue water soluble **(18)**, light green SF yellowish **(13)**, and other stains were substituted for water blue, allowing the colour of the noncornified cells to be changed.



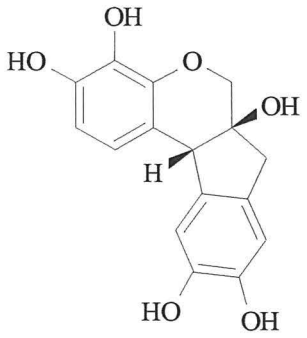
(18) **(13)**



(12) **(25)** **(29)**

Long experimentation with over 400 combinations of various imported and domestic stains led to a number of satisfactory preparations, which were modified for suitable differentiation of cell types. Stains DA15,[254] CY18,[255] DF11,[256] DA19,[257] DF20,[258] and DF32,[259] are more appropriate for staining human vaginal smears. For lower mammals or rodents, a percentage increase of the blue or green basophilic stain secures a better differentiation of the basal cells. Slides are stained as follows:

1) **Fix** smears immediately **before drying** in equal parts of 95% alcohol and ether,[260] using the original method of Stockard and Papanicolaou.[261] Rinse with 70% and 50% alcohols and in distilled water.
2) Stain in Ehrlich's hæmatoxylin for one to two minutes.[251] Harris hæmatoxylin,[262] prepared from hæmatoxylin **(105)** and ammonium alum, gives very good results. Other hæmatoxylin solutions can also be used.[263-270]
3) Rinse in distilled water. Stain in one of the counterstains for two minutes, or five minutes or longer if the smears are thick. Rinse in water.
4) Rinse in dioxane ten to fifteen times until smears are clear. Rinse in absolute alcohol and xylol. **Mount** in clarite, Canada balsam, or gum dammar.



(105)

The use of dioxane instead of solutions with an increasing alcoholic concentration for dehydration, shortens the staining process and improves differential staining in human smears. Dioxane also modifies the staining reaction of mucous cells. The staining time can be shortened by the omission of hæmatoxylin. The nuclei are not as dark and the cell outlines are not as sharp as with hæmatoxylin, but the cornified cells are more prominent and the basal cells are more transparent. Slides are stained as follows:

1) Fix smears in alcohol-ether solution for one to two minutes. Rinse well in 70% and 50% alcohols, and in distilled water.
2) Stain in DF 50,[271] or in one of the other counterstains for two to three minutes. Rinse in water.
3) Rinse in dioxane until clear, then in absolute alcohol, xylol, and mount.

Nuclei are stained red, erythrocytes orange, cornified cells red, pink, or orange, and basophilic cells green or blue. When a prompt examination of a smear is desired for diagnostic purposes this method is of particular advantage, requiring only about 3 minutes. The differential staining of the cornified cells is of value in interpreting human vaginal smears in both normal and pathologic cases.

In the normal cycle, the orange shades tend to predominate during the second half of the cycle. Some local irritations or infections, abortions, and certain pathologic conditions associated with bleeding are characterised by a prevalence of pink or red shades. Cornified cells, like the so-called **eosinophilic** cells of rodents, usually appear in large numbers after *trichomonas* infection, and show a preference for orange shades. *Candida* (*monila*) are stained red and *trichomonas* grey-green. A similar preference is shown by the cornified cells following prolonged estrogenic treatment, such as during premenstruum or during pregnancy.[248]

Drying of the cells modifies their staining reactions. Dry smears cannot therefore be satisfactorily stained with the aforementioned methods. Most dry cells take an orange or yellowish colour, and is often a source of error in smears dried before fixation.

The general adoption of a standard staining method for cervical smears was advocated by Schneider.[272] Each method offers an advantage over the others. The use of the DF 50 stain without hæmatoxylin shortens the procedure, but does not stain nuclei as well as those including hæmatoxylin. It keeps leucocytes in the background and makes the epithelial cells stand out more distinctly. It may be used in the evaluation of estrogenic or other endocrine treatments, and of cellular changes occurring during the various phases of the sex cycle. For gynæcologic diagnostic work, other stains, such as DF 32, DF20, or CY18 after hæmatoxylin, would be preferable, as they bring out more cytological details.

Traut and Papanicolaou realised that certain modifications and improvements in their procedure for diagnosing malignant tumours and other pathological conditions of the female genital tract was necessary.[249] Methods which proved to be successful in other applications,[248,273-276] were not found to be entirely satisfactory in this particular work. The cellular staining was too deep to permit a sharp definition of their outlines in smears that were relatively thick or bloody. In most cases of carcinomas and in many other pathological conditions, there is a profuse vaginal discharge frequently mixed with blood, which forms a heavy film on the slides. In such smears there is considerable crowding and overlapping of cells which, when deeply stained, can not be well differentiated. This applies more particularly to the small endometrial cells which are often found in menstrual and other uterine bleedings, and have a pathognomonic value in the diagnosis of adenocarcinomas of the fundus,[249] and of other gynæcological conditions. These cells are also important in the evaluation of the normal menstrual cycle of sterility and of estrogenic and other endocrine therapy.

Much greater transparency and an equally good visual differentiation of the cells can be obtained by the use of stain solutions in 95% alcohol instead of aqueous solutions. Various alcoholic stains were developed.[277-281] The advantages offered are:[279]

1) The epithelial cells and the erythrocytes are more transparent. Overlapping cells can be more easily differentiated.
2) The colour of the acidophilic cells varies from red to orange, helping in the identification of certain smear types. Basophilic cells stain blue / blue-green.
3) Cells or fragments of tissue penetrated with blood take a characteristic orange or orange-green colour, which permits an easier recognition of small amounts of blood, even when erythrocytes are not distinctly seen.
4) Smears, which have been subjected to partial or complete drying can be stained fairly satisfactorily. The differential colouring is not entirely lost.

The pink or green staining is affected by various conditions such as inflammation, so that distinction between cornified and non-cornified cells is more reliably made using nuclear criteria, *i.e.* **pyknotic** nuclei for cornified and **vesicular** nuclei for non-cornified cells. Staining procedures can be automated,[282,283] although manual cytological inspection is still necessary.

3.6.2 Shorr staining techniques

Smears stained by Foot's modifications of Masson's trichrome method showed a very good differentiation of the cornified cells.[284] A double colouration of these cells was obtained. Some cells were pink, some were orange, and others took on intermediate shades. The method was, however, too long and complicated to be adopted as a standard method. Another disadvantage was its failure to secure a differential staining of the mucous cells, which are characteristic of the early estrous stage in some lower mammals, more particularly in the guinea pig. This disadvantage would not apply to human vaginal smears or to smears of animals in which mucous cells are absent.

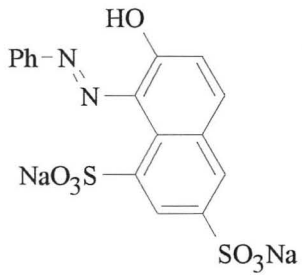
Shorr described three staining methods for vaginal smears,[273-275] based on a modification of the Masson trichrome method by Foot.[285] The three methods are very satisfactory for human smears and are considerably shorter than the original method. In preparing the smear, it is important not to allow the secretion to dry before fixing, as drying alters the morphology and the staining properties of the cells. Very thick smears are also undesirable, since it is difficult to remove the excess stain from the thick areas except by prolonged washing in the alcohols.

The advantages of the trichome stain over the original hæmatoxylin-eosin-waterblue method of Papanicolaou lay in its specific and reliable detection of cornification, and in its production of a sequence of contrasting colour changes in the cells under the influence of estrogens, whether endogenously produced or administered as therapeutic agents. These changes resemble a chemical titration in their sharpness. The addition of the colour changes to the morphological alterations in the vaginal secretion contributes to the ease and certainty of the interpretation of the smear. The changes are seen most strikingly in menopause or amenorrhea when, as a result of estrogenic therapy, an **atrophic smear** is transformed to the estrous or follicular type. The cells of the typical atrophic smear usually stain lavender or pale blue with the trichome stain. In less atrophic smears, the prevailing tint is a pale greenish-blue.

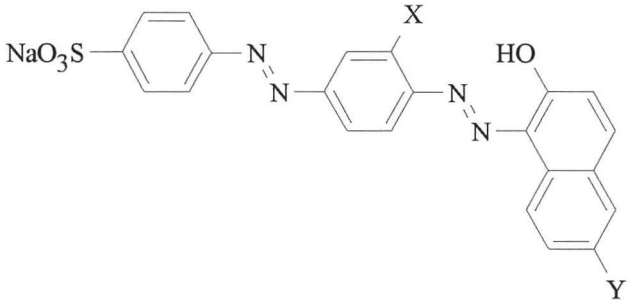
Following the administration of estrogens, the cells, in addition to undergoing morphological changes, become progressively greener. This definite greenish colouration persists up to cornification, at which stage the cells abruptly change to a brilliant orange red. During the normal menstrual cycle, and in amenorrhea following

the use of gonadotropic hormones, similar sharp transitions occur and permit the ready detection of ovulatory reactions. The desirable modifications in adapting the trichrome stain to the vaginal smear were an increase in staining time with light green to 8 minutes and a reduction in the strength of acetic acid from 1% to 0.25%.[286]

Analysis of the action of individual components indicated the possibility for simplification of the stain.[273] Ponceau de xylidine **(71)**, acid fuchsin **(12)** and orange G **(25)** were all taken up by the cornified cells and contributed to their final colour. The most satisfactory staining of the cornified cells, using these dyes separately, was obtained with ponceau de xylidine. Orange G was, however, desirable because of its staining of red cells. Phosphotungstic acid was essential due to its action as a mordant in fixing and intensifying the colour produced by the preceding solution. Light green acted as a counterstain for the non-cornified cells. When used alone it will stain cornified cells a more intense green, but is unable to displace other dyes once they have entered the cell. The nuclei are stained red if hæmatoxylin staining is omitted.

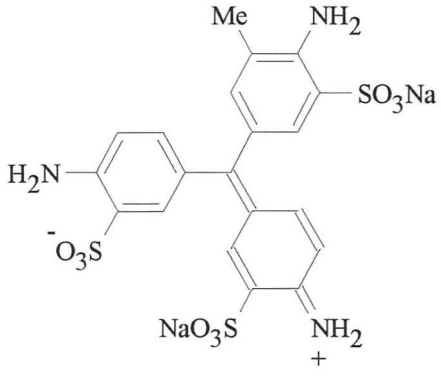


(25)

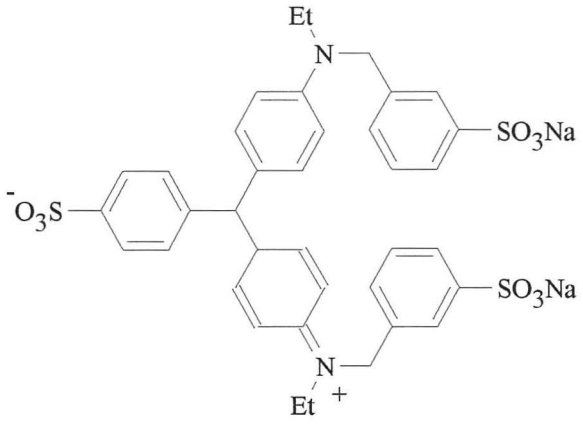


(71) X = H, Y = SO_3Na

(70) X = SO_3Na, Y = H



(12)



(81) X = SO_3Na

Shorr attempted to simplify the technique as well as to avoid the necessity for relying on imported stains such as ponceau de xylidine and light green SF yellowish.[274] Lillie reported that domestic Biebrich scarlet **(70)** and fast green FCF **(81)** may be substituted for ponceau de xylidine and light green SF yellowish respectively, and that a mixture of equal parts of 5% phosphomolybdic and phosphotungstic acids gave adequate mordanting in one minute,[287] simplifying and shortening the technique.[288]

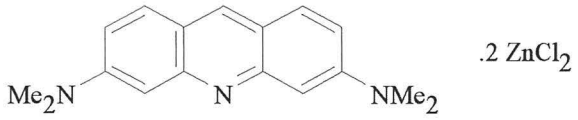
Although this staining method is simple and rapid, the number of solutions required renders it cumbersome for general clinical use. Shorr reported a single differential stain S3,[275,289] providing a sharp differentiation between cornified and non-cornified elements. The former stain a brilliant orange-red; the latter taking on a green stain, which is deeper in the younger cells, and paler in those more advanced. The staining is delicate and reveals cytoplasmic and nuclear details. Other constituents, such as leucocytes, erythrocytes, bacteria and spermatozoa, are satisfactorily differentiated.

3.6.3 Other cervical staining techniques

i) Modified May-Grunwald-Giemsa technique[290,291]

Some workers prefer one of the Romanowsky-type stains for examinations of sputum and serous fluid smears, showing *Trichomonas vaginalis* parasites particularly clearly. With serous fluids, adenocarcinoma cells are well picked out, especially if any intracytoplasmic vacuolation is present. Nuclei are stained purple, cell cytoplasms blue to mauve and red blood cells pink.

ii) Acridine orange technique[292-298]



(106)

Fluorescent staining of exfoliated material, particularly that of the female reproductive tract, has enjoyed some popularity due to the clarity with which RNA fluorescence is seen, as rapidly growing malignant cells have an increased cytoplasmic RNA content.[299-302] The technique is based on the **metachromatic**[303] properties of acridine orange **(106)**, giving differential staining of cellular DNA and

RNA.[304] Impermanence of the preparations and expertise needed in the interpretation, particularly of squamous carcinoma cells, are problems,[305] requiring a pH of 6.0 for the differential staining of RNA and DNA. Formalin-fixed materials do not stain satisfactorily, so alcohol is used as a fixative. With a fluorescence microscope, DNA is yellow-green, while RNA and some mucins are red. Malignant cells in the female genital tract are characterised by bright orange to flaming red fluorescence of the cytoplasm, and often by greenish-yellow hyperchromatic nuclei. *Trichomonas vaginalis* organisms have brownish-red cytoplasm and small, yellow nuclei.[297] The mycelia and spores of *Monila albicans* fluoresce bright red, while bacteria appear red to reddish brown. Erythrocytes do not appear to fluoresce, possibly due to the haemoglobin iron content, exerting a quenching effect on fluorescence.[306,307] Haemorrhagic smears present no obstacle as erythrocytes appear transparent.

iii) Haematoxylin-eosin-indigocarmine[308,309]

Nuclei of epithelial cells appear grey, while cytoplasm is coloured violet. Cornified cells and erythrocytes are stained brilliant pink and pink respectively.

iv) KH Stains[310,311]

The colours of cornified and non-cornified cells stained with KH stains are:

Table 3.2 Staining of cornified and non-cornified cells with KH stains.

Stain	Cornified cells (keratin)	Non-cornified cells (collagen)
KH1	Purple	Blue
KH8	Reddish-orange	Blue
KH9	Brilliant red	Pale blue

v) MF4 Stain[307,310,312,313]

The MF4 stain developed by Asscher *et al*, stains cornified cells yellow, orange or red, while non-cornified cells are stained blue.

vi) PX Stain[307,310,314]

Gurr's PX stain, stains nuclei red. Cornified cells are stained red, pink or orange, while basophilic cells are stained blue or green. Erythrocytes are stained orange.

vii) *Erythrosin - Water blue - Hæmatoxylin*[315,316]

Nuclei are stained dark blue, while cytoplasm is stained red, blue or pale purple depending on cell condition. The proportion of red to blue cytoplasm depends on the stage of menstrual cycle, with pale purple cytoplasm found in smears before menstruation.

viii) *Greenstein's five dye stain*[317,318]

Greenstein's five-dye stain can be applied to vaginal and cervical smears fixed in ether-alcohol (1:1) or isopropyl alcohol. Nuclei are stained a dark greyish purple to a greyish brown. Cytoplasm is stained light greenish yellow to greenish brown, while collagen is stained bright green. Erythrocytes are stained orange-yellow, muscle cytoplasm is stained greyish green to olive brown, and mucus, pale green to turquoise.

ix) *Victoria blue - Kernechtrot - Light Green*[319,320]

Elementary bodies, such as bacteria, fungi and viruses, *e.g. herpes simplex* are stained blue. Cell nuclei and protoplasm are stained pink and green respectively.

x) *Leuco methylene blue for differentiating cancerous and normal tissues*[321-323]

Nuclei of normal cells stain dark blue while the nuclear structures are quite distinct. Nuclei of cancerous cells remain almost colourless or assume a diffuse light blue or pale pink colour. Nucleoli of malignant cells appear more intensely coloured than the almost colourless nuclei, although their blue colour is far below than of normal nucleoli. Cytoplasm of normal and malignant cells is stained red by acid fuchsin. Malignant cells are unable to transform the colourless hyposulphite leuco base of methylene blue into the coloured form.

xi) *Histochemical methods*

The work of Kasdon *et al*[324] drew attention to the elevated levels of β-glucuronidase in human cervical cancer and led to a number of attempts to use this as a diagnostic criterion for distinguishing early malignancy.[325] Unfortunately the results were not decisive.[326] Elevated activity of 6-phosphogluconate dehydrogenase has been used successfully for the chemical diagnosis of human uterine cancers.[327] It was the

histochemical demonstration of exceptionally high activity of 6-phosphogluconate dehydrogenase in cancer[328] which led to a biochemical test and a histochemical stain for assisting in the cytological diagnosis of cervical cancer.[329]

3.7 SUMMARY

The cellular evolution, epidemiology and causes of cervical cancer have been examined. The advantages and disadvantages of the cervical cancer screening using smears, colposcopy and biopsies has also been discussed, together with a discussion on treating the various forms of cancer. Finally the Papanicolaou, Shorr and other cytological methods of cervical cancer screening have been reviewed. The cytochemical methods used by cytologists in the detection of cervical cancer have not changed since the 1940s, when the alchemic ingredients were obtained by "trial and error" empirical procedures.

Although most errors in cervical smear examinations tend to be misinterpretation on the part of the cytologist, the main drawback of the current Pap test, is that it is not preventative, *i.e.* the cancer cells must be formed before they can be visualised. Although, sensitive serological tests for antibodies to type-specific glycoproteins of HSV-1 and HSV-2 have been developed, similar serological tests for CMV-DNA in cervical cancer cells have yielded negative or inconclusive results. Such serological and histochemical tests could in theory be performed in conjunction with the cheaper and simpler "Pap" test. As there are a number of staining procedures specific for certain types of bacteria and viruses, it might be possible to find a stain specific for HSV, HPV or CMV, enabling the agents causing cervical cancer to be identified in a "healthy" smear, before the initial stages of pre-cancer. Excellent detection rates could be obtained.

The calculation of the electronic wavelength maximum, λ_{max}, of dyes using various semi-empirical and density functional theory methods will be discussed in the main body of the thesis, allowing the colour of any stain to be predicted. Initial studies of stain-cervical cell interactions will also be examined, which may be used in the rational design of more efficient Papanicolaou stains.

CHAPTER 4
PREDICTING ELECTRONIC WAVELENGTH MAXIMA

4.1 AIMS

The primary aim of the thesis is to calculate the colour and intensity of cytological stains, such as those used in the Papanicolaou cervical smear test, discussed in *Chapter 3*. The colour of a dye can be obtained from knowledge of the electronic absorption spectrum, *e.g.* the electronic wavelength maximum, λ_{max}, while the intensity is related to the extinction coefficient, ε.

The calculation of the electronic wavelength maximum, λ_{max}, has long been of interest to organic chemists. Empirical methods, such as Woodward's rules for polyenes and enones have been used for nearly fifty years. Empirical methods for simple polyenes, enones and aromatics will be discussed, together with solvent effects on the wavelength maximum. QM calculations must be used for more complicated systems. The use of π-electron calculations, such as the free electron molecular orbital (FE MO), the Hückel molecular orbital (HMO), and the Pariser-Parr-Pople (PPP) methods, to calculate electronic wavelength maxima of planar systems will then be discussed. For non-planar systems, methods based on neglect of differential overlap must be used, as π-electron methods can only be used for planar systems. Results obtained using these methods will be presented in *Chapters 6* and *7*.

4.2 EMPIRICAL METHODS FOR CALCULATING λ

Electronic transitions have been classified according to the electron excitation mechanisms,[330,331] *e.g.* **electron transfer** (E.T.) and **local excitation** (L.E.) bands.

4.2.1 Alkenes

Alkene electronic spectra were classified as shown below in *Table 4.1*.[332] Only type *h* alkenes were clearly distinguishable, with an absorption greater than 200 nm, as shown in *Table 4.2*. Diagnosis of alkene absorption could be accomplished by looking at the intensity of absorption at 210 nm, as shown in the final column of *Table 4.2*.[333]

Table 4.1 Classification of alkenes.[332]

No double bond	*a*
Cis disubstituted	*b*
Trans disubstituted	*c*
Exo methylene	*d*
Trisubstituted and *exo*- to ring	*e*
Tetrasubstituted:	
Not exocyclic	*f*
Exo- to 1 ring	*g*
Exo- to 2 rings	*h*

Table 4.2 Classification of alkene absorption.[332,333]

	Type	λ_{max} / nm†	ε_{max} / dm^3 mol^{-1} cm^{-1}	ε_{210} / dm^3 mol^{-1} cm^{-1}
R-CH=CH-R	*b,c*	196	6200-7500	200-1900
$R_2C=CH_2$	*d*	195-196	8700-9300	2200-2500
$RCH=CR_2$	*e*	195-196	8100-9700	1400-3200
$R_2C=CR_2$	*f*	196-198	7300-9000	3400-4700
RingC=Cring	*h*	204	11,300-12,200	10,000-10,700

† Recorded using ethanol as a solvent.

A second non-conjugated double bond will, generally, give rise to a spectrum obtained by summation of the contributing absorption bands. This principle of summation (and subtraction) of spectra of non-conjugated chromophores has been used in many natural product studies. Maximum shifts of polar groups (N~S>Hal>O) in the electronic spectra of substituted alkenes are illustrated in *Table 4.3*. The presence of methoxyl groups produces a moderate shift of 28 nm to the red,[335] while the SMe and piperidyl groups display a large shift of 60 nm to the red.[336] Polar groups also act as auxochromes, intensifying the absorption - a **hyperchromic** effect.

Table 4.3 Effect of polar groups on alkene absorption.

Compound	λ_{max} / nm	ε / dm^3 mol^{-1} cm^{-1}	Reference
$CH_2=CH_2$	162.5	15,000	334
CH_2=CH-OMe	190	10,000	335
CH_2=CH-Cl	185	10,000	335
CH_2=CH-SMe	228	8,000	335
EtCH=CH-pip†	228	10,000	336

† pip = piperidyl

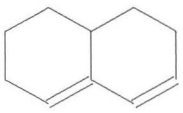
4.2.2 Woodward's rules for polyene absorption

The absorption of butadiene at 217 nm (ε 21,000) corresponds to a transition of a π-electron to an antibonding orbital.[337] Parent butadiene and alkylated derivatives exist

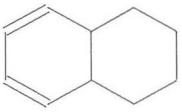
in the preferred *S-trans* conformation. Increasing electron repulsion in the *S-cis* conformer is shown by a red shift of 40 nm in cyclohexadiene, which absorbs at 256 nm.[337] The absorption intensity of *S-cis* dienes is lower (ε 10,000-15,000) than that of fully transoid dienes (ε 20,000-30,000) due to increased electron repulsion.[338]

Heteroannular dienes such as **(107)** are congeneric with butadiene whilst homoannular dienes such as **(108)** are derived from cyclohexadiene. A regular and significant variation of λ with substitution pattern was adumbrated during 1935-40 by Gillam,[337] Fieser,[339] and Dannenberg,[340] later codified with powerful effect by Woodward,[341] in the form of two rules for the prediction of diene absorption maxima, based on a standard wavelength of 217nm for butadiene. The original Woodward rules are:

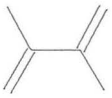
1) Each alkyl substituent or ring residue attached to the diene chromophore displaces λ_{max} by 5 nm towards longer wavelength.
2) Each exocyclic double bond causes a further displacement of 5 nm, the effect being twofold if the bond is exocyclic to two rings.



(107)



(108)



(109)

Using 2,3-dimethybutadiene **(109)** as an example, good agreements are obtained:
Calculated λ = 217 + 2 × 5 (alkyl substitution) = 227 nm.
Observed λ = 226 nm.[339]

Fieser's modification[342] of the diene rules, shown in *Table 4.4*, based on a large number of steroid spectra, lowers the "basic diene" value by 3 nm. A value of 253 nm is assigned to the unsubstituted homoannular system. Two important works on diene maxima are those of Dorfman[343] and Fieser and Fieser.[342] Diene rules are applicable to all acyclic and heteroannular dienes but only to homoannular dienes contained in a 6-membered ring. Deviation from the 256 nm (cyclohexadiene) maximum is found in the alicyclic dienes of increasing size, a minimum being reached at the 9-ring diene.

Table 4.4 Rules of diene absorption.[342]

Parent acyclic and heteroannular dienes	214
Parent homoannular diene	253
Increments for:	
Double bond extending conjunction	30
Alkyl substituent or ring residue	5
Double bond exocyclic to one ring	5
Exocyclic to two rings simultaneously	10
Polar groups:	
OAc	0
OAlk	6
SAlk	30
Cl, Br	5
$NAlk_2$	60

4.2.3 Polyenes

Several equations have been deduced *via* energy level differences relating polyene chain length and absorption maxima. For polyenes where $n < 6$, the Lewis-Calvin equation applies, where k is a constant:[344]

$$\lambda^2 = k\,n \qquad \ldots \text{Equation 4.1}$$

When $n > 6$, the polyene series converges:[345-347]

$$\lambda = a\,\{1 - b\cos[\pi/n - 1]\}^{0.5} \qquad \ldots \text{Equation 4.2}$$

$$\lambda = 1.23 \times 10^{-4} / [18.8\,(2n + 1)^{-1} + V_0\,(1 - \tfrac{1}{2}n)^{-1}] \qquad \ldots \text{Equation 4.3}$$

$$\lambda = D\,(1 - a^{2n-2}) \qquad \ldots \text{Equation 4.4}$$

Table 4.5 Longest wavelength absorption of all-*trans* polyenes $R[CH{=}CH]_nR$.

R	N	λ_{max} / nm	$\varepsilon \times 10^{-3}$ / dm^3 mol^{-1} cm^{-1}	Reference
H	3	267.5†	56	352
Me	3	274.5†	30	348
H	4	304†	-	352
Me	4	310†	76.5	348
H	5	334†	118	353
Me	5	342†	122	348
Me	6	380‡	146.5	348
Me	7	401‡	-	348
Me	8	411‡	-	348

† Recorded using hexane as solvent. ‡ Recorded using chloroform as solvent.

Although none of these equations provide an accurate fit for the observed maxima of dimethylpolyenes, the Lewis-Calvin equation reproduces experimental data.[348] The absorption intensity of the longest wavelength band, approaches a limiting value obeying simple chromophore-area theory.[349-351] For simple polyenes, $\varepsilon_{max} \approx 20{,}000\ n$.

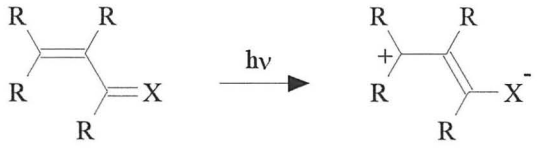
4.2.4 Diynes and poly-ynes

The spectra of conjugated acetylenes are characterised by a series of medium intensity bands (ε 200-300) regularly spaced at intervals of 2300 cm^{-1} and, for $n > 2$, a set of high intensity (ε ≈ 105) peaks separated by 2600 cm^{-1} intervals. The ε values of the first maximum of the high intensity band approach a value of 80,000 *n*. Spectral data for a series of dimethylpoly-ynes, $Me(C{\equiv}C)_nMe$ in ethanol are shown below.

Table 4.6 Spectral data for a series of dimethylpoly-ynes, $Me(C{\equiv}C)_nMe$.[353]

n	λ_1 / nm	ε_1 / $dm^3\ mol^{-1}\ cm^{-1}$	λ_2 / nm	ε_2 / $dm^3\ mol^{-1}\ cm^{-1}$
2			250	160
3	207	135,000	306	120
4	234	281,000	354	105
5	260.5	352,000	394	120
6	284	445,000		

4.2.5 Woodward's rules for α,β-unsaturated compounds



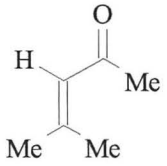
(110) **(111)**

The conjugation of unlike chromophores as in **(110)** results in the possibility of an electron transfer, such as the transition responsible for the intense **electron transfer** (E.T.) band corresponding to the dipole state **(111)**. In some cases it is also possible to locate the weak $n \rightarrow \pi^*$ **local excitation** (L.E.) band of the C=X moiety. The conjugation of a double bond with a carbonyl group leads to an intense absorption (log ε = 3.9 to 4.3) corresponding to transfer of a π-electron from the ethylenic orbitals to the C=O orbitals. The $\pi \rightarrow \pi^*$ E.T band is found between 220 and 250 nm in simple enones. A second characteristic absorption band (ε 50-100) appears at 310 to 330 nm and corresponds to the displaced $n \rightarrow \pi^*$ band of the C=O group. The E.T.

band is affected in a regular way by both alkyl and electron-donating substituents. The L.E. band shows no such dependence but exhibits a complementary set of shifts induced by conformational changes at the γ-position. The reliable prediction of the maxima of the intensely absorbing E.T. bands of α,β-unsaturated ketones can be made using Woodward's rules:

1) The average value for mono-alkylated enones is 224 nm
2) The substitution for hydrogen of an alkyl group on the α or β position of an α,β-unsaturated ketone causes a red shift of about 11nm.
3) A further 5nm shift is observed for an exocyclic double bond.

Using mesityl oxide, 4-methyl-3-penten-2-one **(112)**, as an example:



(112)

Calculated λ = 215 + 2 × 12 (alkyl substitution β position) = 239 nm.

Observed λ = 237 nm (ε 12,600).[354]

Table 4.7 Solvent corrections for enone absorption.[340,355]

Solvent	Correction (nm)
Ethanol	0
Methanol	0
Dioxane	+5
Chloroform	+1
Diethyl ether	+7
Water[356]	-8
Hexane	+11
Cyclohexane	+11

The position and intensity of the E.T. band is solvent dependent (*Section 4.2.6*), but correction factors are available, as shown above. By studying the UV absorption of a large number of steroids, Fieser formulated an extensive range of parameters to deal with many combinations of enone and dienone systems. The Woodward rules for enones, using ethanol as a solvent, are shown below.

Table 4.8 Rules of enone absorption - $\pi \rightarrow \pi^*$ transitions (ε 4500 - 20,000).

$\beta\ \alpha \qquad\qquad \delta\ \gamma\ \beta\ \alpha$

C=C-C=O and C=C-C=C-C=O

6-Ring and acyclic parent enone	215			
5-Ring parent enone	202			
Aldehydes	207			
Acids and esters	197			
Increments for:				
Double bond extending conjugation	30			
If second double bond is homoannular with first	39			
Double bond exocyclic to one ring	5			
If exocyclic to two rings simultaneously	10			
	α	β	γ	δ
Alkyl group, ring residue	10	12	17	17
OH	35	30	30	50
OAc	6	6	6	6
OMe	35	30	17	31
SAlk		85		
Cl	15	12	12	12
Br	25	30	25	25
NR_2		95		
Solvent correction	(see *Table 4.7*)			

The intensity of the E.T. band is usually 10,000 - 15,000 unless steric hindrance to coplanarity is present. Homoannular dienones may be treated in the same way as the heteroannular, *i.e.* an increment of 30 nm for a double bond extending conjugation, with an additional factor of 39 nm for the homodiene component. The values of increments used for polar substituents differ from author to author.[342,357]

Enone absorption rules apply to acyclic and six membered ketones and acylcycloalkenes. For cyclopentenones it is clear that the maxima predicted by the enone rules are consistently 10 to 15 nm higher than the observed values. Endocyclic 5-ring unsaturated ketones have a parent value of 202 nm. Exocyclic cyclopentenones conform to the parent value of 215 nm. The rules can also be applied for trienones.

The E.T. band of α,β-unsaturated aldehydes is affected by alkyl substitution in the expected sense, as shown in *Table 4.9*. Taking 207 nm as the basic chromophore, application of the usual alkyl increments (α = 10, β = 12) gives reliable agreement. The increment for the introduction of a second β-alkyl substituent in unsaturated aldehyde and ketone systems is often greater than 12 nm.[358]

Table 4.9 Spectral data for a series of α,β-unsaturated aldehydes.[358]

Compound	Calculated λ / nm	Observed λ / nm	ε / dm^3 mol^{-1} cm^{-1}
CH_2=CH-CH=O	207	207	11,200
CH_3.CH=CH-CH=O	219	217	17,900
CH_2=CMe-CH=O	217	216	11,000
CH_3.CH=CMe.CH=O	229	226	16,100
Me_2C=CMe.CH=O	241	245	13,000
CHex=CH.CH=O†	236	238	16,000

† CHex = Cyclohexyl

The position of λ can be calculated with fair accuracy for simple unsaturated acids and their alkylated derivatives. Taking the monosubstituted value for simple unsaturated acids as 208 nm, based on the E.T. band of *trans*-crotonic acid at 205 nm (ε 14,000),[359] Nielsen summarised the absorption of α,β-unsaturated acids: [359]

Table 4.10 Absorption of α,β-unsaturated acids (λ ± 5 nm).[359]

Parent acid	198
Monosubstituted - β	208
Disubstituted - α,β or β,β	217
Trisubstituted - α,β,β	225
C=C extending conjugation	30
C=C exocyclic	5

4.2.6 Solvent effects

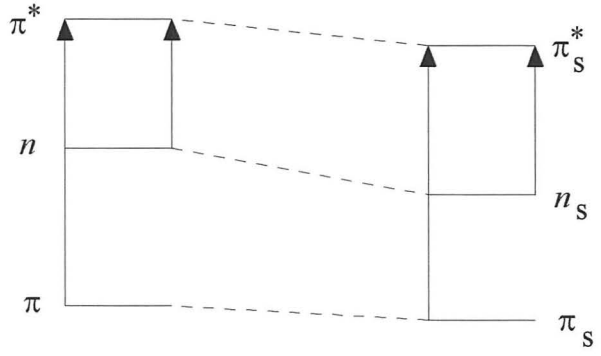
The position and intensity of an absorption band may shift when the spectrum is recorded in different solvents. Although non-polar polyenes and aromatic hydrocarbons experience very little solvent shift, polar enones show two different shifts on increasing solvent polarity:

i) the $\pi \rightarrow \pi^*$ band moves to longer wavelength (red shift), while

ii) the $n \rightarrow \pi^*$ band moves to shorter wavelength (blue shift).

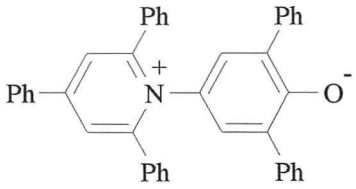
Solvation by a polar solvent stabilises π, π*, and *n* orbitals. The stabilisation of nonbonding orbitals is particularly pronounced with hydrogen bonding solvents such as water or ethanol. The more polar π* orbitals undergo greater stabilisation upon solvation than the π orbitals. As shown in *Figure 4.1* below, the energy of the $\pi \rightarrow \pi^*$ and $n \rightarrow \pi^*$ transitions decrease and increase respectively with solvation.

Figure 4.1 **Effect of solvation on the relative energies of orbitals and transitions in α,β-unsaturated carbonyl compounds.**



2,6-Diphenyl-4-(2,4,6-triphenylpyridinio)phenolate **(113)**, Dimroth's dye or Reichardt's dye, a **solvatochromic** dye, absorbs at markedly different wavelengths in different solvents, and is used in the E_T (transition energy) scale of solvent polarity,[360] where λ is the measured wavelength of absorption in nm:

E_T (in kcal mol^{-1}) = 28 591 / λ



(113)

4.2.7 Substituted benzenes

A well-defined set of rules is not available for calculating aromatic maxima, even on the most empirical basis. The most characteristic band of the substituted benzene ring, corresponding to a transfer of an electron from the ring to a *–M*-substitutent, or to the ring from a *+M*-substituent with available *p*-electrons, is mainly only slightly affected by changes in the nature of substitution. In spite of attempts to classify shifts of maxima engendered by introduction of polar substituents,[361,362] the steric and electronic interplay in situations more complex than disubstituted derivatives, makes searching for a close analogy of a chromophoric system the only alternative.

The spectrum of benzene displays considerable fine structure in the vapour state and in hexane, with L.E. bands at 184 nm (ε 60,000), 203.5 nm (ε 7400), and 254 nm (ε 204).[363] Increasing alkyl substitution causes a bathochromic shift of the 254 nm band due to inductive effects or hyperconjugation.[364] Effects of various +*M*- and –*M*-substituents on the absorption bands of benzene were discussed by Doub and Vandenbelt[361,362] who attempted to relate the effect of the nature and pattern of substitution on the 203.5 nm band of benzene. The availability of *p*-electrons in the groupings -OR, -NR_2 *etc.*, and polarizability of moieties such as C=O, CN, NO_2 allow operation of an E.T. transition characteristic of the composite molecule R-S where R and S represent different chromophores separated by a single bond.[365,366]

By consideration of disubstituted benzenes, Doub and Vandenbelt arrived at $\Delta\lambda$ values for some sixteen substituents in hydroxylic solvent, where:

$$\Delta\lambda = \lambda - 203.5 \text{ nm} \quad \text{... Equation 4.5}$$

Table 4.11 $\delta\lambda_0$ Values for benzene substitution.

Substituent	$\delta\lambda_0$	Substituent	$\delta\lambda_0$
H	24.05	CN	43.9
NH_3^+	25.6	CO_2^-	44.7
CH_3	29.6	CO_2H	50
Cl	29.3	NH_2	50
Br	31.5	O^-	55
OH	36.1	CMeO	63.3
OMe	36.8	CHO	68.9
SO_2NH_2	37.5	NO_2	91.6

Straight line relationships were found for plots of $\Delta\lambda$ for certain series of *para*-disubstituted benzenes against products of contributing substitutional increments, $\Delta\lambda'$ and $\Delta\lambda''$, defined in *Equation 4.7*. Refinements were made using $\lambda_0 = 180$ nm.

$$\Delta\lambda_0 = \lambda - 180 \quad \text{... Equation 4.6}$$

$$\delta\lambda_0' \times \delta\lambda_0'' = 24.05 \times \Delta\lambda_0 \quad \text{... Equation 4.7}$$

A plot of $\Delta\lambda_0$ against products of $\delta\lambda_0$ showed a widely applicable straight-line relationship for *para*-disubstitution, while *ortho*- and *meta*-derivatives are plotted

against the sum of $\delta\lambda_0$ values. This approach is only applicable to substitution of opposite electronic effect involving an *ortho-para*-directing group in the presence of a *meta*-directing group.

For *o*-, *m*- and *p*-hydroxyacetophenones:

Using *Equation 4.6*: $\lambda = 180 + 36.1(OH) + 63.3(COMe) = 279.4$ nm

Using Equations *4.6* and *4.7*: $\lambda = 180 + (36.1 \times 63.3 / 24.05) = 275$ nm.

The experimentally determined wavelengths are 251 nm, 251 nm and 276 nm for the *ortho*-, *meta*- and *para*-isomers respectively.[367] The two methods are therefore satisfactory for *para*- but not for *ortho*- and *meta*-derivatives.

4.2.8 Aromatic carbonyls

The E.T. absorptions of substituted benzeldehydes and acetophenones are sharply divided into transitions characteristic of *ortho-/meta*-substitutions, and to the *para*-disubstitution.[361,362,368,369] The *ortho*- and *meta*-isomers give rise to two E.T. bands in the 230 to 300 nm region, whereas the corresponding *para*-compound shows only one intense E.T. band at longer wavelengths. *ortho*-Substituents may have steric requirements which lead to a decrease in the absorption intensity and / or a wavelength shift.[370] A unifying MO treatment of the "*ortho*" phenomenon has appeared.[371] Murrell[366] pointed out that the E.T. band of a given molecule may be recognised by its susceptibility to the loss of coplanarity by steric hindrance. The fundamental difference between *ortho*-, *meta*- and *para*-transitions can be explained, and appropriate transition energies calculated by QM treatment of the simple disubstituted benzene case.[372]

The substitution of acetophenone by a variety of groupings in the *ortho*-, *meta*- and *para*- positions is reflected in the position of the 246 nm band. Intensity changes become important where operation of the "*ortho*" effect depresses ε to such an extent, that only a shoulder remains at the expected absorption wavelength. The extent of loss of coplanarity on electron transfer is revealed for 2,4,6-trimethylacetophenone, λ_{max} 246 nm (ε 3600),[373] with the absorption intensity obeying an approximate $\cos^2\theta$ law.

Table 4.12 Rules for calculating the principal E.T. band of substituted benzene derivatives, Ar-COR.

		$\lambda^{*\dagger}$ / nm
Parent chromophore: Ph-COR		
R = alkyl or ring residue		246
R = H		250
R = OH, OAlk		230
Increment for each substituent:		
Alkyl or ring residue	*o-*, *m-*	3
	p-	10
OH, OMe, OAlk	*o-*, *m-*	7
	p-	25
O^-	*o-*	11
	m-	20
	p-	78‡
Cl	*o-*, *m-*	0
	p-	10
Br	*o-*, *m-*	2
	p-	15
NH_2	*o-*, *m-*	13
	p-	58
NHAc	*o-*, *m-*	20
	p-	45
NHMe	*p-*	73
NMe_2	*o-*, *m-*	20
	p-	85

† $\lambda^* = \lambda_{max}$ in ethanol. ‡ Markedly decreased by steric hindrance to coplanarity.

Table 4.13 E.T. Bands of substituted acetophenones in ethanol.

Compound	Calculated λ^* / nm	Observed λ^* / nm	$\Delta\lambda^*$ / nm	Reference
2-Me	249	245	-4	374
4-Me	256	256	0	361
2,4,6-tri-Me	262	242	-20†	373
2-OH	253	252	-1	367
3-OH	253	250.5	-2.5	362
4-OH	271	275	+4	361
2,6-di-OH	260	270	+10†	375
2,4,6-tri-OH	285	285.5	+0.5†	376
2-NH_2	259	259	0	377
3-NH_2	259	262	+3	377
4-NH_2	304	311	+7	361
2,4-di-Cl	256	252	-4	378
2-OH-4-NH_2	311	325	+14‡	378

† *Ortho* effect predicted. ‡ At pH 6 in aqueous solution - solvent effects.

The increments in *Table 4.12* have been found to provide predicted wavelengths within 5 nm for many examples of aromatic ketones (*Table 4.13*). A hydroxylic

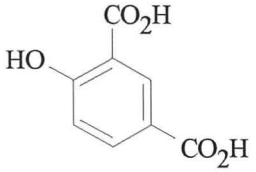
solvent (usually ethanol) is taken as reference, as the E.T. band under review is least susceptible to intramolecular effects in this solvent.[367,379,380]

Substituted benzaldehydes may be treated similarly, with the parent chromophore value of 250 nm, as shown in *Table 4.14*. In non-polar solvents such as hexane, the operation of intramolecular chelation and structural changes becomes evident in the shift of the second E.T. band in *ortho*-substituted aldehydes and ketones.[367,368,381]

Table 4.14 E.T. Bands of substituted benzaldehydes in ethanol.

Compound	Calculated λ* / nm	Observed λ* / nm	Δλ* / nm	Reference
2-Me	253	253	0	377
3-Me	253	251	-2	377
4-Me	260	259	-1	377
2-OH	257	256	-1	367
4-OH	275	282	+7	377
2-NMe_2	270	271	+1	377
4-NMe_2	335	337	+2	377
2-OMe	257	254	-3	377
3-OMe	257	255	-2	377
4-OMe	275	280	+5	377

Taking 230 nm as the parent chromophore of the substituted benzoic acids, esters and lactones, application of the above increments produces a reliable method of predicting the principal E.T. bands of this series. Agreement of calculated and observed wavelengths is satisfactory, except for 2,6-dimethyl- and 2,6-dichloro-benzoic acids[382] (no E.T. bands), and amino acids such as 2,4-diaminobenzoic acid.[378,383-388]



(114)

The presence of two (*–M*) substituents in the benzene ring leads to a cross-conjugation of chromophores. Calculation of E.T. maxima on the basis of the operation of the stronger chromophore, such as the *p*-hydroxybenzoic acid unit in 4-hydroxyisophthalic acid **(114)**, with no increment for the weaker *ortho*-transition leads to the observed value of 255 nm.[389] *trans*-Cinnamic acid absorbs at 273 nm

(ε 21,000) in ethanol, representing an increment of 43 nm.[390] The *cis*-isomer of shorter chromophoric length absorbs at 264 nm (ε 9500).[390] Additional unsaturation increases the maximum by ~25 nm for each double bond.[391]

4.2.9 Structure-activity relationships

The 260 nm band of substituted benzenes is due to a forbidden transition. Substitution causes an increase in the intensity and a shift of the absorption wavelength. The effect of substitution on the intensity of the 260 nm band was treated quantitatively by the use of MO theory, and good agreements with experiment were found.[392,393] Wavelength shifts and intensity changes of this band in a number of substituted benzenes were computed relative to benzene.[394]

The effect of substitution on the 200 nm band of benzene was studied by Doub and Vandenbelt,[361,362] the ratio of the wavelengths of the 260 nm band and the 200 nm band being ~1.25 for most monosubstituted benzenes. Substituent groups were divided into electron releasing (*ortho-/para*-directing) and electron withdrawing (*meta*-directing) groups and arranged in the order of increasing wavelengths of the 200 nm band. Two series of substituents were found:[361,362]

ortho-/para-directing groups: $CH_3 < Cl < Br < OH < OCH_3 < NH_2 < O^-$

meta-directing groups: $NH_3^+ < SO_2NH_2 < CO_2^- = CN < CO_2H < COCH_3 < CHO < NO_2$

Table 4.15 Comparison of λ and spectroscopic moments with orientation in substitution.

Group	% meta on nitration	$\lambda^{361,362}$ / nm (200 nm band)	Spectroscopic moment[395]
OH	2	210	34
Cl	0.2	209.5	6
Br	0.2	210	4
CH_3	4	206.5	7
CH_2Cl	16		-3
$CHCl_2$	34		-11
$COCH_3$	55	245.5	
CO_2H	80	230	-28
CN	81	224	-19
NO_2	93	268.5	

Spectroscopic moments, shown above, have been employed to correlate and predict intensity changes and wavelength shifts in a large number of benzene derivatives.[395] These additivity rules do not hold for groups which interact strongly with the ring, *e.g.* $-NH_2$ or $-NO_2$. The spectroscopic moments were modified by only accounting for the substituent perturbations, and ignoring the vibrational contributions.[396]

The effect of various combinations of these groups in *para*-disubstituted benzenes on the wavelength of the 200 nm band is very interesting.[361,362,397,398] If both substituent groups are of the same electronic nature, the observed shift is similar to the shift in the most displaced monosubstituted benzene. When both substituents are complementary (*ortho-/para*-directing *versus meta*-directing), a large bathochromic shift of the 200 nm band occurs. This is illustrated in *p*-dinitrobenzene[361,362] (λ_{max} 266 nm, ε 14,500), and *p*-nitroaniline[361] (λ_{max} 381 nm, ε 13,500), noting that nitrobenzene[361] and aniline[361] absorb at 269 nm (ε 7800) and 230 nm (ε 8600). Complementary disubstitution has the greatest effect in the *para*-position.[361,362,397,398] *p*-Nitrophenol[361] absorbs at 318 nm (ε 10,000), while the *m*- and *o*-derivatives[361,362] respectively absorb at 274 nm (ε 6000) and 279 nm (ε 6600), indicating the importance of polar resonance forms in determining shifts of the 200nm band.[361,362,397,398]

The displacing effect of a substituent group on the 200 nm band of benzene is proposed to be associated with the directional displacement of electrons to, or from the benzene ring, depending on the electronic nature of the group.[361,362,397,398] The electronic interaction responsible for the displacement of the 200 nm band has been associated mainly with resonance rather than inductive effects. Rao correlated the position of the 200 nm band in monosubstituted and *para*-disubstituted benzenes with Hammett reactivity constants and resonance parameters. For *para*-complementary disubstituted benzenes, the relation with σ-constants is of the form:[397-401]

$$\lambda = \lambda_0 + \rho\sigma \qquad \text{... Equation 4.8}$$

where ρ varies systematically with the electrical property of the invariant group. A similar equation with Taft resonance parameters is not very satisfactory.[401] A general correlation of the type:

$$\lambda = \lambda_0 + | \sigma' - \sigma'' | \rho \quad \text{... Equation 4.9}$$

has also been proposed for complementary disubstituted benzenes where σ' and σ'' are substituent constants. The correlations have been useful in estimating the reactivity constants and resonance parameters of groups.[402-405] A relationship between the IR and UV absorption frequencies has been found in some benzoic acid derivatives.[406]

The UV absorption maxima of *para*-halogen substituted benzene derivatives are found to be in the order I > Br > Cl > F,[397,398,407-412] the reverse order to the accepted trend of halogen resonance parameters. The 200 nm band in *para*-halonitrobenzenes, acetophenones, phenols, anisoles and anilines was studied in different solvents, and interpreted the results in terms of the polarisability of the halogen substituent and solvent orientation and polarisation effects.[413-415] The importance of the expansion of halogen valence shells in determining λ has been pointed out.[416]

The substituent constants σ and σ^+ for halogens in *para*-disubstituted benzenes have been evaluated, where other substituents are electron withdrawing or electron donating.[401] The halogens become increasingly electron donating, in the order F, Cl, Br, I, when the other *para*-substituent is electron withdrawing, and *vice versa*. Effective *para*-F, Cl , Br, and I σ values were found to be -0.05, -0.11, -0.14 and -0.25 in the first case and 0.03, 0.11, 0.11 and 0.15 in the second case. These σ-values satisfactorily correlate the UV absorption data for *para*-halogen substituted benzenes.

The electronic absorption spectra of alkyl substituted benzenes have been interpreted in terms of hyperconjugation. Schubert pointed out the importance of polarisability and relief of strain on excitation in the case of the *p-n*-pentyl group in explaining the UV spectra of *para*-alkylnitrobenzenes.[417,418] The σ constants for Me, Et, *i*-Pr and *t*-Bu groups have been derived using the λ_{max}-σ correlations for complementary disubstituted benzenes.[401] In the correlations of λ_{max} of *para*-complementary disubstituted benzenes with σ constants, the sign and magnitude of the slope, ρ, varies systematically with the electrical property of the invariant group of the series:

Table 4.16 σ And ρ values for disubstituted benzenes.

$R\text{-}C_6H_4\text{-}X$	X	σ	ρ
Nitrobenzenes	*p*-NO_2	0.78	-159.4
Acetophenones	*p*-$COCH_3$	0.50	-118.2
Benzoic acids	*p*-CO_2H	0.45	-93.7
Anisoles	*p*-OCH_3	-0.27	89.5
Anilines	*p*-NH_2	-0.66	109.9

UV absorption spectra of trisubstituted benzenes have been studied in terms of the disubstituted components.[419] The UV absorption spectra of several series of disubstituted benzenes have been extensively studied, interpreting them in terms of the resonance and steric interactions of substitutents.[407-412] Monoaromatic hydrocarbon UV absorption spectra have been surveyed for a wide range of structure and substitution.[420] The empirical regularities in the UV absorption spectra of substituted benzenes have been summarised as follows:[421]

For conjugative substitution:

i) A red shift of the 200 nm and 260 nm bands is produced.

ii) The 180 nm band shifts slightly to the blue or remains insensitive to substitution.

iii) The 200 nm band undergoes the greatest red shifts.

iv) The long wavelength band is intensified, but the sum of the 180 nm band intensities remains insensitive to substitution.

v) For certain very strongly conjugative substituents like $-NH_2$, the correspondence to benzene breaks down.

For inductive substitution:

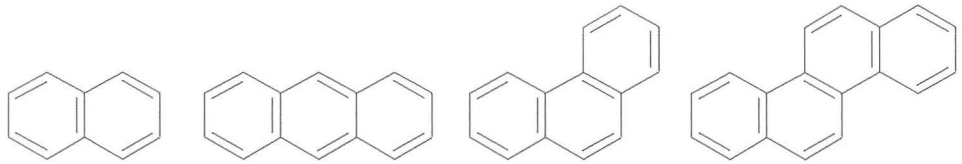
i) The 200 nm and 260 nm bands are only slightly shifted, the 180 nm band undergoes stronger shifts to the blue.

ii) The long wavelength band is greatly intensified, more than in the case of conjugative substitution.

For mixed conjugative and inductive substitution:

i) While the rules for mixed effects are less clear-cut, the conjugative effect, in general, controls the shifts of the long wavelength (260 nm) band.

4.2.10 Condensed-ring benzenoid hydrocarbons



(115) **(116)** **(117)** **(118)**

All condensed ring aromatic hydrocarbons, such as naphthalene **(115)**, anthracene **(116)**, phenanthrene **(117)** and chrysene **(118)**, absorb at longer wavelengths than benzene. Variations of the α, *p* and β bands of the acenes and phenes with the number of benzene rings have been analysed.[422] There is a progressive displacement of the absorption maxima to longer wavelengths with increasing number of fused benzene rings.[423,424] The effect of various substituents on the UV absorption spectrum of anthracene have been studied by Jones[425] and the band shifts have been correlated with the type and position of substituents. The observed transition energies of the primary band (^1L_a or *p*) of aromatic hydrocarbons (*cata*- as well as *peri*-condensed) can be correlated with the Hückel transition energies.[426]

4.2.11 Non-alternant hydrocarbons

Non-alternant hydrocarbons can be divided into two classes: one with $(4n + 2)$ conjugated atoms where n is an integer and the other with $4n$ conjugated atoms. The former exhibits the α, *p*, and β bands similar to alternant hydrocarbons, while the latter class show spectra similar to those of the vinyl derivatives of the corresponding hydrocarbon with $(4n - 2)$ conjugated atoms.[427] Azulene bands in the visible region are of low intensity ($\varepsilon \sim 100$-350), and are displaced by alkyl substitution, the nature and the magnitude of the shift depending on the position.[428] The wavelength shifts of the visible absorption bands in polyalkyl azulenes are found to be additive.[429] Alkyl substitution causes bathochromic shifts of the high intensity ($\varepsilon \sim 1000$-50,000) UV bands. Steric hindrance to resonance also causes bathochromic shifts of the absorption bands.[430] Hückel calculations correctly predict the features of the spectra of azulene and its alkylated derivatives. The Hückel method and other simple π-electron QM methods for calculating electronic wavelength maxima of planar organic compounds are discussed below.

4.3 COMPUTATIONAL METHODS FOR CALCULATING λ

Various empirical methods for calculating wavelength maxima of simple organic compounds have been discussed. This section will show how the FE MO, HMO and PPP π-electron methods, are used to calculate wavelength maxima of planar systems. For non-planar systems, calculations based on neglect of differential overlap must be used, as π-electron methods can only be applied to planar systems. Results obtained using NDO based methods are discussed in *Chapters 6* and *7*.

4.3.1 Free-electron molecular orbital (FE MO) method

A conjugated chain molecule is approximated as a 1-D box, of length l, giving:

$$e_i = n_i^2 h^2 / 8 m_e l^2 \qquad \text{where } n_i = 1, 2, ... \qquad \text{... Equation 4.10}$$

A π-electron HOMO→LUMO transition gives the lowest-frequency absorption as:

$$1/\lambda = (e_{n\pi/2+1} - e_{n\pi/2}) = h\,(n\pi + 1) / 8 m_e c l^2 \qquad \text{... Equation 4.11}$$

For conjugated polyenes, CH_2=[CH–CH=]$_k$$CH_2$, l is taken as the zigzag length along the chain. Alternatively the direct distance between the end carbons can be taken. SCF, MP2, MP4, CI-SD, and CCSD calculations with large basis sets including polarisation functions and geometry optimisation, predict that *S-trans* is the most stable conformation.[431,432] For *S-trans*-1,3-butadiene, λ_1 and λ_2 are 1.46 Å (C–C) and 1.34 Å (C=C) respectively. By adding a $\frac{3}{4}l_1 + \frac{1}{4}l_2$ distance at each end, the π electrons are allowed to move beyond the ends of the chain, allowing for the fact that the MOs are not confined to the regions between the bonding atoms, giving:

$$l = kl_1 + (k+1)l_2 + \tfrac{3}{4}l_1 + \tfrac{1}{4}l_2 + \tfrac{3}{4}l_1 + \tfrac{1}{4}l_2 = (n+1)(l_1 + l_2)/2 \qquad \text{... Equation 4.12}$$

$$\lambda = 2 m_e c h^{-1} (l_1 + l_2)^2 (n+1) = (n+1)(64.6\text{ nm}) \qquad \text{... Equation 4.13}$$

The observed longest-wavelength electronic absorption bands for the conjugated polyenes increases monotonically from λ = 162.5 nm to 447 nm as k increases from 0 to 9.[433] The FE MO equation predicts increasing wavelength with increasing chain length, but quantitative agreement with experiment is poor (± 110 %).

In polymethine ions, $[Me_2N{=}CH(–CH{=}CH)_k–NMe_2]^+$, the C–C and C=C bonds are of equal length, due to equivalent valence bond resonance structures. Assuming the C–C and C–N bond lengths to be 1.4 Å, adding an extra bond length at the end of the chain, so that $l = (2k + 4)(1.4\ Å)$ and $n = 2k + 4$, gives:

$$\lambda = 8m_e c(1.96\ Å^2)(2k + 4)^2/h(2k + 5) = (2k + 4)^2(64.6\ nm)/(2k + 5) \quad \ldots \text{Equation 4.14}$$

Observed λ values increase monotonically from 224 nm to 848 nm as k goes from 0 to 6.[434] Calculated values are in good agreement with experiment (± 12 %).

4.3.2 Hückel molecular orbital (HMO) method

For conjugated polyenes containing n carbon atoms, the HMO secular equation gives:

$$e_j = \alpha + 2\beta \cos(j\pi/(n + 1)) \quad \text{where } j = 1, \ldots n \qquad \ldots \text{Equation 4.15}$$

For the HOMO and LUMO respectively, $j = ½n$ and $½n + 1$. HMO theory predicts the longest wavelength band of the polyene electronic absorption spectrum to occur at:

$$1/\lambda = (-4\beta/hc) \sin[\pi/(2n + 2)] \qquad \ldots \text{Equation 4.16}$$

The bond integral β is a semi-empirical parameter adjusted to reproduce experimental data. Using λ = 217 nm for butadiene, it was found that β = -4.62 eV. Predicted wavelengths increase with increasing n, but agreement with experiment is poor. When two polyene bond integrals, β_1 = -3.32 eV and β_2 = -4.20 eV, are used for the C–C and C=C bonds respectively, agreement is improved.[435] Predicted wavenumbers for HOMO→LUMO transitions in conjugated hydrocarbons are given by:

$$1/\lambda = |\beta|\Delta x/hc \qquad \ldots \text{Equation 4.17}$$

where Δx is the difference in x values $\{x = (\alpha - e_k)/\beta\}$ for the two MOs.

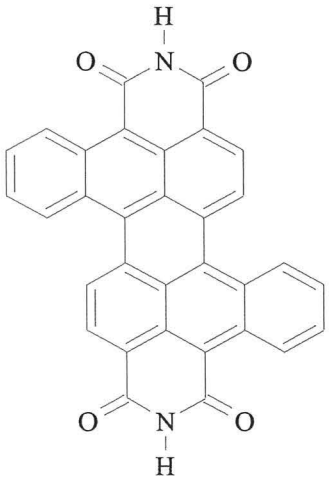
Due to the orbital degeneracy of the first excited term of benzene, naphthalene is used obtain a good β for fitting purposes. For naphthalene, and Δx = 1.236 and

$1/\lambda = 34{,}700\ cm^{-1}$. This gives $|\beta|/hc = 28{,}000\ cm^{-1}$ and $|\beta| = 3.48$ eV. Comparison of predicted and observed longest-wavelength absorptions for a large number of benzenoid hydrocarbons shows only fair agreement with experiment, with many deviations, σ, of a few thousand cm^{-1}. Improved agreement ($\sigma \sim 600\ cm^{-1}$) is obtained if $1/\lambda = |\beta|\Delta x/hc + a$ is used. A least-squares fit of benzenoid hydrocarbon data gives $a = 8200\ cm^{-1}$, $|\beta|/hc = 21900\ cm^{-1}$ and $|\beta| = 2.72$ eV.[426]

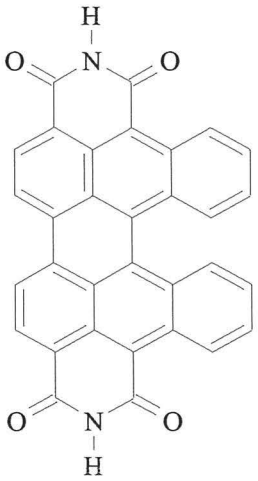
4.3.3 Modern approaches to the calculation of wavelength maxima

a) The Pariser-Parr-Pople (PPP) method

The two-electron PPP method gives a good account of the electronic spectra of many aromatic hydrocarbons.[42,43] The PPP method was recently found to be an important tool for predicting the absorption maxima of near-infrared photogenerator pigments,[436,437] which absorb at the gallium aluminium arsenide (GaAlAs) or indium phosphide (InP) diode laser wavelengths. Near-IR pigments have high photosensitivity in the visible, low dark decay, and excellent cyclic stability,[438] and can be used in xerographic printing, optical recording, and laser filters.



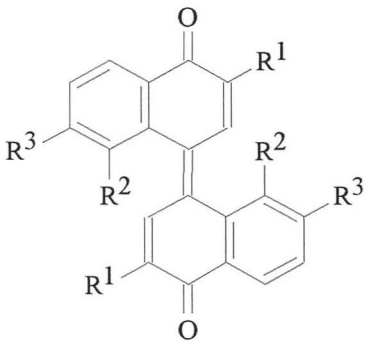
(119)



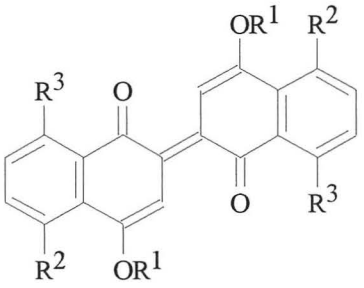
(120)

Aceanthrene green (λ_{max} 678 nm) exists as a 6.8:1 mixture of *trans* **(119)** and *cis* **(120)** isomers. Calculated wavelengths of 605 nm and 615 nm were obtained for the *trans* and *cis* isomers.[436]

The PPP method has also been recently used to calculate the electronic spectra of various 4,4'-binaphthylidene-1,1'-diones **(121)**, and 2,2'-binaphthyliden-1,1-diones **(122)**,[439] where R^1, R^2, and R^3 are alkyl groups.



(121)



(122)

The potentially tautomeric acid azo dye CI Acid Red 138 and its homologues substituted by different alkyl groups have been investigated by UV-VIS and PPP-MO techniques.[440] The calculated π-electronic transition for the hydrazone form corresponds very closely to the experimental absorption maximum. The tautomeric equilibrium of 2-phenylazo-1-naphthol in thirteen different solvents was also analysed using a previously developed quantitative approach.[441] PPP-MO calculations were in good agreement with the experimental spectral data for both tautomeric forms.

The solvatochromic and halochromic behaviour of a series of red to violet thienyl-2-azo disperse dyes in solution have been investigated using PPP-MO calculations with reasonable accuracy.[442] As expected, the attachment of increasingly electron-withdrawing groups to the thiophene ring resulted in bathochromic and hyperchromic shifts.

Electronic absorption spectra for a series of monoazo dyes derived from 2-amino-4-chloro-5-formylthiazole have been calculated without the inclusion of the 3*d* atomic orbitals on the sulphur atom by the PPP molecular orbital method.[443] The electronic absorption spectra of a series of 3-*N*,*N*-dialkylcarbamoyl-4,6-dinitro-phenylazo dyes have been discussed on the basis of sigma-frame-polarisation PPP-CI calculations.[444]

Additionally, *Pisystem for Windows*, a commercial program based on a modified PPP SCF-CI-MO model for the rapid and convenient calculation of excited state properties, especially colour, of planar π-conjugated dyes, has been reviewed.[445]

b) *Investigations of solvatochromism*

A method for calculation of the solvent effect on the molecular electronic ground state structure of ions and molecules has been presented using the Langevin dipole Monte Carlo approach.[446] Additionally, the shape of absorption bands of aggregates formed by two, four, and nine molecules of a polymethine dye was calculated by the Monte Carlo method.[447] Electronic structure evolution of an organic dye, coumarin 120, coupled to polar solvation dynamics, has been examined by combining *ab initio* electronic structure calculations and molecular dynamics (MD) simulations.[448] Different levels of solvent representation of a strongly polar merocyanine dye were considered:[449] a self-consistent reaction field (SCRF) approach; a QM/MM mixed model, and a solute-solvent supermolecular approach, each using ZINDO/S to obtain the electronic spectra. Only the supermolecular approach was able to predict the observed blue shift produced by the effect of the first solvation shell over the zwitterionic resonant form that predominates over the quinone in both systems.

c) *Configuration Interaction*

CS (Conformation Spectra) INDO MRCI (Multi-Reference Configuration Interaction) calculations have been used to study solvent induced changes of *trans-cis* isomerisation efficiency and electronic structure of the excited state of the DCM dye,[450] and the spectroscopic properties of 2,4,6-triaryl pyrylium cations.[451] Atom superposition and electron delocalisation (ASED-MO) calculations were applied to both unprotonated and protonated forms of 1-methyl-2-(4-hydroxystyryl)pyridinium betaine.[452] The lowest electronic transition was found to be $\pi \rightarrow \pi^*$, which is higher for MH^+ than for M. AM1-SCF and AM1-SDCI (minimised using AM1 with subsequent single point CI calculations including spin-adapted singly and doubly excited configurations with respect to the closed-shell Hartree-Fock determinant) calculations have been carried out for indigo and three conformers of *N,N'*-diacetylindigo in their *cis* and *trans* conformations, providing explanations for the phototropic behaviour of *N,N'*-diacetylindigo and the photostability of indigo.[453]

d) Ab initio calculations

A theoretical study of the electronic spectra of indigo, bispyrroleindigo and the H-chromophore, thought to be the basic structure leading to the intense absorption in the visible spectrum of the indigoid dyes, has been performed by means of the CASSCF (Complete Active Space SCF) method, and the multiconfigurational second-order perturbation theory (CASPT2).[454]

4.4 SUMMARY

Various methods for calculating electronic spectra of simple organic compounds have been addressed. These methods may be empirical, as in Woodward's rules for calculating electronic spectra of polyenes and α,β-unsaturated compounds. The effect of substitution on aromatic absorption has been examined, with particular emphasis on relationships with Hammett reactivity constants. These empirical methods are only valid for simple systems, necessitating the use of QM methods such as the FE MO, HMO and PPP π-electron methods. These π-electron methods are only reliable for planar molecules. For non-planar systems, such as cytological dyes, semi-empirical methods based on neglect of differential overlap must be used, the results of which will be discussed in *Chapters 6* and *7*. Results of higher-level *ab initio* calculations are reported in *Chapter 8*.

One of the most important turning points in the theory of colour was the introduction of **the triad theory of organic compounds** by Dähne. The theory of the ideal polymethine state equally existing between the aromatic state and the polyene state, triad theory, provides a comprehensive understanding of the chemistry and physics of conjugated organic compounds, and will be the focus of *Chapter 5*. Triad theory can be used in the design and modelling of new colorants (as shown in this thesis), organic metals and organic materials with nonlinear optical properties. It will be shown in *Chapters 6*, *7*, and *8*, that organic compounds can be separated into three groupings – the same identical groupings as those proposed by triad theory.

CHAPTER 5
THE TRIAD THEORY OF ORGANIC COMPOUNDS

5.1 AIMS

Chapter 4 has described how electronic wavelength maxima can be predicted, from the earliest empirical methods, such as Woodward's rules for polyenes and enones, to modern approaches such as PPP, MRCI and DFT methods. One of the most important turning points in the theory of colour was the introduction of the **triad theory of organic compounds** by Dähne. The theory of the ideal polymethine state equally existing between the aromatic state and the polyene state, **triad theory**, provides a comprehensive understanding of the chemistry and physics of conjugated organic compounds. Triad theory can be used in the design[455] and modelling of new colorants, organic metals and organic materials with nonlinear optical properties. Although triad theory is quintessential in explaining the correlations developed in *Chapters 6*, *7*, and *8*, the topic has received little or no attention in literature of recent years. This is especially true in organic chemistry textbooks, where the polymethine state is not discussed, although aromaticity and polyene theory are.

5.2 POLYMETHINES

As early as 1876, Witt[456] put forward his fundamental theory of chromophoric and auxochromic groups, later extended by Wizinger in 1927 to include antiauxochromic groups,[457] which still allows the description of the colour properties of aromatic compounds.[458] It was mainly due to König that the polymethine structure was recognised in 1926 as the true cause of colour. Organic compounds always intensely absorb at long wavelengths if they contain a polymethine structural unit:[459]

$\ldots (p \pm 1)\pi \ldots$ **(123)**

X—(CR)$_n$—X'

$n = p - 2 = 1, 3, 5 \ldots$; X, X' = atoms from groups 14, 15, or 16 (*i.e.* IV, V, or VI); R = any substituent; (CR)$_n$ = grouping replaced by atoms from groups 15 or 16.

A polymethine chain containing an odd number of carbon atoms, which may carry almost any substituents, and which has two atoms from groups 14, 15, or 16 at its extremities,[460] contains one π electron more or, one π electron less than the number of atoms in the chain. This particular electronic structure leads not only to intense absorption at long wavelengths, but also to an alternating π-electron distribution with simultaneous equalisation of the bond lengths in the polymethine chain. For a given chain length, the absorption wavelength increases with decreasing substituent electronegativity, and with increasing symmetry of the π-electron distribution. The alternating π-electron density distribution associated with deep colours is equivalent to alternating charge distribution in the polymethine chain proposed by König. This alternation follows from QM calculations, *e.g.* from the FE MO method,[461] or the LCAO-MO method,[462] and has been experimentally determined by ^{1}H nmr measurements.[463,464] The alternation of charge, *i.e.* of the π-electron density, occurs irrespective of whether or not the molecule is charged.

The definition of the polymethines essentially includes all known colour theories. Thus in order to obtain $(p + 1)$ or $(p - 1)$ π electrons over a total of p chain atoms, the chain ends X and X' must have the properties of auxochromic (electron donating) and antiauxochromic (electron accepting) substituents. The definition also includes the idea proposed by von Bayer[465] and Willstätter[466] that the colour of meroquinoid compounds (*e.g.* di- and triphenylmethane dyes, which, according to conventional notation contain a quinoid and hydroquinoid benzene ring) is due to oscillation of valences between groups linked to one another.[467]

Two other groups of organic compounds containing π electrons can with equal right stand alongside the polymethines; these being aromatic compounds and polyenes.[468] Compounds of both groups can be coloured if π-system is sufficiently large. Aromatic compounds and polyenes always absorb at lower wavelengths than polymethines containing the same number of sp^2-hybridised atoms. Aromatic compounds and polyenes contain $(4n+2)$ π electrons distributed over $4n+2$ atoms, *i.e.* in general p π electrons over a total of p atoms **(124)**. Polyenes, however, differ from aromatic compounds due to bond length alternation.

$X=(CH\text{-}CH=)_mX'$ ... **(124)**

m = 0, 1, 2 ...; X, X' = atoms of groups 14, 15, or 16.

In the ideal state, the polyenes and aromatic compounds differ from polymethines in that their π-electron densities do not alternate. The possibilities of transitions between the three types of compound are practically unlimited.[467,468] A polymethine becomes more polyene-like as the symmetry of the π-electron distribution decreases, *i.e.* as it approaches one of the possible limiting structures **(125)** or **(126)** as a result of differences in the donor and acceptor actions of the substituents X and X';[469] polymethines change directly to polyenes on uptake or loss of one electron, or to aromatic compounds on formation of rings containing 6 π electrons, and *vice versa*.

$$\underset{\textbf{(125)}}{X=(CH\text{–}CH=)_mCH\text{–}X'} \longleftrightarrow \underset{m = 0, 1, 2 \ldots}{X\text{–}(CH\text{–}CH\text{–})_mCH\text{–}X'} \longleftrightarrow \underset{\textbf{(126)}}{X\text{–}(CH\text{=}CH\text{–})_mCH\text{=}X'}$$

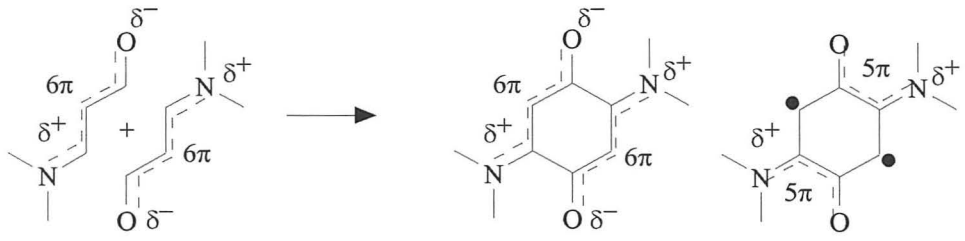
$$\overset{..}{\qquad} \quad \overset{\ldots\ldots\ (2m+4)\pi\ \ldots\ldots}{\qquad\qquad\qquad} \quad \overset{..}{\qquad}$$

5.2.1 Coupled polymethines

For a general colour system, it was found necessary to extend the term "polymethines" to include compounds that contain more than one polymethine structural unit having the general formula **(123)**. If these structural units interact with one another across π bonds, leading to coupling effects, this determines the colour and other physiochemical properties of the molecules. Compounds of this type are known as **coupled polymethines**. A special form of this coupling is the combination of two dipolar meropolymethine structural units, in which X = X' and n = 1, 3, 5 ..., giving rise to quadrupolar meropolymethines, with the compounds obtained from **(123)** with $X = NR_2$ and X' = O being quadrupolar merocyanines.

The physiochemical behaviour of 2,5-diamino-*p*-benzoquinones **(127)** can be fully understood only if the molecule is assumed to consist of two trimethine-merocyanine structural units **(128)**, which can couple *via* the C^1–C^2 and C^4–C^5 bonds. In principle, however, coupling of two dimethine radicals in accordance with formula **(129)** would also be feasible in 2,5-diamino-*p*-benzoquinone. Nevertheless, when there is a possibility of coupling either between two polymethine dyes [**(123)**; n odd] or

between two polymethine radicals [**(123)**; *n* even], resonance within the odd chains, *i.e.* within the "conjugated carbonyl" and amino groups, always predominates.



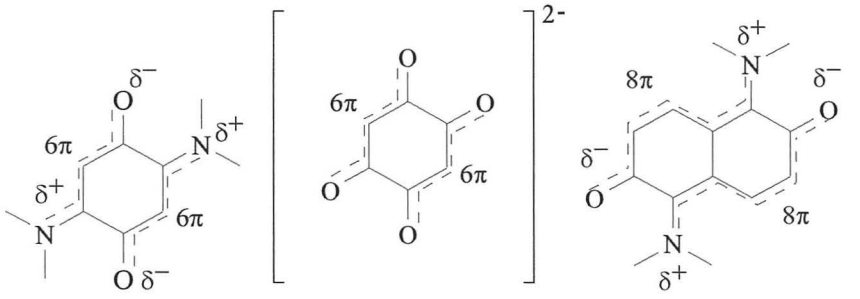
(128) **(128)** **(127)** **(129)**

The coupling concept was first confirmed on the basis of spectroscopic data.[470] It also follows directly from LCAO-MO calculations.[470d] If the calculation is carried out in the usual manner,[471] with equal values of the resonance integral β, the predicted wavelength is too high, and a remarkably low π-bond order is found for the bonds effecting the coupling of the polymethine structural units. The decrease in π-bond orders is accounted for by reducing the resonance integrals to ~0.5 β. From the coupling concept, the bond lengths within the polymethine chains should be largely equalised, and the coupling bonds lengthened. The predicted interatomic distances[470d] in 2,5-diamino-*p*-benzoquinione have in fact been largely confirmed.[470e] The tendency to form polymethines, particularly in more extensive chromophores, is more or less strongly opposed by a tendency towards aromatisation and the possibility of polyene formation. The coupling concept is a useful heuristic principle, with many dyes previously defying classification, readily fitting into the polymethinic system.

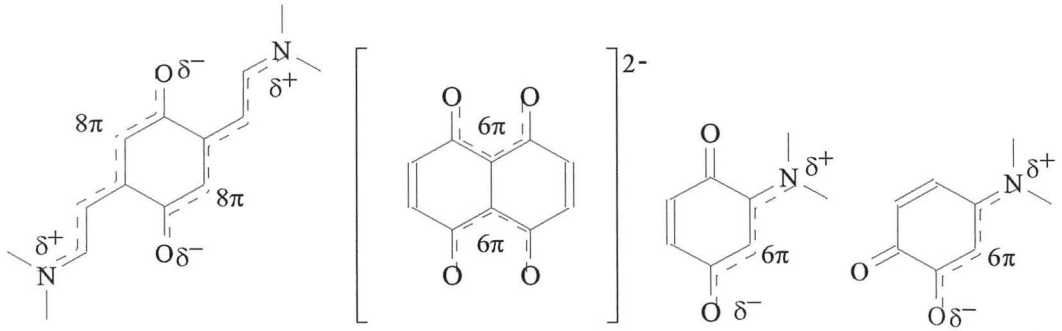
Table 5.1 LCAO-MO Predicted wavelengths of polymethinic dyes.

Compound	Experimental λ / nm	Calculated λ / nm
(127)	336, 488[a]	362, 485[b]
(130)	315, 496[c]	319, 455[d]
(131)	321, 581[e]	303, 572[b]
(132)	537, 645[f]	591[d]
(133)	585, 614[g]	550, 653[h]
(134)	258, 497[i]	280, 530[d]
(135)	266, 314, 503[i]	309, 510[d]

a) $-NH_2$ in DMSO.[472] b) Resonance loss = 0.5β.[470d] c) In 0.1 N NaOH.[473]
d) Resonance loss = 0.5β.[473] e) $-NH_2$ in DMSO.[470b] f) $-NEt_2$ in dioxane.[474]
g) In 0.1 N NaOH.[475] h) Without resonance loss.[473] i) –NMePh in DMSO.[473]



(127) (130) (131)



(132) (133) (134) (135)

The coupling principle is not restricted to compounds with polymethine structural units. Similar coupling effects occur between polymethinic and polyenic structural units. Another important property of the polymethines can be deduced directly from the concept of coupling: On gain or loss of one electron, simple polymethines are converted into polyenic or aromatic substances. Accordingly, the gain or loss of two electrons in coupled polymethines should yield less strongly coloured substances with aromatic or polyene-like behaviour. An example is indigo, the reduction and oxidation products of which (leuco- and dehydroindigo) have been studied in detail.[476]

5.3 TRIAD THEORY

As a theory concerning only dye chemistry, the polymethine concept existed for four decades almost unnoticed side by side with resonance theory, and the theory of the energetically stabilised aromatic state. The various theories were united by the concept of the existence of three ideal states for unsaturated organic compounds.[477-479] An essential preliminary was the introduction of the modern notation for the general polymethine structural unit ((**123**), *p* odd), representing a separate limiting structure

within the framework of resonance theory, making it possible to extend the definition of a polymethine to coloured (*e.g.* Wurster blue) radicals (**(123)**, p even),[477,480,481]

$$\ldots (p \pm 1)\pi \ldots \qquad \ldots \textbf{(123)}$$
$$X—Y_{p-2}—X'$$

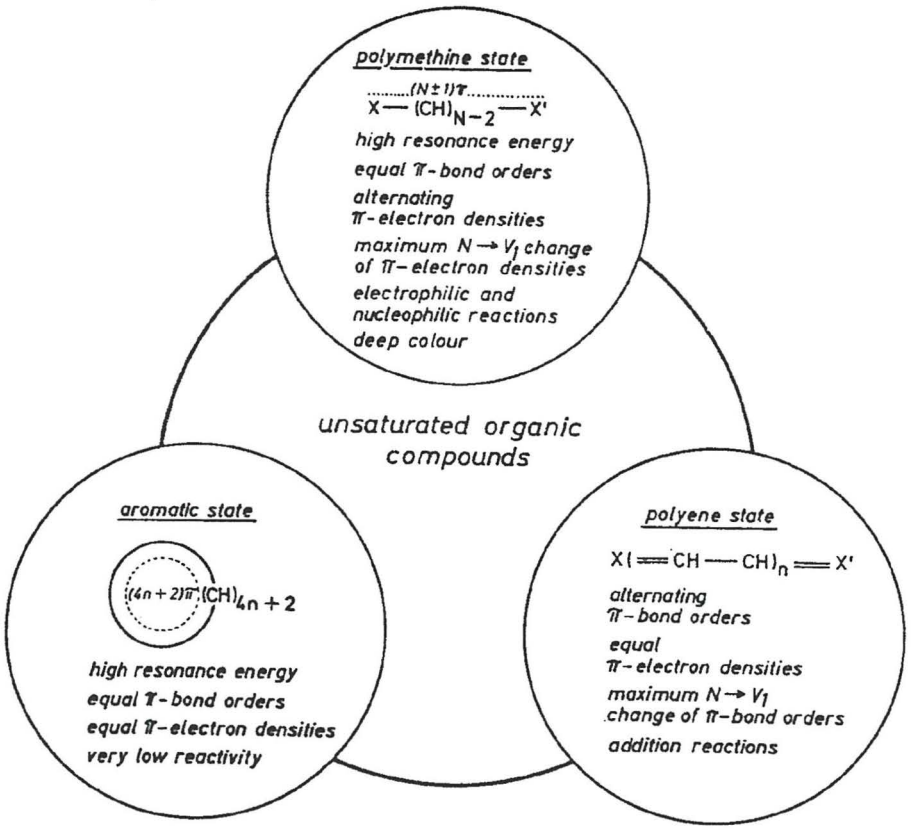
where X, X', Y = atoms from groups 14, 15, or 16.

The new interpretation was founded on the empirical deduction in 1966, based on experimental and HMO results, of an energetically stabilised ideal polymethine state, existing on a par with the energetically stabilised aromatic state.[479] Whereas resonance theory explained the colour of organic compounds in terms of substituent effects in aromatic and polyenic basic structures, general relationships between colour and constitution on a physiochemical basis could be derived in terms of three ideal states: the aromatic state, the polymethine state, and the polyene state.[482] This interpretation was justified by Zahradník,[483] who obtained three different bonding types by comparing experimental and HMO transition energies of several hundred unsaturated hydrocarbons (*Section 5.4*) The diversity of unsaturated organic compounds can be interpreted in terms of intermediates between these three ideal states, this approach being **triad theory**,[482] as depicted overleaf.

The properties of the ideal aromatic state are reduced to those in Hückel's original definition. The state is characterised by a high degree of stability and very low chemical reactivity, and is realised in cyclic hydrocarbons of $(4n + 2)$ atoms. Typical features are identical π-bond orders between all atoms corresponding to a bond order of about 1½ and identical π-electron densities of ~1 for all atoms.

The ideal polyene state is characterised by alternating double and single bonds in chain-shaped compounds exhibiting a relatively low degree of conjugation (resonance) in the ground state. On light absorption and electronic transition from the ground state to the first excited state ($N \rightarrow V_1$), a maximum change in the π-bond orders occurs, so that double bonds acquire increased single-bond character and *vice versa*. Polyenes, like aromatics, have a π-electron density per atom of ~1.

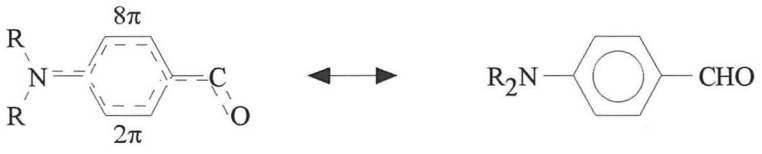
Figure 5.1 **Outline of triad theory[482] - *n* denotes zero and positive integers, while *N* stands for positive integers from 2 upwards.**



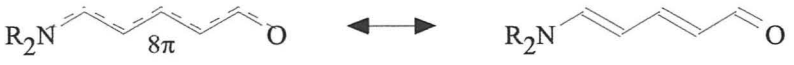
The ideal polymethine state, like the aromatic state, is characterised by high resonance energy and identical π-bond orders of about 1½ attributable to a maximum delocalisation of π-electrons in the ground state. In contrast to the aromatic state and the polyene state, polymethines feature a strongly alternating π-electron density distribution along the chain (allowing nucleophilic and electrophilic substitution[484]), redistributing on light excitation, in that atoms with a negative excess charge acquire a positive charge and *vice versa*. Where the numbers of π-electrons in compounds of comparable size are identical, compounds in the polymethine state will be deepest in colour; *i.e.*, they combine maximum transition probabilities,[485] with relatively low transition energies. Experimentally, cyanine dyes with odd-numbered methine chains, *e.g.* **(123)** with p = 3, 5, 7 ..., and Y = (CH), symmetrically terminated by nitrogen atoms according to X = X' = $NR_2^{0.5+}$, are particularly close to the ideal polymethine state.

A physical model for the ideal polymethine state, similar to the HMO model of the aromatic state, has been developed.[480-481] It is well known that unsaturated cyclic systems consisting of $2N$ carbon atoms will be aromatic, and thus energetically stabilised, only if the number of atoms is $4n + 2$. Antiaromatic cycles consist of $4n$ atoms. If a generalised ring system is dissected, and the available $2N$ π-electrons distributed unevenly over the two partial systems, two polymethinic chain shaped fragments are obtained, one containing $(N + 1)$ π-electrons and the other $(N - 1)$ π-electrons. On bisecting aromatics with $(4n + 2)$ atoms, the two polymethinic chains contain $2n + 1$ (odd) atoms and $2n$ or $2n+2$ (even) π-electrons. On the other hand, when antiaromatics with $4n$ atoms are bisected, less stable polymethine radicals having an even number of atoms and an odd number of π-electrons result.

Qualitatively, the concept of alternant and nonalternant polymethines enables organic dyes to be classified by their structural features. If polymethine charge alternation is maintained on branching, *e.g.* in the diphenylmethane dyes, the resulting alternant polymethine usually absorbs at slightly shorter wavelengths than those found for the unbranched basic chromophore. If polymethine charge alternation is perturbed in such a way that two centres of high and low electron density lie close together, nonalternant polymethines are formed, which always absorb at much longer wavelengths than the constituent basic polymethine chromophore.



(136)



(137)

Studies of intermediates between ideal polymethines, aromatics or polyenes cover a wide field of modern colour theory. These intermediates can be produced by an increasingly unsymmetrical π-electron distribution within an ideal polymethine unit,

e.g. by replacing one N atom at the chain end by an O atom as in (**136**) and (**137**). Such intermediate structures exhibit all the features that would be expected, *e.g.* reduced charge alternation and a hypsochromic shift of the longest-wavelength absorption band in conjunction with a diminished polymethinic character, accompanied by growing bond alternation, especially in the enones **(137)**.

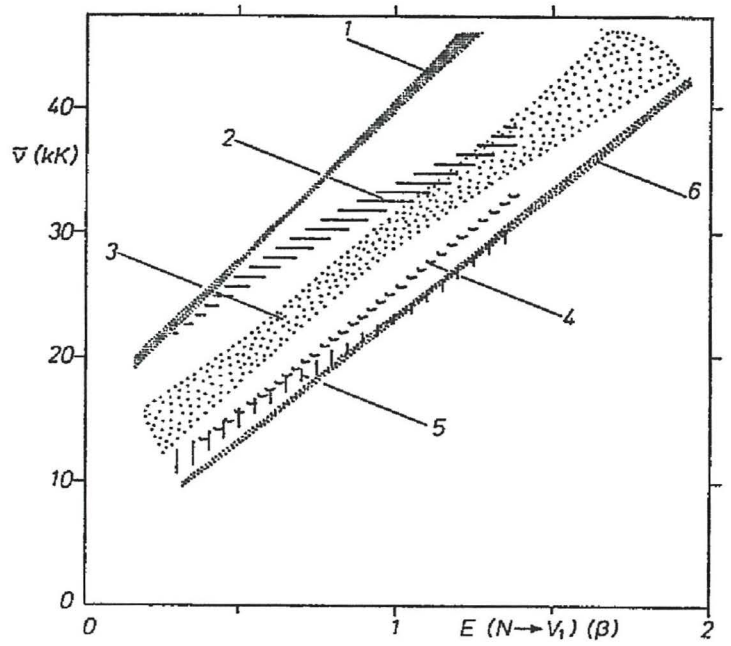
5.4 TRIAD THEORY AND ABSORPTION MAXIMA

General relationships between colour and constitution on a physiochemical basis can be derived in terms of three ideal states – polymethine, aromatic and polyene – triad theory. Statistically important correlations between calculated and experimental maxima were initially found by Heilbronner,[486] for dozens of conjugated organic compounds. Experimental and HMO calculated transition energies of several hundred unsaturated hydrocarbons were compared by Zahradník.[483] The results of the dependence of the wavenumbers (1kK = 10^3 cm^{-1}) of the first intensive absorption peaks on the energies of the $N \rightarrow V_1$ transitions are shown overleaf.[483] An almost linear dependence was noted for polyenes, aromatic (benzenoid) hydrocarbons, their heteroanalogues and for cationic hydrocarbons. The data showed splitting into subgroups even for formally very homogeneous groups of compounds (*e.g.* benzenoid hydrocarbons).[487] An almost linear dependence also occurred in vinylogous rows, derived from systems of manifold types. However, with non-alternating hydrocarbons and their analogues a correlation between experimental and theoretical data only existed in exceptional cases. Three structure types could be distinguished, corresponding to polyenic, aromatic and polymethinic compounds respectively:

(i) polyenes with distinct bond alternation,
(ii) aromatic (benzenoid) hydrocarbons, and
(iii) cationic hydrocarbons with delocalised charge.

Transitional hydrocarbons having both polyenic and aromatic character gave a sigmoidal curve located between the polyenic and aromatic correlations. Analogous behaviour was noted for odd α,α,ω,ω-tetramethylpolyene cations (4), odd α,ω-diphenylpolyene cations (5), and for tropylium and perinapthylium based cations (6), in agreement with the predictions of Koutecký and Paldua.[488]

Figure 5.2 Dependence of HMO calculated wavenumbers of the first intensive absorption maxima on the energies of the $N \rightarrow V_1$ transitions.[483]



Notation: (Translated from the original article)

1) Linear systems with alternating bonds - analogues and derivatives of polyenes.
2) Transitional compounds, *e.g.* α-phenylpolyenes, α,ω-dimethylpolyenes, cyclopolyenes and annulenes.
3) Cyclic hydrocarbons, *e.g.* aromatic hydrocarbons, hetero-analogues and derivatives.
4) Odd linear systems in ionic form - polyenyl cations.
5) Transitional compounds such as α,ω-diphenylpolyenyl cations.
6) Cyclic cations such as tropylium and perinaphthylium.

An explanation for the distribution of the data into groups corresponding to structure types was given by Koutecký *et al* using a limited CI method.[489] The contribution of electron repulsion to the theoretical excitation energy was found to largest for polyenes, smaller for benzenoid hydrocarbons and smallest for cationic hydrocarbons with delocalised charge. For a satisfying interpretation of the experimental data, electron repulsion has to be considered beside bond alternation.[490]

5.5 ELECTRON REPULSION

Simple MO theory at the Hückel level provides the organic chemist with a useful picture of phenomena related to electronic excitation. According to this theory, the ionisation potential, IP, given by IP = $-E_H$ (energy of the HOMO); the electron affinity, EA, given by EA = $-E_L$ (energy of the LUMO); and the excitation energy of the HOMO→LUMO transition, is related to the HOMO-LUMO energy difference.

Naively, one might expect the excitation energy to be equal to the orbital energy difference, but it has been known for some time that this is not so.[426,471,483,486-489] The linear correlation between excitation energies of the singlet L_a band of alternant hydrocarbons, corresponding to the HOMO-LUMO excitation, and the Hückel orbital energy differences does not pass through the origin.[426,471,483,486-489]

Explicit introduction of average electron repulsion is needed to remove this discrepancy. This is achieved in the SCF model, in which an orbital energy corresponds to the energy of an electron, which feels the field of the nuclei as well as time-averaged fields of the other electrons. In accordance with Koopmans' theorem,[87] IP = $-E_H$ and EA = $-E_L$. Excitation occurs in two steps, the first corresponding to removal of an electron from HOMO to ∞, requiring energy equal to IP. In the second step, the electron is brought back and placed in the LUMO. If the first step has not taken place, this would release an amount of energy equal to EA, with the total excitation energy being IP – EA. However, as the first step has been performed, a hole is present in the HOMO while the electron is being brought back from infinity. One electron repulsion contribution will thus be missing, facilitating the delivery of the electron from infinity to the LUMO. The amount of energy saved will be $J_{H,L}$ for a triplet state, or $J_{H,L} - K_{H,L}$ for a singlet state, with a singlet-triplet splitting of $2K_{H,L}$.

Comparison of azulene with anthracene is also instructive since their orbital energies are virtually identical; the relative small difference in triplet excitation energies is due to J being smaller in the larger molecule; the much larger difference in singlet energies is mostly due to the large difference in $2K_{H,L}$. More correctly, the difference in the lowest singlet-singlet excitation energies, and thus in colour and the small singlet-triplet separation can be both explained by the low value of $K_{H,L}$ for azulene (the spin-forbidden singlet-triplet absorption is weak and does not noticeably affect colour. The HOMO and LUMO are largely localised in different parts of space (small overlap and $K_{H,L}$) in the azulene molecule, only possible because azulene is non-alternant (polymethinic in structure), while they are localised in the same parts of the space (larger overlap and $K_{H,L}$) in the alternant anthracene molecule.

Table 5.2 Electron repulsion integrals in azulene and anthracene.[491]

	IP	EA	T	S	2K – J	J	2K
Anthracene	7.4[492]	0.6[493]	1.8[494]	3.3[494]	-3.5	5.0	1.5
Azulene	7.4[495]	0.7[493]	1.3[496]	1.8[494]	-4.9	5.4	0.5

Experimental values of ionisation potential (IP), electron affinity (EA), and triplet and singlet transition energies (T, S) in eV. Electron affinities are generally not known very precisely, but relative values are probably reliable for the above molecules. The values for 2K – J, J, and 2K, were determined from the experimental numbers according to:

$$2K - J = EA - IP + S, \qquad J = IP - EA - T, \qquad \text{and} \qquad 2K = S - T$$

HOMO-LUMO energy differences correlate linearly with experimental HOMO→LUMO singlet excitation energies, if $2K_{H,L} - J_{H,L}$ is a constant, or if it is linearly related to the HOMO-LUMO energy difference. $J_{H,L}$ is fairly constant, decreasing with increasing molecular size. For neutral alternant hydrocarbons, $2K_{H,L}$ also changes very little, and in such a way that $2K_{H,L} - J_{H,L}$ remains almost constant. For other hydrocarbons, 2K can be quite different, destroying the constancy of the $2K_{H,L} - J_{H,L}$ term, causing a deviation from the regression line - the theoretical benzenoid transition energies for example are too high. Polyenes and similar species with strong bond length alternation have unrealistically low excitation energies. It is relatively simple to correct for this deficiency, by allowing the resonance integral β to vary as a function of bond length (estimated from the calculated bond order). For triplet HOMO→LUMO excitation energies, where J is only a factor, only one linear regression line is obtained for various hydrocarbons, including non-alternants.[497]

5.6 SUMMARY

Dähne's theory of an ideal polymethine state existing equally between the aromatic state and the polyene state, triad theory, was justified by Zahradník, by comparing HMO calculated and experimental transition energies with each other. Electronic repulsion contributions to theoretical excitation energies have been discussed, with experimental excitation energies of the order polymethinic>aromatic>polyenic. The same generalisations can be made for NDO based HOMO-LUMO calculations performed on more complicated non-planar systems, as described in the next chapter.

CHAPTER 6
SEMI-EMPIRICAL WAVELENGTH MAXIMA

6.1 AIMS

The primary aim of the thesis was to calculate the colour and intensity of cytological stains, such as those used in the Papanicolaou cervical smear test. The colour of a dye can be obtained from knowledge of the electronic absorption spectrum. We have seen how wavelength maxima of simple enones and polyenes can be predicted with an accuracy of 5 to 10 nm. The use of π-electron methods to calculate wavelength maxima of planar systems has also been reviewed. As most of the dyes in this thesis are non-planar, NDO based methodologies will initially be examined.

This chapter describes how HOMO-LUMO gaps were calculated for a training set of several hundred compounds. Using perturbation theory, the singlet excitation energy is approximated by $(E_L - E_H) + (2K - J)$, where $E_L - E_H$ is the HOMO-LUMO gap. The semi-empirical methods contained in this chapter are not as drastic as the HMO ZDO based π-electron method used by Zahradník. As the 2K – J repulsion terms are ignored; the NDO based excitation energies are overestimated, leading to an underestimation of the wavelength maxima. The trends between calculated and experimental wavelength maxima can be interpreted in terms of Dähne's triad theory. The wavelength maximum of any cytological stains can be calculated, once its spectral character has been identified, with an accuracy of about 25 nm after correction. For accurate colour prediction, the initial model described in this chapter will need to be refined, so that an accuracy of about 10 nm or less is achieved.

6.2 GENERAL COMPUTATIONAL PROCEDURES

CNDO, INDO, MINDO/3, MNDO and AM1 QM calculations were performed using *HyperChem* on a Silicon Graphics 4D-25 Personal Iris Workstation (IRIX 3.4.0). PM3 and ZINDO/S QM calculations were performed using *HyperChem* running under *Microsoft Windows 3.1* on a 66 MHz 486-DX2 IBM-compatible PC. The molecules were built using sketching tools in the normal manner.

Ideally, if our method were to be used by cytologists, it would need to be as computationally inexpensive as possible, with a high turnover, and with the possibility of importing into a molecular modelling package such as *HyperChem*. Spectral transitions are sensitive to small changes in geometry, particularly on excitation or solvation.[498] Indeed, a 5 pm change in the Cu-S-Cys bond length of the blue copper protein plastocyanin results in a change in the *ab initio* multiconfigurational second-order perturbation theory CASPT2 excitation energies by as much as 2000 cm^{-1} (0.25 eV).[499] If semi-empirical methods were used for geometry optimisation as well as electronic structure prediction, more reliable correlations could be obtained. However, due to slow running hardware at Bangor, semi-empirical computational times were seriously compromised. For our initial model, molecules were therefore minimised using the MM+ force field and Polak-Ribiere conjugate gradient method to a final gradient of 0.003 kcal mol^{-1} $Å^{-1}$. This presents a serious defect in our methodology, due to the poor reproduction of unsaturated molecular structures. It is envisaged that with the proper hardware, improved reproduction of molecular geometries, with concurrent improvement in excitation energies, can be obtained using either MMP2 minimised molecules, or ideally, unsaturated organic molecules minimised using semi-empirical methodologies. A comparison of AM1 wavelengths for a representative sample of both MM+ and AM1 minimised compounds will be given in *Section 6.5*.

Larger systems such as **(13)**, were minimised by an initial steepest descents method to a gradient of 0.1 kcal mol^{-1} $Å^{-1}$, followed by Polak-Ribiere conjugate gradient minimisation to a final RMS gradient of 0.003 kcal mol^{-1} $Å^{-1}$. Single-point semi-empirical calculations were performed on the MM+ minimised molecules, and the energies of the HOMO, E_H and LUMO, E_L, noted. The wavelength maximum, λ_{max}, was calculated from the HOMO-LUMO gap, $E_L - E_H$ in eV, using the equation:

$$\lambda_{max}\,(nm) = 1240 / (E_L - E_H) \qquad \ldots \text{Equation 6.1}$$

In the following section, HOMO and LUMO energies for a series of molecules will be compared with ionisation potentials (*Table 6.1*) and electron affinities (*Table 6.2*) respectively to estimate the reliability of the HOMO-LUMO gaps.

Table 6.1 Comparison of semi-empirical HOMO energies with ionisation potentials, IP (both in eV).

Compound	IP / eV	CNDO	INDO	MINDO	MNDO	AM1	PM3	ZINDO
Benzene	9.25[500]	-13.9	-13.4	-9.2	-9.5	-9.6	-9.7	-9.2
Benzaldehyde	9.49[500]	-12.9	-12.0	-9.7	-9.9	-10.1	-10.1	-9.5
Acetophenone	9.26[500]	-12.5	-11.6	-9.3	-9.8	-10.0	-10.0	-9.4
Toluene	9.0[500]	-12.9	-12.4	-9.0	-9.4	-9.4	-9.4	-8.8
Ethene	10.5[500]	-15.7	-15.5	-10.2	-10.1	-10.3	-10.5	-10.3
(*E*)-But-2-ene	9.1[500]	-12.9	-12.6	-9.5	-9.8	-9.6	-9.6	-9.0
1,3-Butadiene	9.0[500]	-12.5	-11.9	-8.7	-8.9	-9.1	-9.1	-8.4
1,3-Pentadiene	8.7[500]	-11.7	-11.2	-8.5	-8.8	-8.8	-8.9	-8.1
1,3,5-Hexatriene	8.3[500]	-10.8	-10.2	-7.9	-8.3	-8.4	-8.5	-7.6
1,3,5,7-Octatetraene	7.8[500]	-9.8	-9.2	-7.4	-7.9	-8.0	-8.6	-7.5
Cyclopropane	10.0[500]	-15.5	-14.9	-10.3	-11.6	-11.5	-11.8	-12.3
Methane	14.4[501]	-19.6	-19.4	-13.3	-13.8	-13.3	-13.5	-15.9
Ammonia	10.9[501]	-16.5	-14.3	-10.0	-11.3	-10.8	-10.0	-11.7
Water	12.6[501]	-17.8	-16.3	-12.7	-12.2	-12.5	-12.3	-12.8
HF	16.0[501]	-21.0	-19.8	-14.6	-14.5	-13.8	-15.9	-16.1
Fluoroethyne	11.3[501]	-17.0	-16.2	-10.6	-11.3	-11.4	-11.6	-11.1
Cyclopentadiene	8.6[500]	-11.8	-11.2	-8.5	-8.9	-8.9	-8.9	-8.0
Cyclohexadiene	8.4[500]	-10.9	-10.5	-8.4	-8.8	-8.6	-8.7	-7.8
Furan	8.95[500]	-11.7	-11.0	-8.2	-9.4	-9.4	-9.5	-8.0
Pyrrole	8.2[500]	-11.7	-11.0	-8.3	-8.8	-8.9	-9.2	-7.9

Table 6.2 Comparison of semi-empirical LUMO energies with electron affinities, EA (both in eV).

Compound	EA / eV	CNDO	INDO	MINDO	MNDO	AM1	PM3	ZINDO
Ethene	-1.78[502]	5.3	5.9	2.1	1.3	1.4	1.2	1.6
1,3-Butadiene	-0.62[502]	2.9	3.4	0.9	0.2	0.2	0.0	0.4
Benzene	-0.72[503]	4.0	4.6	1.3	0.4	0.6	0.4	1.1
Styrene	0.15[504]	2.4	2.9	0.5	-0.3	-0.2	-0.3	0.2
Naphthalene	-0.19[502]	1.9	2.5	0.3	-0.5	-0.4	-0.6	0.0
Biphenyl	0.13[504]	2.2	2.8	0.4	-0.4	-0.3	-0.5	0.2
Phenanthrene	0.31[503]	1.8	2.4	0.2	-0.6	-0.6	-0.7	0.0
Pyrene	0.56[503]	0.8	1.4	-0.3	-1.2	-1.1	-1.3	-0.6
Anthracene	0.56[505]	0.6	1.2	-0.3	-1.1	-1.1	-1.2	-0.6

6.3 RELIABILITY OF ORBITAL ENERGIES

Koopmans' theorem[87] equates the ionisation potential, IP, of an electron with the orbital energy E_H of the HOMO. It implies that not only are relativistic and correlation energies the same in the molecule and ion, but that there is no reorganisation of electronic structure or distribution on ionisation. Such approximations are unacceptable, but are commonly used. Koopmans' theorem[87] also equates the electron affinity, EA, of an electron with the orbital energy E_L of the LUMO. In addition to the assumption involved in Koopmans' theorem, this

approximation is made even less acceptable as there is no definite physical meaning for the virtual orbitals, the only mathematical constraint upon them being one of orthogonality between themselves and with occupied orbitals.

Ionisation potentials have been obtained from UV photoelectron spectroscopy (adiabatic IPs from the highest electron energy within a group, and vertical IPs from the energy of the most abundant electrons, *i.e.* from spectral maxima), photoionisation and electron impact studies.[501,502] Ionisation potentials obtained in *Table 6.1* were compared with more recent values, and were found to be reliable.[506] Various methods have been used to obtain experimental values of electron affinities of molecules. In the early 1960s, the first experimental determination of gas phase electron affinities for several aromatic hydrocarbons were reported based on the molar response of the electron capture detector,[507,508] later refined by measuring the temperature dependence of the molar response, and applying a kinetic model to the reactions.[509] The ECD method has been applied to the determination of the electron affinity of about a hundred molecules. In the 1980s, relative electron affinities of over a hundred molecules, including aromatic hydrocarbons, were obtained by measuring equilibrium constants for thermal charge transfer reactions.[510-514] Electron affinity data from the 1960s are considered unreliable. However, the similarity of the calculated orbital energies for anthracene and pyrene leads to the conclusion that the electron affinity for anthracene can be considered to be reliable.

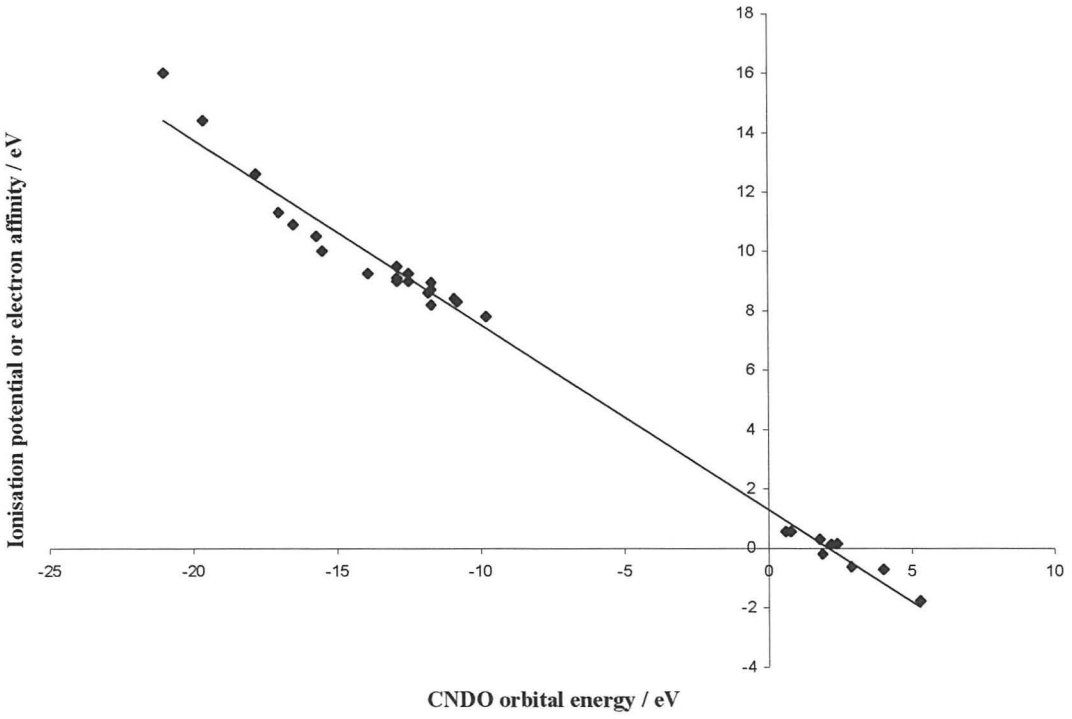
It is evident from *Tables 6.1* and *6.2* that CNDO and INDO orbital energies differ from ionisation potentials and electron affinities by a systematic error of ~3-4 eV, while MINDO/3, MNDO, AM1, PM3 and ZINDO/S orbital energies differ by ~0.2-0.5 eV. Random scatter in CNDO and INDO orbital energies (1 to 1.5 eV) is also larger than the scatter obtained using the more recent methods (0.2 to 0.4 eV). This trend of systematic and random errors is to be expected, with the superior methods involving less neglect of interelectronic repulsions through differential overlap. Ionisation potentials and electron affinities were correlated against HOMO and LUMO energies respectively, as shown below in *Graphs 6.1* to *6.7* (pages 122-125). Simple linear regression was used to obtain straight-line relationships for each method (*Table 6.3*), the ideal case being $r^2 = 1$ for a perfect linear correlation with no scatter, together with a gradient, $\beta = 1$, and an intercept through the origin, $\alpha = 0$. The

regressions were determined to be statistically significant at the 0.001% level. The MINDO/3 method is the most accurate (closeness of α and β to the ideal case) and predictive (closeness of r^2 to the ideal case), followed by the PM3, MNDO and AM1 methods. The unexpected higher accuracy of the MINDO/3 method may be due to the choice of data, as the PM3, AM1 and MNDO methods, should produce more accurate fits by neglecting less differential overlap.

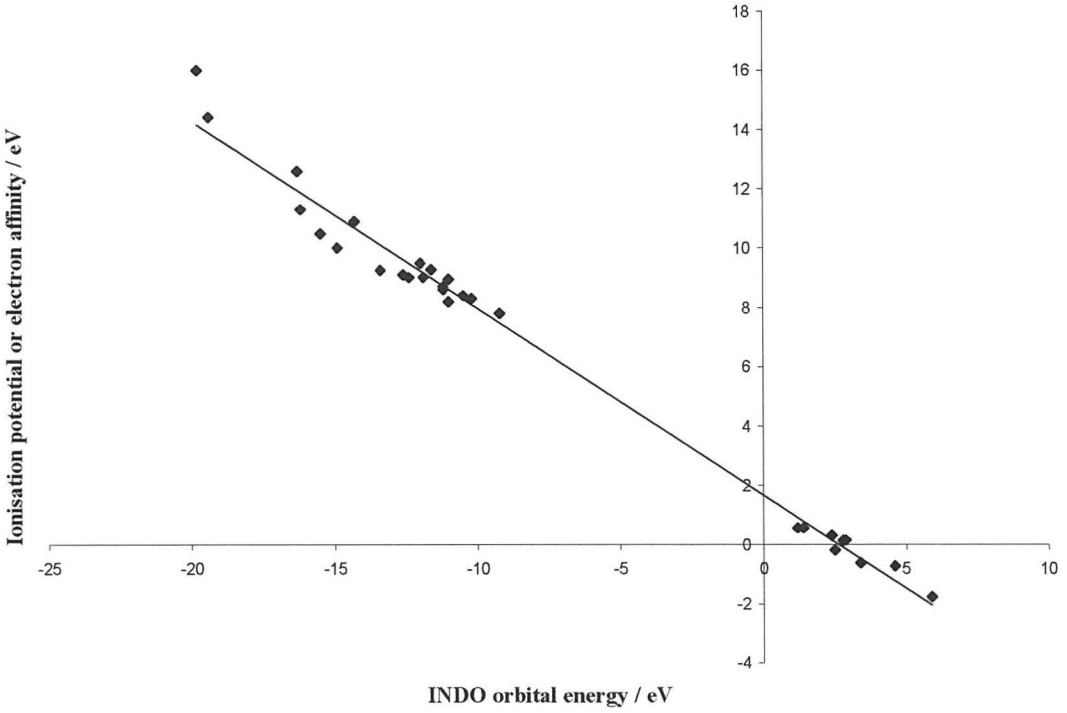
Table 6.3 Correlations between calculated HOMO and LUMO energies, and ionisation potential and electron affinity respectively (n = 29).

Method	α	β	r^2	σ	F
CNDO	1.302 (± 0.147)	-0.624 (± 0.013)	0.990	0.53	2573
INDO	1.669 (± 0.156)	-0.631 (± 0.014)	0.988	0.58	2134
MINDO/3	0.305 (± 0.128)	-1.001 (± 0.016)	0.994	0.42	4127
MNDO	-0.513 (± 0.166)	-1.038 (± 0.019)	0.991	0.50	2870
AM1	-0.470 (± 0.200)	-1.039 (± 0.023)	0.987	0.61	1961
PM3	-0.566 (± 0.169)	-1.028 (± 0.020)	0.991	0.51	2773
ZINDO/S	0.298 (± 0.194)	-0.961 (± 0.023)	0.985	0.63	1808

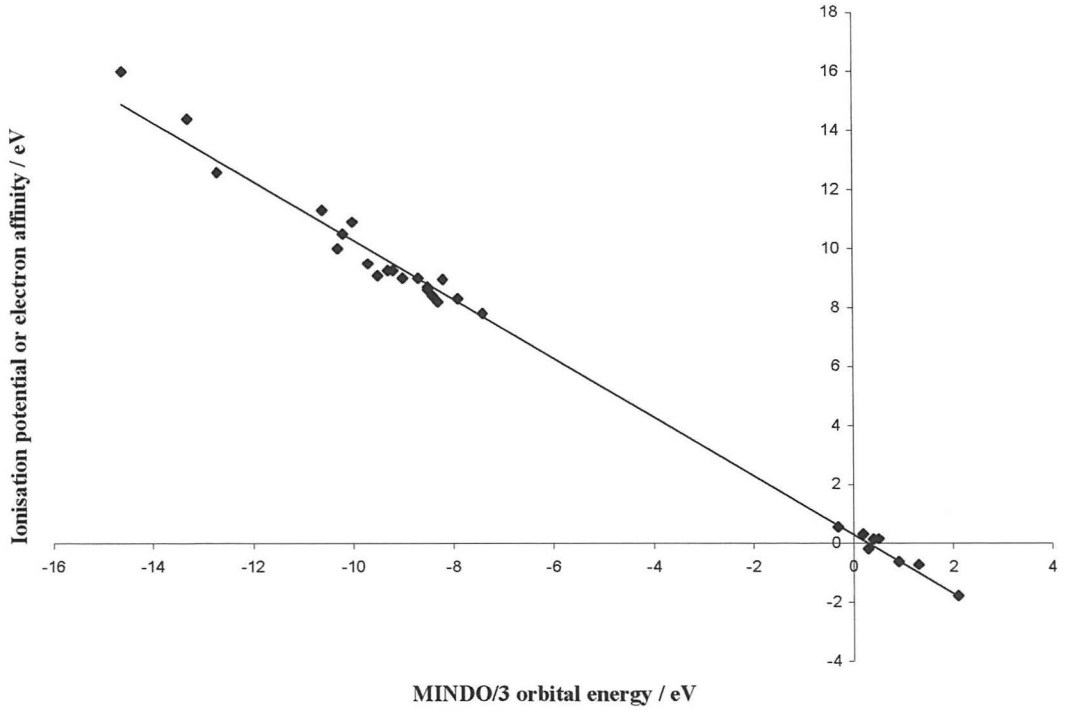
Graph 6.1 Correlation of CNDO orbital energies with ionisation potentials and electron affinities.



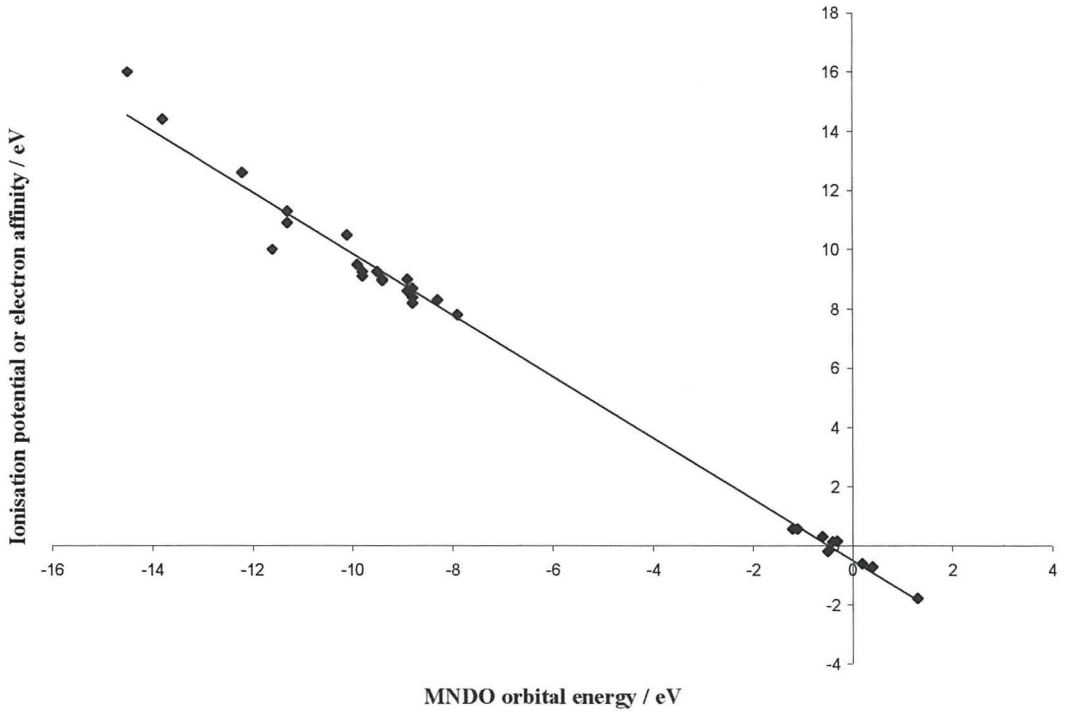
Graph 6.2 **Correlation of INDO orbital energies with ionisation potentials and electron affinities.**



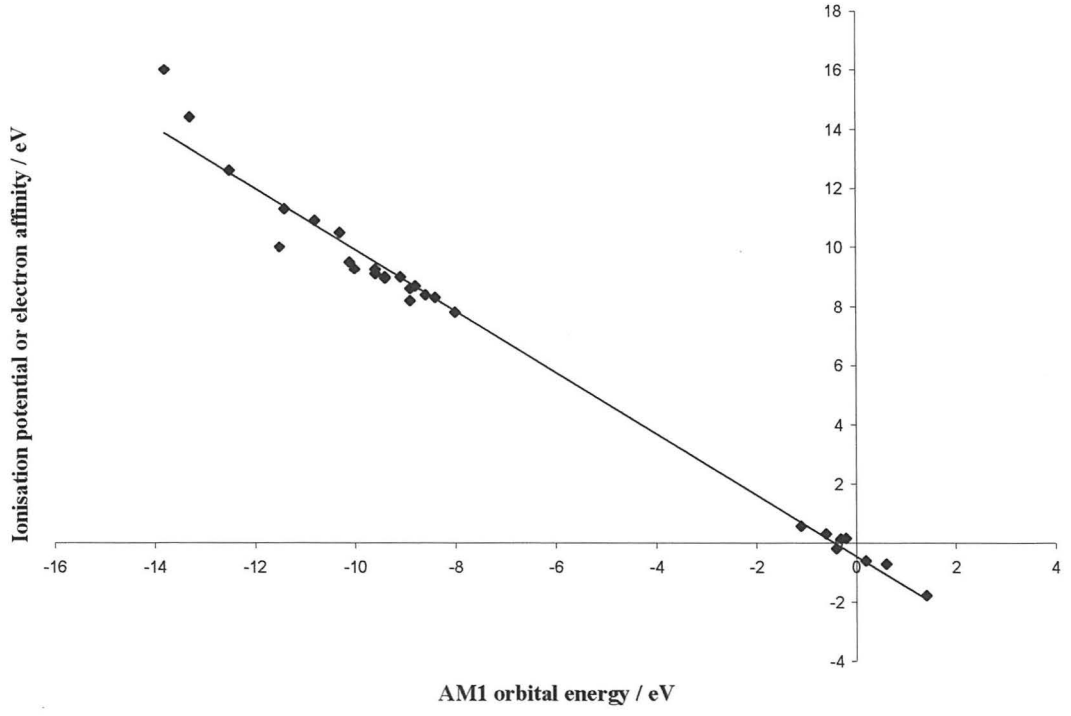
Graph 6.3 **Correlation of MINDO/3 orbital energies with ionisation potentials and electron affinities.**



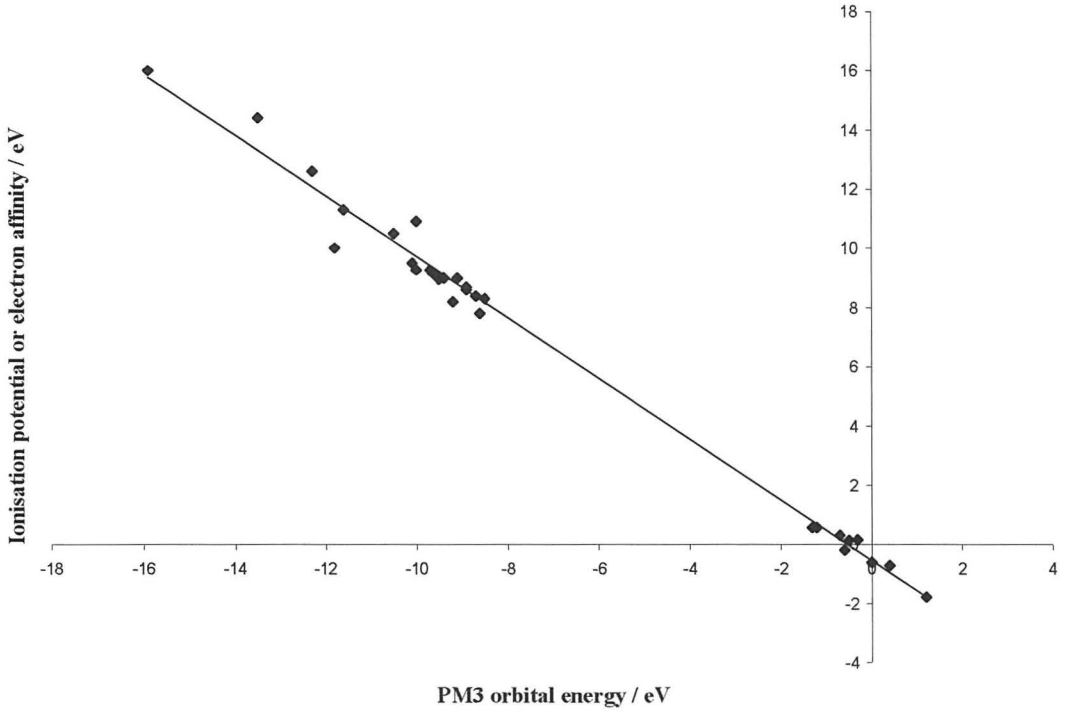
Graph 6.4 **Correlation of MNDO orbital energies with ionisation potentials and electron affinities.**



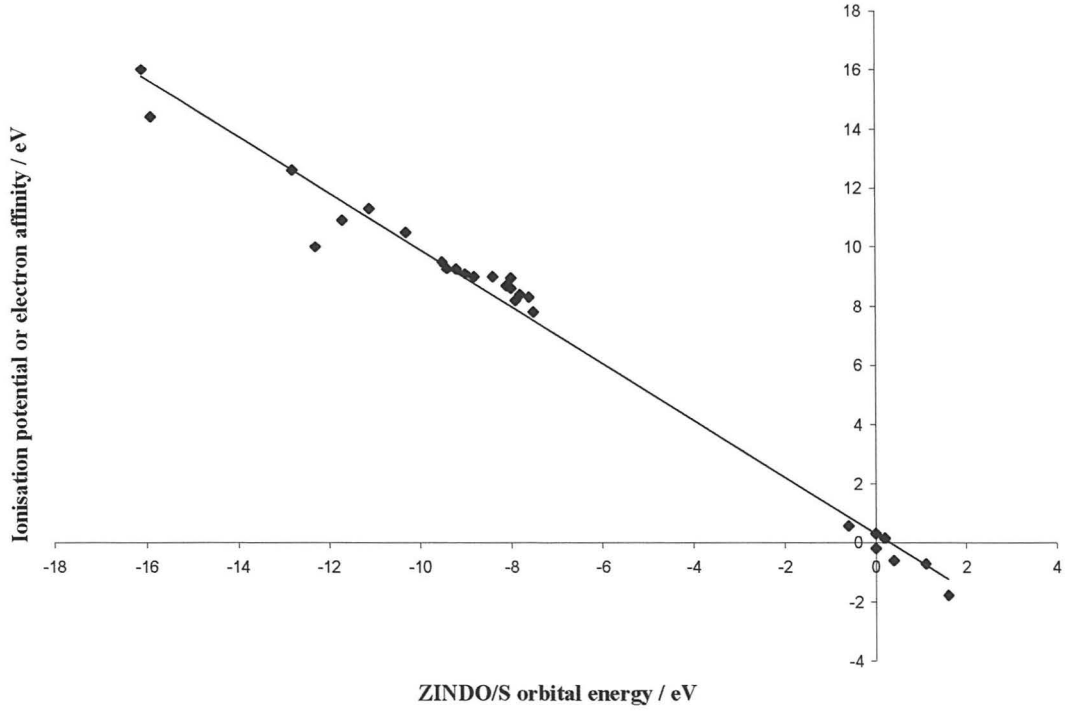
Graph 6.5 **Correlation of AM1 orbital energies with ionisation potentials and electron affinities.**



Graph 6.6 **Correlation of PM3 orbital energies with ionisation potentials and electron affinities.**



Graph 6.7 **Correlation of ZINDO/S orbital energies with ionisation potentials and electron affinities.**



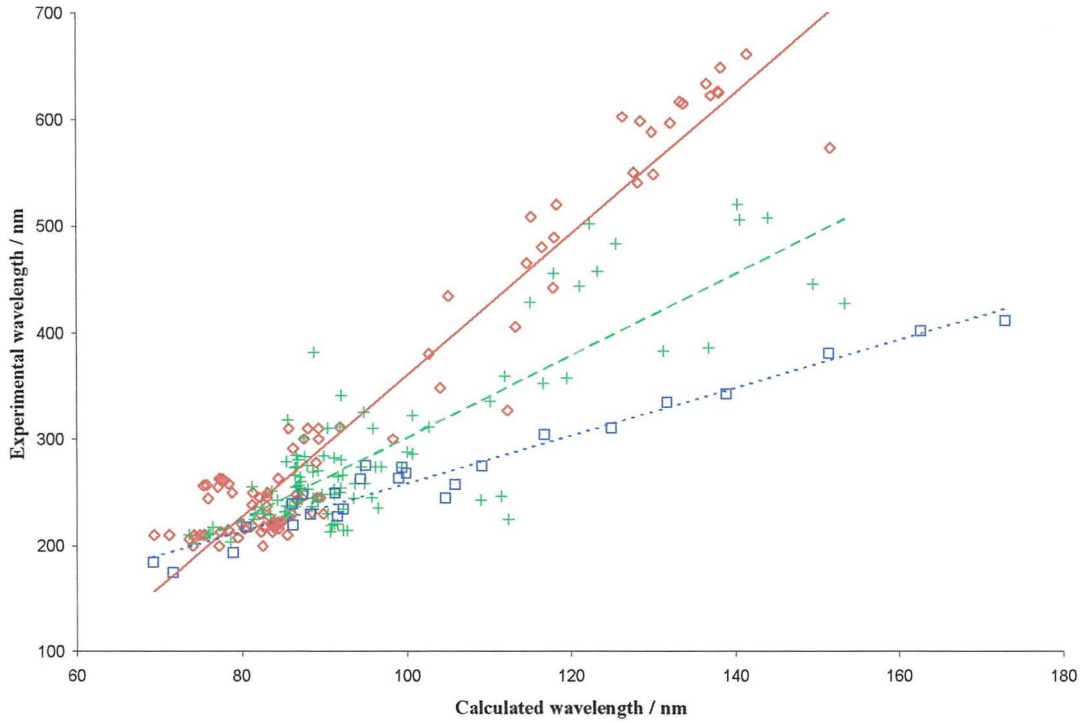
As the CNDO and INDO methods are not very accurate for calculating orbital energies, they should give poor HOMO-LUMO gaps, and therefore poor excitation energies and wavelength maxima. The MINDO/3, MNDO, AM1 and PM3 methods produce correlations with similar least-squares regression equations, with excellent accuracy and predictability, and should produce the most reliable wavelength maxima from HOMO-LUMO gaps. This hypothesis will be tested in the following section.

6.4 TRAINING SET DATA

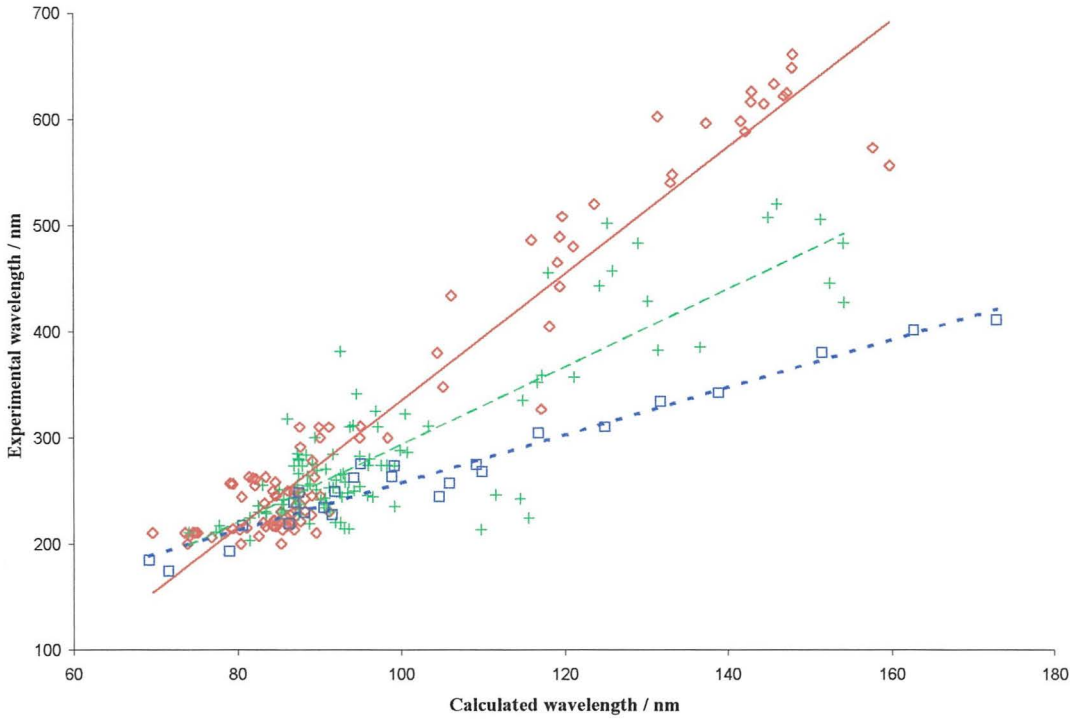
A training set of several hundred organic compounds (*Appendix 1*) was developed in order to predict trends between calculated wavelengths obtained from HOMO-LUMO gaps, and the experimental wavelength maximum. Calculated wavelength maxima, λ_{max}, obtained from semi-empirical HOMO-LUMO gaps, $E_L - E_H$ in eV, using *Equation 6.1*, are tabulated in *Appendix 2*. The experimental wavelength maximum, λ_{max}, for each training set molecule, was correlated against the calculated wavelength, as shown below in *Graphs 6.8* to *6.14*.

$$\lambda_{max} = 1240 / (E_L - E_H) \quad \text{... Equation 6.1}$$

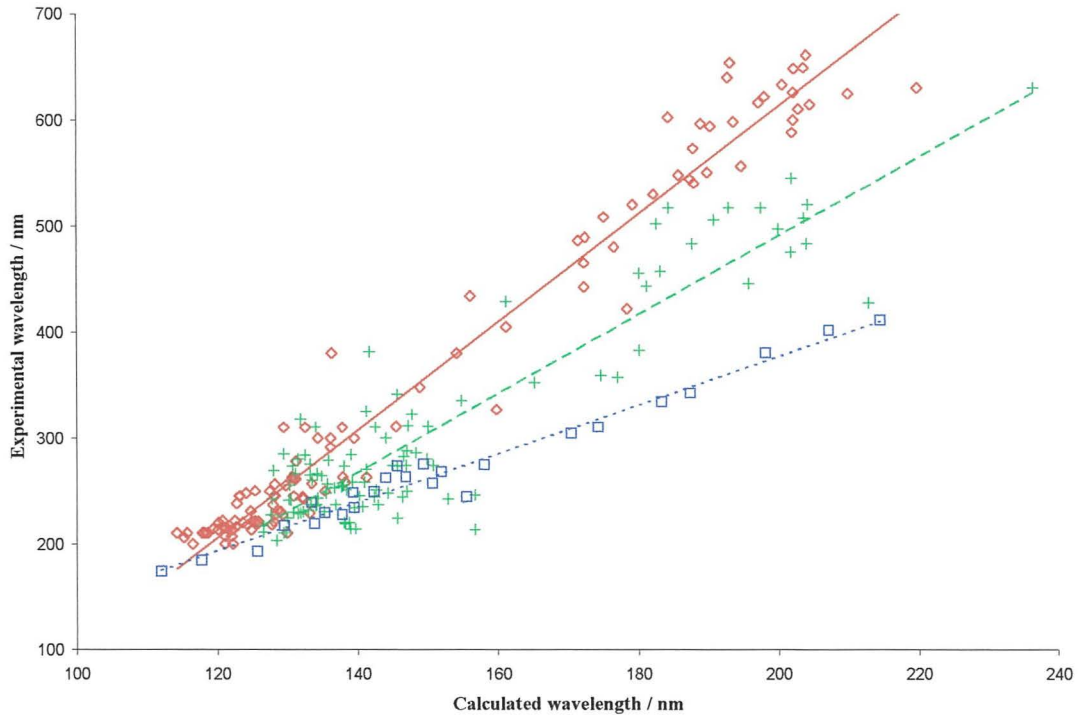
Graph 6.8 CNDO Training set.



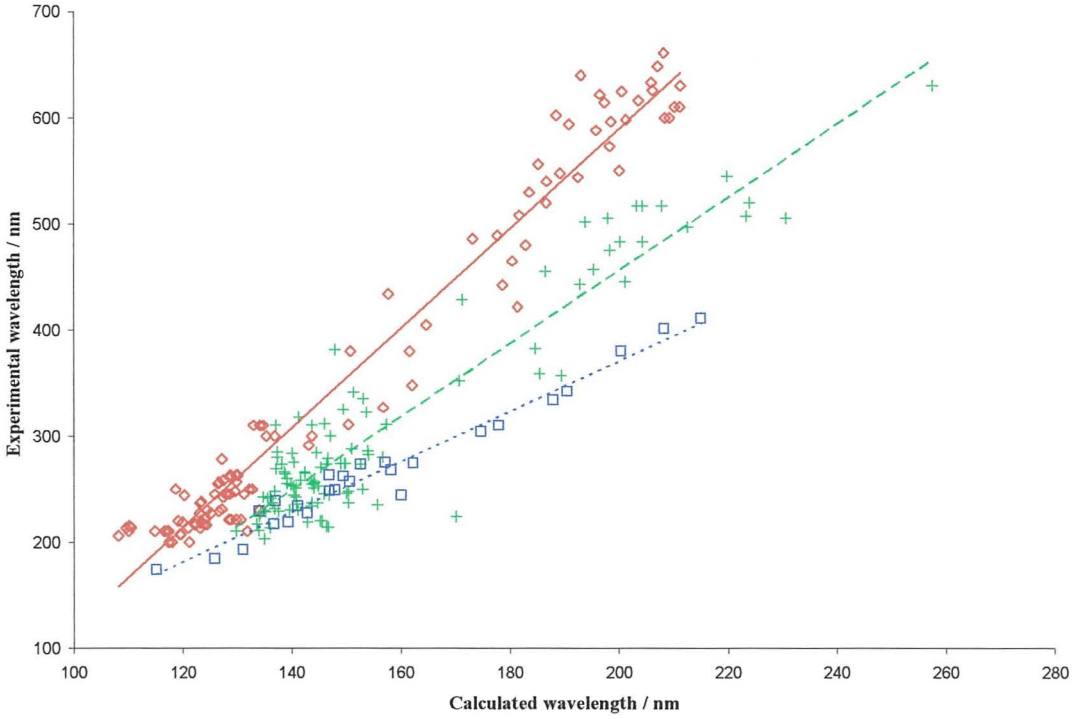
Graph 6.9 INDO Training set.



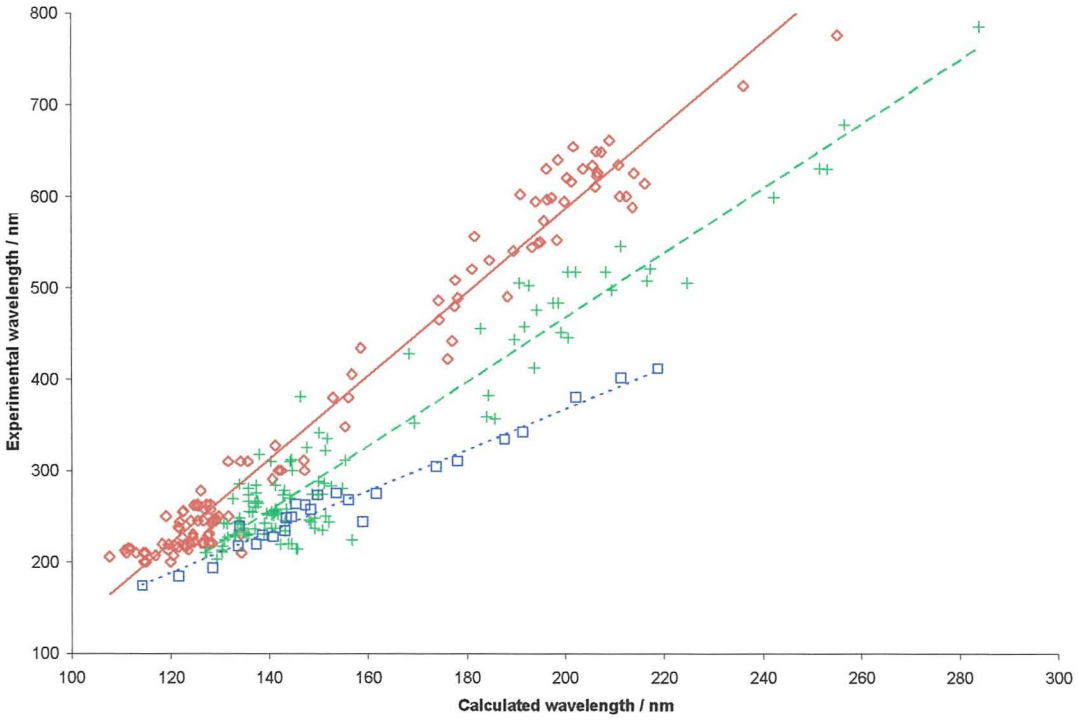
Graph 6.10 MINDO/3 Training set.



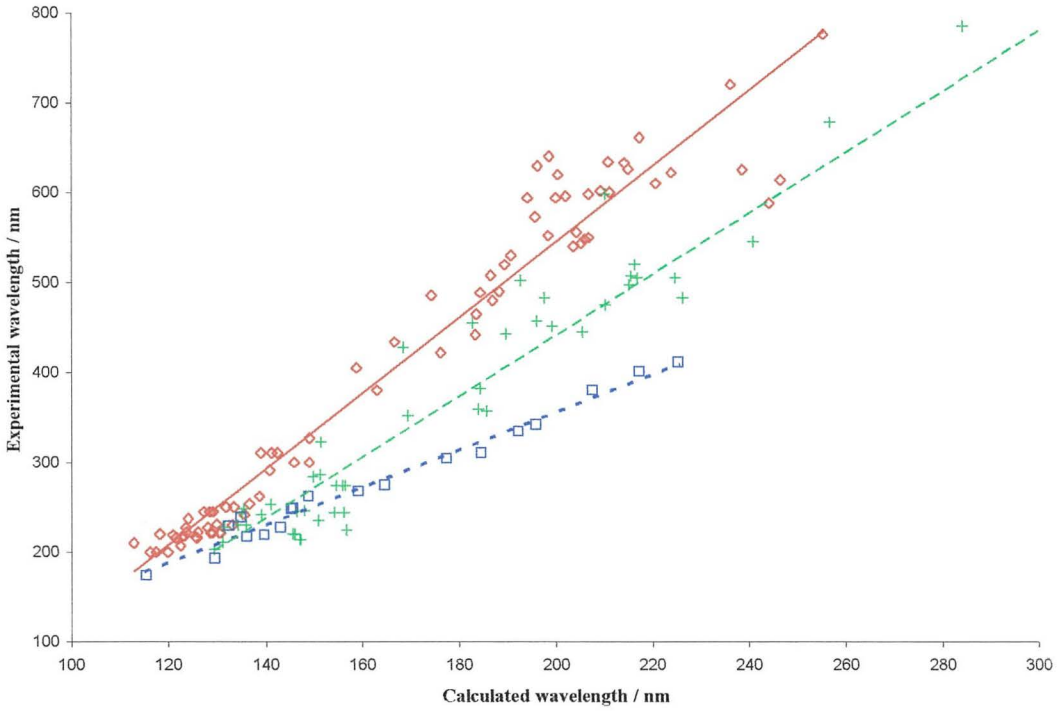
Graph 6.11 MNDO Training set.



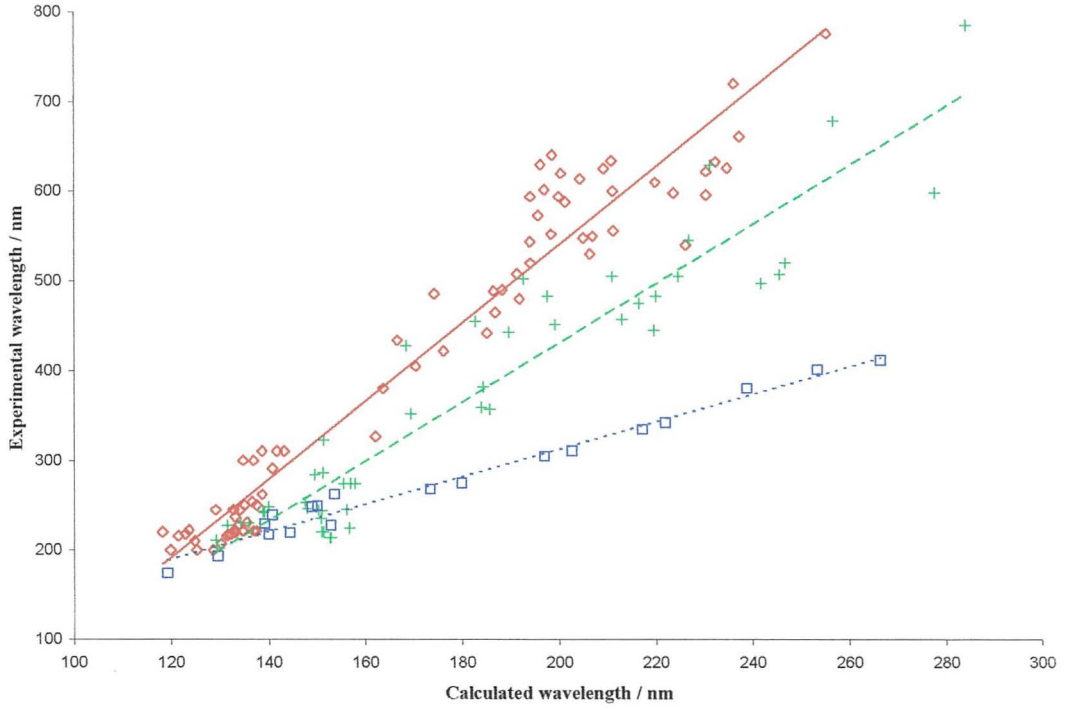
Graph 6.12 AM1 Training set.



Graph 6.13 PM3 Training set.



Graph 6.14 ZINDO/S Training set.



In complete agreement with the HMO results of Zahradník, three main correlations can be identified using Dähne's triad theory, namely due to the presence of polymethinic (◊), aromatic (+), and polyenic (□) compounds. Additionally, the existence of transitional compounds exhibiting various degrees of polyfunctional characteristics, can be also be identified, again in agreement with the HMO results of Zahradník. Polymethinic (◊) dyes have the lowest calculated wavelengths, while polyenic (□) dyes have the largest calculated wavelengths. Differences between the three correlations are small in the UV region, due to statistical scatter, but steadily increase at higher wavelengths. Differentiating between the characteristics of UV absorbers is therefore more difficult. While phenol, indene and benzoic acid obey the aromatic (+) regression, nitrobenzene exhibits polymethinic characteristics, probably due to the zwitterionic nature of the nitro group. Additionally, wavelength maxima of complementary disubstituted benzenes, such as *p*-nitroaniline and *p*–nitrophenol lie close to the polymethinic regression. Simple linear regression was performed to obtain least-squares regression equations for the prediction of experimental wavelengths, as shown in *Table 6.4* below.

Table 6.4 Correlations between experimental and calculated wavelengths.

Method	Character	α	β	r^2	n	σ	F
CNDO	Polymethinic	-301	6.60	0.934	97	37.4	1348.7
	Aromatic	-83.7	3.74	0.708	100	41.6	237.8
	Polyenic	33.4	2.25	0.957	26	13.2	529.6
INDO	Polymethinic	-276	6.13	0.945	97	33.8	1643.5
	Aromatic	-74.7	3.68	0.712	100	41.3	242.8
	Polyenic	32.8	2.25	0.957	26	13.2	531.9
MINDO/3	Polymethinic	-402	5.07	0.968	110	28.7	3260.2
	Aromatic	-253	3.71	0.843	107	38.7	565.6
	Polyenic	-81.8	2.29	0.971	26	10.9	795.8
MNDO	Polymethinic	-349	4.69	0.967	110	28.8	3168.8
	Aromatic	-234	3.45	0.878	107	34.6	752.8
	Polyenic	-103	2.37	0.966	26	11.7	688.4
AM1	Polymethinic	-326	4.55	0.973	119	28.0	4293.8
	Aromatic	-236	3.52	0.922	113	33.5	1310.2
	Polyenic	-81.9	2.25	0.967	26	11.6	702.1
PM3	Polymethinic	-297	4.21	0.953	78	38.0	1541.1
	Aromatic	-237	3.39	0.902	50	46.1	440.0
	Polyenic	-64.6	2.10	0.979	19	10.4	786.1
ZINDO/S	Polymethinic	-331	4.35	0.945	75	41.2	1248.6
	Aromatic	-230	3.30	0.895	49	47.6	400.7
	Polyenic	4.3	1.56	0.981	19	9.6	921.3

The experimental wavelength maximum of a dye can be extrapolated from the training set data, once the HOMO-LUMO gap has been calculated. The most reliable methods, with the largest r^2 values are the MINDO/3, MNDO, AM1 and PM3 methods. As the MINDO/3, MNDO and PM3 training sets are smaller than the AM1 training set, the reliability of these methods is diminished. It should be noted that electronic spectra are usually not composed of a series of sharp lines, but consist of broad bands, with band widths of ~50 nm, due to unresolved vibrational and rotational transitions. Predicted wavelengths obtained by extrapolation are accurate to within 25 nm for most compounds, and within 50 nm for transitional compounds. This accuracy is acceptable due to the band structure of electronic transitions, although further refinement may be possible. Defining the absorption characteristics of a dye can sometimes be more difficult than originally anticipated, *e.g.* the azo dye, methyl orange **(68)**, exhibits polymethinic character due to its complementary *para*-disubstitution, whereas most other azo dyes exhibit aromatic character.

The effect of molecular geometries on wavelength maxima will be described in the next section, followed by a discussion of possible applications of our HOMO-LUMO method to the design of UV absorbing sunscreens (*Section 6.6*), cytological dyes (*Section 6.7*), and near-IR absorbing pigments (*Section 6.8*).

6.5 EFFECT OF GEOMETRY ON WAVELENGTH MAXIMA

Spectral transitions are sensitive to small changes in geometry.[498] A 5 pm change in the Cu-S-Cys bond length of plastocyanin results in changes in CASPT2 excitation energies of by as much as 2000 cm^{-1} (0.25 eV).[499] A flaw in our MM+ approach is the poor reproduction of bond lengths in conjugated π-systems. The central bond in butadiene is ~1.47Å long, while the two terminal $CH{=}CH_2$ bonds are ~1.34 Å. If butadiene is modelled using the MM+ force field, each bond will be assigned the same bonding parameters, and in the equilibrium structure all C–C bonds will be almost identical in length. A similar situation arises for naphthalene.

One way of alleviating this problem would be to use a molecular mechanics method that accurately reproduces geometries of unsaturated systems, such as the MMP2 force field,[515] an extension of the MM2 force field, which incorporates a PPP MO

calculation into the force field. CIS-CNDO/S calculations have been performed on MMP2 and PM3 optimised geometries of [*n*]-helicenes (*n* = 3 to 9), with good agreement for long wavelengths.[498c] Ideally, the same semi-empirical method used in the electronic SCF calculation should also be used for geometry optimisation. Due to slow running hardware at Bangor, semi-empirical computational times were seriously compromised, necessitating the use of MM+ geometries as an initial model.

To illustrate the dependency of molecular geometries on excitation energies and wavelength maxima, single-point AM1 calculations were performed on MM+ and AM1 minimised geometries, as shown in *Table 6.5* below. The largest wavelength shifts can be seen for larger π-systems, which also have the largest wavelength maxima. It can also be seen that wavelength shifts are highest for polyenes, where bond length alternation is not reproduced for the MM+ minimised polyenes. In contrast, the terminal AM1 calculated $CH=CH_2$ bond length is 1.33 Å, with the central bond being 1.45 Å long, in excellent agreement with experimental results. Aromatic and polymethinic systems have less pronounced wavelength shifts. It is almost certain that new, perhaps more reliable correlations can in future be developed.

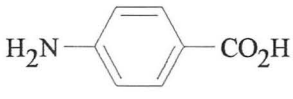
Table 6.5 Effect of geometry on AM1 calculated wavelength maxima.

Compound	Calculated wavelengths / nm		Δλ / nm
	MM+ geometry	AM1 geometry	
Acrolein	120.5	117.5	3.0
Cyclopentenal	120.7	119.6	1.1
Cyclohexenal	122.2	120.2	2.0
Anthraquinone	147.4	142.0	5.4
Azulene	177.7	173.6	4.1
Indigo	191.1	181.6	9.5
1,4,5,8-Tetraaminoanthraquinone	206.4	191.4	15.0
Benzene	121.6	121.5	0.1
Naphthalene	153.7	147.0	6.7
Phenanthrene	157.6	151.1	6.5
Anthracene	182.7	170.4	12.3
Tetracene	208.2	190.1	18.1
Ethene	104.7	103.4	1.3
(*E*)-1,3-Butadiene	133.8	126.8	7.0
Cyclohexadiene	137.9	134.5	3.4
1,3,5-Hexatriene	156.2	143.1	13.1
1,3,5,7,9-Decapentaene	187.8	163.2	24.6

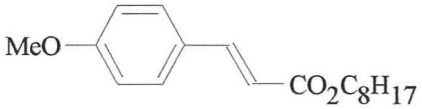
6.6 PREDICTING UV MAXIMA OF COSMETIC SUNSCREENS

Overexposure to sunlight risks short-term skin damage in the form of sunburn or long-term damage in the form of skin cancer. Most UV-A (320-400 nm) radiation, the least biologically harmful type, reaches the surface.[516] Most of the UV-B (280-320 nm) and higher energy UV-C (200-280 nm) rays are filtered out by the ozone layer, which is being constantly eroded by chlorofluorocarbons. Overexposure to UV-B causes most skin cancers, and has also been linked to a suppression of the human immune system. Sunscreens that block only UV-B and not UV-A do not adequately protect us from the risk of skin cancer.[516-518]

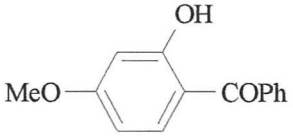
Melanocytes which produce melanin, an ultraviolet absorbing black pigment, act as the skin's natural barrier.[519,520] Artificial methods of protection are accomplished by the topical application of chemical sunscreens. Substances such as zinc oxide and titanium dioxide, physical absorbers, are opaque enough to reflect and scatter the incident radiation. Photon absorbing agents absorb radiation in the ultraviolet region, then undergo very rapid vibrational relaxation back to the ground state.[521-524] Any molecule for which vibrational relaxation to the ground state is the fastest pathway can act as a sunscreen. Most chemical sunscreens are aromatic compounds with a high degree of conjugation.[521,525,526] The first commercially available photon absorbing chemical sunscreen was *para*-aminobenzoic acid (PABA), which absorbs strongly in the UV-B region.[524] PABA esters, which are more soluble in oily lotions, cause a photoallergic response in 1-2 % of the population,[523] so PABA-free sunscreen preparations are being developed by cosmetic companies. Cinnamate esters, oxybenzone and avobenzone are also highly oil soluble and also strongly absorb in the UV-A range, offering better overall protection.[521,524] Salicylates have lower molar absorptivities, ε, meaning that a higher concentration is needed.[522]



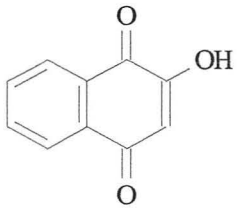
(138)



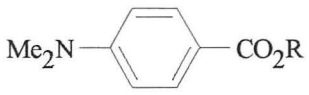
(139)



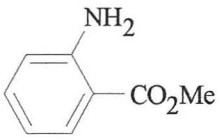
(140)



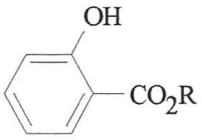
(141)



(142) R = $CH_2CH(Et)Bu$



(143)



(144) R = Ph

(145) R = Octyl

Wavelength maxima were calculated for PABA **(138)**, octyl *p*-methoxycinnamate **(139)**, oxybenzone (benzophenone-3) **(140)**, lawsone **(141)**, padimate O **(142)**, methyl anthranilate **(143)**, phenyl and octyl salicylates **(144)** and **(145)** respectively, using the HOMO-LUMO method developed in *Section 6.4*, the results of which are shown in *Table 6.6*. Good agreements (accurate to within 30 nm) are found between the predicted wavelengths obtained by extrapolation of the polymethinic regression equations, and the experimental wavelengths for the compounds. Although most of the above sunscreens obey the polymethinic regression, the extra conjugating double bond of the cinnamate **(139)** causes a deviation to aromatic behaviour. Improved agreement of sunscreen maxima was obtained by Walters *et al*,[527] using a single determinant configuration interaction method, discussed in *Chapter 7*.

Table 6.6 Semi-empirical calculations of sunscreen wavelength maxima.

Absorber	Observed λ^{526} / nm	Calculated λ / nm			Predicted λ / nm		
		MINDO/3	MNDO	AM1	MINDO/3	MNDO	AM1
(138)	284	139.1	144.6	143.1	303.2	329.2	325.1
(139)	311	150.0	157.4	155.5	303.5	309.0	311.4
(140)	325	141.2	149.5	147.7	313.9	352.2	346.0
(141)	335	154.9	153.2	151.9	383.3	369.5	365.1
(142)	310	142.5	137.1	144.3	320.5	294.0	330.6
(143)	341	145.7	151.4	150.2	336.7	361.1	357.4
(144)	300	144.0	147.1	144.8	328.1	340.9	332.8
(145)	310	134.0	143.8	140.4	277.4	325.4	312.8

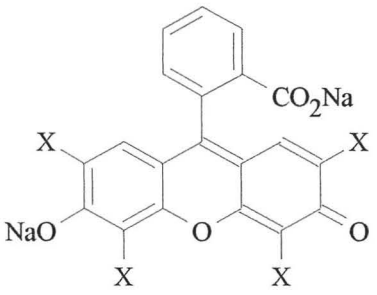
6.7 PREDICTING MAXIMA IN THE VISIBLE REGION

6.7.1 Initial studies of Papanicolaou dyes

Wavelength maxima of cytological dyes, which absorb in the visible (400 to 700 nm) region of the electromagnetic spectrum will be calculated in this section. Particular emphasis will be given to light green SF yellowish **(13)**, eosin Y **(29)**, orange G **(25)**,

and Bismarck brown Y **(54)**, which are used in the Papanicolaou cervical smear stain. Calculations on hæmatoxylin compounds will be discussed separately in *Chapter 9*.

Initial semi-empirical calculations (*Table 6.7*) were performed for molecules minimised using the MM+ force field to a final RMS gradient of 0.003 kcal mol^{-1} $Å^{-1}$. As a large number of valence orbitals were required for eosin Y, due to the presence of the bromine atoms, calculations were also performed for uranine **(146)**, the sodium salt of fluorescein, together with the fluorine **(147)** and chlorine **(148)** analogues of eosin Y. Good agreements, accurate to within 30 nm, were found between the experimental and predicted wavelengths of the Papanicolaou dyes. The polymethinic regression was used to predict the wavelength of light green SF yellowish, while the aromatic regression was used to predict the wavelengths of the other Papanicolaou dyes. Predicted maxima are all shifted to shorter wavelengths than expected.



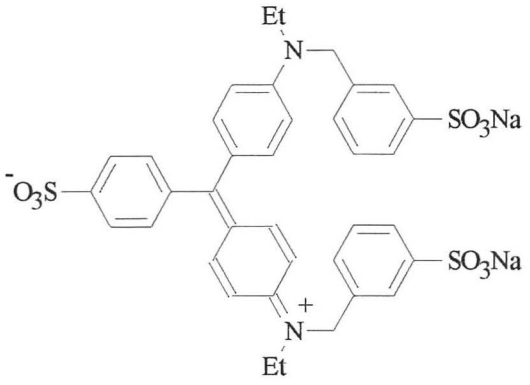
(146) X = H **(147)** X = F **(148)** X = Cl **(29)** X = Br

Table 6.7 Calculated wavelengths for MM+ minimised Papanicolaou dyes.

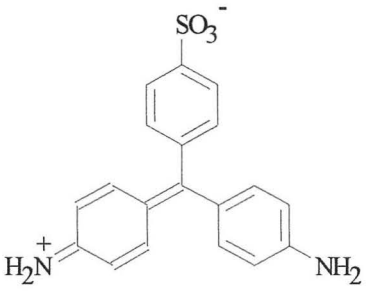
Dye	Observed λ / nm	Calculated λ / nm			Predicted λ / nm		
		MINDO/3	MNDO	AM1	MINDO/3	MNDO	AM1
(54)	457[528]	183.3	195.5	191.8	427.0	440.5	439.1
(146)	491[528]	184.4	203.4	200.7	431.1	467.7	470.5
(147)	517[529]	197.6	208.0	208.4	480.1	483.6	497.6
(148)	517[529]	193.1	204.4	202.4	463.4	471.2	476.4
(29)	517[529]		208.4	200.1		485.1	468.3
(13)	630[528]	219.9	211.4	203.9	712.9	642.5	601.7
(25)	475[528]	201.9	198.4	194.4	496.0	450.5	448.3

Due to the large size of the light green SF yellowish molecule, it was decided to use a simpler model **(149)** for the chromophore, as the *meta*-sulphonated aryl groups are insulated from the central chromophore by a methylene linkage, and should not affect the chromophore. The results obtained using the central chromophore are different from those obtained using **(13)**. The larger light green model **(13)**, typical of other

triarylmethanes, displays polymethinic characteristics, while the smaller central chromophore **(149)** obeys the aromatic regression. The change in electronic character on removal of the insulated sulphonate colligators cannot readily be explained.



(13)



(149)

Table 6.8 Calculated wavelengths for different light green models.

Model	Observed λ / nm	Calculated λ / nm			Predicted λ / nm		
		MINDO/3	MNDO	AM1	MINDO/3	MNDO	AM1
(149)	630[528]	236.6	257.7	251.9	624.8	655.1	650.7
(13)	630[528]	219.9	211.4	203.9	712.9	642.5	601.7

The calculations have also been applied to other cervical screening dyes used by Shorr, namely acid fuchsin **(12)**, Biebrich scarlet **(70)**, ponceau de xylidene **(71)**, and fast green FCF **(81)**. Predicted wavelength maxima accurate to about 40 nm are obtained by extrapolation of the aromatic regression equation. The difference in predicted wavelength maxima for **(70)** and **(71)** is especially noted. The predicted experimental wavelength for fast green FCF is approximately 80 nm greater than expected using the polymethinic regression equation. This may either be due to incomplete minimisation, or that fast green is a transitional compound - a mixture of "ideal" polymethinic and aromatic character.

Table 6.9 Predicted for cervical screening dyes used by Shorr.

Dye	Observed λ / nm	Calculated λ / nm			Predicted λ / nm		
		MINDO/3	MNDO	AM1	MINDO/3	MNDO	AM1
(12)	545[528]	202.0	219.8	211.5	496.4	524.3	508.5
(70)	505[528]	220.3	230.7	224.9	564.3	561.9	555.6
(71)	~480[530]	191.3	198.2	194.3	456.7	449.8	447.9
(81)	622[528]	†	†	226.4	†	†	704.1‡

† Calculations not performed.

‡ A value of λ = 555.6 nm is obtained using the aromatic regression equation.

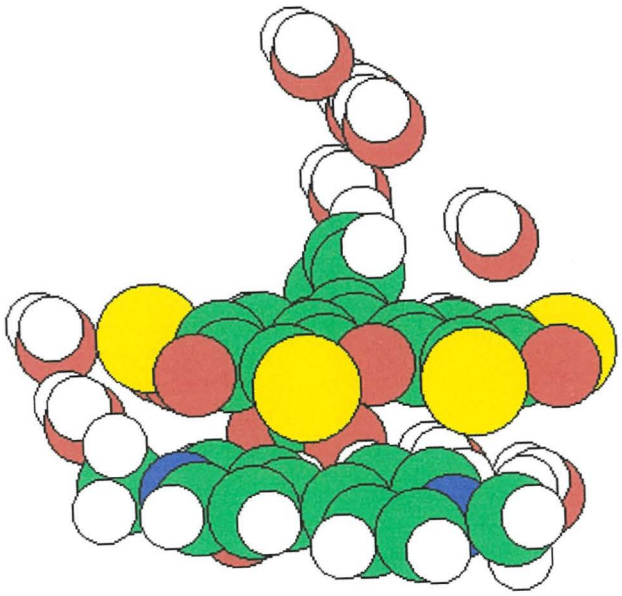
6.7.2 Crystallographic models

We have already noted that small changes in molecular geometries can have profound effects on calculated excitation energies and wavelength maxima. A search of the QUEST 3D Cambridge Structure Database[531] of the Chemical Database Service[532] at the Daresbury Laboratory, yielded crystallographic structures (see pages 138 to 140) for *N,N'*-dimethyl-4,4'-bipyridinium eosin tridecahydrate **(150)**,[533] *O*-tetramethyl-hæmatoxylin **(151)**,[534] diammonium Orange G tetrahydrate **(152)**,[535] bis(tetra-aqua-lithium) Orange G **(153)**,[535] hexa-aqua-magnesium Orange G dihydrate **(154)**,[535] DL-lysine Orange G dihydrate **(155)**,[536] L-lysine Orange G sesquihydrate,[536] DL-histidine Orange G trihydrate **(156)**,[536] and L-histidine Orange G trihydrate.[536] The downloaded Cambridge (*.cssr*) files were converted into *HyperChem .hin* files using the file translation program BABEL.[537]

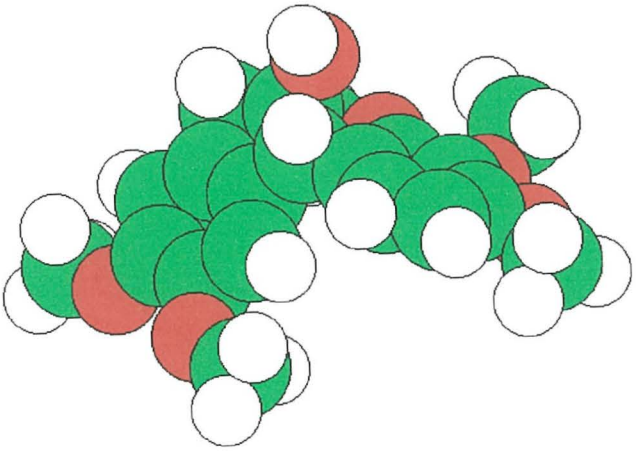
Table 6.10 Single-point AM1 wavelength maxima for crystallographic models.

Model	AM1 Calculated wavelengths / nm		$\Delta\lambda$ / nm
	MM+ geometry	Crystallographic geometry	
(150)	201.1	188.7	12.4
(151)	Not calculated	139.6	
(152)	194.4	173.0	21.4
(153)	194.4	175.7	18.7
(154)	194.4	174.1	20.3
(155)	194.4	173.9	20.5
(156)	194.4	174.1	20.3

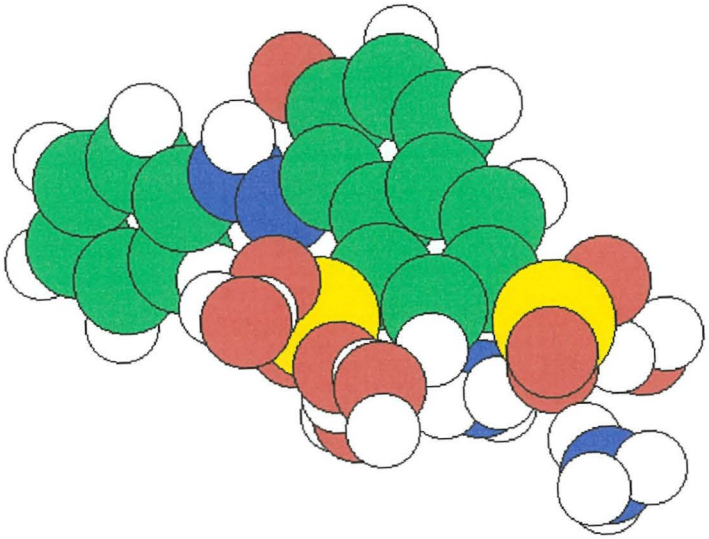
When single-point AM1 calculations were performed on supermolecular assemblies, such as **(150)**, the HOMO was localised on the dye, *e.g.* eosin Y, while the LUMO was localised on the other fragment, *e.g.* the bipyridinium cation. In order to calculate the energy of the LUMO for the dye, all other atoms including those of solvent molecules were deleted, leaving a residual dye molecule. Single-point AM1 calculations were therefore performed on these unoptimised residual molecules. In agreement with results obtained in *Section 6.5*, an improved molecular geometry causes a decrease in the calculated AM1 wavelengths. In particular, Orange G exists primarily as the keto tautomer, whereas our initial model assumed the existence of the enol tautomer. Improved predicted wavelengths should be obtained for the above dyes on application of new regression equations for AM1 minimised systems. Different charge-compensating cationic species do alter the Orange G wavelengths.



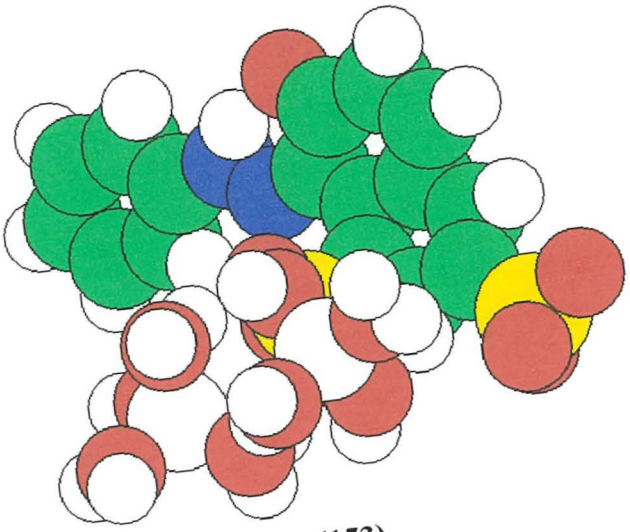
(150)



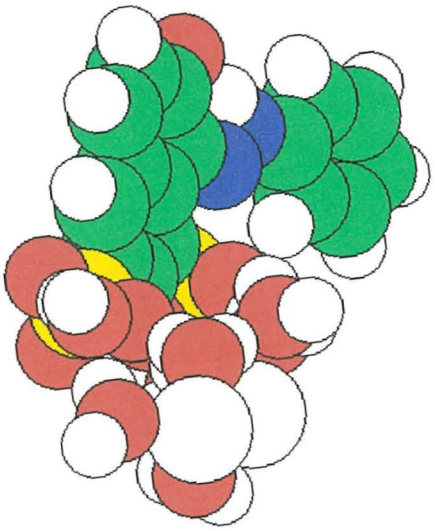
(151)



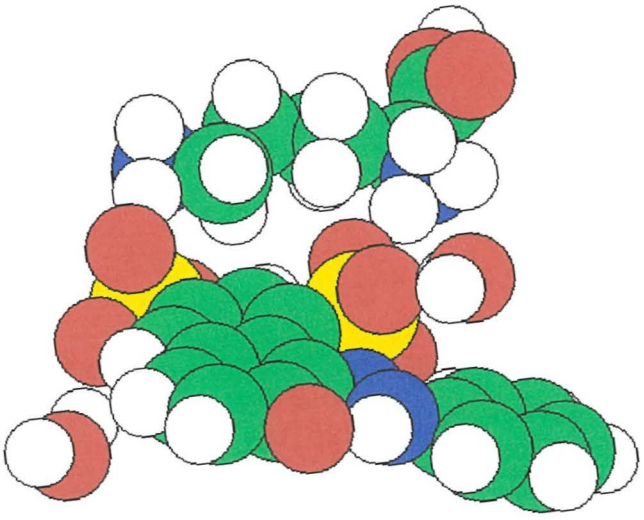
(152)



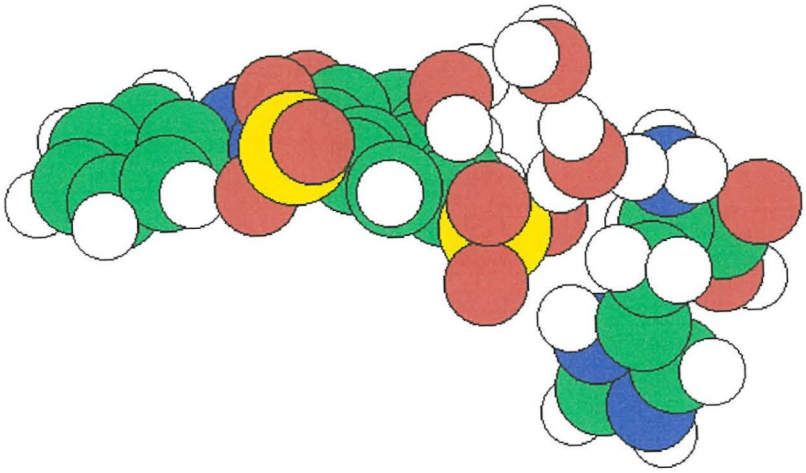
(153)



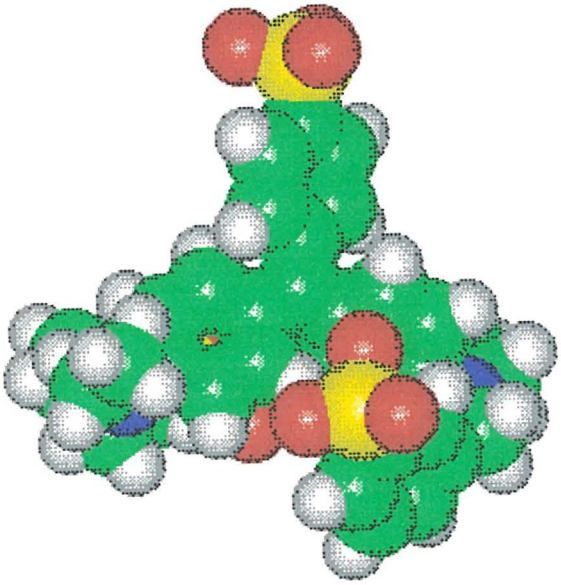
(154)



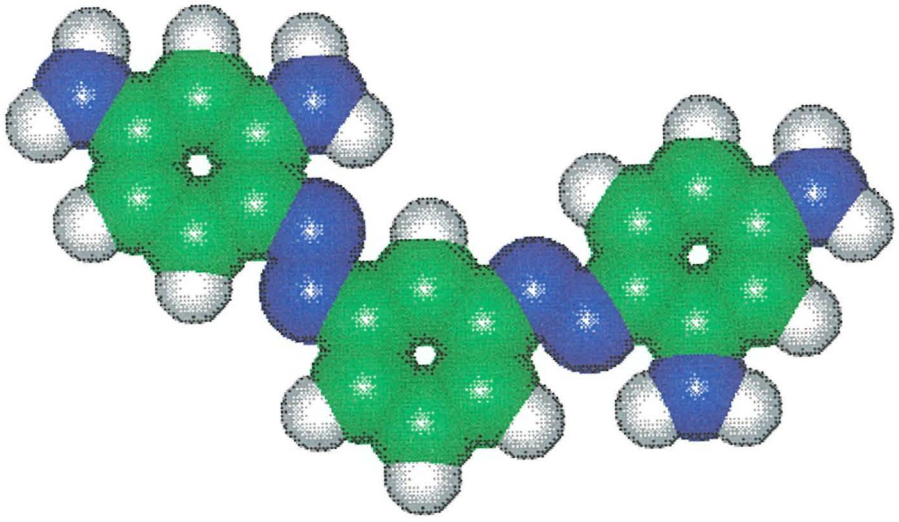
(155)



(156)



(13)



(54)

6.7.3 Other cytological dyes

Wavelength maxima of other important cytological dyes - acridine orange **(106)**,[538] acridine yellow G **(157)**, alizarin red S **(158)**,[539] auramine O **(159)**,[540] Azure A **(160)**,[541] Azure B **(161)**,[542] Azure C **(162)**, brilliant cresyl blue ALD **(163)**,[543] brilliant green **(95)**,[544] congo red **(85)**,[545] cresyl violet acetate **(164)**,[546] crystal violet **(55)**, guinea green B **(165)**, indigo **(4)**, malachite green oxalate **(166)**,[547] methylene blue **(11)**,[548] methyl green **(31)**,[549] methyl orange **(68)**, neutral red **(167)**,[550] orange II **(168)**, pararosaniline chloride **(33)**,[551] pyronin Y **(169)**,[552] safranin O **(56)**,[553] Sudan III **(44)**,[554] Sudan IV **(45)**, Sudan black B **(170)**,[555] thionin **(171)**,[556] and toluidine blue O **(172)**,[557] were calculated using the semi-empirical HOMO-LUMO method:

Table 6.11 Calculated wavelengths for other cytological dyes.

Dye	Observed λ[528] / nm	Calculated λ / nm			Predicted λ / nm		
		MINDO/3	MNDO	AM1	MINDO/3	MNDO	AM1
(4)	602	184.4	188.6	191.1	532.9	535.5	543.5
(11)	661	204.2	208.4	209.3	633.3	628.4	626.3
(31)	629	*	*	253.5	*	*	656.3
(33)	544	187.5	192.6	193.4	548.6	554.3	554.0
(44)	507	203.8	223.4	216.8	503.1	536.7	527.1
(45)	520	204.3	224.0	217.4	505.0	538.8	529.2
(55)	588	202.1	196.1	213.9	622.6	570.7	647.2
(56)	530	182.3	183.7	184.7	522.3	512.5	514.4
(68)	505	190.9	198.1	190.9	565.9¶	580.1¶	542.6¶
(85)	497	200.1	212.7	209.7	489.4	499.8	502.1
(95)	625	210.1	200.7	214.2	663.2	592.3	648.6
(106)	489	172.5	177.8	178.3	472.6	484.9	485.3
(157)	442	172.3	178.7	177.2	471.6	489.1	480.3
(158)	556	194.9	185.3	181.8	586.1	520.1	501.2
(159)	434	156.1	157.8	158.7	389.4	391.1	396.1
(160)	633	200.7	206.1	205.9	615.5	617.6	610.8
(161)	648	202.4	207.3	207.6	624.2	623.3	618.6
(162)	616	197.4	203.7	201.5	598.8	606.4	590.8
(163)	622	198.2	196.6	206.7	602.9	573.1	614.5
(164)	596	189.1	198.7	196.6	556.7	582.9	568.5
(165)	618	255.6	234.8	226.4	695.3¶	576.1¶	560.9¶
(166)	614	204.7	197.5	216.4	635.8	577.3	658.6
(167)	540	188.1	186.9	189.7	551.7	527.6	537.1
(168)	483	204.2	204.4	198.8	504.6	471.2	463.8
(169)	548	185.9	189.3	194.7	540.5	538.8	559.9
(170)	598	*	*	242.5	*	*	617.6
(171)	598	193.8	201.4	197.5	580.6	595.6	572.6
(172)	626	202.3	206.4	206.9	623.7	619.1	615.4

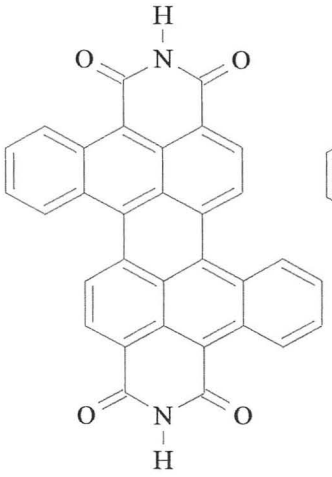
* Wavelength not calculated.

¶ Transitional character halfway between ideal polymethinic and ideal aromatic.

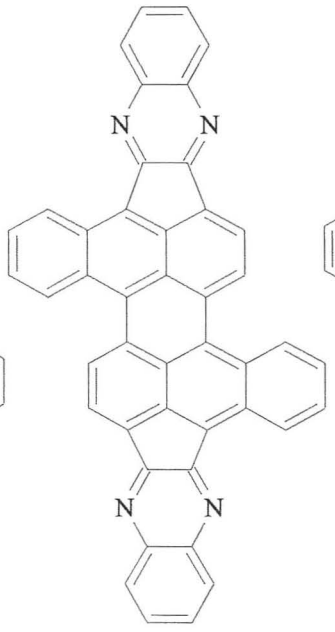
Predicted aromatic and polymethinic wavelengths are accurate to within 30 nm for most compounds. The existence of outliers is more notable for the polymethines. For example, several triarylmethanes, archetypal polymethines, such as methyl green **(31)** and guinea green B **(165)**, have wavelength maxima typical of aromatic character. This behaviour has already been noted for the central chromophore of light green **(149)**, and fast green FCF **(81)**, possibly to incomplete minimisation.

6.8 THE DESIGN OF NEAR-IR ABSORBERS

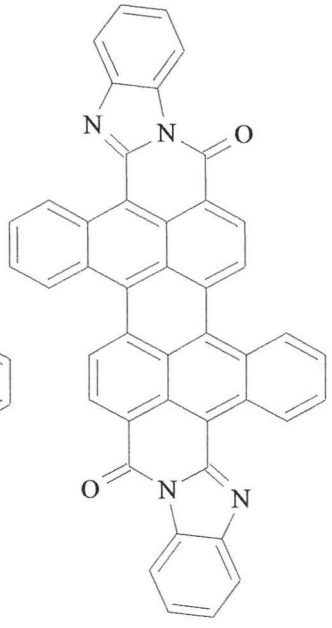
Dyes which absorb at wavelengths greater than about 750 nm, the near-IR region, such as at the GaAlAs or InP diode laser wavelengths, are required for xerographic printing, optical recording and laser filters.[436,437] PPP-MO Calculations have been performed on near-IR absorbing pigments related to aceanthrene green **(119)** and **(120)**.[436] Near-IR absorbers **(119)**, **(173)** and **(174)**, prepared by the respective KOH fusion of 1,9-anthracenedicarboxylic imide,[436] aceanthra[1,2-b]quinoxaline,[558] and the benzimidazole derivative of 1,9-anthracenedicarboxylic anhydride.[436] Maxima of **(119)** and **(174)** recorded in trichlorobenzene can be predicted with an accuracy of about 20 nm upon extrapolation of the aromatic regression equation. Maxima of **(119)** and **(173)** recorded in concentrated H_2SO_4, can be predicted with less accuracy (due to solvent effects) by extrapolation of the polymethinic regression equation.



(119)



(173)



(174)

Table 6.12 AM1 calculated wavelengths for near-IR absorbers.

Absorber	Observed λ / nm	Calculated λ / nm	Predicted λ / nm
(119)	867[436‡] 687[436†]	256.9	842.9[‡‡] 668.3[††]
(173)	840[558‡]	272.6	914.3[‡‡] 723.6[††]
(174)	785[436†]	284.3	764.7[††]

† Recorded in 1,3,5-trichlorobenzene. ‡ Recorded in conc. H_2SO_4.
†† Aromatic regression equation. ‡‡ Polymethinic regression equation.

6.9 AROMATIC HYDROCARBON EXCITATION ENERGIES

Several experimental studies have been performed linking orbital energies with excitation energies, and are used below to analyse trends in HOMO-LUMO gaps.

6.9.1 Vertical and adiabatic IPs and EAs

The electron affinity EA on an electron-acceptor molecule A is equal to the difference in energy between the anion A- and the uncharged molecule A + e^-. Similarly, the ionisation energy I of an electron-donor molecule D is equal to the difference in energy between the cation D^+ + e^- and the uncharged molecule D. The difference in energy between the lowest vibrational level of the ground state of the neutral molecule and the corresponding level of the ion is termed the **adiabatic** ionisation energy or electron affinity respectively. The energy difference between the ground state and that part of the potential curve of the ionised state to which, by applying the Franck-Condon principle, transition is most likely to occur, is denoted as **vertical** ionisation energy or electron affinity respectively.

6.9.2 Ionisation potentials

For polycyclic aromatic hydrocarbons, it follows from HMO theory that:

$E_i = -I = \alpha_0 + f_H \beta$, and ... Equation 6.2

$E_j = -EA = \alpha_0 + f'_H \beta$ ... Equation 6.3

where α_0 is the Coulomb integral, β is the resonance integral for two neighbouring atoms, and f_H is the Hückel coefficient which expresses the contribution of β to the total energy E. The electron overlap integral is assumed to be zero. The energy difference between the HOMO and LUMO is given by:

$E_j - E_i = \nu_0 = (f' - f)\,\beta$, so that: ... Equation 6.4

$E_j = \alpha + [f / (f' - f)]\nu$ ... Equation 6.5

It has been shown that $f / (f' - f)$ is approximately constant for several aromatic hydrocarbons.[559] As decreasing electron repulsion on transition from D to D^+ affects α_0,[560] Ehrenson[561] expressed α in units of β:

$\alpha = \alpha_0 - (1-1/n)\omega\beta$ (α, α_0 and β are negative; $|\alpha| < |\alpha_0|$) ... Equation 6.6

where ω is an empirical proportionality factor and n is the number of conjugated π-electrons in the aromatic hydrocarbon. Substituting α for α_0 in *Equation 6.2* and using *Equation 6.6*, it can be shown that:

$-E_i = I = -\alpha_0 - \{f_H - [\omega(n-1)/n]\}\beta = -\alpha_0 - \chi\beta$ ... Equation 6.7

When α = -9.878 eV, β = -2.11 eV and ω = 1.4,[562] *Equation 6.7* gives satisfactory values for the ionisation energy of aromatic hydrocarbons,[561] generally given by:[563]

$I = C + \gamma\nu_0$ ... Equation 6.8

Matsen[559] selected for ν_0 the frequency of the longest wavelength absorption maximum as recommended by Clar,[564] and obtained:

$I = 4.39 + 0.857\nu_0$ [eV] ... Equation 6.9

Becker and Wentworth[565] obtained using the 0,0-frequency:[566]

$I = 5.156 + 0.775\nu_{00}$ [eV] ... Equation 6.10

While Briegleb[567] obtained using ν_0 frequencies:

$I = 5.11 + 0.701\nu_0$ [eV] ... Equation 6.11

6.9.3 Electron affinities

According to Pople and Brickstock,[568] the energy of the HOMO and LUMO of an aromatic hydrocarbon are:

$E_i = -I = \alpha' + F$ ... Equation 6.12

$E_j = -EA = \alpha' - F$ ... Equation 6.13

F[569] can be calculated by the semi-empirical ASMOSC method.[570] Electron affinities have been calculated from *Equation 6.13* by Pople and Brickstock[568] and by Hedges and Matsen.[571] α' was normalised by Hedges and Matsen to the ionisation energy of naphthalene (8.12 eV) to $\alpha' = -3.87$ eV. The calculations[568,571] were carried out by Hoyland and Goodman[572] with even fewer assumptions. They also took into account changes in electron density following removal or capture of an electron, as well as the effect of the σ-electron energy. The ω-approximation of Ehrenson[561] gives an expression for the electron affinity corresponding to *Equation 6.7*:

$-E_j = EA = -\alpha_0 - \{f_H - [\omega(n + 1)/n]\}\beta = -\alpha_0 - \chi\beta$ ... Equation 6.14

The ω-approximation gives satisfactory values for the ionisation energies of aromatic hydrocarbons when $\alpha = -9.878$ eV, $\beta = -2.11$ eV and $\omega = 1.4$. Electron affinities which are too high are obtained from *Equation 6.13* using $\omega = 1.4$. To obtain good agreement with the electron affinities calculated from *Equation 6.13* by Hedges and Matsen,[571] Pople,[568] and Hoyland and Goodman,[572] as well as with the experimental values of Wentworth and Becker,[573] the very high value of $\omega = 3.77$ must be used in *Equation 6.13*.[574-576] A value of $\omega = 3.73$ must be used for phenanthrene, benzo[*c*]phenanthrene, and triphenylene.[576] No theoretical explanation is available for the fact that the calculations of I using *Equation 6.6* and of EA using *Equation 6.13* require different values of ω.

It follows from *Equation 6.6* and *Equation 6.13* that:

$(EA + I)/2 = \alpha'$ ... Equation 6.15

On the basis of the electron affinities determined experimentally for a few aromatic hydrocarbons,[573] and the experimental values for the first ionisation energies, an average value of 8.14 eV is obtained for the sum (I + EA) in agreement with Becker and Wentworth.[565] From *Equation 6.10* and *Equation 6.14*, using $\alpha' = 4.07$:[567]

$$EA = 3.03 - 0.701\nu_0 \quad \text{... Equation 6.16}$$

The differences between the electron affinities calculated by the ASMOSC method, those obtained from ν_0, and the few experimental EA values can be quite considerable. By averaging all EA values, EA values are obtained. Their absolute accuracy is never better than ± 0.2 eV. *Equation 6.17* obtained[567] in this manner is in good agreement with *Equation 6.16*, which was derived directly from experimental data and the expression of Becker and Wentworth,[565] *Equation 6.18*:

$$EA = 2.991 - 0.700\nu_0 \quad \text{... Equation 6.17}$$

$$EA = 2.92 - 0.762\nu_{00} \quad \text{... Equation 6.18}$$

6.9.4 HOMO-LUMO gaps

$$E_j - E_i \approx I - EA = 5.11 + 0.701\nu_0 - (3.03 - 0.701\nu_0)$$
$$= 2.08 + 1.402\nu_0 \text{ [eV]} \quad \text{... Equation 6.19}$$

If the HOMO-LUMO gap, $E_j - E_i$ (calculated from *Equation 6.11* and *Equation 6.16* using Koopmans' theorem as an initial approximation) is correlated against the frequency of the absorption maximum, ν_0, both expressed in eV, new straight-line relationships can be obtained. An excellent agreement with *Equation 6.19* is seen for the AM1 polyenic regression equation, but not with the expected benzenoid aromatic regression, probably due to unreliable electron affinities, bond alternation, or electronic effects upon ionisation / electron capture.

$$E_j - E_i = 4.4 + 0.83\nu_0 \text{ [eV]} \quad \text{... Polymethinic}$$

$$E_j - E_i = 4.1 + 0.77\nu_0 \text{ [eV]} \quad \text{... Aromatic}$$

$$E_j - E_i = 2.2 + 1.2\nu_0 \text{ [eV]} \quad \text{... Polyenic}$$

6.10 CONCLUSION

Electronic wavelength maxima were calculated from semi-empirical HOMO-LUMO gaps, for a training set of several hundred unsaturated organic compounds. The energies of the HOMO and LUMO were obtained *via* single-point calculations performed upon molecular mechanics minimised molecules. For each compound, the wavelength maximum is substantially underestimated, with larger deviations occurring at longer wavelengths. The underestimation of the wavelength maximum is caused by an overestimation of the singlet excitation energy, which, according to perturbation theory, is smaller than the HOMO-LUMO gap by an amount approximately equal to J – 2K. Empirical correlations between calculated and experimental wavelength maxima were found. These correlations were interpreted in terms of Dähne's triad theory, allowing the electronic properties of any unsaturated organic compound to be described by varying degrees of polymethinic, aromatic and polyenic character. Upon application of a correction factor obtained by least-squares fitting, wavelength maxima can, in general, be predicted to within an accuracy of about 25 nm. However, for accurate (<10 nm) colour prediction, the initial model described above will need to be refined. Predictability may, for example, be improved by the use of molecular geometries obtained by semi-empirical minimisation, as the MM+ force field does not adequately reproduce the bond alternation found in polyenes or aromatics. Nevertheless, electronic absorption maxima can be predicted with good success for a wide range of applications, namely UV absorbing cosmetic sunscreens, cytological dyes and near-IR xerographic pigments. Wavelength maxima for the Papanicolaou dyes can, for example, be predicted to within 30 to 40 nm.

In the SCF method of this chapter, the electron-electron interactions are averaged with a consequent loss of electron correlation, a constant but large term. Hartree-Fock (HF) theory is therefore unsatisfactory for describing electronic excitation. Post HF theory, *e.g.* configuration interaction (CI) and Many-Body Pertubation Theory (MBPT) approaches, seek to improve the description of electron-electron interactions using HF theory as a reference. Additionally, the above HOMO-LUMO method does not allow the prediction of absorption intensity. The following chapter will demonstrate how smaller estimated values for the molecular exchange integral K can be used to explain the weaker absorption intensities of polymethines. The CI approaches of the following chapter can also be used to obtain qualitative insights into absorption intensities.

CHAPTER 7
POST HF TREATMENT OF EXCITATION ENERGIES

7.1 AIMS

In the last chapter, Hartree-Fock SCF semi-empirical calculations of HOMO-LUMO gaps were used as an initial approximation of singlet electronic excitation energies. Singlet excitation energies are themselves approximated by $(E_H - E_L) + (2K - J)$, where $(E_H - E_L)$ is the HOMO-LUMO gap. By neglecting the evaluation of the Coulomb and exchange electron integrals, J and K, overestimation of the singlet excitation energies occurred, giving wavelength maxima which were seriously underestimated. In HF SCF theory, electron-electron interactions are averaged with a consequent loss of electron correlation, a constant but large term. There are some applications for which the neglect of electron correlation, or the assumption that the error is constant (and so will cancel) are not warranted, one example being electronic excitation. Post Hartree-Fock methods seek to improve the description of the electron-electron interactions, using HF theory as a reference point. Improvements to HF theory can be made in a variety of ways, including the method of configuration interaction (CI) and by use of many-body perturbation theories (MBPT).

In this chapter singlet excitation energies will be calculated using two different methods. In the first method, heats of formation of the singlet ground state, and first excited triplet and singlet states will be calculated, the relative values producing singlet and triplet excitation energies. In the second approach, single determinant configuration interaction (CIS) will be used to calculate not only excitation energies, but also oscillator strengths, giving an indication of the intensity of an electronic transition.

7.2 THE HALF ELECTRON METHOD

For open-shell RHF calculations using *HyperChem*, the half-electron method[577] recommended by Dewar[578] is used. For molecules with an even number of electrons and a closed-shell ground state this means that triplet and first excited singlet states have the same orbitals - both being computed by a pair of ½-electrons in the HOMO and LUMO orbitals. The excited triplet state has a lower energy than the first excited

singlet state upon application of a correction term due to exchange effects. In order to obtain accurate ground state heats of formation, the molecules were initially minimised using the MM+ force field and Polak-Ribiere conjugate gradient method to a gradient of 1×10^{-4} kcal mol^{-1} Å^{-1}, followed by subsequent AM1 minimisation to a final RMS gradient of 0.01 kcal mol^{-1} Å^{-1}. Single-point AM1 calculations were then performed and the heats of formation of the singlet ground state, first excited singlet and triplet states noted from the *.log* file. The resultant singlet and triplet excitation energies were compared with experimental values, as shown in *Table 7.1* below. As minimisation using the half-electron method can produce errant geometries, ground state geometries were used for the calculation of excited state properties. Accurate excited state heats of formation may therefore not have been obtained.

Table 7.1 AM1 Excitation energies (ΔE) for a series of test compounds.

Compound	Observed ΔE / kcal mol^{-1}		AM1 excitation energy / kcal mol^{-1}		2K / eV	J / eV
	$S_0 \rightarrow S_1$	$S_0 \rightarrow T_1$	$S_0 \rightarrow S_1$	$S_0 \rightarrow T_1$		
Acrolein	74[579]	70[579]	69.2 (3.00)	63.8 (2.77)	0.23	7.79
Cyclopentenal	83[579]	74[579]	73.4 (3.18)	68.6 (2.97)	0.21	7.40
Cyclohexenal	80[579]	75[579]	75.0 (3.25)	69.4 (3.01)	0.24	7.3
Azulene	42[494]	30[496]	48.8 (2.12)	33.1 (1.44)	0.68	5.7
Benzene	115 [579]	85[579]	125.6 (5.45)	70.7 (3.07)	2.38	7.14
Naphthalene	90 [579]	61[579]	103.2 (4.48)	61.8 (2.70)	1.78	5.74
Anthracene	76[579]	42[579]	87.1 (3.78)	47.4 (2.06)	1.72	5.22
Phenanthrene	83[580]	62[580]	102.9 (4.46)	65.4 (2.84)	1.62	5.37
Tetracene	61[580]	29[580]	75.9 (3.29)	38.5 (1.67)	1.62	4.85
Ethene	120[579]	82[579]	146.1 (6.34)	62.7 (2.72)	3.62	9.27
Butadiene	80[579]	60[579]	129.9 (5.63)	55.3 (2.40)	3.23	7.38
Cyclohexadiene	75[579]	54[579]	101.1 (4.38)	49.5 (2.15)	2.23	7.10
Hexatriene	70[579]	48[579]	109.5 (4.75)	49.4 (2.14)	2.61	6.53

$2K - J = EA - IP + S$, $J = IP - EA - T$, and $2K = S - T$

1 kcal mol^{-1} = 349.76 cm^{-1}; 1 eV / molecule = 23.061 kcal mol^{-1}

Most polymethinic singlet and triplet excitation energies are accurate to ~5 kcal mol^{-1}. Aromatic and polyenic singlet excitation energies are overestimated by ~20 and ~40 kcal mol^{-1} respectively, probably due to relaxation phenomena. Estimates of the Coulombic and exchange repulsion integrals (see *Section 5.5*) can be obtained by using AM1 minimised ground state HOMO-LUMO gaps as an approximation for the term IP – EA. An increase in the exchange integral, K, with a concurrent decrease in the Coulombic integral, J, is seen for compounds absorbing at higher wavelengths. Small values of K are obtained for polymethinic compounds, with intermediate and

large values for aromatic and polyenic compounds respectively. As described in *Section 5.5*, the HOMO and LUMO of azulene are largely localised in different parts of space (small overlap and $K_{H,L}$ – as azulene is non-alternant), while they are localised in the same parts of the space (larger overlap and $K_{H,L}$) in anthracene. Small values of exchange integrals are also connected with a low transition probability for the excitation, since the transition dipole, D, is relatively small. Since the oscillator strength, f, proportional to the integral to the integral of the area under the absorption band, is given by:

$$f = 1.085 \times 10^{-5} \Delta E D^2 \qquad \ldots \text{Equation 7.1}$$

where $\Delta E = E_L - E_H - J_{H,L} + 2 K_{H,L}$

D must be sufficiently large to compensate for the decrease in the oscillator strength resulting from a small value of ΔE.

The lowest triplet state of enones may be n,π^* or π,π^* depending on substituents. These two configurations will be comparable in energy. As a result, solvent may have a major influence on which configuration best corresponds to the triplet state. Similar σ/π ordering effects have been observed in acrolein by Rzepa.[581]

7.3 CONFIGURATION INTERACTION

Basic CI theory has been discussed in *Chapter 1*. An SCF computation yields a configuration describing orbital occupancy. Other configurations may be generated from this reference configuration by excitation. A CI calculation yields a set of improved states represented by a linear combination of configurations. Full CI is important because it is the most complete treatment possible within the limitations imposed by the basis set. It is common practice however to limit the excited states considered. It is common practice to limit the excited states considered, *e.g.* CIS (only wavefunctions differing from the HF wavefunction by a single spin orbital are included), CID (double substitution) and CISD (both single and double substitution). It can be shown that most of the correlation error in HF theory, namely that associated with pairs of electrons in the same orbital, may be corrected using CISD. The truncation of the CI expansion introduces a small, correctable size consistency error.

Excited state energies and wavefunctions are automatically obtained from CI calculations. However, the quality of the wavefunctions is more difficult to achieve. The equivalent of the HF description for the ground state requires CIS. Singly excited configurations do not affect the energy of the ground state by mixing with the HF determinant, Brillouin's theorem,[582] and provide a reasonable description of electronic excitation processes. A description for an excited state equivalent to the CISD description of the ground state requires all singles, double, and many of the triple excitations. Some of these may be added by perturbation theory. Quite accurate electronic transition energies and transition dipole and optical rotary strengths may be calculated at this level of theory.

CI calculations start with a set of computed RHF orbitals, known as the reference state. From this reference state, a set of microstates (different single determinants) is generated by promoting electrons from orbitals closest to the HOMO into unoccupied orbitals closest to the LUMO. A more accurate set of states is then obtained by taking appropriate linear combinations of these microstates, the coefficients of which are obtained by diagonalising the Hamiltonian matrix - this is the CI calculation. Transitions in the computed electronic spectrum are represented as vertical lines contained in a spectral display, with the wavelengths marked along the top of the display. The top set of lines represents all the excited electronic states (both singlet and triplet), and the bottom set represents only states that are spectroscopically active.

When studying electronic spectra of conjugated organic molecules, it is often sufficient to include only a few orbitals in the CI. For instance, UV absorption maxima of *p*-aminobenzoic acid are calculated just as accurately using two occupied and two unoccupied orbitals as using ten of each.[583] The extra orbitals yield more lines in the high-energy region (beyond the range of UV spectrometers) but do not affect the lines of interest found at longer wavelengths. As with many computational methods, qualitative rather than quantitative trends are found. ZINDO/S[53] calculations on several substituted benzenes underestimate the UV wavelength maxima by about 20 nm, although relative values are very well reproduced.[583]

7.4 GENERAL COMPUTATIONAL DETAILS

CIS calculations were mainly carried out either using *HyperChem for Windows*, running on a 66MHz 486-DX2 IBM-compatible personal computer. Similar results were obtained using the *ZINDO* module, part of the *Insight II* molecular modelling system running on a RISC-6000 IBM workstation. Solvent effects, *e.g.* rhodamine equilibria,[584] can be investigated using the conductor-like screening solvation model (COSMO)[585] or self-consistent reaction field (SCRF)[586] methods.

Initial geometries were obtained by minimisation to a gradient of 0.003 kcal mol^{-1} $Å^{-1}$ using the MM+ force field and Polak-Ribiere conjugate gradient method, again a serious defect as small changes in molecular geometries affect wavelength maxima. Singly excited CIS calculations were carried out using default overlap weighting factors.[587] Default values of 1.267 for σ–σ,[588] which adjusts the weighting for σ-σ atomic orbital overlap, and 0.640 for π–π,[589] which adjusts the weighting for π-π atomic orbital overlap, were used. This value is appropriate for transition metal complexes. A value of 0.585 for π-π[583] can be used to calculate electronic spectra of conjugated organic molecules (*Section 7.8*).

When choosing the CIS method, orbital criteria can be specified which set the range of orbitals to and from which electrons are excited to generate a set of interacting configurations. *Occupied* sets the range of occupied orbitals starting from the HOMO from which electrons are excited, while *Unoccupied* sets the range of virtual orbitals starting from the LUMO to which electrons are excited. In the following sections, an $n \times m$ calculation indicates n occupied and m unoccupied configurations. The longest wavelength transition for ethene calculated using the PM3 CIS method is 207 nm, comparing favourably with the experimental value of 190-200 nm.[583]

7.5 EFFECT OF CIS MATRIX SIZE

The effect of varying CIS matrix size on polyene {R–$(CH{=}CH)_n$–R} wavelength maxima, can be seen below. For a 1×1 matrix only 2 configurations occur, a singlet and a triplet; while 128 configurations occur for an 8x8 matrix. As the size of the CIS matrix increases, not only does the wavelength maximum slightly increase, the number of lower wavelength maxima also greatly increases.

Table 7.2 Effect of CIS matrix size on polyene wavelength maxima of *trans* polyenes, R-(CH=CH)n-R.

n	*R*	λ / nm	Calculated wavelength maxima / nm for various matrix sizes							
			1×1	2×2	3×3	4×4	5×5	6×6	7×7	8×8
1	Me	174^{590}	164.8	166.4	167.0	168.4	168.7	169.1	169.7	170.4
2	H	217^{337}	203.2	203.8	204	207.6	207.8	207.9	211.6	214.7
2	Me	227^{517}	223.5	223.7	224.9	226.1	227.5	228.0	228.6	229.4
3	H	268^{352}	258.2	258.2	260.2	261.9	262.3	263.4	263.7	264.1
3	Me	275^{348}	275.3	275.5	276.9	277.9	278.4	280.2	280.5	280.7
4	H	304^{352}	306.6	307.6	307.6	309.8	311.0	311.5	312.2	313.2
4	Me	310^{348}	321.3	323.0	323.1	324.4	324.9	325.8	327.3	328.4
5	H	334^{353}	349.9	353.1	353.1	353.1	355.3	355.7		357.5
5	Me	342^{348}	362.8	367.0	367.1	368.2	368.6	369.1	371.0	371.3
6	Me	380^{348}	400.2	407.9	408.4	408.4	409.4	409.4	409.9	411.0
7	Me	401^{348}			447.2	447.2	447.2	448.3		
8	Me	411^{348}			483.7	483.8	483.8	484.5		

7.6 TRAINING SET DATA

Using the same methodology as that addressed in *Chapter 6*, a smaller training set of organic molecules (*Appendix 1*) was used to establish the existence of polymethinic (◊), aromatic (+) and polyenic (□) correlations (*Graph 7.1* – page 155) between CIS calculated maxima (4×4 matrix) and experimental maxima, in accordance with Dähne's triad theory. The correlations were analysed by least-square fitting, as shown below. Uncorrected maxima accurate to ~20 nm are obtained for UV absorbers before extrapolation. Large systematic errors between CIS and experimental maxima are again obtained for polymethinic compounds, although these are smaller than those obtained using the HOMO-LUMO method of *Chapter 6*. Improved maxima are obtained for aromatic and polyenic compounds, even up to 400 nm, as *HyperChem* uses the singlet ground state of ethene as a reference, generates a series of singly excited configurations, computes the Hamiltonian matrix elements between them, and diagonalises the matrix to get the electronic spectrum.[583]

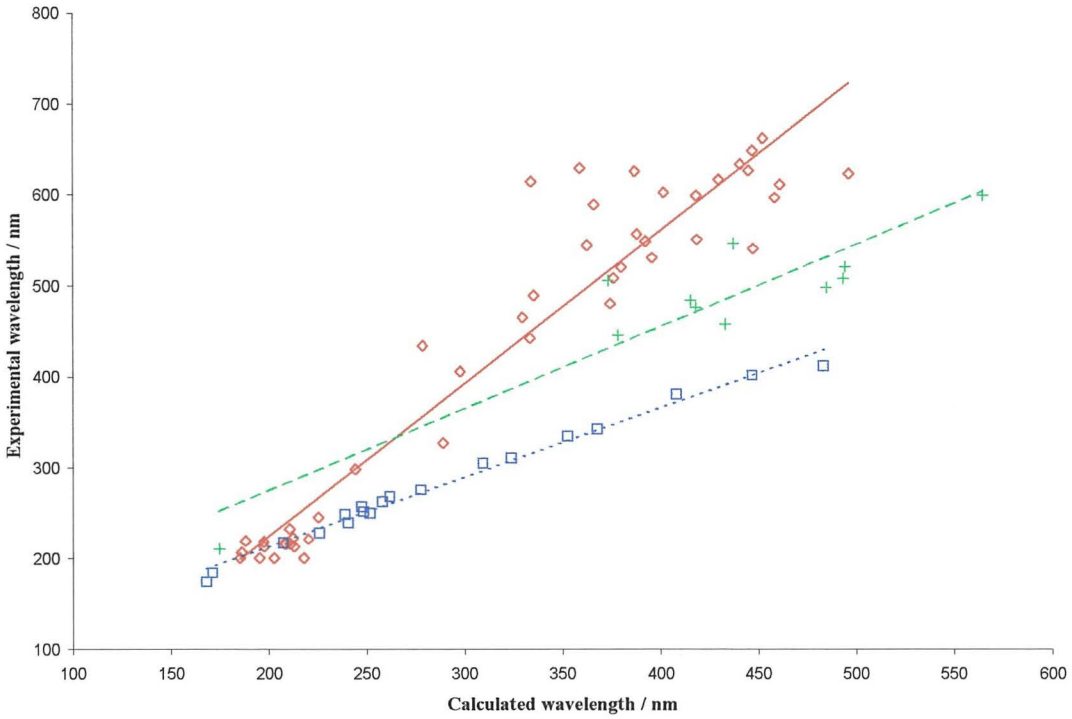
Table 7.3 Least-squares correlations between experimental and CIS calculated maxima.

Training set	α	β	r^2	*n*	σ	*F*
Polymethinic	-112	1.62	91.6	46	51.2	481.5
Aromatic	94.2	0.90	85.2	11	39.7	51.9
Polyenic	59.8	0.77	98.7	20	7.6	1409.4

Table 7.4 ZINDO/S Training set data.

Molecule	λ / nm	Matrix size							
		1×1	2×2	3×3	4×4	5×5	6×6	7×7	8×8
(175)	174[590]	164.8	166.4	167.0	168.4	168.7	169.1	169.7	170.4
(176)	217[337]	203.2	203.8	204	207.6	207.8	207.9	211.6	214.7
(177)	227[590]	223.5	223.7	224.9	226.1	227.5	228.0	228.6	229.4
(176)	268[352]	258.2	258.2	260.2	261.9	262.3	263.4	263.7	264.1
(179)	275[348]	275.3	275.5	276.9	277.9	278.4	280.2	280.5	280.7
(180)	304[352]	306.6	307.6	307.6	309.8	311.0	311.5	312.2	313.2
(181)	310[348]	321.3	323.0	323.1	324.4	324.9	325.8	327.3	328.4
(182)	334[353]	349.9	353.1	353.1	353.1	355.3	355.7		357.5
(183)	342[348]	362.8	367.0	367.1	368.2	368.6	369.1	371.0	371.3
(184)	380[348]	400.2	407.9	408.4	408.4	409.4	409.4	409.9	411.0
(185)	401[348]			447.2	447.2	447.2	448.3		
(186)	411[348]			483.7	483.8	483.8	484.5		
(187)	327[591]		212.3	209.2	289.4	295.4	296.4	296.6	296.8
(188)	405[591]				298.0				
(189)	465[591]				330.0				
(190)	508[591]				376.3				
(191)	520[591]				380.4				
(192)	550[591]				418.9				
(193)	480[591]				374.7				
(194)	610[591]	455.5	455.5	461.4	461.4	467.5	474.7	479.5	479.8
(195)	239[337]	236.7	237.0	240.6	240.9	241.8	243.1	244.5	246.3
(196)	256[337]	245.8	246.0	247.6	247.8	249.0	250.7	251.8	252.9
(197)	200[592]	215.3	215.4	215.4	217.9	218.4	218.7	222.5	224.3
(198)	200[593]	200.4	201.2	201.2	202.9	203.2	204.0	205.7	207.6
(199)	219[354]		184.5	185.8	188.1	189.7	190.4	190.5	191.3
(200)	207[358]		184.5	185.6	186.1	189.6	191.7	194.7	198.5
(201)	232[594]		208.5	210.0	210.6	211.6	212.7	213.3	214.5
(202)	245[358]		220.2	222.5	225.4	226.7	227.0	229.0	229.3
(203)	218[595]		193.7	195.1	197.4	198.0	200.3	201.9	202.2
(204)	200[359]		182.2	184.1	185.4	186.2	186.6	188.3	192.8
(205)	213[359]			210.8	213.0	213.9	214.6		
(206)	216[359]	202.0	204.8	206.7	208.8	209.1	210.0		
(207)	216[359]			209.3	211.1	212.0	212.9		
(208)	221[359]			218.6	220.4	221.6	222.4		
(209)	222[359]			210.7	212.5	213.1	213.6		
(210)	217[359]			206.7	208.1	208.7	208.9		
(211)	213[596]				197.7				
(212)	200[596]				195.3				
(213)	184[363]	179.6	169.2	171.1	171.7	171.7	172.0	172.4	173.7
(214)	211[361]				175.0				
(215)	248[597]				239.0				
(216)	251[597]				248.7				
(217)	262[598]				258.2				
(218)	249[598]				252.1				

Graph 7.1 CI Training set.



7.7 OSCILLATOR STRENGTH

It was hoped that a relationship could be found between the oscillator strength, *f*, and the absorbance, *A*, or extinction coefficient, ε, of a dye, which would enable the extinction coefficient at the wavelength maximum, ε_{max}, to be calculated. The oscillator strength, *f*, is related to the area under the absorption band. As the extinction coefficient, ε, is related to the absorption or height of the band, the two are linearly related only if the bands widths are the same and if the bands are symmetrical. The oscillator strength, *f*, often gives poor correlations with ε within a particular compound class and can only be used as a qualitative tool.[599] Oscillator strengths, *f*, were calculated for the training set using the ZINDO/S CIS method with default values. Different straight-line relationships were found between *f* and ε for polyenes and dimethylpolyenes. No relationship could be found between *f* and ε for other compounds. Large and small oscillator strengths are obtained for polyenes and polymethines respectively due to their large and small values of the exchange integral.

Table 7.5 Oscillator strengths, *f* for the training set.

	$\varepsilon \times 10^{-3}$	1×1	2×2	3×3	4×4	5×5	6×6	7×7	8×8
(176)	21^{337}	1.261	1.223	1.214	1.125	1.125	1.115	1.124	1.151
(178)	56^{352}	1.794	1.801	1.736	1.704	1.698	1.672	1.668	1.655
(179)	30^{348}	1.827	1.847	1.795	1.779	1.771	1.743	1.739	1.741
(181)	76.5^{348}	2.346	2.403	2.399	2.347	2.342	2.230	2.309	2.290
(182)	118^{353}	2.769	2.861	2.853	2.854	2.783	2.781		2.758
(183)	122^{348}	2.825	2.923	2.921	2.880	2.870	2.866	2.838	2.834
(184)	146.5^{348}	3.274	3.413	3.418	3.418	3.380	3.378	3.373	3.354
(195)	3.4^{337}	0.294	0.301	0.273	0.266	0.255	0.260	0.255	0.250
(196)	8^{337}	0.460	0.461	0.435	0.435	0.417	0.400	0.401	0.401
(199)	3.6^{354}		1.009	0.897	0.847	0.798	0.793	0.787	0.785
(200)	11.2^{358}		1.073	1.027	0.991	0.904	0.925	0.877	0.852
(201)	12.5^{594}		0.892	0.851	0.832	0.780	0.762	0.753	0.740
(202)	13^{358}		0.906	0.872	0.825	0.786	0.781	0.752	0.750
(203)	9.5^{595}		0.979	0.893	0.841	0.817	0.780	0.739	0.738
(204)	10^{359}		0.992	0.851	0.835	0.836	0.831	0.841	0.800
(205)	12.5^{359}			0.721	0.685	0.674	0.664		
(206)	9^{359}			0.761	0.731	0.733	0.717		
(207)	12^{359}			0.836	0.792	0.783	0.769		
(208)	9.7^{359}			0.760	0.725	0.712	0.698		
(209)	10.5^{359}			0.739	0.714	0.710	0.705		
(210)	10^{359}			0.774	0.748	0.739	0.744		
(213)	60^{363}	0.589	1.250	1.091	1.091	1.091	1.071	1.065	1.078

Table 7.6 Effect of changing the π-π atomic orbital overlap on CIS maxima.

Compound	λ_{exp} / nm	π-π = 0.640 (4×4)	π-π = 0.585 (4×4)
(176)	217^{337}	207.6 (1.125)	219.1 (1.057)
(177)	227^{590}	226.1 (1.185)	236.5 (1.108)
(178)	268^{352}	261.9 (1.704)	274.8 (1.568)
(179)	275^{348}	277.9 (1.779)	290.1 (1.648)
(181)	310^{348}	324.4 (2.347)	328.1 (2.177)
(183)	342^{348}	368.2 (2.880)	382.0 (2.703)
(184)	380^{348}	408.4 (3.418)	424.3 (3.152)
(185)	401^{348}	447.2 (3.887)	463.6 (3.583)
(186)	411^{348}	483.8 (4.335)	500.6 (3.990)
(195)	239^{337}	240.9 (0.266)	252.6 (0.259)
(196)	256^{337}	247.8 (0.435)	258.7 (0.417)
(199)	219^{354}	188.1 (0.847)	200.0 (0.785)
(200)	207^{358}	186.1 (0.991)	197.8 (0.912)
(201)	232^{594}	210.6 (0.832)	222.1 (0.752)
(202)	245^{358}	225.4 (0.825)	235.2 (0.772)
(203)	218^{595}	197.4 (0.841)	208.2 (0.788)
(205)	213^{359}	213.0 (0.685)	225.8 (0.639)
(206)	216^{359}	208.8 (0.731)	221.4 (0.680)
(208)	221^{359}	220.4 (0.725)	233.2 (0.676)
(215)	248^{597}	239.0 (0.853)	253.7 (0.778)
(216)	251^{597}	248.7 (0.932)	262.9 (0.849)
(217)	262^{598}	258.2 (0.582)	272.1 (0.565)

7.8 EFFECT OF CHANGING THE π-π OVERLAP

The effect of changing π-π, the weighting for π-π atomic orbital overlap, for a series of 4×4 CIS calculations matrices is shown overleaf in *Table 7.6*. Oscillator strengths, *f*, are given in parentheses. Previous calculations used a default value appropriate to transition metals of 1.267 and 0.640 for σ-σ and π-π respectively. A π-π value of 0.585 has been used to calculate organic electronic maxima, and causes a bathochromic shift of no more than 20 nm. Although the two methods have similar maxima, improved maxima are obtained using the new value of 0.585 for π-π.

7.9 CI APPROACHES USING OTHER HAMILTONIANS

It has been shown that small changes in molecular geometries can have large effects on calculated maxima. The above calculations have all been calculated using ZINDO/S CIS calculations based on MM+ minimised geometries. Although ZINDO/S is specifically parameterised to reproduce electronic spectra, other semi-empirical Hamiltonians can be used to calculate electronic spectra. Ideally, one would use the same semi-empirical method for geometry optimisation and CIS calculation. Initial results using other Hamiltonians are presented below, although a further study is envisaged. The AM1 and ZINDO/S CIS (σ-σ = 0.585) calculations were performed using AM1 minimised molecules, while the PM3 CIS calculations were performed using PM3 minimised molecules.

Table 7.7 CIS maxima for a various Hamiltonians.

Compound	AM1 λ / nm		PM3 λ / nm		ZINDO/S λ / nm	
	$S_0 \rightarrow S_1$	$S_0 \rightarrow T_1$	$S_0 \rightarrow T_1$	$S_0 \rightarrow T_1$	$S_0 \rightarrow S_1$	$S_0 \rightarrow T_1$
Azulene	557.6	772.3	527.1	791.8	527.1	791.8
Naphthalene	344.6	529.7	345.1	566.7	310.0	646.0
Anthracene	375.6	715.9	365.7	680.6	335.8	816.8
Butadiene	249.4	629.3	247.8	589.5	208.7	731.5

1 eV / molecule = 23.061 kcal mol^{-1}

7.10 CONCLUSION

As HF theory does not adequately describe electronic excitation processes, initial studies involving two post-HF methods, *i.e.* $\Delta\Delta H_f$ and CIS approaches, have been presented in the above chapter. The two methods not only allow the calculation of

singlet and triplet excitation energies, but also allow the prediction of absorption intensities.

In the first instance, the half-electron RHF method was used to calculate heats of formation for ground and first excited (singlet and triplet) states. Excitation energies derived from the relative heats of formation were compared with experimental values. Although polymethinic excitation energies were found to be accurate to within 5 kcal mol^{-1}, aromatic and polyenic excitation energies were overestimated by 20 to 40 kcal mol^{-1}. It is possible that the biradical nature of the aromatic and polyenic excited states was not adequately described using the RHF half-electron method. The small values of K predicted for polymethinic compounds could be used to explain why these compounds absorb at higher wavelengths, usually with low oscillator strengths. Similarly, the high values of K for polyenes can be used to explain why they absorb at lower wavelengths, usually with much higher oscillator strengths.

Configuration interaction (CIS) approaches were also used to predict singlet and triplet excitation energies. The majority of the results presented in this chapter were obtained using the ZINDO/S Hamiltonian. Increasing the CIS matrix size caused a slight increase in the wavelength maximum, with a concomitant increase in the number of absorption maxima at lower wavelengths. A smaller training set of unsaturated organic molecules was used to investigate absorption maxima using the ZINDO/S Hamiltonian (4 x 4 matrix). Polymethinic CIS maxima showed good agreement in the UV region, with deviations at higher wavelengths, probably due to the use of ethene as a reference. Aromatic compounds showed the best agreement in the UV region, together with short chain polyenes. As the polyene chain length increases, CIS maxima deviate due to the tailing off of experimental maxima. The effect of various Hamiltonians on the wavelength maximum is as yet inconclusive.

It was hoped that electron correlation and repulsion effects would be implicitly defined using a first-principles approach. The drawback of *ab initio* methods such as STO-6-31G**, is the large number of integrals to be calculated, increasing computational effort, and necessitating the use of small molecules. Density functional theory (DFT) is a first-principles method that can be successfully applied for larger systems, and will be used to predict electronic spectra in the next chapter.

CHAPTER 8
DFT CALCULATIONS OF WAVELENGTH MAXIMA

In the previous two chapters, electronic wavelength maxima of unsaturated organic compounds have been calculated using semi-empirical based approaches. Singlet excitation energies can be approximated by the expression $(E_L - E_H) + (2K - J)$. As the $(2K - J)$ repulsion terms were ignored in *Chapter 6*, singlet excitation energies were overestimated. It was hoped that better description of electron repulsion effects leading to better excitation energies could be obtained using *ab initio* methods. The majority of this work was carried out as a three-month project under the supervision of Dr. Erich Wimmer, Director of the *European Centre for Computational Science and Technology*, *Molecular Simulations Inc.* (formerly *Biosym Technologies SARL*), Orsay, France. Results using the JMW and VWN local functionals were obtained using *DMol* (*version 2.3*), Becke, LYP and BP non-local calculations were performed using *DMol* (*version 3.0*) at the *Centre for Computational Chemistry* at Bangor.

8.1 INTRODUCTION

DMol is a first-principles *ab initio* software package, used to perform accurate theoretical calculations on biological, organometallic and inorganic systems. *DMol* calculates variational self-consistent solutions to the density functional theory (DFT) equations, expressed as a numerical atomic orbital basis, the solutions of which produce molecular wavefunctions and electron densities.

8.1.1 Density functional theory

DFT began with a theorem by Hohenberg and Walter Kohn (1998 Nobel Chemistry Laureate),[600] generalised by Levy,[601] which stated that all ground-state properties are functions of the charge density, ρ. The total energy E_t may be expressed as:

$$E_t[\rho] = T[\rho] + U[\rho] + E_{xc}[\rho] \quad \text{... Equation 8.1}$$

where $T[\rho]$ is the kinetic energy of a system of noninteracting particles of density ρ, $U[\rho]$ is the classical Coulombic electrostatic energy, and $E_{xc}[\rho]$ includes all many-body contributions to the total energy, particularly exchange and correlation energies.

In accordance with other MO methods,[602-604] the wavefunction is taken to be a Slater determinant, an antisymmetrised product of one particle MOs.

$$\Psi = A(n)\,|\phi_1(1)\,\phi_2(2)\,\ldots\,\phi_n(n)| \qquad \ldots \text{Equation 8.2}$$

The MOs must also be orthonormal, *i.e.*:

$$\langle\phi_i|\phi_j\rangle = \delta_{ij} \qquad \ldots \text{Equation 8.3}$$

The charge density is given simply as the sum over all occupied MOs ϕ_i:

$$\rho(\mathbf{r}) = \Sigma|\phi_i(\mathbf{r})|^2 \qquad \ldots \text{Equation 8.4}$$

The MOs can be occupied by spin-up (α) electrons or by spin-down (β) electrons. Spin-restricted calculations use the same ϕ_i for both α and β electrons; while, spin-unrestricted or spin-polarised calculations use different ϕ_i for α and β electrons. Total charge densities and spin densities are given respectively by the sum and difference for α and β MOs. This is analogous to restricted and unrestricted HF calculations.[605]

The kinetic energy is essentially a second derivative of the wavefunction; while the electrostatic potential energy is made up of three terms: electron-nucleus attraction, ρV_N, electron-electron repulsion, $\frac{1}{2}\rho V_e$, and nucleus-nucleus repulsion, V_{NN}.

$$T = \langle\Sigma\phi_i|-\tfrac{1}{2}\nabla^2|\phi_i\rangle \qquad \ldots \text{Equation 8.5}$$

$$\begin{aligned} U = {} & \Sigma\Sigma\,\langle\phi_i(\mathbf{r})|-Z/(\mathbf{R}_\alpha-\mathbf{r})|\phi_i(\mathbf{r})\rangle + \tfrac{1}{2}\Sigma\Sigma\,\langle\phi_i(\mathbf{r_1})\phi_j(\mathbf{r_2})\;1/(\mathbf{r_1}-\mathbf{r_2})\;\phi_i(\mathbf{r_1})\phi_j(\mathbf{r_2})\rangle \\ & + \Sigma\Sigma Z_\alpha Z_\beta/|\mathbf{R}_\alpha-\mathbf{R}_\beta| \\ \equiv {} & \langle-\rho(\mathbf{r_1})V_N\rangle + \langle\rho(\mathbf{r_1})\tfrac{1}{2}V_e(\mathbf{r_1})\rangle + V_{NN} \end{aligned} \qquad \ldots \text{Equation 8.6}$$

The exchange-correlation energy requires some approximation. A simple approximation is used, called the local density approximation (LDA). This is based on

the known exchange-correlation energy of the uniform electron gas.[606-608] Analytical representations have been made.[606,607,609-611] The LDA approximation assumes that the charge density varies slowly on an atomic scale, so that each region of a molecule looks like a uniform electron gas. The total exchange-correlation energy, ε_{xc} is obtained by integrating the uniform electron gas result:

$$\varepsilon_{xc}[\rho] \equiv \int \rho(\mathbf{r})\varepsilon_{xc}[\rho(\mathbf{r})]d\rho \quad \ldots \text{Equation 8.7}$$

The default local spin-density (LSD) functional used in *DMol 3.0* is the form derived by Vosko *et al*,[610] denoted the VWN functional. Other LSD functionals available are those developed by von Barth and Hedin (BH),[608] Janak, Morruzi and Williams (JMW), and Perdew and Wang (PW).[610]

The next step in improving the LSD model is to take into account the inhomogeneity of the electron gas, occurring in any molecular system. This can be accomplished by a density gradient expansion, referred to as the nonlocal spin-density (NSLD) approximation. The gradient corrected exchange-correlation energy $E_{xc}[\rho, d(\rho)]$ is necessary for studying thermochemical processes.[612-614] *DMol 3.0* uses the Perdew and Wang[610] (PW) generalised gradient approximation for the correlation functional and the Becke[615] (B) gradient corrected exchange functional – this approximation of the NLSD Hamiltonian referred to as the BP functional. The gradient corrected correlation functional of Lee, Yang, and Parr[616] (LYP) is also available.

To determine the actual energy E_t, variations in E_t must be optimised with respect to variations in ρ, subject to the orthonormality constraints.[617] This leads to a set of coupled equations involving the exchange-correlation potential μ_{xc}.

$$E_t[\rho] = \sum \langle\phi_i|-\tfrac{1}{2}\nabla^2|\phi_i\rangle + \langle\rho(\mathbf{r_1})[\varepsilon_{xc}\{\rho(\mathbf{r_1})\} + \tfrac{1}{2}V_e(\mathbf{r_1}) - V_N]\rangle + V_{NN} \quad \ldots \text{Equation 8.8}$$

$$\{-\tfrac{1}{2}\nabla^2 - V_n + V_e + \mu_{xc}[\rho]\}\phi_i = \varepsilon_i\phi_i \quad \ldots \text{Equation 8.9}$$

For the local-spin density approximation, μ_{xc} and E_t are given by:

$$\mu_{xc} = \partial/\partial\rho\ (\rho\varepsilon_{xc}) \quad \text{... Equation 8.10}$$

$$E_t = \Sigma\varepsilon_i + <\rho(\mathbf{r_1})[\varepsilon_{xc}[\rho] - \mu_{xc}[\rho] - \tfrac{1}{2}V_e(\mathbf{r_1})]> + V_{NN} \quad \text{... Equation 8.11}$$

The MOs are given by $\phi_i = \Sigma C_{i\mu}\chi_\mu$, where χ_μ are atomic orbital basic functions, and $C_{i\mu}$ are LCAO-MO coefficients. Several choices of basis set are possible including Gaussian,[618] Slater[619] and numerical orbitals. *DMol* uses numerical orbitals. The AOs are not orthonormal, leading to a reformulation of the DFT equation in the form:

$$HC = \varepsilon SC \quad \text{... Equation 8.12}$$

where:

$$H_{\mu\nu} = <\chi_\mu(\mathbf{r_1})|\text{-}\tfrac{1}{2}\nabla^2 - V_N + V_e + \mu_{xc}\rho(\mathbf{r_1})|\chi_\nu(\mathbf{r_1})> \quad \text{... Equation 8.13}$$

and:

$$S_{\mu\nu} = <\chi_\mu(\mathbf{r_1})|\chi_\nu(\mathbf{r_1})> \quad \text{... Equation 8.14}$$

As H depends upon C, $HC = \varepsilon SC$ must be solved by an iterative technique:

1) Choose an initial set of $C_{i\mu}$.
2) Construct an initial set of MOs ϕ_i.
3) Construct ρ.
4) Using ρ, construct V_e and μ_{xc}.
5) Construct $H_{\mu\nu}$.
6) Solve *Equation 8.12* for a new set of $C_{i\mu}$.
7) Construct a new ϕ_i and new ρ.
8) If $\rho_{new} = \rho_{old}$ evaluate E_t and stop. Otherwise if $\rho_{new} \neq \rho_{old}$ then return to step 4.

For organic molecules about ten iterations are typically required to obtain convergence at $|\rho_{new} - \rho_{old}| < 10^{-6}$. For metallic systems many more iterations are frequently required.

8.1.2 Numerical basis sets

The basic functions χ_μ are given numerically as values on an atomic-centred spherical-polar mesh, rather than as analytical functions (*i.e.* Gaussian orbitals). The angular part of each function is the appropriate spherical harmonic $Y_{lm}(\theta,\phi)$. The radial part $F(\mathbf{r})$ is obtained by solving the atomic DFT equations numerically. A reasonable level of accuracy is obtained by using a range of 300 radial points from the nucleus to an outer distance of 10 Bohrs (approximately 5.3 Å). The radial functions are stored as a set of cubic spline coefficients for each of the 300 sections. The $-\nabla^2/2$ terms required to evaluate the kinetic energy are also stored as spline coefficients. There are several advantages to the use of exact DFT spherical atomic orbitals. The molecule can be dissociated exactly into its constituent atoms. Basis set superposition effects are minimised allowing weak bonds to be described.[612]

There are 4 types of numerical basis set used in *DMol*:

a) MINIMAL (MIN)

This uses one atomic orbital for each occupied orbital in the free atom.

b) DOUBLE NUMERIC (DN)

Uses approximately two atomic orbitals for each occupied in the free atom.

c) DNP

Uses DN basis function together with polarisation functions. The polarisation functions have angular momentum one higher than that of the highest occupied orbital in the free atom. The polarisation function for carbon is 3*d*, and 2*p* for hydrogen.

d) DND

Identical to DNP, except that no *p* functions are used for hydrogen.

Minimal basis sets are generally inadequate for anything but qualitative results. The DN basis set is analogous to the Gaussian double-zeta (DZ) set, while the DND and DNP basis sets are analogous to Gaussian 6-31G* and G-31G** basis sets respectively. DNP basis functions are usually the most reliable. For first-row atoms, functions from divalent cations provide a reasonable double basis set. A hydrogenic $3d$ orbital obtained for a nucleus of $Z = 5$, provides a good polarisation function for each of these atoms. A hydrogenic $2p$ function for $Z = 1.3$ is used for hydrogen. Various nuclear charges are also used to generate polarisation functions, analogous to the variation of zeta in Gaussian basis sets. For metals, $4p$ polarisation functions are generated by solving the atomic equations for a $4s \rightarrow 4p$ excited state. Basis set quality has been analysed in detail.[620] A frozen-core approximation may be used where core functions are frozen at values for free atoms, and valence orbitals are orthogonalised to them, reducing the size of the secular equation without much loss in accuracy.

8.1.3 Numerical integration

Evaluation of the integrals in *Equations 8.13* and *8.14* must be accomplished by a three-dimensional numerical integration procedure, due to the nature of the basis functions. The matrix elements are approximated by the finite summations:

$$H_{\mu\nu} \equiv \sum \chi_\mu(\mathbf{r_i})\, H_{\text{eff}}(\mathbf{r_i})\, \chi_\nu(\mathbf{r_i})\, w(\mathbf{r_i}) \qquad \text{... Equation 8.15}$$

$$S_{\mu\nu} \equiv \sum \chi_\mu(\mathbf{r_i})\, \chi_\nu(\mathbf{r_i})\, w(\mathbf{r_i}) \qquad \text{... Equation 8.16}$$

Summations run over several numerical integration points $\mathbf{r_i}$. H_{eff} is the value of the integrand at point $\mathbf{r_I}$, and $w(\mathbf{r_i})$ is a weight associated with each mesh point. The integration points are generated in a spherical pattern around each atomic centre. Radial points are taken in shells from the nucleus out to a distance of 10 Bohrs. The number of shells is a function of atomic number, so heavier elements have a greater number of points. The number of radial points N_R for a nucleus of charge Z is:

$$N_R = 14(Z+2)^{1/3} \qquad \text{... Equation 8.17}$$

The spacing between points is logarithmic so that points are spaced more closely near the nucleus where oscillations in the wavefunction are more rapid. The choice of

integration points can have a significant impact on the final results.[620-622] The use of an **xcoarse** mesh size results in a serious compromise in the calculation, but can be used for debugging, qualitative results and generation of data for plotting. **Medium** mesh sizes give a reasonable compromise between computational expense and numerical precision, while an **Xfine** mesh results in a nearly saturated mesh. Angular integration points N_θ are generated at each of the N_R radial points, creating a series of shells around each nucleus. Angular points are selected by quadrature schemes.[623-625] The variation of N_θ with mesh size is shown in *Table 8.1*. Partition functions are used to increase numerical integration convergence by avoiding nuclear cusps.[615,620,626]

Table 8.1 Variation of N_θ with mesh size.

Mesh size	N_θ
XCOARSE	50
COARSE	110
MEDIUM	194
FINE	302
XFINE	434

8.1.4 Evaluation of the effective potential

Evaluation of the exchange-correlation energy ε_{xc} and potential μ_{xc} is accomplished by solving the uniform electron gas model. This requires numerical evaluation of the charge density $\rho(\mathbf{r})$ at many points in space. The use of numerical basis functions simplifies this process, as all the required quantities are already available on a grid.[627] The Coulombic potential is evaluated by solving the Poisson equation for the charge density, in a completely numerical (non basis set) manner,[620] giving greater precision:

$$-\nabla^2 V_e(\mathbf{r}) = 4\pi e^2 \rho(\mathbf{r}) \quad \text{... Equation 8.18}$$

rather than evaluating the Coulombic term as:

$$V_e(\mathbf{r_1}) = \int \rho(\mathbf{r_2})/|\mathbf{r_2} - \mathbf{r_1}|\, d\mathbf{r_2} \quad \text{... Equation 8.19}$$

8.1.5 SCF Procedure

Before the SCF procedure can be used, a step is required analogous to the HF evaluation of integrals over AOs - the interpolation of the numerical atomic basis set

onto the molecular grid. Neglecting symmetry and frozen core approximations, this step requires a computational effort on the order of $N \times P$ for N atomic orbitals and P integration points. The overlap matrix and constant portion of the effective Hamiltonian (the kinetic and nuclear attraction terms) are also constructed, each requiring $N \times (N + 1) \times P$ operations. Once $\rho(\mathbf{r})$ is known, evaluation of the exchange-correlation potential μ_{xc} requires only P operations. Construction of the Coulombic potential requires only $P \times M$ operations where M is the number of multipolar functions (9 functions atom^{-1} for $l = 2$ or 16 for $l = 3$). Constructing the Hamiltonian matrix elements are among the most time consuming steps, requiring $N \times (N + 1) \times P$ operations per iteration. Solving the secular equation requires N^3 operations.

8.1.6 Damping

Simple damping is necessary to ensure smooth convergence:

$$\rho = d\rho_{new} + (1 - d)\rho_{old} \quad \text{... Equation 8.20}$$

where d is the damping factor, ρ_{old} is the density that was used to construct the secular matrix, ρ_{new} is the density constructed from the new MO coefficients without damping, and ρ is the density used in the next iteration. The keywords Mixing_Alpha (FASCF) and Mixing_Beta (FBSCF) specifies the mixing coefficients for the charge density and spin density respectively.

$$\rho = \rho_{old}\,(1 - \text{Mixing\_Alpha}) + \rho_{new}\,\text{Mixing\_Alpha} \quad \text{... Equation 8.21}$$

$$\rho = \rho_{old}\,(1 - \text{Mixing\_Beta}) + \rho_{new}\,\text{Mixing\_Beta} \quad \text{... Equation 8.22}$$

For systems with many low-lying excited states, *e.g.* metal clusters, or for oscillatory behaviour, Mixing_Alpha and Mixing_Beta can be reduced from 0.25 to 0.005. The Poisson equation requires an auxiliary density representation Υ, differing from ρ by:

$$\delta\rho = \Upsilon - \rho \quad \text{... Equation 8.23}$$

To minimise $\delta\rho$, a total-energy formula is used which is second order in $\delta\rho$,[620,628,629] the total energy calculated in the first SCF cycle being the Harris approximation.[630]

8.1.7 Optimisation

Analytic energy derivatives are available for HF,[631] CI,[632] and MBPT[633] theories. Energy gradient formulas for the HF-Slater method were first derived by Satako,[634] later implemented using Slater[619] and Gaussian[618] basis sets. The first derivative of the total energy with respect to a nuclear perturbation of atom α can be shown to be:

$$E^{\alpha}_{t} = \sum \langle \phi_i^{\alpha} | h - \varepsilon_i | \phi_i \rangle + E^{\alpha}_{HF} \quad \text{... Equation 8.24}$$

where E^{α}_{HF} is the Hellmann-Feynmann term, the derivative in the absence of orbital relaxation. The minimum does not correspond exactly to point with zero derivative.[619]

Geometries can be optimised for both Cartesian and internal co-ordinates using the EF algorithm.[635] Optimisation is efficient if a reasonable estimate of the Hessian is available for the starting geometry.[636] By following the lowest Hessian mode, the EF algorithm can locate transition structures. GDIIS can also be used for optimisation, which is based on the DIIS technique for accelerating SCF convergence.[637] Mode following involves modification of the standard Newton-Raphson step:

$$\mathbf{h} = -\mathbf{H}^{-1} \; [\equiv \{\nabla^2 F(\mathbf{x})^{-1}\} \nabla F(\mathbf{x})] \quad \text{... Equation 8.25}$$

by introducing a shift parameter λ[638] so that:[639]

$$\mathbf{h} = -(\mathbf{H} - \lambda \, \mathbf{1})^{-1} \mathbf{g} \quad \text{... Equation 8.26}$$

8.1.8 The electron gas model

Parameterised equations for the electron gas exchange-correlation energy are presented below. An excellent review of the electron gas model appears in *Appendix E* of Parr and Yang.[640] *DMol* is based on the work of von Barth and Hedin.[601]

$\mathbf{r}_s$ is defined as:

$$\mathbf{r}_s^3 = 3/(4\pi\rho) \quad \text{... Equation 8.27}$$

The exchange energy is given as:

$\varepsilon_x = -C_x/\mathbf{r}_s$ ... Equation 8.28

where:

$C_x = \frac{3}{4}\,(9/\{4\pi^2\})^{1/3}$ ... Equation 8.29

For a restricted computation, the correlation contribution is:

$\varepsilon_c = -c_0 F[\mathbf{r}_s/\mathbf{r}_0]$ ... Equation 8.30

where: $c_0 = 0.0225$ $\mathbf{r}_0 = 21$

and $F[Z] = (1 + Z^3)\ln(1 + 1/Z) + \frac{1}{2}Z - Z^2 - 1/3$ ... Equation 8.31

For a spin-unrestricted computation,

$\varepsilon_x = -c_0 F(\mathbf{r}_s/\mathbf{r}_0) + [(2^{1/3} - 1)C_x/\mathbf{r}_s - c_1 F(\mathbf{r}_s/\mathbf{r}_1) + c_0 F(\mathbf{r}_s/\mathbf{r}_0)]f(\Delta\rho/\rho)$ ... Equation 8.32

where:

$c_1 = \frac{1}{2}.c_0$ $\mathbf{r}_1 = 2^{4/3}.\mathbf{r}_0$ $\rho = \rho^\alpha + \rho^\beta$ $\Delta\rho = \rho^\alpha - \rho^\beta$

and:

$f(z) = [2^{4/3} - 2]^{-1}[(1+z)^{4/3} + (1-z)^{4/3} - 2]$ ... Equation 8.33

8.2 POLYENES

DMol works well for metallic systems with constant local charge density. Polyenes were considered ideal model systems for initial study, as the polyene backbone is essentially a long chain of carbon atoms, with fairly constant local charge density.

8.2.1 Effect of symmetry

DMol has a symmetry option, reducing the computational time. The secular equation becomes block diagonal, reducing the size of the matrix diagonalisation, speeding up the calculation. Fewer numerical integration points also need to be generated. Although one SCF cycle for but-2-ene takes 3½ minutes and 1 minute for C_1 and C_{2h} symmetry respectively, one SCF cycle for the dimethyl analogue of dodecahexaene takes 41 minutes and 9 minutes for C_1 and C_{2h} symmetry respectively.

8.2.2 Initial calculations

The *trans* polyenes, Me-(CH=CH)$_n$-Me, were built using *BUILDER*, and minimised using *DISCOVER* (CVFF forcefield, steepest descent, followed by conjugate gradients and BFGS) to a final gradient of 0.0003 kcal mol^{-1} Å^{-1}. JMW functional, DND basis, medium mesh single point calculations were performed. Orbital energies were converted to eV using the fact that 1 a.u. (Hartree) = 27.211396 eV. The wavelength maximum was then calculated using *Equation 6.1*.

Table 8.2 Single point JMW DND medium mesh wavelength maxima.

n	Experimental λ / nm	Calculated λ / nm
1	174[590]	251.9
2	227[590]	363.1
3	275[348]	481.7
4	310[348]	602.0
5	342[348]	722.2
6	380[348]	843.9

DFT HOMO-LUMO gaps are underestimated, leading to an overestimation of wavelength maxima, in sharp contrast with results obtained in *Chapter 6*.

8.2.3 Effect of optimisation

The Hessian matrix needed for optimisation was generated using *DISCOVER 3*. The optimisation was carried out *via DMol*, using JMW functional, DND basis and a medium mesh until a final gradient of 0.0003 kcal mol^{-1} Å^{-1} was achieved. The general effect of optimisation for polyenes is the reduction of λ towards the experimental value. The lowering of E_H is far greater than the lowering of E_L for these polyenes, accounting for the larger energy gap and hence lower λ after optimisation.

Table 8.3 Effect of optimisation on polyene wavelength maxima.

n	Experimental λ / nm	Calculated λ / nm		
		Non optimised	Optimised	Δλ / nm
1	174[590]	251.9	251.5	0.4
2	227[590]	363.1	348.7	14.4
3	275[348]	481.7	448.5	33.2
4	310[348]	602.0	548.4	53.6
5	342[348]	722.2	647.3	74.9
6	380[348]	843.9	744.6	99.7

8.2.4 Basis set variation

Single point calculations were carried out using the JMW functional, a medium mesh size, while allowing the basis set to vary. It is clear that basis set choice plays a significant role in the calculated wavelength maximum.

Table 8.4 JMW functional, medium mesh calculations for *trans* polyenes Me-(CH=CH)$_n$-Me using various basis sets.

n	Experimental λ / nm	Calculated λ / nm			
		MIN	DN	DND	DNP
1	174[590]	227.6	247.1	251.9	260.1
2	227[590]	347.1	357.7	363.1	367.8
3	275[348]	469.8	474.7	481.7	486.3
4	310[348]	593.3	593.3	602.0	606.0
5	342[348]	716.5	710.9	722.2	725.6
6	380[348]	840.8	828.5	843.9	847.0

8.2.5 Mesh size variation

Mesh size does not appear to greatly affect polyene wavelength maxima, possibly due to a combination of small molecular size and C_{2h} symmetry reducing the number of mesh points needed to describe the local nature of the repeating carbon backbone.

Table 8.5 JMW functional, DND basis calculations for *trans* polyenes Me-(CH=CH)$_n$-Me using various mesh sizes.

n	Experimental λ / nm	Calculated λ / nm		
		XCOARSE	MEDIUM	XFINE
1	174[516]	250.8	251.9	251.8
2	227[516]	362.8	361.1	362.8
3	275[347]	480.7	481.7	481.7
4	310[347]	603.6	602.0	601.2
5	342[347]	725.6	722.2	722.2
6	380[347]	848.6	843.9	842.3

8.3 ACCURACY OF THE HOMO AND LUMO EIGENVALUES

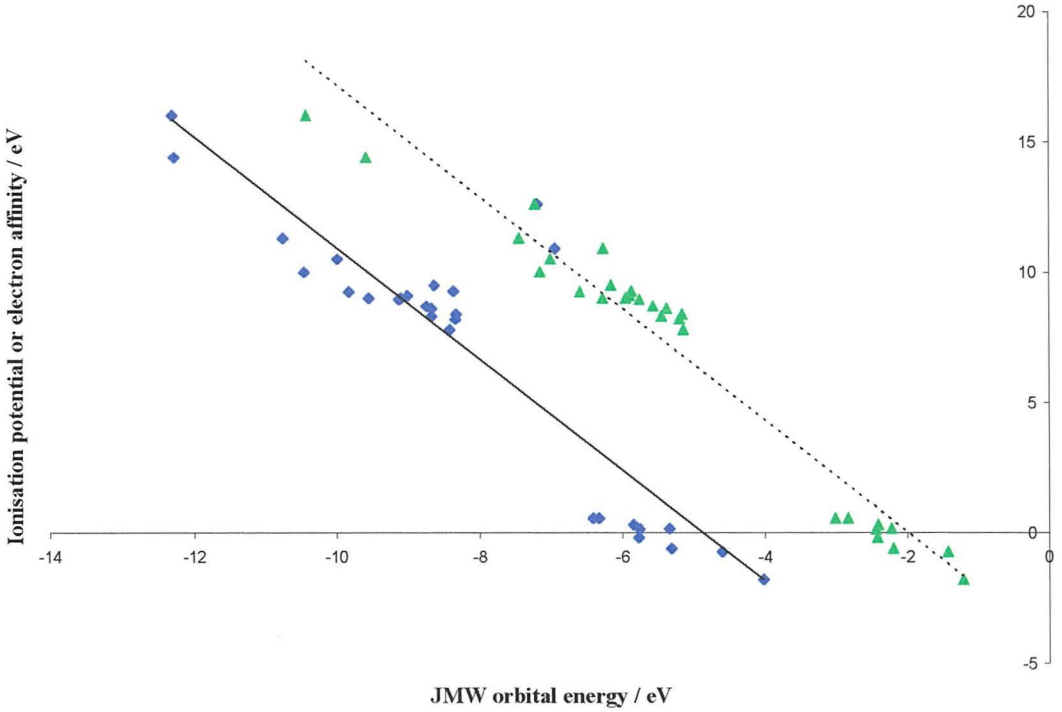
In a similar approach to that used in *Chapter 6*, DFT HOMO and LUMO energies were compared with ionisation potentials and electron affinities to find the most reliable method for reproduction of HOMO-LUMO gaps. Orbital energies were calculated (*Tables 8.6 to 8.9*) with a variety of local (JMW,[641] VWN[642]) and non-local (BP,[643] BPS,[644] BLYP,[645] BLS,[646] BVWN,[647] and BVS[648]) functionals for minimal (MIN) and DND basis sets respectively. The results have been compared graphically in *Graphs 8.1* and *8.2* for the local JMW and VWN functionals.

The effective one-electron potential in the Kohn-Sham equations is intimately related to the physics of electron correlation. It is useful to break down the exchange-correlation part of the potential into a part directly related to the total energy (the hole potential or screening potential), and a part related to the response of the exchange-correlation hole to density change. The latter part is poorly represented by the generalised gradient approximation, leading to the incomplete cancellation of self-interaction in the Coulomb integral by that in the exchange integral, and poor reproduction of orbital energies.[649] Higher r^2 valued least-squares regression equations are obtained (*Table 8.10*) with the inclusion of polarisation functions. Orbital energies for BVWN and VWN functionals are identical, as the Becke B88 gradient-corrected exchange functional is applied after the SCF calculation, changing only thermodynamic variables. Orbital energies are only affected when the Becke B88 correction is applied self-consistently. Although only a few calculations have been performed on non-local functionals, little difference can be seen in the orbital energies on application of nonlocal functionals, although differences may be seen when the functionals are also used for optimisation as molecular mechanics based geometries are used in this study.

Table 8.10 Correlations between calculated HOMO and LUMO energies, and ionisation potential and electron affinity respectively (n = 29).

Basis	Functional	α	β	r^2	σ	F
MIN	JMW	-10.35 (± 1.71)	-2.130 (± 0.20)	0.800	2.33	108
MIN	VWN	-10.21 (± 1.69)	-2.137 (± 0.21)	0.801	2.33	109
DND	JMW	-4.25 (± 0.52)	-2.150 (± 0.09)	0.953	1.33	546
DND	VWN	-4.04 (± 0.51)	-2.146 (± 0.09)	0.953	1.33	552

Graph 8.1 **Correlation of MIN (diamonds) and DND (triangles) JMW orbital energies with ionisation potentials and electron affinities.**



Graph 8.2 **Correlation of MIN (diamonds) and DND (triangles) VWN orbital energies with ionisation potentials and electron affinities.**

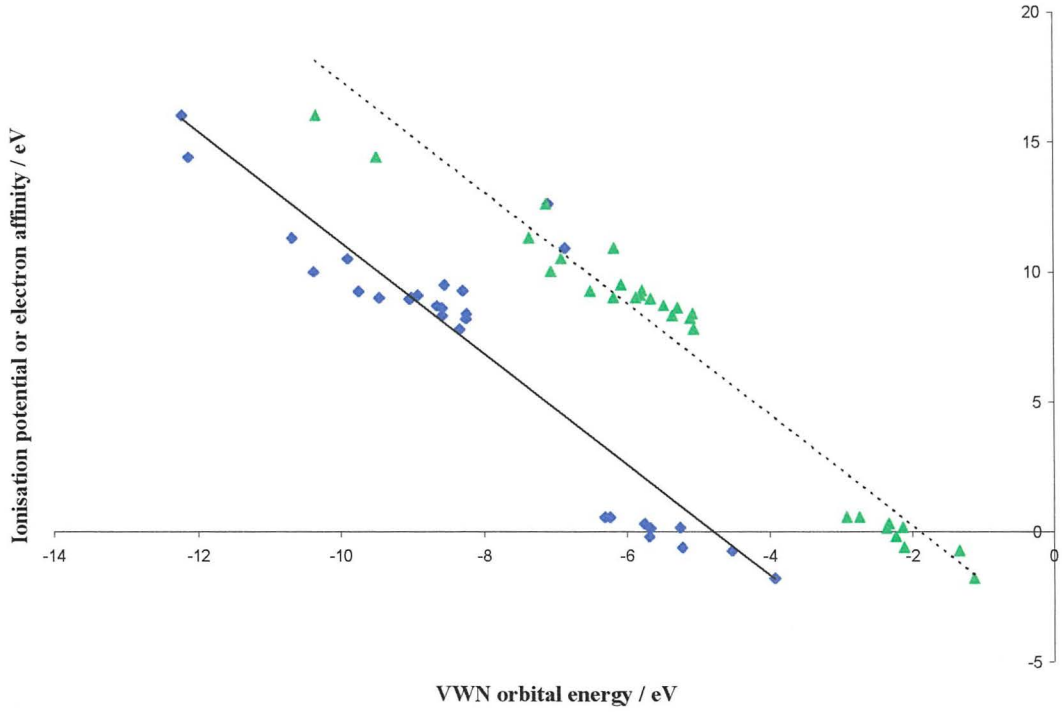


Table 8.6 Comparison of MINIMAL basis, medium mesh HOMO energies with ionisation potentials, IP (both in eV).

Compound	IP/ eV	JMW	VWN	BP	BPS	BLYP	BLS	BVS
Benzene	9.25[500]	-9.82	-9.74	-9.74	-9.57	-9.3	-9.49	-9.85
Benzaldehyde	9.49[500]	-8.62	-8.54					
Acetophenone	9.26[500]	-8.36	-8.28					
Toluene	9.0[500]	-9.54	-9.45					
Ethene	10.5[500]	-9.98	-9.89					
(*E*)-But-2-ene	9.1[500]	-9.00	-8.91					
1,3-Butadiene	9.0[500]	-9.09	-9.00	-9.00	-8.55	-8.56	-8.78	-9.15
1,3-Pentadiene	8.7[500]	-8.73	-8.64					
1,3,5-Hexatriene	8.3[500]	-8.66	-8.57					
1,3,5,7-Octatetraene	7.8[500]	-8.41	-8.33					
Cyclopropane	10.0[500]	-10.45	-10.37	-10.37	-10.35	-9.92	-10.20	-10.60
Methane	14.4[501]	-12.28	-12.12					
Ammonia	10.9[501]	-6.93	-6.85					
Water	12.6[501]	-7.18	-7.09					
HF	16.0[501]	-12.3	-12.21	-12.21	-12.37	-11.74	-12.32	-12.87
Fluoroethyne	11.3[501]	-10.75	-10.67	-10.67	-10.48	-10.21	-10.43	-10.81
Cyclopentadiene	8.6[500]	-8.66	-8.57					
Cyclohexadiene	8.4[500]	-8.32	-8.23					
Furan	8.95[500]	-9.12	-9.03					
Pyrrole	8.2[500]	-8.33	-8.24					

Table 8.7 Comparison of MINIMAL basis, medium mesh LUMO energies with electron affinities, EA (both in eV).

Compound	EA / eV	JMW	VWN	BP	BPS	BLYP	BLS	BVS
Ethene	-1.78[502]	-4.02	-3.93					
1,3-Butadiene	-0.62[502]	-5.31	-5.22	-5.22	-5.05	-4.78	-5.04	-5.41
Benzene	-0.72[503]	-4.60	-4.52	-4.51	-4.31	-4.08	-4.32	-4.67
Styrene	0.15[504]	-5.33	-5.25					
Naphthalene	-0.19[502]	-5.77	-5.68					
Biphenyl	0.13[504]	-5.75	-5.67					
Phenanthrene	0.31[503]	-5.84	-5.75					
Pyrene	0.56[503]	-6.32	-6.23					
Anthracene	0.56[505]	-6.40	-6.31					

Table 8.8 Comparison of DND basis, medium mesh LUMO energies with electron affinities, EA (both in eV).

Compound	EA / eV	JMW	VWN	BP	BPS	BLYP	BLS	BVS
Ethene	-1.78[502]	-1.21	-1.12					
1,3-Butadiene	-0.62[502]	-2.20	-2.11	-2.11	-1.85	-1.64	-1.75	-2.13
Benzene	-0.72[503]	-1.42	-1.33	-1.33	-1.06	-0.86	-0.95	-1.33
Styrene	0.15[504]	-2.22	-2.13					
Naphthalene	-0.19[502]	-2.42	-2.23					
Biphenyl	0.13[504]	-2.44	-2.36					
Phenanthrene	0.31[503]	-2.41	-2.33					
Pyrene	0.56[503]	-2.83	-2.74					
Anthracene	0.56[505]	-3.01	-2.92					

Table 8.9 Comparison of DND basis, medium mesh HOMO energies with ionisation potentials, IP (both in eV).

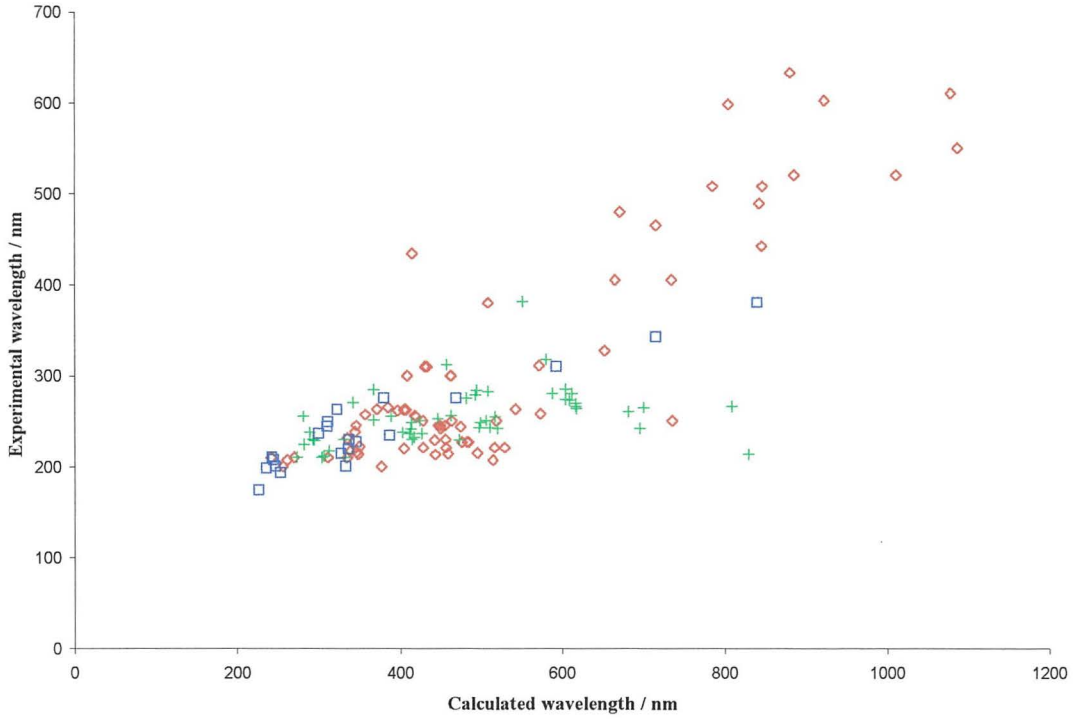
Compound	IP / eV	JMW	VWN	BP	BPS	BLYP	BLS	BVS
Benzene	9.25[500]	-6.59	-6.50	-6.50	-6.26	-6.02	-6.11	-6.49
Benzaldehyde	9.49[500]	-6.15	-6.07					
Acetophenone	9.26[500]	-5.86	-5.78					
Toluene	9.0[500]	-6.27	-6.18					
Ethene	10.5[500]	-7.00	-6.91					
(*E*)-But-2-ene	9.1[500]	-5.89	-5.80					
1,3-Butadiene	9.0[500]	-5.95	-5.86	-5.85	-5.63	-5.38	-5.49	-5.89
1,3-Pentadiene	8.7[500]	-5.56	-5.47					
1,3,5-Hexatriene	8.3[500]	-5.44	-5.35					
1,3,5,7-Octatetraene	7.8[500]	-5.14	-5.05					
Cyclopropane	10.0[500]	-7.14	-7.05	-7.05	-6.88	-6.57	-6.77	-7.20
Methane	14.4[501]	-9.58	-9.49					
Ammonia	10.9[501]	-6.26	-6.17					
Water	12.6[501]	-7.21	-7.12					
HF	16.0[501]	-10.42	-10.34	-10.34	-10.50	-9.85	-10.43	-11.00
Fluoroethyne	11.3[501]	-7.44	-7.36	-7.36	-7.15	-6.87	-7.05	-7.47
Cyclopentadiene	8.6[500]	-5.37	-5.28					
Cyclohexadiene	8.4[500]	-5.15	-5.07					
Furan	8.95[500]	-5.75	-5.66					
Pyrrole	8.2[500]	-5.19	-5.10					

8.4 TRAINING SETS

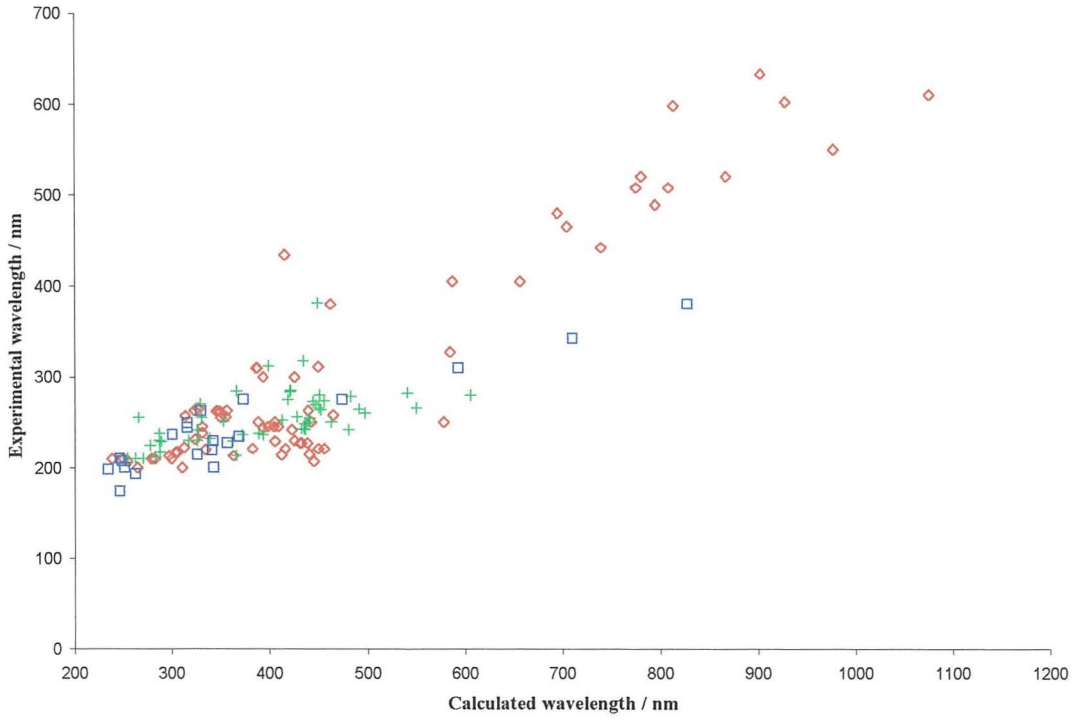
A training set of organic compounds (*Appendix 1*) was again developed in order to predict trends between wavelength maxima obtained from HOMO-LUMO gaps, and the experimental wavelength maxima. Wavelength maxima obtained from single point JMW functional, medium mesh calculations performed on molecular mechanics minimised geometries using various bases are tabulated in *Appendix 3*.

Again, the existence of polymethinic (◊), aromatic (+), and polyenic (□) correlations as predicted by Dähne's triad theory can be confirmed graphically as shown overleaf in *Graphs 8.3* to *8.6*. While semi-empirical excitation energies were overestimated, DFT calculated HOMO-LUMO gaps are underestimated, producing very high wavelength maxima. Low r^2 values were obtained for the aromatic correlation as azo dyes and other dyes ($\lambda > 400$ nm) obeying this correlation failed to converge. DFT is not particularly useful for polar molecules due to the failure of the local density approximation. The larger standard deviations lead to en effective merging of the correlations, notably for the MIN basis. Smaller standard deviations and improved r^2 values were obtained for the DN, DND and DNP bases on application of polarisation functions. DFT predicted wavelength maxima are less accurate than AM1 values.

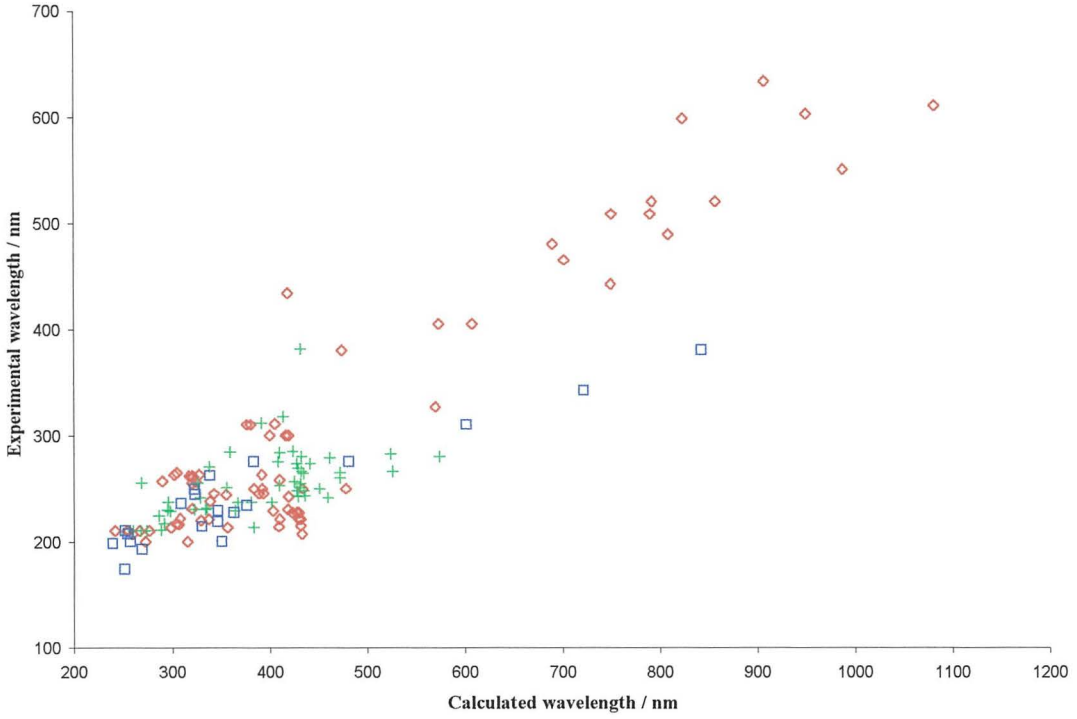
Graph 8.3 MINIMAL Basis training set (JMW functional, medium mesh).



Graph 8.4 DN Basis training set (JMW functional, medium mesh).



Graph 8.5 DND Basis training set (JMW functional, medium mesh).



Graph 8.6 DNP Basis training set (JMW functional, medium mesh).

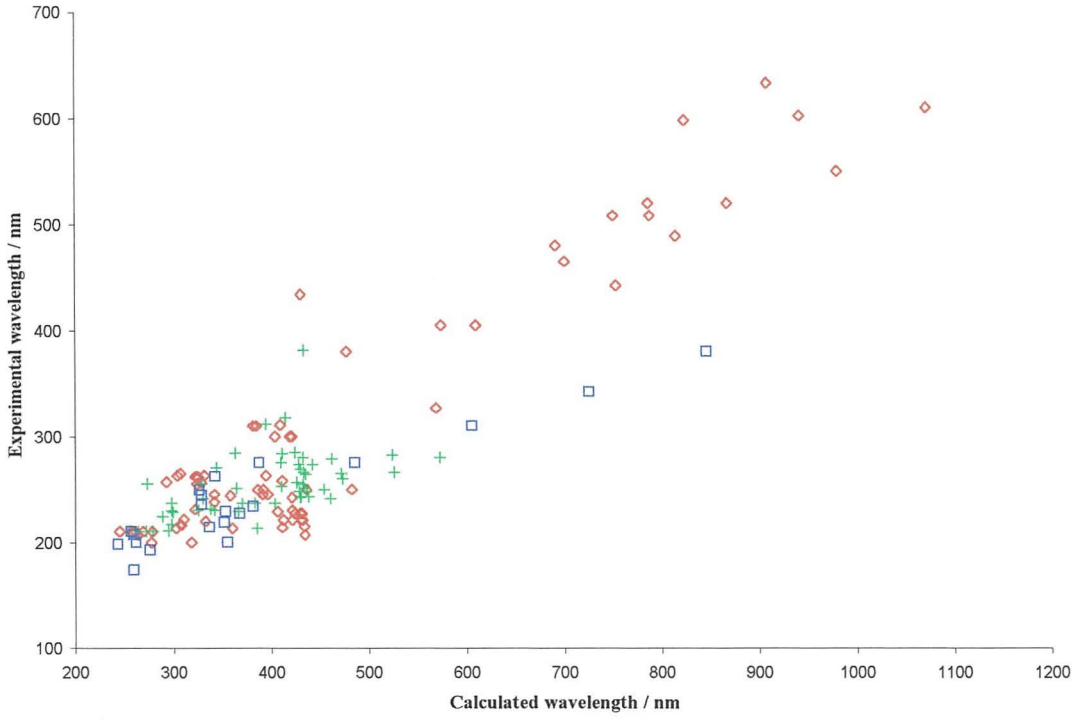


Table 8.11 Least-squares correlations between experimental and DFT JMW functional, medium mesh calculated wavelengths.

Basis	Class	α	β	r^2	n	σ	F
MIN	Polymethinic	34.5	0.52	77.1	74	57.7	241.8
	Aromatic	212	0.09	15.8	53	28.3	9.6
	Polyenic	130	0.30	88.9	21	17.5	152.7
DN	Polymethinic	37.8	0.57	85.6	74	45.6	429.0
	Aromatic	171	0.21	31.1	53	25.6	23.1
	Polyenic	127	0.31	88.3	21	18.1	142.7
DND	Polymethinic	47.2	0.56	87.1	74	43.2	485.9
	Aromatic	165	0.23	28.5	53	26.1	20.3
	Polyenic	126	0.30	88.5	21	17.9	146.9
DNP	Polymethinic	42.5	0.57	87.0	74	43.4	481.9
	Aromatic	163	0.23	28.5	53	26.1	20.3
	Polyenic	124	0.31	88.5	21	17.9	146.7

8.5 MESH SIZE VARIATION

Due to time restrictions, only the effects of mesh size on JMW, DND calculations for a small proportion of training set compounds were studied as shown in *Table 8.12.* Again, the greatest deviations occur for polar molecules as the LDA breaks down, as a large number of mesh points are needed when the charge density varies too quickly.

8.6 OPTIMISATION

Only molecules with a high level of internal symmetry were optimised. A high level of symmetry reduces the number of modes needed for optimisation, reducing computational times. The effect of optimisation on the value of λ is shown in *Table 8.13.* Surprisingly, better wavelengths are obtained before optimisation!

8.7 OTHER FUNCTIONALS

DFT maxima have largely calculated using the JMW local functional. Differences between wavelength maxima calculated using the JMW and VWN local functionals are shown in *Table 8.14.* The calculated wavelength maxima are essentially identical, although the calculated orbital energies are different – VWN orbital energies are 0.003 Hartrees lower than JMW orbital energies. It would be interesting to evaluate maxima using non-local functionals, particularly for polar molecules.

Table 8.12 Effect of mesh size on JMW DND maxima.

Compound	Experimental λ / nm	Calculated λ / nm		
		XCOARSE	MEDIUM	XFINE
(213)	184[363]	241.5	240	240
(219)	193[363]	268.2	270.1	270.1
(220)	200[653]	257.6	257.9	257.7
(221)	207[361]	254.9	255	255
(222)	210[653]	252.7	252.9	253
(223)	210[653]	260	260.7	260.7
(224)	210[653]	265.1	268.4	268.2
(225)	210[653]	268.5	274.3	274.2
(213)	211[361]	285.5	289.1	289.1
(226)	213[653]	362.5	383.6	383.6
(176)	214[337]	332.4	330.9	331.2
(227)	217[361]	286.6	292.3	292.3
(228)	219[590]	347.3	346.8	346.8
(229)	224[361]	286.4	287.1	287.1
(230)	229[590]	346.5	347.3	347.3
(231)	230[361]	328.8	347.3	347.3
(232)	230[382]	298.2	323.0	323.2
(233)	234[590]	373.2	376.9	376.6
(234)	236[653]	309.2	309.2	309.2
(235)	242[650]	366.3	429.1	429.5
(236)	243[368]	369.6	429.1	429.1
(215)	244[597]	324.8	324.1	323.9
(218)	249[597]	325.5	323.2	323.2
(237)	251[651]	343.1	356.0	356.3
(238)	255[652]	268.2	269.2	269.3
(217)	262[597]	338.8	339.1	339.3
(239)	269[361]	369.3	429.1	429.5

Table 8.13 Effect of optimisation on JMW DND medium mesh maxima.

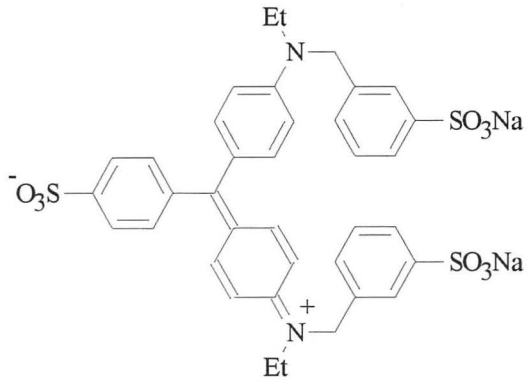
Compound	Experimental λ / nm	Single point λ / nm	Optimised λ / nm
(213)	184[363]	240.0	238.6
(223)	210[653]	260.7	260.7
(214)	211[361]	289.1	286.4
(176)	214[337]	330.9	315.6
(228)	219[590]	346.8	322.1
(229)	224[361]	287.1	289.5
(231)	230[361]	334.3	322.5
(232)	230[382]	323.0	308.9
(234)	236[653]	309.2	313.0
(235)	242[650]	429.1	428.7
(236)	243[368]	429.1	416.9
(215)	244[597]	324.1	340.6
(238)	255[652]	269.2	267.4
(217)	262[597]	323.2	327.6
(239)	269[361]	429.1	395.6

Table 8.14 DND medium mesh calculated maxima for different functionals.

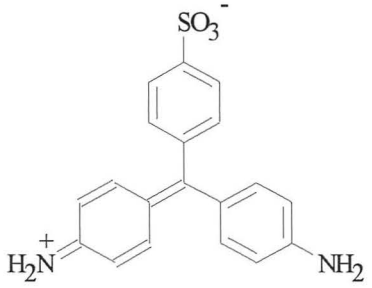
Compound	Experimental λ / nm	Calculated λ / nm	
		JMW	VWN
(219)	193[363]	270.1	270.0
(221)	207[361]	255	255
(214)	210[361]	289.1	289
(176)	214[337]	330.9	331.2
(227)	217[361]	292.3	292.1
(228)	219[590]	346.8	346.8
(229)	224[361]	287.1	287.0
(177)	227[590]	363.1	362.8
(230)	229[590]	347.3	347.6
(232)	230[382]	323.0	322.7
(233)	234[590]	376.9	376.6
(234)	236[653]	309.2	309.2
(215)	248[597]	324.1	323.9
(218)	249[597]	323.2	323.4
(238)	255[652]	269.2	269.2
(217)	262[597]	339.1	338.8
(239)	269[361]	429.1	428.7
(179)	275[348]	481.7	481.7
(181)	310[348]	602.0	601.2
(183)	342[348]	722.2	722.2
(184)	380[348]	843.9	842.3
(187)	327[591]	570.3	568.2
(240)	380[528]	474.7	474.2
(189)	465[591]	702.1	702.1
(193)	480[591]	690.4	690.4
(192)	550[591]	988.5	990.6
(194)	610[591]	1082.4	1085.0

8.8 PREDICTION OF LIGHT GREEN WAVELENGTH MAXIMA

The first model of a cytological dye studied using *DMol* was the smaller central portion **(149)** of the light green SF yellowish molecule **(13)**. The sulphonated aryl groups are insulated from the central portion by the methylene groups, and should not affect the chromophore. The central core was minimised using *DISCOVER* to a gradient of 0.0003 kcal mol^{-1} $Å^{-1}$ using the CVFF forcefield. The resulting model had C_s symmetry. Several DND basis, JMW functional, medium mesh calculations were carried out on a model, and the results shown below.



(13)



(149)

Table 8.15 Attempts to calculate maxima for the central core of light green.

Attempt	Mixing_alpha	Mixing_beta	Observation
1	0.3	0.3	Oscillates every 2-3 cycles
2	0.2	0.2	Oscillates every 2-3 cycles
3	0.1	0.1	oscillates every 4-5 cycles
4	0.05	0.05	oscillates every 4-5 cycles
5	0.025	0.025	oscillates every 4-5 cycles
6	0.0125	0.0125	Oscillated

Mixing_Alpha (FASCF) and Mixing_Beta (FBSCF) are the mixing coefficients for charge and spin densities respectively. Low values result in smoother convergence at the expense of computational time. Reducing the values of Mixing_Alpha and Mixing_Beta did not seem to help, as the calculations were now taking over 100 cycles without reaching the desired level of self-consistency. On the sixth attempt, the SCF oscillated at a convergence of 0.0004 to 0.0005. Unlike other runs, a list of orbital eigenvalues was obtained, some of which are shown below.

Table 8.16 Selected eigenvalues from attempt 6.

Eigenvalues / a.u.	Symmetry	Occupancy	
-0.23174	A'	2	
-0.21403	A"	2	
-0.20703	A'	2	
-0.19335	A'	2	
-0.18215	A"	2	HOMO
-0.18190	A"	0	LUMO
-0.10212	A"	0	
-0.08492	A'	0	

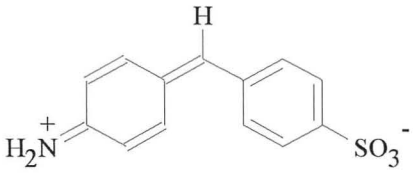
In an attempt to remove this degeneracy, charge smearing was introduced, allowing electrons to be smeared out among all orbitals within a given ΔE (expressed in a.u.) of

the Fermi level. The procedure improves convergence of the SCF procedure by allowing orbitals to relax more rapidly. JMW DN basis, medium mesh results using Mixing_Alpha = Mixing_Beta = 0.05 and SMEAR = 0.03 a.u. are shown below.

Table 8.17 Charge smearing of 0.03 a.u.

Eigenvalues / a.u.	Symmetry	Occupancy	
-0.24219	A'	2	
-0.22510	A"	2	
-0.21570	A'	2	
-0.20977	A'	1.94	
-0.20391	A"	1.55	
-0.18841	A"	0.51	HOMO
-0.10534	A"	0	LUMO
-0.10042	A'	0	

Another attempt at removing the degeneracy was to use a model with C_1 symmetry, accomplished by methylating the quaternary nitrogen. Again, the SCF equations also oscillated when no smearing was introduced. The neutral aryl functionality was removed to give a simpler core **(241)**, but again the solutions oscillated. Both restricted and unrestricted calculations with other functionals gave little success.



(241)

Calculations have also been performed for other dyes using the VWN medium mesh calculations. Although wavelength maxima can be predicted with varying degrees of accuracy (20 nm to 100 nm) for some polymethinic dyes, most other dyes tend to have HOMO and LUMO orbitals less than 1 eV apart. Local DFT treatment cannot be used to accurately predict wavelength maxima of polar or charge molecules.

8.9 CONCLUSIONS

The majority of the results presented in this chapter have been obtained using density functional theory employing the JMW local density functional. As DFT works best for metallic systems with constant local charge density, calculations were initially

performed for polyenes, with a relatively constant charge density in the backbone. In complete contrast to the semi-empirical results (*Chapter 6*), DFT wavelength maxima were overestimated, due to an underestimation of the singlet excitation energy. Optimisation plays an important role in wavelength maxima, followed by basis set quality. Mesh size differences only affect molecules with rapidly changing charge density. Orbital energies are poorly reproduced, due to the failing representation of the response of the exchange-correlation hole to density change by the GGA.[649] A training set of unsaturated organic molecules was employed to investigate trends in calculated maxima, the results of which could again be interpreted in terms of Dähne's triad theory. DND or DNP bases must be employed for polar molecules, together with a large numerical mesh size, due to the breakdown of the LDA. Although JMW and VWN wavelength maxima are essentially identical, it would be interesting to evaluate whether non-local wavelength maxima are different. Maxima of polar or charged dyes such as light green SF yellowish could not be calculated due to non-convergence.

8.10 WHICH METHOD IS BEST FOR PREDICTING ELECTRONIC WAVELENGTH MAXIMA?

Empirical and π-electron methods can be used to predict maxima of simple systems. Semi-empirical NDO based methodologies employing a HOMO-LUMO approximation to the singlet excitation energy can be used to calculate wavelength maxima. However, as the molecular J and K integrals are not explicitly evaluated, the singlet excitation energies are overestimated, producing low values for wavelength maxima. Upon correction, maxima can be predicted to within 25 to 50 nm depending on the system. Improved maxima can be obtained using semi-empirical minimised geometries. DFT orbital energies employing local functionals are poorly reproduced, yielding high values of wavelength maxima. Unfortunately, the LDA fails for highly polar / ionic species, such as dyes. On reflection, the faster semi-empirical HOMO-LUMO methods can be used for initial results, followed by CI approaches for improved maxima and intensities (given by oscillator strengths).

Elucidation of the smear test staining mechanism will be attempted in the next chapter, so that more intense/selective stains can be designed.

CHAPTER 9
ELUCIDATION OF THE CERVICAL SMEAR STAINING MECHANISM

9.1 INTRODUCTION

Considerable effort and research has been directed toward achieving the tinctorial qualities[654] to which most cytopathologists and cytotechnologists have become accustomed, and on which their evaluation of specimens is partially based. Problems frequently occur when a component of a stain is out of stock, such as the recent world shortage of light green SF yellowish **(13)**, where some stockists substituted fast green FCF **(81)** for light green in their solutions.[655] By understanding how a stain interacts with cellular components, it should be possible to design more intense and possibly cheaper analogues, which may provide better differentiation between tissues.

Little is known about the mechanism of the Papanicolaou stain, due to the empirical method employed by Papanicolaou in its development. Hypotheses concerning the staining mechanisms of Bismarck brown Y, eosin Y, orange G, and light green SF yellowish will be presented. For example, we believe that light green SF yellowish binds to the collagen rich cytoplasm of parabasal cells, but does not bind to the keratin rich cytoplasm of superficial cells. The complex chemistry of hæmatoxylin based compounds will be reviewed, together with an analysis of hæmalum-nucleic acid interactions, leading to several hæmalum-nucleotide models. The absorption maxima of unbound and bound hæmalums will be calculated using semi-empirical HOMO-LUMO methodologies described in *Chapter 6*. Dyes with similar staining properties as the original Papanicolaou dyes will be presented as possible analogues.

9.2 MECHANISM OF THE PAPANICOLAOU STAIN

Little is known about the mechanism of the Papanicolaou stain due to its conglomeration of alchemic ingredients. Absorption spectra of normal glandular cell and tumour cell nuclei have been published by Wied *et al*,[656] while Holmquist *et al*[657] have presented nuclear and cytoplasmic spectra of orangeophilic, cyanophilic and unspecified cervical cells. Mean absorption spectra (*Table 9.1*) of then cervical cell types stained with Harris hæmatoxylin, OG-6, and EA-50 stains were measured by

Galbraith *et al.*[658] Maxima at ~ 530 nm, 595 nm and 635 nm result from eosin Y **(29)**, Harris hæmatoxylin, and light green SF yellowish **(13)** respectively. Secondary nuclear maxima from most normal cells, except superficial cells, occur at ~ 625 nm, while secondary nuclear maxima of abnormal cells occur at ~ 595 nm, probably due to a different amounts of bound hæmalum and light green SF yellowish. In all spectra the primary peak is located at ~ 531 nm, and is the wavelength of choice for isolating cells from the background,[657,659,660] as the absorbance differences *(Table 9.2)* between nuclei and cytoplasm reach a maximum at 531 nm. A more effective measure of the nuclear-cytoplasmic contrast is the maximum absorbance ratio *(Table 9.2)*. The optimal wavelength for distinguishing between nuclei and cytoplasm varies with cell type, but falls in the range between 560 nm and 605 nm with a mean of 575 nm.[658] Aggarwal and Bacus,[660] Holmquist *et al*[657] and Wied *et al*[656] used wavelengths of 569 nm, 570 nm and 590 nm in the image processing of Papanicolaou stained cells.

Table 9.1 Spectral maxima found in ten types of cervical cells[658] after staining with Harris hæmatoxylin, OG-6 and EA-50 stains.

Cell type	Main spectral peak		Secondary peak	
	λ / nm	Absorbance	λ / nm	Absorbance
Nucleus				
Superficial	529.8	1.139	585.3	0.499
Intermediate	531.0	0.858	628.7	0.627
Parabasal	531.0	1.012	622.9	0.528
Metaplastic	530.4	0.906	623.7	0.419
Endocervical	531.0	0.670	628.7	0.494
Other normal	529.2	1.873		
Dysplastic	531.0	1.665	594.1	0.534
Carcinoma *in situ*	530.4	1.165	591.1	0.554
Invasive carcinoma	530.4	2.058	594.8	1.052
Other abnormal	529.2	1.879	599.3	0.852
Cytoplasm				
Superficial	528.0	0.370		
Intermediate	530.4	0.196	638.8	0.211
Parabasal	531.0	0.527	636.3	0.341
Metaplastic	529.8	0.517	636.3	0.279
Endocervical	531.0	0.360	638.0	0.359
Other normal	528.6	0.874	627.0	0.171
Dysplastic	531.0	0.795	638.0	0.133
Carcinoma *in situ*	529.8	0.668	634.6	0.205
Invasive carcinoma	530.4	0.826	635.4	0.345
Other abnormal	529.8	1.029	636.3	0.459

Table 9.2 Nuclear-cytoplasmic absorbance comparisons.[658]

Cell type	Maximum difference	λ / nm	Maximum ratio	λ / nm
Superficial	0.772	531.0	1.507	605.4
Intermediate	0.662	531.0	5.202	561.1
Parabasal	0.485	531.0	2.553	567.8
Metaplastic	0.389	531.0	2.444	569.9
Endocervical	0.310	531.0	2.368	563.1
Other normal	1.000	529.8	4.413	580.3
Dysplastic	0.870	531.0	5.059	586.7
Carcinoma *in situ*	0.499	531.0	3.784	574.7
Invasive carcinoma	1.231	530.4	4.457	574.7
Other abnormal	0.852	528.6	3.090	571.2

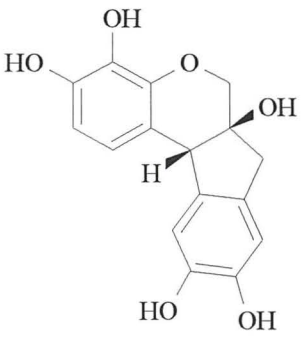
Table 9.3 Spectral maxima of superficial cells stained with a single dye.[661]

	Nucleus		Cytoplasm	
	λ / nm	A	λ / nm	A
Hæmalum	590.4	0.861	594.8	0.210
Orange G	483.4	0.508	483.4	0.303
Light Green SF yellowish (1° peak)	645.8	0.779	645.8	0.697
Light Green SF yellowish (2° peak)	434.0	0.187	434.0	0.182
Eosin Y	530.4	0.313	529.2	0.107

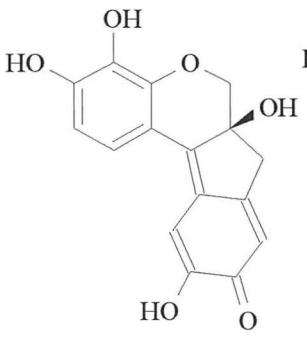
Table 9.4 Binding strengths for Papanicolaou stained cervical cells.[661]

Cell type	Hæmalum	Eosin Y	Light green SF	Orange G
Normal nuclei				
Superficial	0.489	0.731	0.001	0.217
Intermediate	0.513	0.427	0.201	0.083
Parabasal	0.465	0.592	0.140	0.058
Metaplastic	0.373	0.555	0.109	0.104
Endocervical	0.413	0.323	0.167	0.078
Other	0.634	1.259	0.008	0.274
Abnormal nuclei				
Dysplastic	0.533	1.159	0.018	0.107
Carcinoma *in situ*	0.549	0.674	0.039	0.169
Invasive carcinoma	1.038	1.074	0.059	0.226
Other	0.832	1.067	0.080	0.302
Normal cytoplasm				
Superficial	0.036	0.333	0.000	0.101
Intermediate	0.100	0.111	0.132	0.030
Parabasal	0.183	0.361	0.169	0.049
Metaplastic	0.153	0.362	0.148	0.066
Endocervical	0.176	0.213	0.212	0.033
Other	0.184	0.667	0.023	0.142
Abnormal cytoplasm				
Dysplastic	0.111	0.691	0.033	0.058
Carcinoma *in situ*	0.145	0.525	0.077	0.117
Invasive carcinoma	0.233	0.612	0.148	0.127
Other	0.271	0.735	0.212	0.142

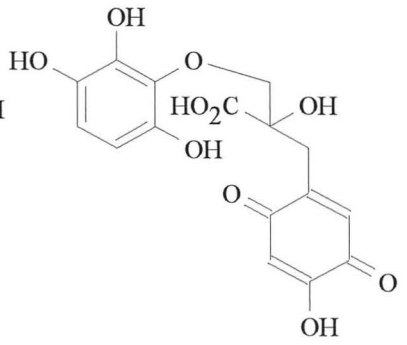
Marshall *et al*[661] quantified dye binding to Papanicolaou stained cervical cells using a spectral subtraction technique already successfully applied to azure B / eosin stained blood cells.[662] Spectra of each component (*Table 9.3*) were subtracted from the original spectrum, enabling binding strengths (*Table 9.4*) to be calculated. Commercial Harris hæmatoxylin solutions contain hæmatoxylin **(105)**, hæmatein **(242)**, hæmalum **(243)** (having a high affinity for chromatin and RNA-rich cytoplasms), and occasionally oxyhæmatein **(244)**.[663,664] Only Harris hæmatoxylin, orange G **(25)**, light green SF yellowish **(13)** and eosin Y **(29)** bind to nuclei and cytoplasm, although diffuse cytoplasmic staining is observed.



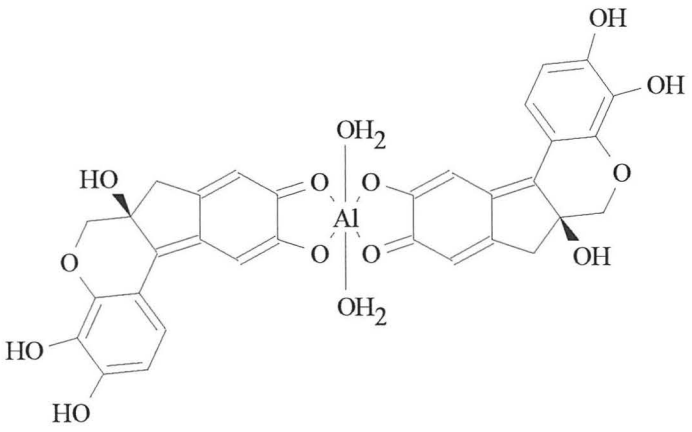
(105)



(242)



(244)



(243)

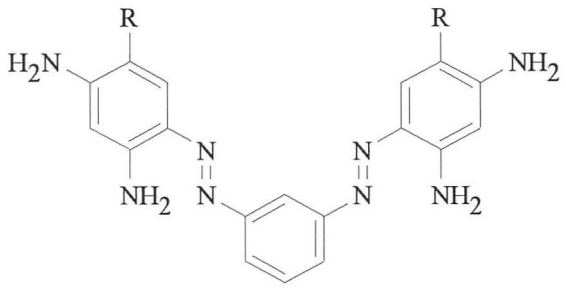
Binding strengths proportional to the dye uptake were found by subtraction. A residual spectrum, with very small baseline area, was obtained after subtraction in the following order: hæmalum, eosin Y, light green SF yellowish and orange G. The following conclusions were drawn:

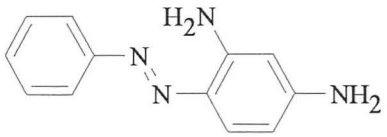
- Compound dyes, such as the falgic acids (*Section 2.5*), produced when light green SF yellowish reacts with acid fuchsin, are not formed.
- Stains such as eosin Y do not dimerise when staining cellular objects.[662]
- The chromophores do not interact, such as the decolourisation produced when orange G reacts with light green SF yellowish (*Section 2.5*).

Eosin Y undergoes dimerisation, its spectra showing both monomer and dimer forms.[665-667] The peak at ~ 490 nm in the absorbance spectrum of nuclei of eosin Y stained superficial cells being due to the dimer.[662] The ratio of dimer to monomer appears to be constant, allowing a single eosin Y spectrum to be used.[662] Dimerisation is not documented for hæmalum, orange G and light green, although a shoulder at 440 nm and a secondary maximum at 434 nm for hæmalum and light green may be due to dimerisation. Orange G shows a maximum at 483 nm perhaps due to a dimer (orange II dimerisation has been proved[668]), and a prominent shoulder at ~ 500 nm, perhaps due to monomer. Hæmalum binds more strongly to nuclei than to cytoplasm, orange G and eosin Y are less specific for nuclei, while light green SF yellowish binds almost equally to nuclei and cytoplasm (does not bind to superficial cells).

9.3 BISMARCK BROWN Y

Some commercially available EA stains[669,670] still retain Bismarck brown Y (prepared by the self-diazotisation of *m*-phenylenediamine) as it stains mucin,[671] a range of glycoproteins produced by mucous membranes, containing many short *O*-linked glycans distributed along polypeptide chains.[671] Mucin staining methods employing their basophilic[672] and metachromatic[673] nature have been investigated. List[674] was the first to use an alcoholic solution[675] of Bismarck brown Y to stain water-soluble mucoproteins and mucopolysaccharides. Marshall[661] however provided objective data that Bismarck brown Y should be omitted from the EA solution as it plays no role in the staining reaction.[676-678] Additionally, Bismarck brown Y may form an insoluble complex with phosphotungstic acid, reducing the concentration of the latter,[677,679,680] affecting both pH and stain performance. Bismarck brown Y **(54)** is used the staining of embryonic mucin, cartilage and goblet cells.[681,682] Bismarck brown R **(245)** can also be used, although it is redder.[683] Chrysoidin Y **(246)** also gives similar results.[683]





(54) R = H **(245)** R = Me **(246)**

9.4 HÆMATOXYLIN

9.4.1 Hæmatoxylin chemistry

Hæmatein-aluminium complexes are important stains used in histology and cytology.[684] A uniform description of the various processes taking place during staining has not been achieved. The structures of hæmatein-aluminium complexes have not been unambiguously elucidated, due to differing experimental conditions and contamination. Pure hæmatoxylin obtained by purification of commercial samples,[663,685-688] can be transformed into pentaacetylhæmatoxylin,[689] or hæmatein.[690]

a) Hæmatoxylin

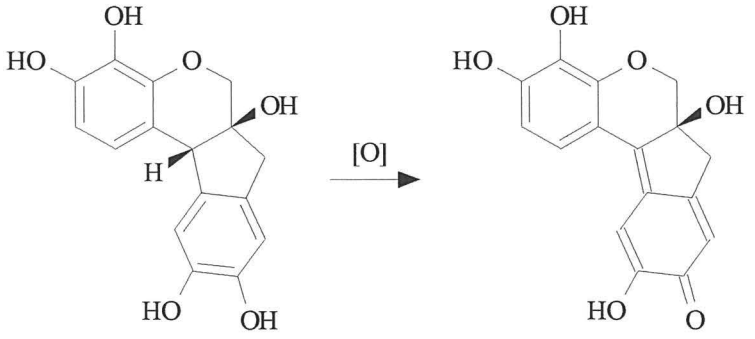
Hæmatoxylin was first described in 1812 by Chevreul,[691] who isolated it as colourless crystals from the corewood of the logwood tree, *Hæmatoxylon campechianum*, which is indigenous to Mexico but cultivated in the West Indies. Elemental analysis was determined by Erdmann,[692] and confirmed by Hesse[693] and others. The structure was elucidated by Perkin and Robinson,[694] synthesised by Dann and Hofmann,[695] and Morsingh and Robinson,[696] and verified by ^{1}H NMR spectroscopy.[697] Crude hematoxylin is cheap and easily prepared by boiling logwood chips with water, filtering off the aqueous extract and evaporating to dryness. Certified hæmatoxylin is obtained by extracting hematoxylin with ether.

The UV absorption spectrum of hæmatoxylin (Htx) has been recorded in methanol.[698] The extinction coefficient ε_{Htx} at the maximum of the absorption, λ_{max} = 292 nm, was concentration dependent and decreased with increasing initial concentration of the sample. This decrease in ε_{Htx} was attributed to intermolecular interactions between the hæmatoxylin molecules. The molar extinction coefficient of free hæmatoxylin at 292 nm was found to be 5700 M^{-1} cm^{-1}. Lalor and Martin[686a] previously found ε_{Htx} in

methanol to be 5148. Morsingh and Robinson[621] reported ε_{Htx} as 4360 at 293 nm for synthetic *dl*-hæmatoxylin and 6790 for their natural product.[699] Freshly prepared aqueous solutions of hæmatoxylin absorb at λ_{max} = 289 nm (ε_{Htx} = 5550). When allowed to stand, a weak band is formed at 445 nm, attributed to hæmatein. The ageing process does not stop at the hæmatein stage, but proceeds further, when an intense, broad absorption band at 390 nm accompanied by a weak shoulder at 445 nm is observed. The spectra of aqueous solutions of hæmatoxylin are also pH-dependent.

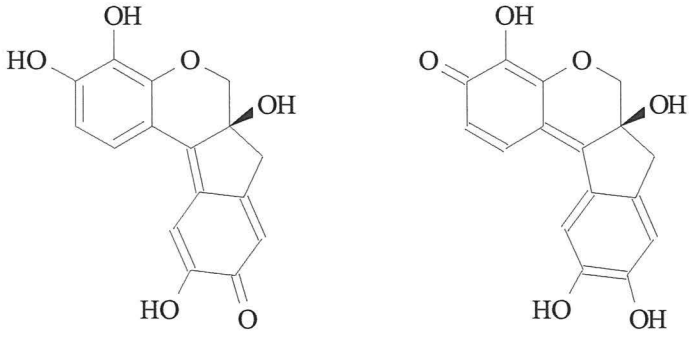
b) Hæmatein

Oxidative formation of hæmatein **(242)** from hæmatoxylin **(105)** by natural ageing in air over a period of months, or with various oxidising agents[663,700] such as hydrogen peroxide, chloral hydrate, potassium permanganate and sodium iodate, proceeds *via* radical intermediates which can be detected in alkaline media by ESR spectroscopy.



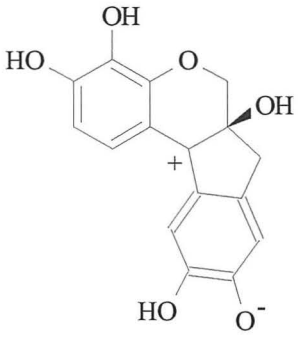
(105) **(242)**

Several isomeric structures, such as **(242)** and **(247)** can be constructed for hæmatein. The ^{1}H-NMR spectrum of a freshly prepared solution of pure hæmatein in d_6-DMSO shows resonances for only one species, **(242)**.[663]

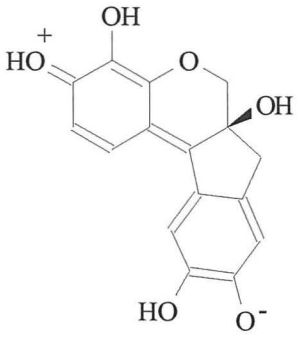


(242) **(247)**

Puchtler *et al*[684] mention that the usual covalent quinoid structure **(242)** does not satisfactorily describe the properties of hæmatein,[701] and proposed the use of ionic valence structures, **(248)** and **(249)**.[702] IR spectra of hæmatein have been recorded as KBr discs and in d_6-ethylene glycol.[703] The C=O stretching vibration is observed at 1602 cm^{-1} in solution, and 1605 cm^{-1} as a KBr disc.[698]

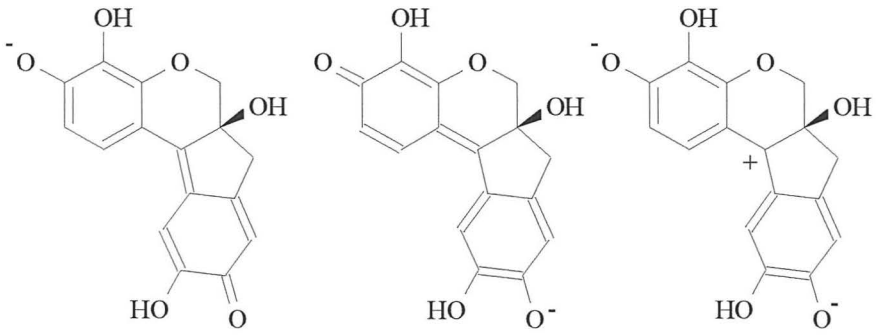


(248)



(249)

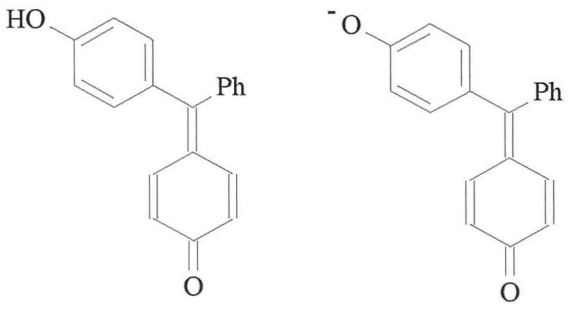
The UV spectrum of analytically pure hæmatein (Hm) in methanol shows a strong absorption band at 445 nm (ε_{Hm} = 35,800).[698] In dilute solutions, ε_{Hm} is independent of the initial concentration. Lalor and Martin[686a] determined ε_{Hm} as 35,490 in methanol. Values of Morsingh and Robinson[696] for their synthetic *dl*-hæmatein (22,400) and natural product (23,900) are smaller.[698] A weak absorption can be observed at 560 nm in the spectrum of hæmatein in methanol, due to the anion, which is in equilibrium with hæmatein at very low concentration. The UV spectra of aqueous solutions of hæmatein are strongly pH-dependent. With increasing pH, the intensity of the absorption of neutral hæmatein **(242)** at 445 nm decreases, while the intensity absorption of the monoanion (Hm^-) **(250)** at 556 nm increases. The colour of the solution at the same time changes from reddish-yellow **(242)** to violet **(250)**. The isobestic point at 486 nm confirms that **(242)** and **(250)** are in equilibrium. A slight deviation from the isobestic point is observed indicative of the formation of the dianion. At 556nm, ε_{Hm^-} and ε_{Hm} are found to be 58,300 and 2100 respectively.[698] At 445 nm, ε_{Hm} is found to be 33,600 and ε_{Hm^-} found to be 7000 in water.[698] Lalor and Martin[686a] determined pKa for the first dissociation step of hæmatein to be 6.86, in good agreement with 6.71 obtained by Bettinger and Zimmermann.[698] In strongly basic media hæmatein ages rapidly and undergoes irreversible changes. Hæmatein is also unstable in strongly acidic media.[686b,690,704]



(250) **(251)** **(252)**

For the ground and first excited state of hæmatein, structures **(242)**, **(248)** and **(249)** need to be considered; **(248)** and **(249)** are higher in energy than **(242)** due to charge separation. Structure **(242)** participates to a large extent in the linear combination for the ground state, while **(248)** and **(249)** contribute more to the first excited state. The latter structures weaken the C=O bond and are involved in the shift of the C=O stretching frequency to longer wavelengths (1602 cm^{-1}). Intramolecular hydrogen bonding between the OH and C=O groups supports the inductive effect, as seen for tropolone (1620 cm^{-1}) and *o*-hydroxyacetophenone (1643 cm^{-1}).[705] In the case of the hæmatein monoanion Hm^-, covalent structures **(250)** and **(251)** are almost equivalent and differ from the higher energy species **(252)**. Structures **(250)** and **(251)** participate with large and nearly equal weights in the ground state.[706]

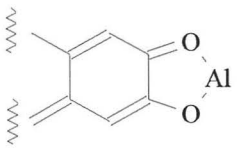
A similar bathochromic shift and hyperchromic effect is also observed for benzaurine. In aqueous solution, neutral benzaurine **(253)** is reddish-yellow, while in alkaline solution, the violet coloured anion **(254)** is formed.[707]



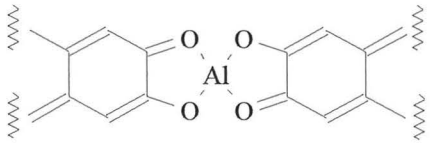
(253) **(254)**

c) Hæmalum

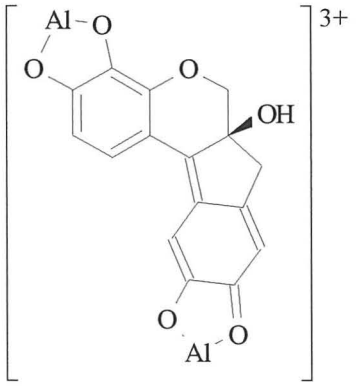
Several reliable staining procedures using hæmatein and polyvalent metal cations such as Al^{3+} and Fe^{3+} have been described by Romeis,[708] although knowledge of the molecular principles involved is incomplete and contradictory. Arshid *et al* were the first to investigate the composition of hæmatein-metal (hæmalum) complexes.[685b] 1:1 HmM complexes **(255)** were formed from hæmatein and metal cations in a first step. The 1:1 HmM complexes were reported to be converted into 2:1 Hm_2M complexes **(256)** in the presence of an excess of metal cations, a process which defies logic. Although the existence of the 1:1 complexes has been confirmed, experimental confirmation of the 2:1 complexes is still lacking. Arshid *et al* assumed that hæmatein forms chelate complexes with metal cations.[685b] Polyvalent metal cations generally bind not only two but also four or six ligands and thus the 2:1 complexes of Arshid *et al* become probable.[685] Hæmatein also forms 2:1 complexes with Sn^{2+} and Sn^{4+}.[690]



(255)



(256)



(257)

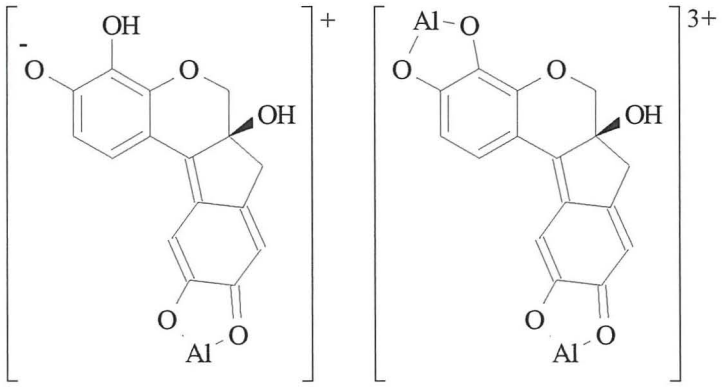
In contrast to Arshid *et al*,[685b] Harms[709] and Baker[710] proposed that in the presence of a large excess of metal cations, a second metal cation is bound to the OH groups in the 3 and 4 positions, giving rise to a 1:2 $HmAl_2$ complex **(257)**.

2:1 Hm_2M complexes should be anions according to Puchtler *et al.*[684] Baker[710] came to the conclusion that hæmalum complexes are cations which favours HmAl and $HmAl_2$ complexes, while Marshall and Horobin,[711] Marshall *et al*,[561] and Puchtler *et al*[684] favour the Hm_2Al stoichiometry. In the course of hæmalum staining and subsequent blueing, all colour tones between red, violet and blue can be observed, corresponding to different molecular species.

Depending on the pH value, hæmatein can exist in aqueous solution either as the neutral, uncharged molecule **(242)**, the monoanion **(250)**, or the dianion Hm^{2-}.[698] The dianions are not stable in aqueous alkaline solutions. In acidic solution, aluminium cations form hexahydrate complexes, $Al(H_2O)_6^{3+}$, while aluminates $Al(OH)_4^-$ are formed in alkaline solution. At pH 5-8, poorly soluble $Al(OH)_3$ precipitates. The pKa value of the first dissociation step of hæmatein is 6.70.[698] At about pH 8.5, the prototropic equilibrium between Hm^- and Hm lies completely to the side of Hm^-. Similarly, at about pH 4.5, the prototropic equilibrium lies completely to the side of Hm. In strongly acidic media hæmatein undergoes an acid-catalysed consecutive reaction.[686b,690,704]

i) Studies in acidic media

The spectra of buffered[712] hæmatein-aluminium solutions have been recorded at pH 4.7.[713] The absorption band of neutral hæmatein **(242)** can be seen at 445 nm. With increasing aluminium concentration, the intensity of the 445 nm band decreases, while a band at 556 nm increases in intensity.[713] The spectra cross at isobestic points at 470 and 488 nm. Two different equilibria exist in dilute solutions at pH 4.70, and two different hematein-aluminium complexes **(258)** and **(259)** are formed. On complex formation, the colour of the solution changes from red-yellow to violet. The maxima of the absorption bands of **(258)** and **(259)** agree with the maximum of the absorption of Hm^- **(250)**.[698] The three species contain the same chromophore, and this is only slightly influenced by the bound aluminium ions. By replacing the proton in one of the hydrogen bonds of **(250)** by Al^{3+}, the 1:1 complex $HmAl^+$ **(258)** is obtained. If the protons of both hydrogen bonds are replaced by Al^{3+}, the 1:2 complex $HmAl_2^{3+}$ **(259)** is obtained. The 1:1 complex **(258)** is formed when hæmatein is in excess, while the 1:2 complex **(259)** is formed when Al^{3+} ions are in excess. The aluminium ions in the Hm–Al complexes are hydrated in aqueous solutions.[714]



(258) **(259)**

The formation of the complexes is shown below:[715]

$Hm \rightarrow Hm^- + H^+$

$Hm^- + Al^{3+} \rightarrow HmAl^+ + H^+$

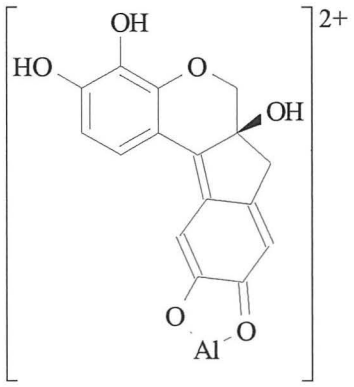
$HmAl^+ + Al^{3+} \rightarrow HmAl_2^{3+} + H^+$

Violet precipitates are formed from hæmalum solutions on standing. These redissolve in acids to give a red colour. The formation of the precipitates can be avoided by acidifying the dye solution after complex formation. In this case, the colour changes from violet to red. The colour change was described by Harms as the decomposition of the 1:2 complex **(259)** into hæmatein Hm, and Al^{3+} ions:[709]

$HmAl_2^{3+} + 3H^+ \rightarrow Hm + 2Al^{3+}$

Hæmatein is immediately transformed into a consecutive product by an acid-catalysed reaction and the dye solution fades. At lower acidity, a further red Hm–Al complex, $HmAl^{2+}$ **(260)** is formed which is important for staining in acidic media. The intense absorption band of the 1:2 complex can be seen at 556 nm.[713] The intensity of this band decreases rapidly with decreasing pH, as the labile 1:2 complex **(259)** is transformed to the stable 1:1 complex **(258)**:[716]

$HmAl_2^{3+} + H^+ \rightarrow HmAl^+ + Al^{3+}$



(260)

At pH ≤ 3.4, a further absorption band, shifted to shorter wavelengths can be seen. At pH 2.6 the maximum of the red 1:1 complex $HmAl^{2+}$, **(260)**, formed by protonation of the 1:1 complex **(258)** in strongly acidic media is at 506 nm ($\varepsilon_{(260)}$ = 13 000):[713] At pH < 2.6 the complex solution rapidly decolourises, producing free hæmatein, which ages and fades in strongly acidic solution.

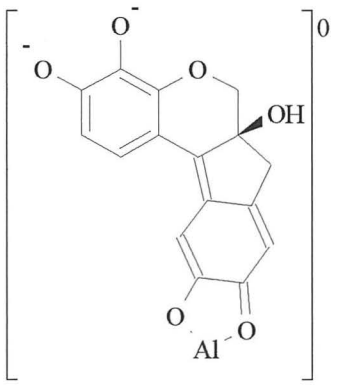
$HmAl^{+} + H^{+} \rightarrow HmAl^{2+}$

$HmAl^{2+} + H^{+} \rightarrow Hm + Al^{3+}$

ii) Studies in alkaline media

Hæmalum staining is generally performed in acidic media. After staining, either water, almost neutral aqueous solutions, or weakly alkaline solutions are employed to produce a deep blue coloration - **blueing**. At pH 6-7, deep-blue precipitates of unknown composition separate from hæmatein-aluminium solutions. Soluble blue dye complexes are observed in weakly alkaline solutions. At pH = 8.7, two pH units above its pKa of 6.7, hæmatein exists as the violet monoanion, Hm^{-}, and aluminium as the $Al(OH)_4^{-}$ anion. When an alkaline aluminium solution at pH 8.7 is treated dropwise with an aqueous solution of hæmatein, the aluminium solution takes on a violet colour (Hm^{-}) in the vicinity of the drop, before turning blue. The intensity of the 556 nm Hm^{-} band decreases with increasing aluminium concentration, with the appearance of a band at 626 nm.[713] The blue $HmAl^{0}$ **(261)** complexes are formed by:

$Hm^{-} + Al(OH)_4^{-} \rightarrow HmAl^{0} + 2H_2O + 2OH^{-}$



(261)

iii) Staining of HeLa cells

In general, hæmalum staining is performed in strongly acidic solutions with high hæmatein and aluminium concentrations. An excess of dye is first bound under kinetic control, and afterwards removed by differentiation with strong acid. Hæmalum staining with very dilute solutions has been reported.[709] HeLa cells[717] can be stained and examined using microspectrophotometry.[718]

After blueing at pH 8.7, an absorption band is observed at 615 nm, whereas the corresponding maximum in the solution spectrum was at 626 nm, assigned to $HmAl^0$ **(261)**.[713] When blueing is carried out at pH 4.7, an absorption band at 595 nm is observed in the microspectrum.[713] The 1:1 and 1:2 complexes **(258)** and **(259)** exist in solution at pH 4.7 with absorption bands at 556 nm. The 1:2 complex $HmAl_2^{3+}$ **(259)** is labile in acidic media, and is formed from $HmAl^+$ **(258)** with an excess of Al^{3+} ions. Blueing removes excess Al^{3+} ions, converting the 1:2 complex into the 1:1 complex **(258)**, which remains bound to the cell. In more strongly acidic solutions the 1:1 complexes $HmAl^{2+}$ **(260)** are stable. A broad band with a maximum at 580 nm is observed at pH 2.7.[688] The short wavelength shoulder can be assigned to the 1:1 complex **(260)** and the long wavelength region to the 1:1 complex **(258)**.

Baker[710] assumed that the dyes are bound to DNA and RNA. Lillie[719] proposed that the dye complexes bind preferentially to proteins. The nuclei of HeLa cells were unstained after DNA digestion,[720-722] although RNA containing nucleoli and cytoplasm were stained. After RNA digestion,[720,721,723] the nuclei were stained, but nucleoli and cytoplasm were not stained. After combined DNA and RNA digestion,[724]

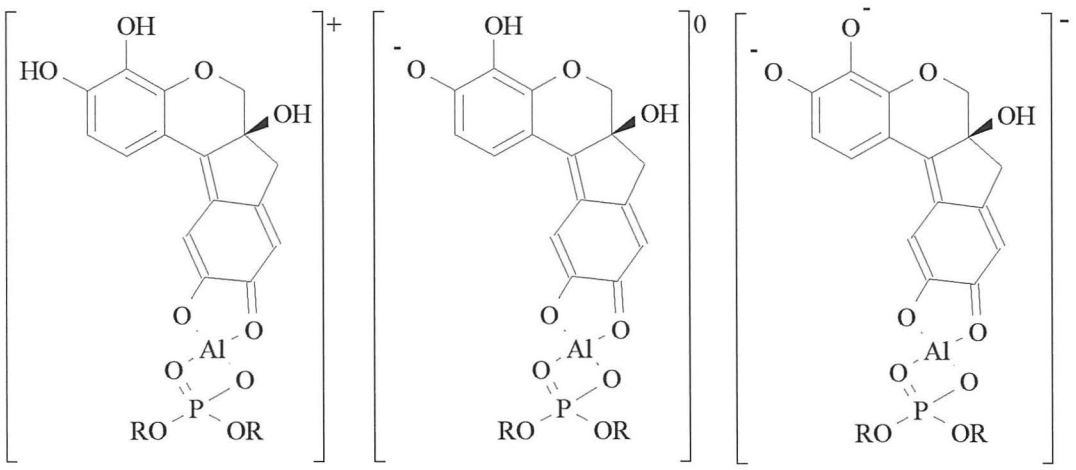
it was impossible to stain nuclei, nucleoli, or cytoplasm. At low dye concentrations the dye complexes are bound to DNA and RNA. Overstaining, *e.g.* protein staining, can occur at higher dye concentrations, but is removed on differentiation. Nucleic acids are stained in acidic media, as the cationic complexes **(258)**, **(259)** and **(260)** bind to DNA and RNA. The uncharged complex **(261)** stains poorly.

The mixture of the dye complexes determines the colour of the solution. The maxima of the absorption bands and the molar extinction coefficients of the most important species of the hæmalum staining are summarised in *Table 9.7*. Acid dye solutions contain the three cationic complexes **(258)**, **(259)** and **(260)**. These are bound to nucleic acids in the staining - nuclei and cytoplasm are stained red to violet in colour. After blueing in alkaline media, the only bound dye species is the neutral dye complex **(261)**. The direct staining by **(261)** in an alkaline medium is negligible.

Table 9.7 The most important dye species of hæmalum staining in aqueous solution and in HeLa cells.[713]

Substance	λ_{max} / nm (H_2O)	ε_{max} (H_2O)	λ_{max} / nm (HeLa cells)
(242) Hm	445	34 000	
(250) Hm^-	556	58 000	
(258) $HmAl^+$	556	22 000	595 **(263)**
(259) $HmAl_2^{3+}$	556	41 000	
(260) $HmAl^{2+}$	506	13 000	530 **(262)**
(261) $HmAl^0$	626	40 000	615 **(264)**

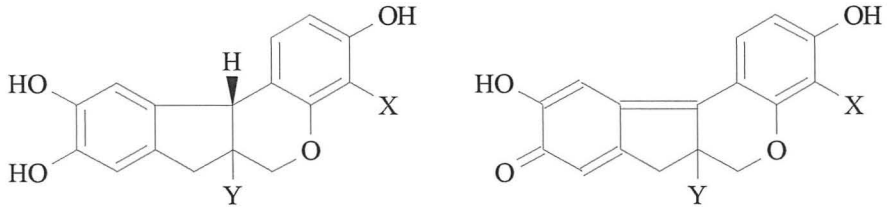
The cationic dye complexes **(258)**, **(259)** and **(260)** are bound to DNA and RNA by covalent bonding and electrostatic interactions. Electrostatic interactions also influence the transport of dye cations to the cells. The bound dye complexes can be mutually interconverted by blueing. The centre for binding and chemical reaction must be different. The dye complexes bind to the phosphodiester residues of DNA and RNA, mediated by the aluminium ions, with the four oxygen atoms forming a distorted tetrahedron around the central aluminium atom. The red 1:1 complex **(260)** dominates in acidic solution forming a cation **(262)** which readily loses a proton, forming **(263).** In alkaline media the blue anion **(264)** is formed.



(262) **(263)** **(264)**

9.4.2 Natural or semi-synthetic analogues

Brazilin **(265)**, a member of the homoflavonoid family, is a natural colouring matter extracted from the heartwood of *Caesalpinia sappan*[725] or Brazil wood, owing its tinctorial properties to the oxidation product, brazilein **(266)**. Brazilin is closely related to hæmatoxylin **(105)** and resembles the latter in its application and staining effects, but its colour is much weaker than that obtained from hæmatoxylin, due to the loss of one of the auxochromic hydroxyl groups. Johansen[726] states that brazilin has come into extensive use as a stain for smears, its employment originated by Belling and refined by Capinpin. Gatenby and Beams[727] mention Hickson's iron brazilin as a nuclear stain. Under appropriate conditions, brazilein acts as an amphoteric dye.

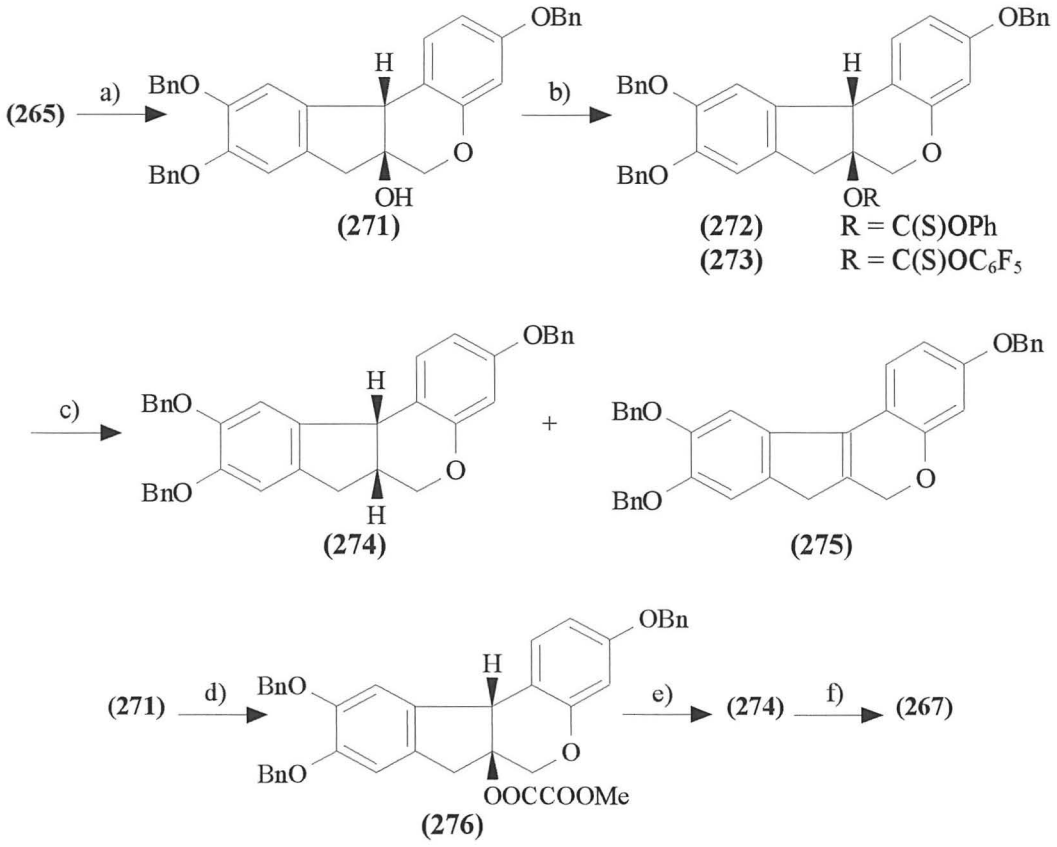


(105)	X = Y = OH	**(242)**	X = Y = OH
(265)	X = H, Y = OH	**(266)**	X = H, Y = OH
(267)	X = H, Y = H	**(268)**	X = H, Y = H
(269)	X = OH, Y = H	**(270)**	X = OH, Y = H

Brazilane **(267)**, a semi-synthetic derivative of brazilin **(265)** has been mentioned as a crude mixture, but its purification and characterisation have never been effected.[728,729] Racemic *O*-tri-methylbrazilane has been obtained either by total synthesis[730,731] or by semisynthesis from brazilin.[728,732] The first synthesis of enantiomerically pure

brazilane by deoxygenation of brazilin has recently been reported (*Scheme 9.1*).[733] Classical deoxygenation of the tertiary alcohol generating an sp^2 carbon could not be used to preserve chirality. Other methods,[734] such as the dissolving metals reduction of ester derivatives,[735] the photolysis of acetates[736] and the stannane reduction of various ester derivatives[737] could be used.

Scheme 9.1 Synthesis of brazilane (267) from brazilin (265).[733]



Reagents and conditions:

(a) BnBr, K_2CO_3, acetone, Δ, 18h, 85 %;
(b) for **(272)**: (i) MeLi, THF, -70 °C to rt; (ii) ClC(S)OPh, -70 °C to rt, 97 %;
for **(273)**: (i) MeLi, THF, -70 °C to rt; (ii) $ClC(S)OC_6F_5$, -70 °C to rt, 64 %;
(c) from **(272)**: Bu_3SnH, AIBN (cat.), C_6H_6, Δ, 0.5 h, 84 %; from **(273)**: idem, 71 %.
(d) (i) MeLi, THF, -70 °C to rt; (ii) ClC(O)COOMe, - 70 °C to rt, 90%;
(e) Bu_3SnH, AIBN (cat.), toluene, 20 mins, 71%;
(f) H_2, 10% Pd/C, AcOEt, rt, 18 h, 100%.

The three phenoxyl groups of brazilin **(265)** were protected as tribenzylbrazilin **(271)** according to Dann *et al.*[738] Robins' modification[739] of Barton and McCombie's deoxygenation reaction[740] was first applied to **(271)**. Deprotonation of **(271)** with MeLi at -70°C, and treatment with phenyl chlorothionoformate gave **(272)**. Reduction of **(272)** was carried out using Bu_3SnH with a catalytic amount of AIBN giving **(274)**

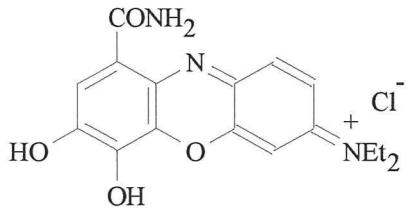
and a very small amount of **(275)**. Absence of AIBN gave **(275)** as the major product. Phenylthionocarbonate **(272)** was thermally unstable and a Chugaev-type elimination occurred. Tribenzylbrazilene **(275)** could also be prepared by refluxing **(272)** in toluene. Pentafluorophenylthionocarbonate[741] **(273)** gave lower yields of **(275)**, while the thiohydroxamate[742] failed.

Reduction of methyl oxalate **(276)** gave **(274)** {verified using NMR techniques, and by comparison with racemic **(±)-(274)**} with no trace of **(275)**.[743] Catalytic hydrogenation of **(274)** afforded brazilane **(267)** in quantitative yield. A configurationally favoured sp^3 radical probably occurs, resulting in the retention of configuration, as inversion would entail an energetically unfavourable ring constrained planar radical intermediate. Hæmatoxylane **(269)** could in principle be synthesised on a commercial scale, together with the oxidation products, brazilene **(268)** and hæmatoxylene **(270)**. However, due to the cost of their manufacture, and less intense colour due to the loss of auxochromic hydroxyl groups, it is unlikely that they will ever be used as substitutes for hæmatoxylin **(105)** and brazilin **(265)**.

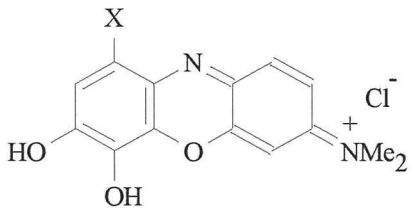
9.4.3 Synthetic analogues

Celestin blue **(277)**, a basic dye of the oxazine group, has been employed as a nuclear stain in place of hæmatoxylin. Proescher and Arkush[744] considered this dye more satisfactory as a nuclear stain than either gallamine blue **(278)** or gallocyanin **(39)**. Lendrum[745] employed celestin blue in conjunction with chromotrope 2R **(279)** as a substitute for hæmatoxylin-eosin for simple diagnosis or for photographic purposes, emphasising the staining of collagen and reticulum. The same paper includes a method using celestin in conjunction with Rubens-Duval orcein and light green for breast carcinoma and skin lesions. Lendrum and McFarlane[746] also employed the dye in conjunction with hæmatoxylin, acid fuchsin, alcoholic orange G, and acetic aniline blue in one of their modifications of Mallory's connective tissue stain, offering a greater degree of specificity for fibrin and erythrocytes in paraffin sections. Gray *et al*[747] recommended a combination of celestin blue and crocein scarlet **(280)** for the simultaneous staining of nuclei and cytoplasm with a single solution, resembling conventional hæmatoxylin-eosin stains. Proescher and Arkush[744] employed the

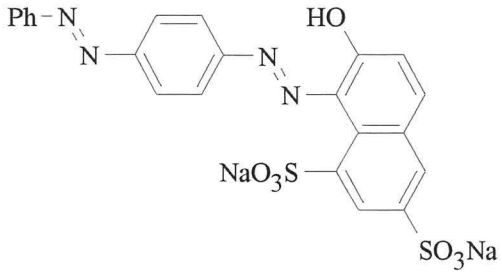
hydrochloride of gallamine blue **(278)**, a violet-blue basic dye of the oxazine group, first introduced by Becher,[748] as a nuclear stain instead of hæmatoxylin.



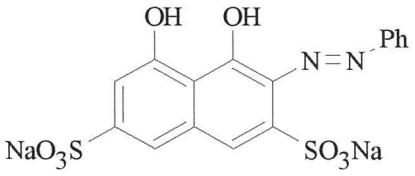
(277)



(278) X = $CONH_2$ **(39)** X = CO_2H

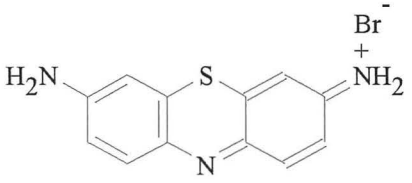


(279)

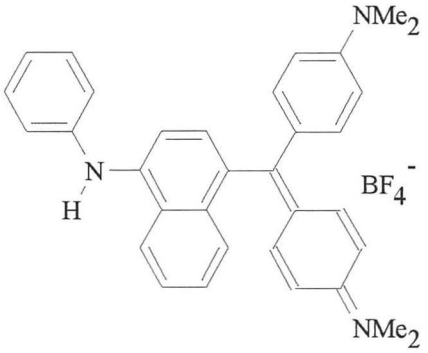


(280)

Wittekind *et al*[749] found that the hæmatoxylin content varied greatly in commercial dye samples.[750] The staining pattern with thionin **(154)** was highly reproducible and differed little from that achieved with the conventional stain.[751] Schulte and Wittekind[752] examined the staining characteristics of five nuclear stains used in a Papanicolaou staining procedure. The pararosaniline Feulgen stain,[753] Victoria blue B **(281)**,[754] a modified gallocyanin chrome alum stain,[755,756] alcoholic thionin,[757] thionin-SO_2,[758] and hæmatoxylin[759] were used. Thionin **(154)** and Victoria blue B **(281)** gave deep-blue nuclei, hæmatoxylin and gallocyanin gave grey-blue nuclei, while thionin-SO_2 did not stain nucleoli. Alcoholic thionin was the most intense nuclear stain, giving a good contrast between nuclei and cytoplasm. Both gallocyanin and hæmatoxylin gave some non-specific cytoplasmic staining. Thionin-SO_2 allowed a quantitative assessment of DNA, but gave a low staining intensity.



(154)



(281)

Nuclear staining with hæmatoxylin is simple and cheap.[760] Hæmatoxylin however stains nuclear and cytoplasmic proteins.[761] Alcoholic thionin and Victoria blue B are pure and stable in solution, do not overstain and do not need differentiation.[754,757] When used in acidified alcoholic solutions, thionin and Victoria blue B bind quantitatively to nucleic acids, but the staining intensity decreases on drying, as in the absence of water, the Coulombic forces between the dye and substrate decrease.[754,758] The modified gallocyanin-chrome alum stain is fast, simple and reproducible,[762] and like hæmatoxylin is insensitive to air drying,[763] but has a short shelf life. Thionin-SO_2,[764,765] a Feulgen reagent where the blue thionin has replaced the red pararosaniline, is stoichiometric for DNA, but is complicated to use.[764,766]

9.4.4 Computational studies

Table 9.8 Calculated electronic spectra for hæmatoxylin based compounds.

Dye	λ_{exp} / nm	Calculated λ / nm			Predicted λ / nm		
		MINDO/3	MNDO	AM1	MINDO/3	MNDO	AM1
(105)	292[698]	136.1	143.2	140.7	288.0§	322.6§	314.2§
(265)	~292†	136.2	142.6	140.6	288.5§	319.8§	313.7§
(267)	~292†	136.2	142.3	140.4	288.5§	318.4§	312.8§
(269)	~292†	136.1	142.3	140.4	288.0§	318.4§	312.8§
(242)	445[698]	195.9	201.2	200.7	473.8¶	460.1¶	470.5¶
(266)	~445‡	194.9	200.2	199.7	470.0¶	456.7¶	466.9¶
(268)	~445‡	194.0	199.8	199.2	466.7¶	455.3¶	465.2¶
(270)	~445‡	194.8	200.8	200.1	469.7¶	458.8¶	468.4¶

† and ‡ Approximate values as the chromophores are similar.
§ Using polymethinic equation. ¶ Using aromatic equation.

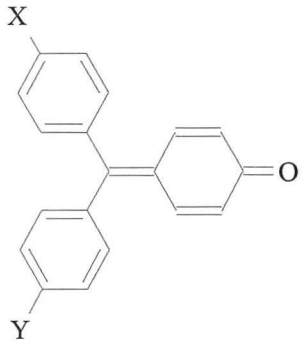
The hæmatoxylin-based compounds were minimised using the MM+ force field and Polak-Ribiere conjugate gradient method to a final gradient of 0.001 kcal mol^{-1} $Å^{-1}$. Electronic spectra were calculated from single point semi-empirical calculations using

the AM1, MNDO and MINDO/3 methods described in *Chapter 6*, the results of which are shown in *Table 9.8*. Wavelength maxima were predicted to within 25 nm for polymethinic hæmatoxylin and aromatic hæmatein based compounds. Spectra of brazilin based compounds are less intense due to the loss of auxochromic OH groups.

Table 9.9 AM1 Calculated electronic spectra for various hæmateins.

Compound	Experimental λ / nm	Calculated λ / nm	Predicted λ / nm
(242)	445[698]	200.7	470.5
(247)	445[698]	203.3	479.6
(248)	445[698]	216.1	524.7
(249)	445[698]	207.5	494.4
(250)	556[698]	220.1	538.8
(251)	556[698]	219.4	536.3
(252)	556[698]	266.1	700.7†
(253)	435[698] (2300 mm^{-1})	205.1	486.0
(254)	567[698] (1765 mm^{-1})	279.1	746.4‡
(282)	381[767] (2625 mm^{-1})	203.8	481.4§
(10)	463[767] (2160 mm^{-1})	206.7	491.6
(283)	536[767] (1865 mm^{-1})	201.0	471.5¶

† Predicted wavelength of 516.8 nm by extrapolating AM1 polyenic equation.
‡ Predicted wavelength of 546.1 nm by extrapolating AM1 polyenic equation.
§ Predicted wavelength of 376.7 nm by extrapolating AM1 polyenic equation.
¶ Predicted wavelength of 588.5 nm by extrapolating AM1 polymethinic equation.



(253) X = H, Y = OH **(254)** X = H, Y = O^- **(282)** X = Y = H

(10) X = Y = OH **(283)** X = Y = O^-

Wavelength maxima were also predicted for whole hæmalums or partial anionic hæmatein fragments using the AM1 aromatic regression equation unless other stated, and are within moderate (~60 nm) agreement with experimental values, as the ionic nature of the Al^{3+} cations or hæmatein molecules affects SCF convergence.

Table 9.10 AM1 Calculated electronic spectra for various hæmalums.

Compound	Observed[713] λ / nm	Calculated λ / nm	Predicted λ / nm
Hm^{2-}	556	246.7	632.4†
(258) $HmAl^{+}$	556	413.9	849.4‡
Hm^{3-}	556	271.6	529.2‡
(259) $HmAl_2^{3+}$	556	195.6	564.0§
Hm-	506	224.5	554.2†
(260) $HmAl^{2+}$	506	266.4	517.5‡
Hm^{3-}	626	260.0	679.2†
(261) $HmAl^{0}$	626	384.6	783.5‡

§ Polymethinic regression. † Aromatic regression. ‡ Polyenic regression.

9.4.5 Nuclear staining models - binding of hæmalum to DNA and RNA

The nuclear staining of Harris hæmatoxylin can be modelled by analysing the binding of hæmalum with DNA and RNA. It is thought that the cationic dye complexes **(258)**, **(259)** and **(260)** are bound to DNA and RNA by covalent bonding and electrostatic interactions. The dye complexes, mutually interconverted by blueing, bind to the phosphodiester residues of DNA and RNA, mediated by the aluminium ions, with the four oxygen atoms forming a distorted tetrahedron around the central Al^{3+} ion. The red complex **(260)** dominates in acidic solution forming a cation **(262)** which readily loses a proton, forming **(263)**. In alkaline media the blue anion **(264)** is formed.

It was envisaged that the binding of hæmalum to a whole portion of DNA or RNA obtained from the *Brookhaven Protein Database* could be modelled. Due to the large size of the DNA or RNA fragment, it was decided to use simpler models, such as **(262)**, **(263)** and **(264)** where R = Me. Models of hæmalum bound to anionic phosphodiester analogues of DNA and RNA were minimised using the MM+ force field to a final gradient of 0.01 kcal mol^{-1} $Å^{-1}$. The predicted maxima of **(262)** and **(263)** obtained using the AM1 aromatic regression equation are accurate to 30 nm. A wavelength maximum ~200 nm too high was predicted for **(264)** using the AM1 polyenic regression equation, due to the very small HOMO-LUMO gap.

Table 9.11 AM1 Calculated electronic spectra for various bound hæmalums.

Compound	Experimental λ / nm	Calculated λ / nm	Predicted λ / nm
(262) R = Me	530[713]	209.6	501.8†
(263) R = Me	595[713]	242.7	618.3†
(264) R = Me	615[713]	390.3	796.2‡

† AM1 aromatic regression equation. ‡ AM1 polyenic regression equation.

The binding of hæmalum to individual nucleotides has also been investigated. Structures of these nucleotides were obtained from the *Nucleic Acids Database* within *HyperChem*, and merged with hæmalum. The resulting hæmalum-nucleotide complexes were then minimised using the MM+ force field and Polak-Ribiere conjugate gradient method, to a final gradient of 0.01 kcal mol^{-1} $Å^{-1}$. Predicted wavelengths for the hæmalum-nucleotide complexes are shown in *Table 9.12*. Only hæmalum-nucleotide complexes resembling the cationic species **(263)** have been investigated due to SCF convergence problems encountered with other species. Only two complexes, **(284)** and **(287)** reached the desired gradient, due to the large size of the hæmalum-nucleotide complexes. Predicted maxima of these two complexes are within 30 nm of the experimental value of 530 nm.[713] The hæmalum-deoxyadenosine monophosphate complex **(284)** obeys the polyenic regression, while the hæmalum-deoxycytidine monophosphate complex **(287)** obeys the aromatic regression.

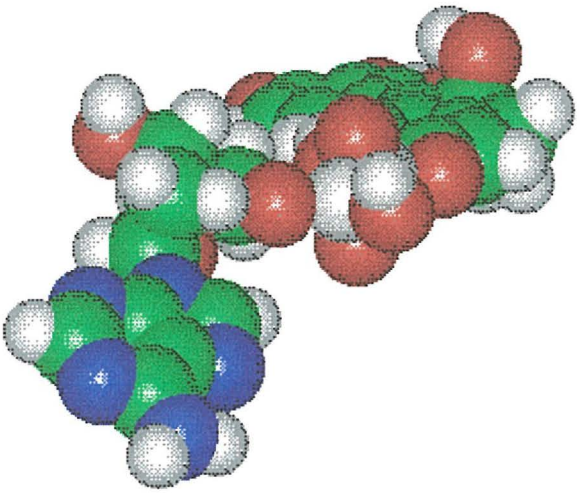
Table 9.12 AM1 Wavelength maxima for nucleotide-bound hæmalums.

Compound	Calculated λ / nm	Predicted λ / nm
(284) dA	268.2	521.6†
(285) dT	‡	
(286) dG	‡	
(287) dC	227.2	563.7§
(288) rA	‡	
(289) rU	‡	
(290) rG	‡	
(291) rC	‡	

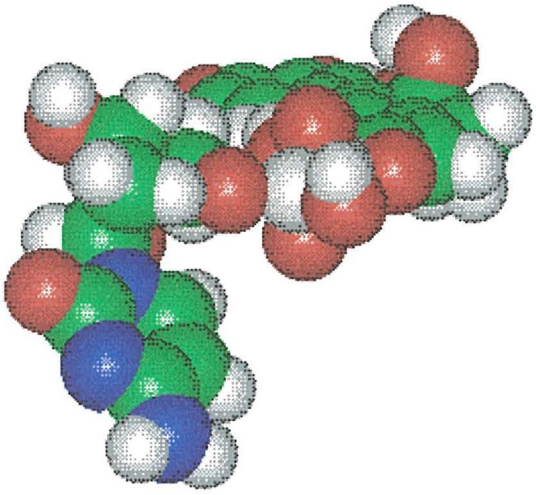
† Polyenic regression. § Aromatic regression. ‡ Not converged.

dA	Hæmalum-deoxyadenosine monophosphate	dT	Hæmalum-thymidine monophosphate
dG	Hæmalum-deoxyguanosine monophosphate	dC	Hæmalum-deoxycytidine monophosphate
rA	Hæmalum-adenosine monophosphate	rU	Hæmalum-uridine monophosphate
rG	Hæmalum-guanosine monophosphate	rC	Hæmalum-cytidine monophosphate

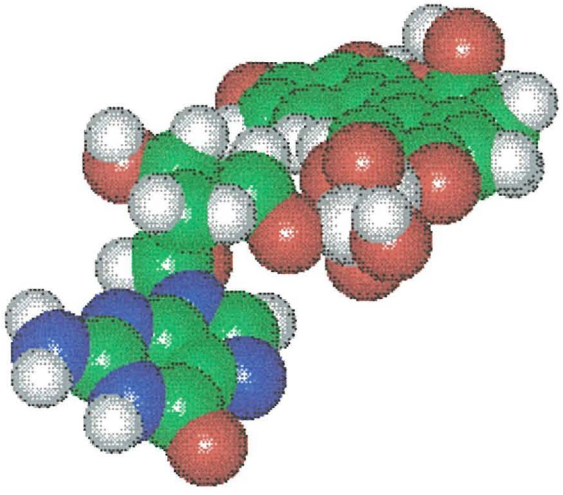
dA



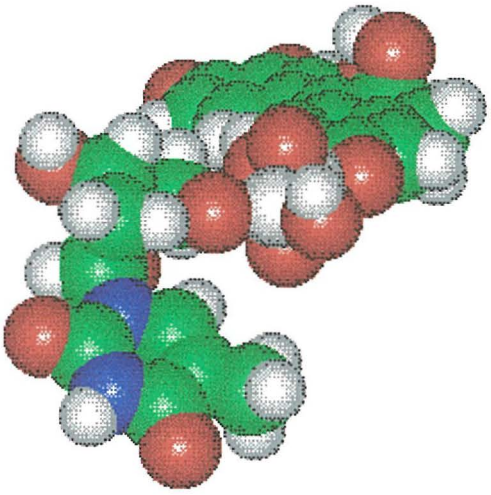
dC



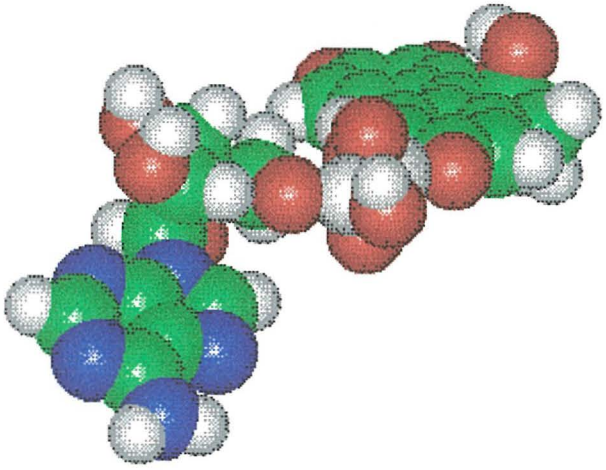
dG



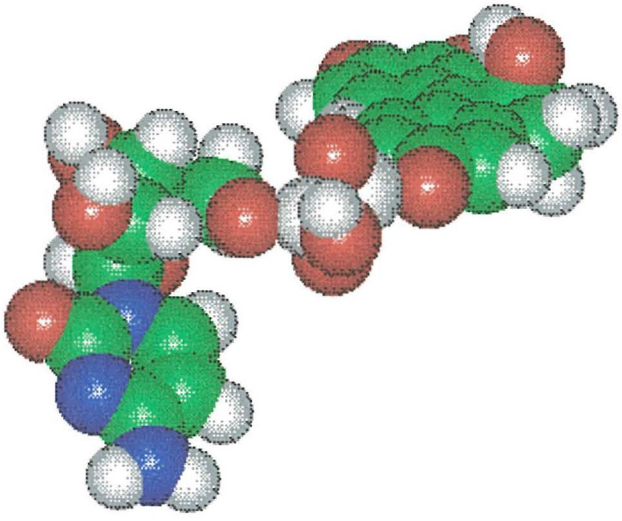
dT



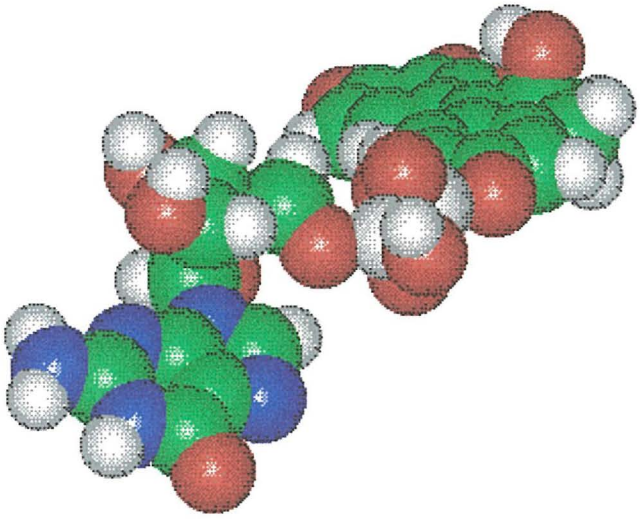
rA



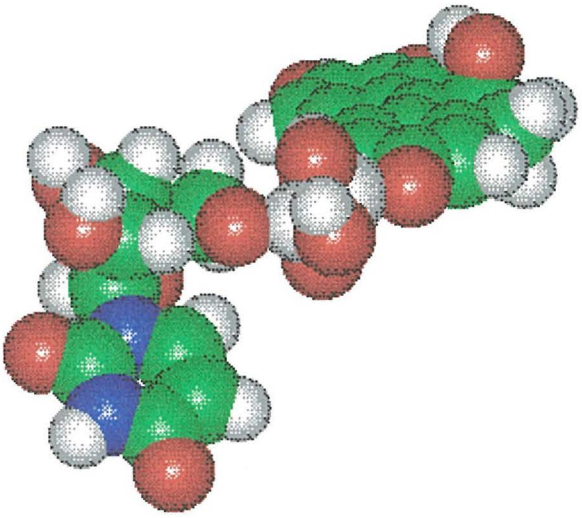
rC



rG

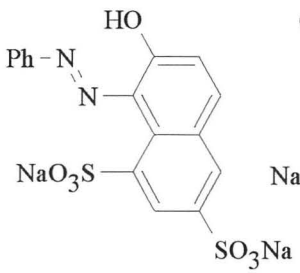


rU

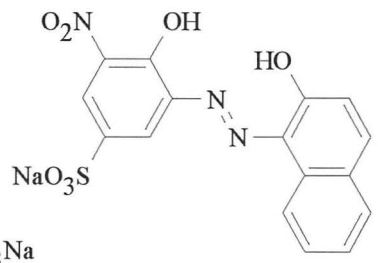


9.5 SUBSTITUTES FOR ORANGE G

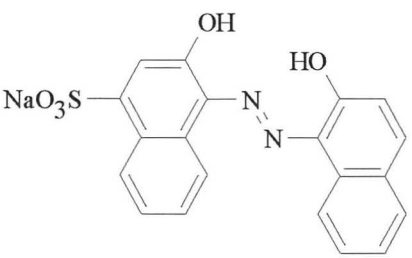
Orange G **(25)** was used to stain vaginal smears by Fuller.[768] Acid alizarin black **(292)**, alizarin blue black **(293)**, azo yellow **(294)**, chlorazol brown G **(295)**, and solochrome red B **(296)** can be used as substitutes for orange G.[683] Chlorazol brown G **(295)** can be used where a brown coloration is desirable, and will not stain keratin due to its high molecular weight.



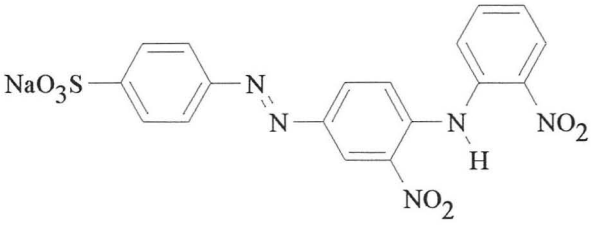
(25)



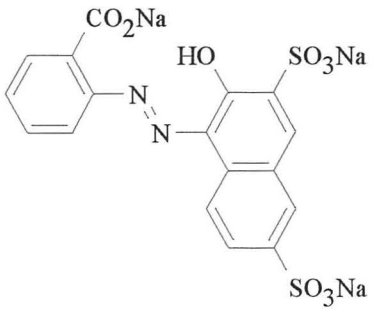
(292)



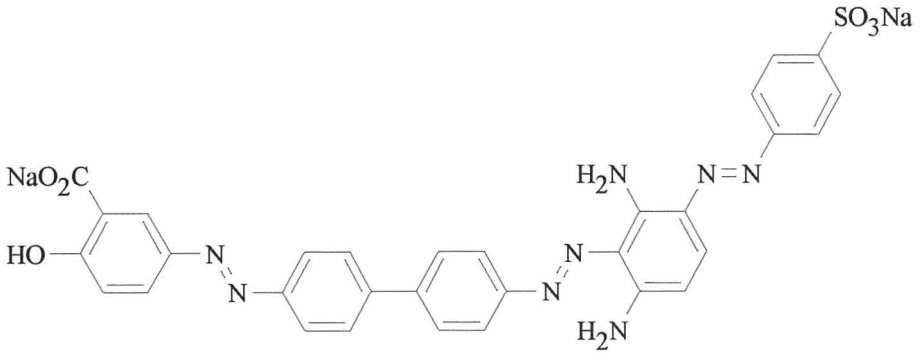
(293)



(294)



(296)



(295)

9.5.1 Possible staining mechanism

Lysine, histidine and arginine complexes with sulphonated azo dyes have been studied[536] as model systems for understanding the biomolecular recognition of glycosaminoglycans[769] by proteins. The diverse biological activities of sulphonated compounds make them potential pharmaceutical agents and important lead compounds for the development of new drugs.[770] Evans blue **(87)** and Congo red **(85)** have been shown to bind to HIV protease and reverse transcriptase, inhibiting viral replication.[771] Congo red is used as a histological stain for amyloid deposits[772] implicated in Alzheimer's disease, as well as in neurodegenerative diseases in sheep and cattle. Congo red inhibits the build-up of protease-resistant plaque in infected cells and inhibits replication of the scrapie infective agent.[773] Suramin, a sulphonated urea derivative, evaluated for anti-HIV activity,[771b,774] is a growth factor receptor poison[775] showing potential as an anti-cancer agent due to its ability to inhibit angiogenesis.[776] Sulphonated polymeric compounds mimic heparin and exhibit antithrombogenic activity.[777] Sulphated carbohydrates have been shown to inhibit HIV infection and replication.[778]

The sulphonate groups of Congo red can be placed within hydrogen bonding distance of two arginine side chains located at both ends of the HIV-1 protease binding cavity.[779] The four sulphonate groups of Evans Blue bind to both arginine and lysine side chains of the protease, leading to higher affinity.[771] A pattern of preferred hydrogen bonding geometries assumed by the sulphonyl anion as it interacts with a variety of hydrogen bonding donors has been published.[780] Binding of sulphonate or sulphate groups to proteins is most likely to occur through the basic side chains of lysine, histidine and arginine as shown in the binding[781] of the antiulcer drug sucrose octasulphate (SOS)[782] to lysine and arginine side chains in acidic fibroblast growth factor (aFGF). The preferred conformations assumed by lysine and arginine side chains in contact with anions has also been reported.[783]

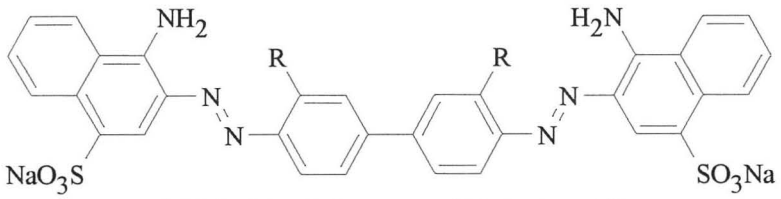
Crystal structures of DL-lysine/Orange G dihydrate **(155)** and DL-histidine/Orange G trihydrate **(156)** have been published.[536] The S(1)-C(1) bond in Orange G is significantly longer than the S(2)-C(3) bond, *e.g.* 1.795(4) Å *vs* 1.759 (4) Å in the calcium salt as a result of steric interaction with the bulky phenylazo substituent on C(8).[784b] An intramolecular hydrogen bond is found in the Orange G molecules, the donor not being the naphthol oxygen, but a hydrazo nitrogen atom, as Orange G exists as an equilibrium mixture of azo and hydrazo compounds in the solid state and solution.[784,785] Sulphonate groups on the dye molecules assume either staggered or eclipsed orientations, with no clear preference for the staggered conformation.[786] Hydrogen-bonded dimers common for protonated carboxylic acids[787] are absent from the crystal structures. Close contacts to the carboxyl groups are made by sulphonate oxygen atoms or by water molecules. Bridging interactions are a favoured mode of contact between sulphonates / sulphates and guanidinium groups, being found in guanidinium alkanesulphonates and arenesulphonates.[788] Sulphonate bridging interactions occur between the sulphonate groups of Orange G and the protonated forms of adenine and cytosine.[784a]

9.6 SUBSTITUTES FOR EOSIN Y

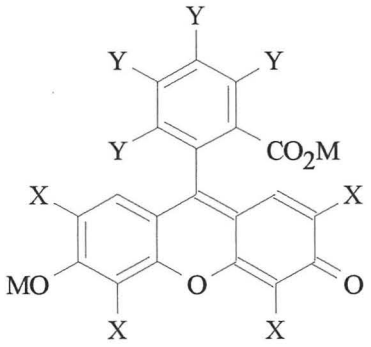
Eosin Y **(29)** is a rose-pink acid dye of the xanthene series.[789] It is the most valuable plasma stain known, and is used as a counterstain after hæmatoxylin.[681-683] Biebrich scarlet water soluble **(70)** is one of the oldest and most frequently used counterstains,[681] after hæmatoxylin, taking second place to eosin. Lillie[790] employed

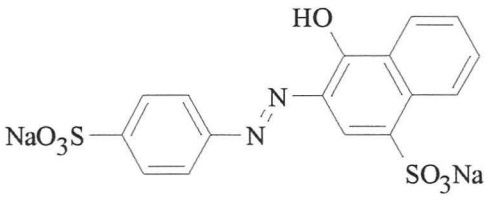
Biebrich scarlet in conjunction with picro aniline blue as a differential stain for connective tissue.[791] Scott[792] recommended the use of Biebrich scarlet as a counterstain after hæmatoxylin staining. Biebrich scarlet and ponceau de xylidene **(71)** are components of Shorr's cervical stain.[273-275]

Benzopurpurin 4B **(297)**, a brownish-red acid dye can be used instead of eosin Y as a plasma stain.[683] Carmoisine L **(298)**, a bluish-red acid dye, "might prove of interest in place of Biebrich scarlet or eosin, after hæmatoxylin and other nuclear stains".[683] Cyanosine B **(299)**, a red acid dye, similar in structure to rose Bengal N, is rarely used.[683] Geschickter[793] employed Erie garnet **(300)** with Azure A for diagnosis of early stages of cancer, affording a polychromatic stain giving colour reactions closely simulating hæmatoxylin and eosin, without overstaining. Erythrosin B **(64)**, a cherry-red dye, is not so easily removed from sections by alcohol, as eosin is, during dehydration.[683] Mercurochrome **(301)**, an acid red dye, used as an antiseptic,[794] is more expensive than eosin. Phloxine B **(65)**, a bluish-red dye, is "the best eosin for use in the eosin-methylene blue stain for paraffin sections of tissues fixed in Zenker fluid".[683] Rose Bengal 2B **(63)** can also be used instead of erythrosin and eosin.[683]



(297) R = Me **(300)** R = H





(29) M = Na, X = Br, Y = H
(63) M = K, X = I, Y = Cl
(64) M = Na, X = I, Y = H
(65) M = Na, X = Br, Y = Cl
(299) M = Na, X = Br, Y = Cl
(301) M = HgOH, X = Br, Y = H

(298)

9.6.1 Possible staining mechanism

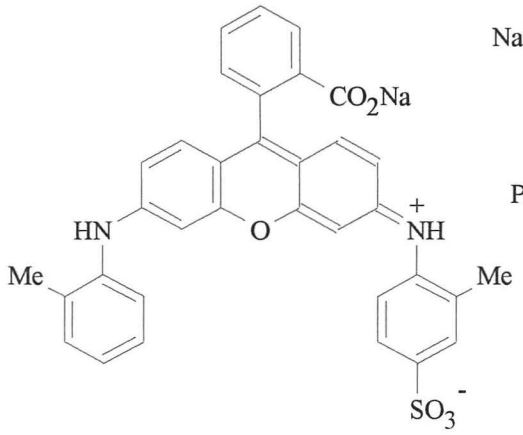
The relationship between ionised molecular weight and staining of keratin or collagen has been discussed in *Section 2.11*. Below an ionised MW of 350 Daltons, no correlation can be found between the acidity or basicity of a dye, and the diffusiveness of staining. Between 350 and 590 Daltons, keratin is stained preferentially, while above 700 Daltons, collagen is the only one of the two types of fibre stained. The intermediate group stains both with equal intensity.

The best stains for cytoplasm are therefore anionic dyes with a molecular weight of between 350 and 590. Orange G has an ionised MW of 406.4 g mol^{-1}, eosin Y has an ionised MW of 645.9 g mol^{-1}, while Biebrich Scarlet **(70)**, ponceau de xylidene **(71)** and acid fuchsin **(12)** used in Shorr's methods have ionised MWs of 510.5, 510.5 and 515.5 g mol^{-1} respectively. Eosin Y with its higher MW can bind to both cornified and non-cornified cells as seen in *Table 9.4*.

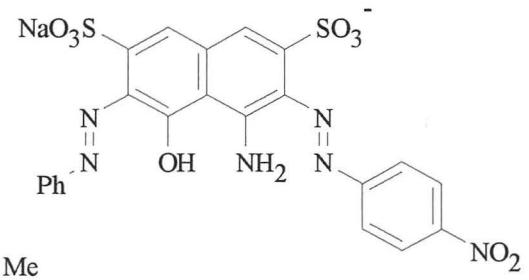
9.7 SUBSTITUTES FOR LIGHT GREEN SF YELLOWISH

Light green SF yellowish **(13)** is an acid dye of the triphenylmethane class and fades fairly rapidly when exposed to light.[795] The dye powder is hygroscopic, and liable to congeal into a solid mass when kept in store for any length of time, unless it is kept in hermetically closed containers. Light green SF yellowish is a component of Twort's stain[146] devised for the demonstration of bacteria in tissues.

Fast green FCF **(81)**, an acid dye of the diphenylnaphthylmethane group may be used for all purposes for which light green is employed. It is considerably faster to light than the latter, and is not hygroscopic. Ollett[796] employed fast green FCF in place of light green SF in combination with neutral red in a modified Twort's stain, which is used as a counterstain to Gram's stain for bacteria in sections. Fast green FCF, together with violamine R **(302)**, naphthol blue black **(303)**, acid fuchsin **(12)**, aniline blue aqueous **(18)** and wool green S **(80)** were among the stains studied by Lillie[797] for the selective staining of collagen with acid dyes. Fast green was among those found to excel as plasma stains in the Masson-Van Gieson and the Masson-Mallory-Van Gieson procedures.

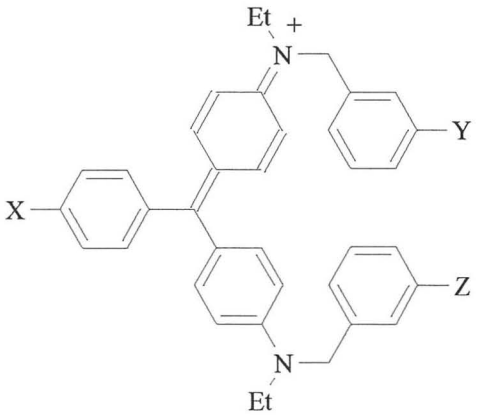


(302)



(303)

Gray[681] describes a large number of techniques such as staining of vaginal smears in which aniline blue **(18)** is employed, together with a large number of recipes such as: Hæmatoxylin-ponceau S-aniline blue for differential staining of connective tissue and muscle; Masson's trichrome stain- hæmatoxylin, ponceau fuchsin, aniline blue for connective tissue and Mallory's aniline blue-orange G and various modifications of the original. Aniline blue aqueous was among the dyes Lillie[797] found gave the best results for collagen fibres, in the Van Gieson, Masson-Van Gieson, and Masson-Mallory-Van Gieson procedures. Guinea green B **(165)** acts more rapidly than either light green SF yellowish **(13)** or fast green FCF **(81)**, and gives a more pleasing coloration. Wool green S **(80)**, an acid dye of the diphenylnaphthylmethane series, was employed by Lillie[797] as a component, instead of fast green FCF, in a modified Masson trichrome stain.[791]



(13) $X = SO_3^-$, $Y = Z = SO_3Na$

(165) $X = H$, $Y = SO_3^-$, $Y = SO_3Na$

9.7.1 Possible staining mechanism

The relationship between ionised MW and the staining of keratin or collagen was discussed in *Section 2.11* and *Section 9.6.1*. The best collagen stains collagen are therefore anionic dyes with a MW greater than 700 Daltons, with light green SF yellowish and fast green FCF having respective ionised MWs of 747 and 763 Daltons.

9.8 EFFECT OF ACIDIFICATION

The effect of acidification and mordanting on staining has been discussed in *Section 2.12* and *Section 2.13* respectively. Picric acid is a yellow dye containing a nitro group, used in conjunction with the Van Gieson stain[798] for connective tissue, as a component of a number of fixatives, and as a mordant. Lillie[797] found that picric acid surpasses phosphomolybdic and phosphotungstic acids in connective tissue staining, and as a mordant between plasma and collagen stains.

The selectivity of violamine R **(302)**, naphthol blue black **(303)**, fast green FCF **(81)**, acid fuchsin **(12)**, wool green S **(80)**, water soluble aniline blue **(18)**, and possibly light green SF yellowish **(13)** is best below pH 2.0. Picric acid, phosphotungstic and phosphomolybdic acids all appear to depend on this principle rather than on any mordanting effect. Acidification of the OG-6 staining solution is also necessary as little staining occurs without addition of phosphotungstic acid.

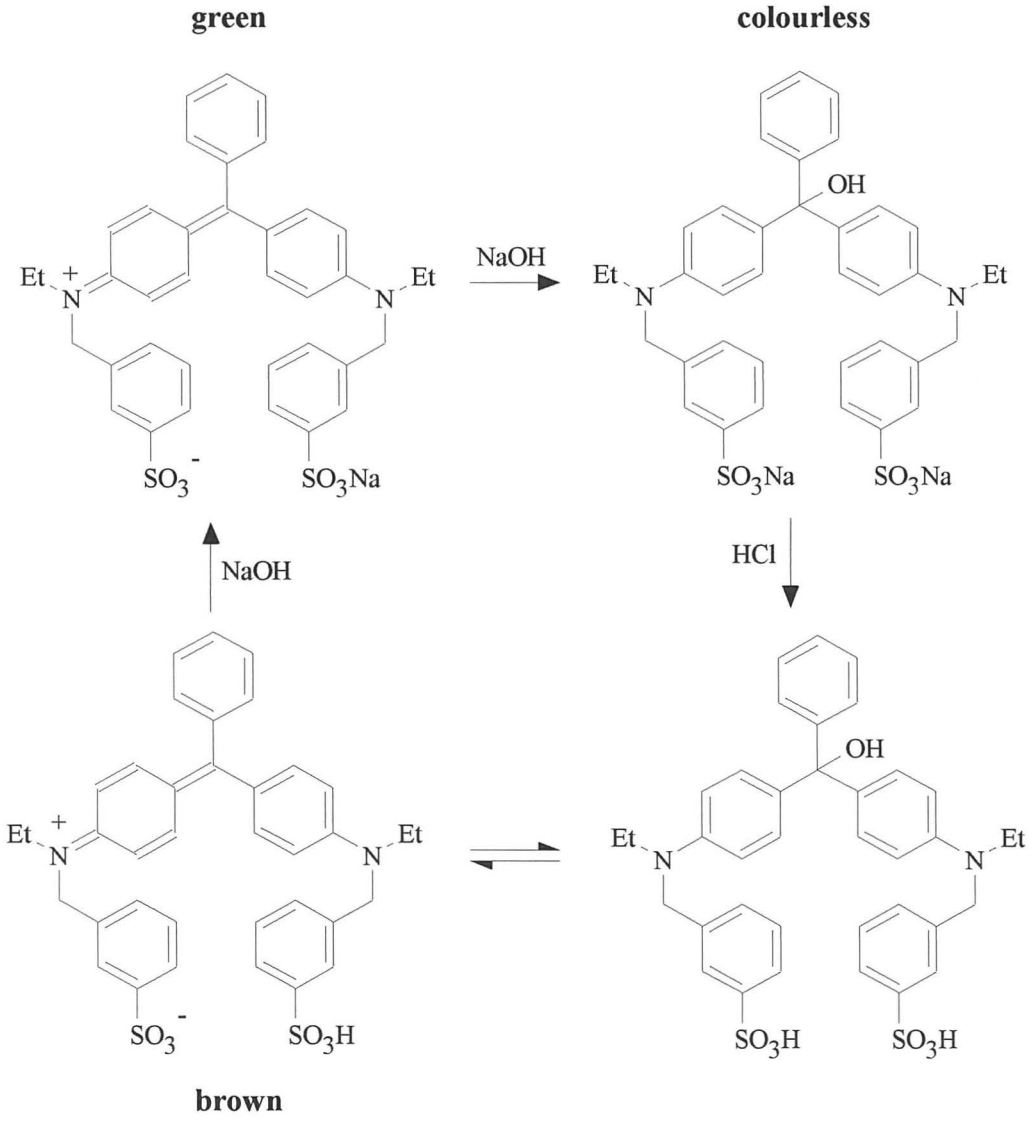
Several factors may be responsible for increasing staining efficiency on acidification:

- There is evidence that acetic acid added to solutions of anionic dyes has a direct effect on the dyes themselves apart from its effect on tissue elements. Acidification of the sodium salts of the dyes may produce free dye acids, which are more reactive to basic tissue elements than are the sodium salts.

- Alternatively, protonation of amino acids, such as lysine, histidine, arginine or serine in proteins found in cells may occur. The positively charged protonated amino acids would engage in electrostatic binding with dye sulphonate colligators.

- Finally, instead of protonation mechanisms, Mo^{3+} and W^{3+} ions from phosphomolybdic and phosphotungstic acids may act as mordants, in the same way as Al^{3+} and other metal ions acts as mordants in hæmalum staining. Mo^{3+} and W^{3+} mediated interactions may occur between proteins and the sulphonate colligators of eosin Y and light green SF yellowish, in the same manner as water mediated interactions occur between the sulphonate groups of Orange G and DL-lysine, and DL-histidine.[536]

The effect of pH on guinea green B **(165)**, a dye closely related to light green SF yellowish **(13)** has been examined, as shown below.[799] Light green SF yellowish should behave similarly, although it has an extra sulphonate group.



9.9 HYPOTHESES CONCERNING THE STAINING MECHANISM

Although the Papanicolaou stain has been in use for over half a century, the staining properties of some of the alchemic ingredients need to be rationalised, due to the empirical methodology used by Papanicolaou when devising the stain. It is unlikely that cheaper substitutes will be found for most of the dyes, such as hæmatoxylin, orange G and eosin Y. As cervical cells are frequently cluttered, causing cellular overlap and lack of differentiation, the engineering of dyes with low cellular binding strengths would, however, be advantageous.

The staining mechanisms for the component dyes of the Papanicolaou stain are unknown, except in the case of hæmalum. The complex chemistry of haematoxylin, haematein and haemalum based compounds has been discussed. Electronic absorption maxima have been predicted to within 25 nm for hæmatoxylin and brazilin based compounds using the AM1 methodologies developed in *Chapter 6*. Wavelength maxima of hæmatein and hæmalum based compounds have been predicted with accuracies of about 50-70 nm, about 20-30 nm less accurate than those of the other Papanicolaou dyes, possibly due to their strained molecular geometries. Hæmalum stains DNA and RNA purple-blue, probably interacting with nucleic acid phosphate groups *via* an Al^{3+} cation of the mordant. Wavelength maxima of bound hæmalums have also been predicted, with accuracies of 30-150 nm, again possibly due strained geometries and ionic character.

Collagens are stained by anionic dyes with a MW above 700 Daltons, such as light green SF yellowish; while keratins are stained by anionic dyes with a MW of 350-590 Daltons. The staining mechanism probably occurs *via* lysine, histidine or arginine interactions. Orange G is also known to stain erythrocytes, although eosin Y does not. This aspect of the staining mechanism cannot be explained at this time. The role of acidification is unclear, but the intensity of staining is increased on acidification. This may be due to protonation of the sodium salts forming the free dye, protonation of the substrate leading to electrostatic interactions, or due to a mordanting effect of the Mo^{3+} or W^{3+} ions. The role of bismarck brown Y in the Papanicolaou stain is unclear, although staining performance is not impeded by its omission.

9.10 THE FUTURE OF CERVICAL SMEAR TESTING

The most advanced breakthrough in cervical cancer research, "THE INFALLIBLE SMEAR TEST", was reported on the front page of the *Daily Mail* on the 8th December 1998. The revolutionary technique,[800] the Campaign test was developed at Cambridge University and Addenbrooke's Hospital, in conjunction with the Cancer Research Campaign. The problem of interpretation leading to potentially fatal false-negative diagnoses (30%) was addressed by using affinity-purified antibodies against human proteins that regulate DNA replication, namely Cdc6 and Mcm5. These antibodies were applied to sections and smears of normal and diseased uterine cervix by using immunoperoxidase or immunofluorescence to detect abnormal precursor malignant cells. Antibodies against Cdc6 and Mcm5 stain abnormal cells with high specificity and sensitivity. Proliferation markers Ki-67 and proliferating cell nuclear antigen are much less effective. Immunostaining of cervical smears can be combined with the conventional Pap stain so that all the morphological information from the conventional method is conserved. Large-scale tests are expected to start in the US in 18 months, with possible commercial availability in three years. The test proved 100% accurate in trials, and is set to end the scandals that have plagued the screening programme. Full Government backing was promised for the new test if trials confirmed its effectiveness. An automatic system could be developed, giving the all-clear to most smears, with the rest double-checked by trained staff, increasing accuracy and cost-effectiveness.

9.11 FURTHER ENVISAGED STUDIES

Semi-empirical calculations were performed on gas phase molecular minimised molecules in *Chapter 6*. Preliminary investigations of solvent effects employing simple supermolecular solvated assemblies were unsuccessful, and are not noted in this thesis. A study of solvation effects using either the SCRF or COSMO models is envisaged as the dyes are used in alcoholic solution. In order to obtain more reliable results, geometries obtained by semi-empirical minimisation will have to be used, as molecular mechanics is unreliable for reproducing structures of π-bonded systems.

The $\Delta\Delta H_f$ approach to calculating singlet excitation energies was briefly investigated in *Chapter 7*. Some doubt remains about the heats of formation of the excited states. It would also be interesting to note the effect of CID or CISD approaches on the predicted spectra. A more intensive CI study employing other Hamiltonians would also be instructive.

The results of DFT calculations employing local functionals have been presented in *Chapter 8*. A further study employing non-local functionals (with concomitant optimisation) should also be attempted, together with a study of the newer hybrid DFT/HF methods such as B3LYP/6-31G(*d*). These may allow the incorporation of electron correlation and repulsion effects.

Homology and comparative modelling of proteins, together with pharmacophore definition, *de novo* ligand design, molecular docking, and other modelling techniques employed by the pharmaceutical industry can also be used to produce more advanced models of stain-cell interactions, *e.g.* between dyes and keratins[801] or collagens.[802]

CHAPTER 10

SEMI-EMPIRICAL PERICYCLIC TRANSITION STRUCTURES

10.1 AIMS

Initial studies of transition structures of various pericyclic cyclopropanic and cyclopropenic reactions using semi-empirical methodologies are presented in this final chapter. Semi-empirical methodologies were chosen as they require less memory and computational time, and can be used for larger systems. Further studies are envisaged using the more expensive HF, DFT or hybrid HF-DFT *ab initio* methods.

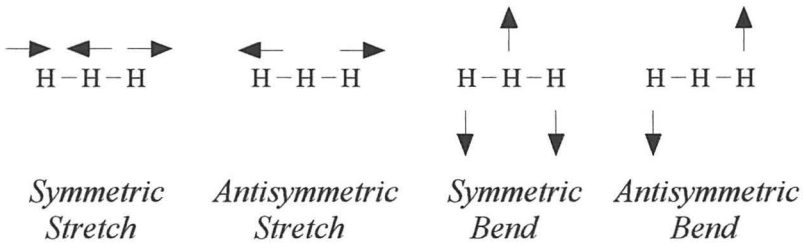
The effectiveness and reliability of various *ab initio* and semi-empirical computational methodologies for locating transition structures and calculating activation energies will be examined. It will be shown that PM3 most accurately reproduces the physical properties of cyclopropanes and cyclopropenes. The electrocyclic ring opening of cyclopropanes and cyclopropenes will be examined, as 3,3-disubstituted-1,2-dichlorocyclopropenes undergo electrocyclic ring opening to produce vinyl carbenes which can be trapped using alkenes. The diimide reduction of cyclopropenes, will be used as an initial model for the biologically important desaturase enzymes, which produce fatty acids. The dimerisation of cyclopropenes *via* an ene reaction, which could be used to synthesise biologically important contiguous oligocyclopropanes, will also be examined.

10.2 TRANSITION STRUCTURES AND TRANSITION STATES

A transition structure corresponds to a saddle point separating two energy minima on a theoretical potential energy surface,[803-806] and is closely related to the transition state, a quasi-thermodynamical state defined in Eyring's transition state theory,[807] situated at a free energy maximum along the reaction path connecting two minima. The free energy, ΔG, differs from the enthalpy, ΔH, by $-T\Delta S$. ΔH is equal to the difference of potential energy E, plus RT, zero-point energy, and thermal corrections. The rate of an elementary reaction is determined by the difference between the free energies of the transition state and the reactants, as well as the rate of passage through the region of the transition state.[807-810] Transition state geometries will be the same as transition structure geometries, if a reaction has a large barrier, or a slowly varying

entropy in the region of the potential energy maximum. The regions of interest on the potential energy surface are stationary points, at which all force constants, second derivatives of the energy with respect to the $3N$–6 atomic co-ordinates, are positive. Transition structures have one, and only one, negative second derivative; the remaining $3N$–7 second derivatives are positive. The negative eigenvalue corresponds to a force constant for the motion along the reaction co-ordinate. This is also referred to as an imaginary vibrational frequency, since the vibrational frequency is proportional to the square root of the force constant.[805,806] For the $H + H_2 \rightarrow H_2 + H$ reaction, the transition state $[H–H–H]^{‡}$ is linear, and has four normal modes of vibration, shown below. The imaginary vibration, the symmetric stretch corresponds to the transition from reactants to products. The energy of the system decreases along the symmetric stretching vibration, giving rise to negative force constants.

Figure 10.1 The normal modes of vibration for the $[H–H–H]^{‡}$ transition state.



10.2.1 Methods for locating saddle points

For simple cases, such as the reaction of chloride with methyl chloride, transition structure geometries can be predicted by inspection. In other cases, a **grid search** can be used to scan the PES to locate the approximate position of the transition state, whereby an analytical expression is fitted to the energy of a set of structures obtained by systematic co-ordinate variation. Grid searching is good for constructing energy surfaces, and investigating reaction dynamics, but is restricted to small systems with a limited number of degrees of freedom.

If the conversion of one minimum energy structure into another occurs primarily along one or two co-ordinates, an approximation to the reaction pathway can be obtained by gradually changing the co-ordinate(s), and allowing the system to relax at each stage by minimisation, constraining the chosen co-ordinate. The point of highest energy on the path is an approximation to the saddle point. **Adiabatic mapping** has

been used to study the energetics of rotation of the aromatic rings of tyrosine and phenylalanine side chains in the interior of proteins.[811] An alternative to moving along a predefined set of co-ordinates is to gradually move up the valley to the saddle point along the **least steep path, or path of shallowest ascent**, as Newton-Raphson and other minimisation methods converge to a transition structure if the starting point is within the quadratic region.[635,812] The **self penalty walk** (SPW)[813,814] and **conjugate peak refinement** algorithms[815,92] have been developed to tackle complex molecules.

10.2.2 The location of transition structures using *ab initio* techniques

It is more difficult to locate transition structures than it is to locate minima.[816,817] A transition structure can be proved by a frequency calculation, but only a complete search of the PES can prove it to be the lowest energy transition structure. Multiconfiguration self-consistent field (MCSCF) calculations are used for systems poorly represented by a single configuration of occupied orbitals,[818-820] and correct for nondynamical correlation energy. The most common MCSCF method is the complete active space (CASSCF) method, which mixes all configurations involving the active orbitals. Active orbitals are those which undergo significant bonding changes during the reaction.[821] For a breaking σ bond, the active orbitals are the σ and σ* orbitals. Pulay and Hamilton described a computational method to choose the active space.[822] For calculations of geometries of stable molecules, correlation energy is generally of little consequence. Nondynamical correlation has a dramatic effect upon the geometries of transition structures of forbidden reactions. Correlation energy corrections are required for the calculation of reasonable activation energies, since correlation energy is larger for transition structures than it is for stable molecules.

10.2.3 Quantities obtained from calculations

Corrections must be applied to obtain energies directly comparable with experimental quantities, the largest of these being the zero point energy $\frac{1}{2}\sum h\nu_i$. At higher temperatures, molecules have an additional vibrational energy $C_p\Delta T$. Addition of $P\Delta V = \Delta nRT$ to the internal energy gives the enthalpy of a molecule. Finally, the entropy can be calculated from the masses, moments of inertia, and vibrational frequencies.[823] Since moments of inertia and vibrational frequencies can be calculated, the former with great accuracy and the latter about 10% too large,[824,825]

enthalpies, entropies and free energies can be readily calculated. Kinetic isotope effects can be calculated,[826-828] and compared with observed values.[829-831] This method is one of the few means of studying the details of potential surfaces experimentally.[832]

Transition structure geometries and electron distributions are essential for understanding reactions in solution, or catalysed by enzymes. The shape and energies of frontier orbitals can be used to predict substituent effects on rates. Bond orders in the transition structure can also be calculated,[833-836] facilitating the understanding of the nature of partial bonds in the transition structure.[837,838] Contributions of intrinsic reaction barriers[839] can be evaluated to understand the origin of substituent effects.[840] MP2/6-31G* optimisations are necessary for heteroatoms to achieve reliable results.

10.2.4 Reliability of calculations

There is no general method to determine experimental geometries of transition states.[841] Calculated bond dissociation energies within a few kcal mol^{-1} of experimental results have been reported.[842] The [1_s,5_s] hydrogen shift in (*Z*)-1,3-pentadiene, the electrocyclic ring opening of cyclobutene, and the Diels-Alder reaction will illustrate the reliability of the various computational methods.

The transition structure corresponding to the [1_s,5_s] hydrogen shift in (*Z*)-1,3-pentadiene is easy to reproduce, with no alternative mechanisms of similar energy. C–H dissociation followed by recombination would require at least 58 kcal mol^{-1}; a 90° rotation, followed by hydrogen transfer would require about 44 kcal mol^{-1},[843] while the experimental activation energy is only 36 kcal mol^{-1}.[844] Transition structures for the degenerate [1_s,5_s] hydrogen shift in (*Z*)-1,3-pentadiene have been calculated at various *ab initio* levels,[845-847] and with several methods of Dewar *et al.* 3-21G Transition structure geometries give bond lengths within several hundredths of an Ångström of MP2/6-31G* values for reactions with no comparable alternatives.[848] As the calculated isotope effect is too low, a tunnelling correction must be applied.[831]

Transition structure geometries for the conrotatory electrocyclic cyclobutene ring opening are very similar regardless of the level of theory employed.[849-851] The largest bond length deviation is seen in for the cleaving C–C bond, which differs by 0.14 Å.

The two-configuration (TCSCF) geometry is the latest, *i.e.* most bond breaking, at 2.238 Å for the cleaving C–C bond, while the RHF/STO-3G geometry is the earliest at 2.102 Å. It is unknown which geometry is the best, as transition structure geometries cannot be experimentally determined.

The Diels-Alder reaction between 1,3-butadiene and ethene has an activation energy of 27 kcal mol^{-1}.[852] The activation barrier for a diradical intermediate is only a few kcal mol^{-1} higher than the experimental activation energy.[853] Transition structures located using semi-empirical methods,[854] *ab initio* HF and MP2 calculations,[855,856] and MCSCF techniques[857] are similar for a concerted mechanism. The forming C–C bonds show the largest variation, ranging from 2.20 Å to 2.28Å. Inclusion of electron correlation lengthens the forming C–C bond in the transition structure. Dewar's early methods predict very asynchronous transition structures, but AM1 and PM3 predict concerted transition structures similar to *ab initio* predictions.[854a]

10.2.5 Reaction energetics

Experimental and calculated activation energies and heats of reaction for the above cases are tabulated overleaf. Experimental activation energies for the $[1_s,5_s]$ hydrogen shift,[844] conrotatory electrocyclic ring opening of cyclobutene,[858] and Diels-Alder reaction between butadiene and ethene,[852] are 36.1 ± 2.2 kcal mol^{-1}, 32.9 ± 0.5 kcal mol^{-1}, and 26.5 ± 2 kcal mol^{-1} respectively. For the $[1_s,5_s]$ hydrogen shift, the 6-31G* calculated $\Delta E^{\ddagger}$ approaches the HF limit of 58-59 kcal mol^{-1}, 20 kcal mol^{-1} above the experimental value. For the cyclobutene ring opening and Diels-Alder reactions, RHF $\Delta E^{\ddagger}$ are ~12 and 20 kcal mol^{-1} higher than the experimental values. The correlation energies of activation, the gap between HF and experimental activation energies are different for different reactions, with smaller errors for cyclobutene ring opening (four electrons involved in bonding changes, compared to six for the other reactions). As more terms in the Møller-Plesset perturbation series are included, the calculated $\Delta E^{\ddagger}$ slowly converges, with good agreements at the MP4(SDTQ)/6-31G*//MP2-6-31G* level. AM1 Activation energies are within 4 kcal mol^{-1} of the experimental values, although enthalpies of reactions are accurate to within 5 to 20 kcal mol^{-1} of experimental values.

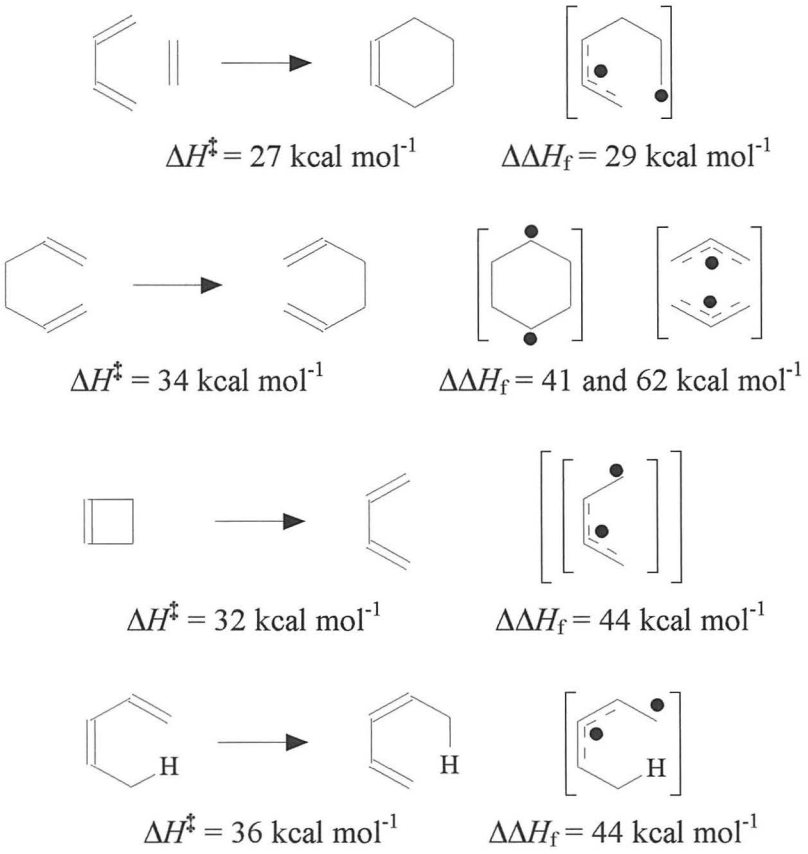
Table 10.1 **Calculated activation energies / kcal mol^{-1} for the [1$_s$,5$_s$] hydrogen shift of (*Z*)-1,3-pentadiene (I), the electrocyclic ring opening of cyclobutene (II), and the Diels-Alder reaction of butadiene with ethene (III).**

	I	II		III	
Computational level	$\Delta E^{\ddagger}$	$\Delta E^{\ddagger}$	ΔE_{react}	$\Delta E^{\ddagger}$	ΔE_{react}
RHF/STO-3G	64.3	79.7	-12.5	36.0	-104.7
RHF/3-21G	54.8	41.6	-18.0	35.9	-43.1
RHF/4-31G		41.9	-20.1	42.6	-34.4
RHF/6-31G	56.5	41.9	-20.3		
RHF/6-31G*	58.7	46.9	-10.6	45.0	
RHF/6-31G*//3-21G	**58.7**	**46.2**	**-13.5**	**45.1**	
RHF/6-31G**	58.1	46.5	-18.4		
MP2/3-21G	39.3	32.7	-22.0		
MP2/6-31G*	36.5	36.8	-7.8	17.6	
MP2/6-31G*//RHF/3-21G	37.6	36.6	-8.1	17.0	
MP2/6-31G**//MP2/6-31G*	36.5	36.8			
MP3//MP2/6-31G*	44.7	40.2		26.9[a]	
MP4(DQ)//MP2/6-31G*	46.3	40.5			
MP4(SDQ)//MP2/6-31G*	46.1	39.8		29.0[a]	
MP4(SDTQ)//MP2/6-31G*	**41.1**	**36.6**		**21.9**[a]	
MCSCF		39.4[b]	-5.0[b]	32.7[c]	
MCSCF		35.8[d]	-6.2[d]	33.3[e]	
MINDO/3-CI[f]		**48.9**	**-1.1**	**28.2**	
MNDO[f]		**49.9**	**-2.0**	**45.4**	**-53.9**
AM1[f]		**35.3**	**-15.9**	**23.8**	**-56.5**
Experimental ΔE	38.8[g]	34.5[h]		25.1[i]	
Arrhenius E_a	**36.1**[g]	**32.9**[h]		**26.1**[i]	
Experimental ΔH_{react}	**0.0**		**-10.9**		**-38.4**

a) RHF/6-31G* geometries. (855b) b) TCSCF (DZ).[j] (849b)
c) CAS2/STO-3G.[k] (857) d) CISD (DZ).[l] (849b)
e) CAS1/4-31G.[k] (857) f) These methods calculate ΔH not ΔE.
g) ± 2.2 kcal mol^{-1}. h) ± 0.5 kcal mol^{-1}. i) ± 2 kcal mol^{-1}.
j) TCSCF calculations using Huzinaga-Dunning[859] double-z basis: C (9s5p/4s2p) and H (4s/2s). For DZ+d basis a higher activation of 42.9 kcal mol^{-1} is obtained.
k) CASSCF wavefunctions with CI expansions derived from 4 electrons in 4 orbitals - π_2, π_3 of butadiene, π, π^* of ethene (CAS1); and six electrons in 6 orbitals (CAS 2) - π_1, π_2, π_3, π_4 of butadiene, π, π^* of ethene.
l) CISD calculations using Huzinaga-Dunning[859] double-z basis. For DZ+d basis a higher activation energy of 42.4 kcal mol^{-1} is obtained.

Thermally allowed pericyclic reactions involving two π systems may follow stepwise mechanisms with biradical intermediates, rather than concerted mechanisms. The energy of concert[853] was defined as the difference between the energy to form a biradical intermediate, and the measured activation enthalpy, shown overleaf. Cycloadditions have small to moderate energies of concert, while sigmatropic hydrogen shifts and electrocyclisations have large energies of concert.

Figure 10.2 Experimental activation energies $\Delta H^{\ddagger}$ compared to estimated energies of biradical intermediates $\Delta\Delta H_f$, relative to reactants.



$\Delta H^{\ddagger} = 27$ kcal mol^{-1} $\Delta\Delta H_f = 29$ kcal mol^{-1}

$\Delta H^{\ddagger} = 34$ kcal mol^{-1} $\Delta\Delta H_f = 41$ and 62 kcal mol^{-1}

$\Delta H^{\ddagger} = 32$ kcal mol^{-1} $\Delta\Delta H_f = 44$ kcal mol^{-1}

$\Delta H^{\ddagger} = 36$ kcal mol^{-1} $\Delta\Delta H_f = 44$ kcal mol^{-1}

10.3 PERICYCLIC REACTIONS

10.3.1 Pericyclic reactions in organic synthesis

Pericyclic reactions were defined by Woodward and Hoffmann[860] as "reactions in which all first-order changes in bonding relationships take place in concert on a closed curve".[861] There has been considerable interest in the nature of the transition structures of such reactions and the theoretical underpinnings of the Woodward-Hoffmann rules,[862] which showed that thermally allowed reactions could occur by concerted mechanisms. Other theoretical treatments have assumed aromatic transition states,[863] but occasionally very unsymmetrical, highly asynchronous transition structures have been advocated.[864] Pericyclic reactions are widely applicable in synthesis.[865-872]

10.3.2 Classification of pericyclic reactions

Woodward and Hoffmann described five categories of pericyclic reactions: cycloaddition reactions, electrocyclic reactions, sigmatropic shifts, cheletropic reactions and group transfer reactions.[861] Ene reactions[873] and dyotropic shifts[874] do not conveniently fit into the classification scheme. The reactions are classified according to the types of bond transformations involved. Some sigmatropic shifts and group transfers do not involve a change in the number of π or σ bonds. 1π-to-1σ transformations include electrocyclic and ene reactions; unstrained reactions of this type are exothermic by ~ 23 kcal mol^{-1}. The 2π-to-2σ transformations include cycloadditions and cheleotropic reactions. 3π-to-3σ Transformations include [2+2+2] cycloadditions. The two types of reactions are exothermic, $\Delta H \approx$ -46 and 69 kcal mol^{-1} respectively. Based upon the Hammond postulate,[875] a gradual shift from a central to an earlier transition state is expected along this enthalpy series.[876]

10.4 COMPUTATIONAL TECHNIQUES

Transition structures and activation energies were obtained using the MINDO/3, MNDO, AM1 and PM3 semi-empirical methods contained in the *MOPAC* package. While simple systems can be studied on a PC by using the DOS version, faster results can be obtained for large systems on a RISC-6000 IBM workstation using the *InsightII* version of *MOPAC*. Transition structures can be obtained by scanning constrained reaction co-ordinates,[877] combining the use of *PCModel*[878] for initial construction of the *.mop* input file,[879,880] and *HyperChem* for structural analysis. To calculate transition structures of more complex reactions, **SADDLE** (an unreliable method - *Section 10.4.1*) or **mode following** (*Section 10.4.2*) can be used. Reaction co-ordinates can also be obtained using the **IRC** method (*Section 10.4.3*).

10.4.1 SADDLE calculations

In a SADDLE calculation,[881] the first geometry, specifying the reactants and the second geometry, specifying the products are defined, using the same format as that for the first geometry. SADDLE often fails, as the two geometries must be related by a continuous deformation of the co-ordinates, frequently due to equivalent dihedral angles in the reactant and product differing by ~ 360° rather than 0°. The calculation can also be run with Cartesian co-ordinates using the XYZ keyword. A reversal of the

gradient direction vector and its cosine, occurs close to the saddle point. SADDLE can be made to run in Cartesian co-ordinates using the XYZ keyword. As SADDLE often fails, another way of calculating transition structures was sought, this time using the TS option within the Eigenvalue Following (EF) routine.

10.4.2 Calculating transition states using TS

The EF (Eigenvector following) routine[635] is a faster alternative to BFGS[2], whose efficiency was evaluated by Agrafiotis and Rzepa,[882] and works well with Cartesian co-ordinates. Within the EF routine, a transition state can be optimised using the TS keyword. In the event that TS does not converge onto a stationary point, convergence can be attained by recalculation of the Hessian every 5 steps in the EF procedure (RECALC=5). In the EF routine, the maximum step size is 0.2 (Ångströms or radians) by default (changed by specifying DMAX=*n.nn* - increasing DMAX leads to faster convergence, but with deterioration of the Hessian for large stepsizes). Reduction of the stepsize to 0.10 or 0.05 is recommended when encountering convergence problems. In extreme cases RECALC=1 and DMAX=0.05 will always find a stationary point, if it exists. Mode following is used in the EF routine. The default MODE=1 is used to locate a transition structure, but if this is incorrect, the vector to be followed can be defined using MODE=*n*. All pericyclic transition structures encountered in this chapter were calculated using the TS method. The method sometimes produces multiple maxima, *i.e.* several negative force constants are obtained on frequency analysis. When this occurs, modifying the geometry, *e.g.* substituent torsional angles, can lead to the true transition structure.

10.4.3 Intrinsic reaction co-ordinate (IRC) calculations

The traditional way to elucidate one of the many reaction paths is to move downhill from a saddle point to the two minima. The Dynamic Reaction Co-ordinate (DRC) is the path followed by all the atoms in a system assuming conservation of energy. As minima and transition states are stationary points, an initial deflection or momentum must be supplied. The Intrinsic Reaction Co-ordinate (IRC) is the path followed by all the atoms in a system assuming all kinetic energy is completely lost at every point, *i.e.* the total energy is always equal to the potential energy.[883] The IRC is intended for use starting with the transition state geometry. A normal co-ordinate is chosen, usually the reaction co-ordinate, and the system is displaced along this co-ordinate. The internal

modes are obtained by calculating the mass-weighted Hessian using FORCE, and translating the resulting Cartesian normal mode eigenvectors to conserve momentum.

A simple steepest descents algorithm with a reasonable step size will usually give a reaction path that oscillates about the true minimum energy path. To determine better reaction pathways, corrective methods must be applied. Each of the following three methods first calculates the gradient at the current point, $\mathbf{x}_k$:

(i) The method of Ishida, Morokumz and Komornicki[884] takes a predefined step size, *s*, to give a new point **x'**. The gradient at the point **x'** is calculated and minimisation performed in the direction along the bisector of the initial gradient at $\mathbf{x}_k$ and the gradient at the new point, giving the next point $\mathbf{x}_{k+1}$ on the approximation to the reaction path.

(ii) The method of Müller and Brown[885] also takes a step size, *s*, along the gradient. The next point on the reaction path is obtained by minimising from this point while keeping a constant distance, *s*, from the original point.

(iii) Gonzalez and Schelgel[886] take a step of *s*/2 along the gradient, giving a new point, **x'**. The next point on the reaction path is obtained by minimising the energy subject to the constraint that the distance between **x'** and the new point on the reaction path $\mathbf{x}_{k+1}$ is *s*/2. The reaction path is then approximated by a circle that passes through both $\mathbf{x}_k$ and $\mathbf{x}_{k+1}$ and whose tangents at those two points are in the directions of the gradients.

10.4.4 Calculating properties of transition states

The Hessian matrix (in millidynes/Å) of second derivatives of energy with respect to displacements of all pairs of atoms in *x*, *y*, and *z* directions, is calculated using the FORCE keyword. Force constants are obtained from the Hessian upon diagonalisation. The force matrix, weighted for isotopic masses, is then used to calculate the vibrational frequencies. The system can be characterised as a ground state or a transition state by the presence of five (for a linear system) or six eigenvalues, which are very small (less than about 30 cm^{-1}). A transition state is further characterised by one, and exactly one, negative force constant. Before a force

calculation is started, a check is made to ensure the existence of a stationary point. Thermodynamic quantities can be calculated for a single temperature, *e.g.* THERMO(298,298), or a range of temperatures. If THERMO is specified on its own, then a temperature range starting at 200 K and increasing in steps of 10 K to 400 K is used. To calculate the rotational contributions to the thermodynamic quantities, the symmetry number of the molecule, ROT=*n*, must be supplied. The symmetry number of a point group is the number of equivalent positions attainable by "pure" rotation, and may be affected by subtle changes, such as isotopic substitution.

Table 10.2 Symmetry number of point groups.

Point group						Symmetry number, *n*
C_1	C_I	C_s	$C_{\infty v}$			1
C_2	C_{2v}	C_{2h}	$D_{\infty h}$			2
C_3	C_{3v}	C_{3h}	S_6			3
C_4	C_{4v}	C_{4h}	D_2	D_{2d}	D_{2h}	4
C_6	C_{6v}	C_{6h}	D_3	D_{3d}	D_{3h}	6
D_4		D_{4d}		D_{4h}		8
D_6	D_{6d}	D_{6h}	T	T_d		12
O_h						24

10.4.5 Animating transition states using *HyperChem for Windows*

Once a transition structure has been verified within *MOPAC*, the resulting structure can be saved as a *.pdb* file, and can be imported into *HyperChem for Windows*. The Hessian can be calculated using *Vibrations*. The highlighted vibration (*e.g.* the imaginary vibration) can be animated using *Vibrational Spectrum* – the eigenvectors (*IR vectors*) for which can be displayed using the *Rendering* option.

10.5 SIGMATROPIC SHIFTS

Sigmatropic shifts involve the migration of a σ bond along one or two π systems. The transition structures show some resemblance to the geometries of radical fragments, although interaction energies between these fragments are substantial, and free diradical intermediates are not generally formed. Sigmatropic shifts were chosen for initial study, as "true" (verified by vibrational analysis) minima of single-molecule systems are readily calculated, in contrast with those of bimolecular assemblies (*e.g.* butadiene and cyclopropene).

10.5.1 [1,2] Sigmatropic shifts

Symmetry-allowed two-electron [1,2]-shifts in carbocations are facile. Bridged nonclassical ethyl and vinyl cations, transition structures between classical cations, are actually minima on potential energy surfaces.[887] Transition structures for the nonaromatic [1,2] hydrogen shift of three-electron hydrocarbon radicals,[888] and the antiaromatic four-electron hydrocarbon anions,[889] are comparable in energy to fragments formed in stepwise dissociation-recombination pathways.

10.5.2 [1,3] Sigmatropic shifts

The [1,3] sigmatropic H-shift in propene has been studied extensively,[845,890-892] although no experimental evidence exists for a thermal, concerted [1,3] H-shift. The allowed [1_s,3_a] transition structure is very contorted, and similar in energy to an allyl radical plus a hydrogen atom, formed *via* a dissociation-recombination mechanism. At the CASSCF/6-31G* level, the partial C---H bond is 1.61 Å long, while the C–C bond length resembles an allyl radical, with pyramidalised terminal carbon atoms maintaining partial bonding to the migrating hydrogen. The forbidden [1_s,3_s] H-shift is unfavourable. Both UHF and MCSCF calculations predict a transition structure resembling a trimethylene diradical, ~ 60 kcal mol^{-1} less stable than propene.[892]

a) [1,3] Sigmatropic hydrogen shifts of radical cations

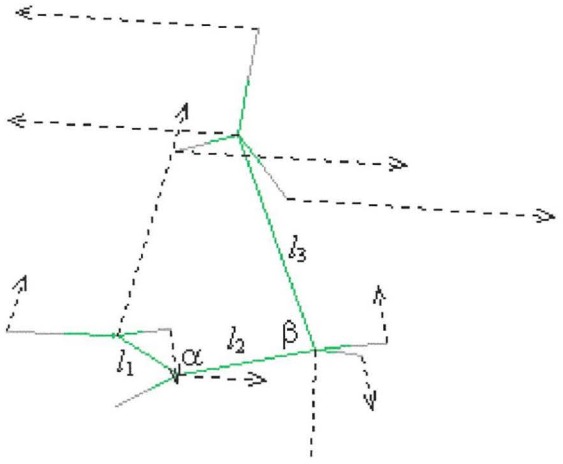
The [1_s,3_s] sigmatropic shift in the radical cation of propene has been studied at the MP2 level with UHF/6-31G* geometries.[893] The transition structure has C_s symmetry resembling the trimethylene radical cation (an energy minimum).[893] The calculated activation energy is 30 kcal mol^{-1}, lower than the 60 kcal mol^{-1} barrier of propene, illustrating that forbidden reactions can be made facile by removal of an electron.

b) [1,3] Sigmatropic methyl shifts

The ground state of 1-butene **(304)** and the transition structure **(305)** of the allowed [1_a,3_s] methyl shift, were calculated using *MOPAC*. The transition structure was probed using the TS keyword, with Cartesian co-ordinates (XYZ), following the first normal mode and recalculating the Hessian after 5 steps. MINDO/3, MNDO, AM1 and PM3 Activation energies of 108.4, 123.4, 106.6 and 111.3 kcal mol^{-1} respectively, are within 10-20 kcal mol^{-1} of the MP2/6-31G*//3-21G value[894] of 96 kcal mol^{-1}, while C–C bond cleavage requires 72 kcal mol^{-1}. Steric interactions between the

methyl hydrogens and the allyl moiety are indicated by very long forming and breaking C---C bond lengths of ~2.4 Å. A concerted symmetry-forbidden [1_s,3_s] sigmatropic shift would be more favourable than a stepwise process, as subjacent orbital interactions stabilise the forbidden transition structure, resembling a 2-methyltrimethylene diradical.[895]

Table 10.3 Geometries of 1-butene and the transition structure for the allowed [1_a,3_s] methyl shift.



(304) & (305)

	(304)					(305)			
	l_1 / Å	l_2 / Å	l_3 / Å	α / °	β / °	$l_1 = l_2$ / Å	L_3 / Å	α / °	β / °
A	1.33	1.49	1.51	126.0	115.9	1.38	2.36	119.2	79.8
B	1.33	1.48	1.51	126.0	114.8	1.38	2.41	122.4	87.0
C	1.34	1.50	1.53	129.9	118.6	1.39	2.93	126.2	77.2
D	1.33	1.48	1.48	134.5	123.2	1.45	1.63	106.8	79.6

A = PM3, B = AM1, C = MNDO, D = MINDO/3

Table 10.4 Thermodynamic data for the [1_a,3_s] methyl shift of 1-butene.

	MINDO/ 3		MNDO		AM1		PM3	
	(304)	**(305)**	**(304)**	**(305)**	**(304)**	**(305)**	**(304)**	**(305)**
ΔH_f^{298} kcal mol^{-1}	0.4	108.8	1.1	124.5	0.8	107.4	2.3	113.6
ν_1 / cm^{-1}	53	-839	-53†	-462	82	-885	80	-1003

† Minimised structure is not a true minimum.

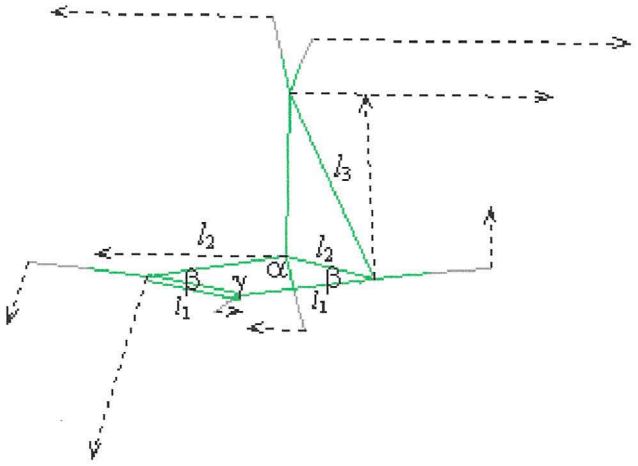
c) [1,3] Sigmatropic shifts involving the methylene group

Many [1,3] methylene shifts have been observed.[896-898] The thermal expansion of a vinylcyclopropane to a cyclopentene ring[897] may be considered as an internal [$_\pi$2 + $_\sigma$2] cycloaddition. Both concerted[899] and diradical[900] pathways have been

proposed. In the transition structure for the allowed [1_a,3_s] vinylcyclopropane to cyclopentene rearrangement, the cyclopropane ring is nearly open, with the breaking and forming C–C bonds equal to 2.21 Å and 2.56 Å respectively. The π system resembles an allylic radical with C–C bond lengths of 1.38 Å and 1.39 Å. Inversion of the migrating methylene carbon also takes place. MCSCF calculations yield a lower barrier than the RHF value of 84 kcal mol^{-1}, with a transition structure even more diradical in nature, implying a stepwise mechanism comparable in energy to a concerted path.[901] Semi-empirical calculations for the vinylcyclopropane to cyclopentene rearrangement were attempted using *MOPAC* using several starting geometries. The presence of multiple maxima (several negative force constants) implied that true transition structures were not obtained.

In the transition structure for the walk rearrangement of bicyclo[2.1.0]pentene,[902] the partial C---C single bonds are 2.24 Å long, with partial double bonds similar to those of the allyl radical. The cyclobutene ring is planar, with little pyramidalisation of the allyl termini or bonding with the migrating group. The ground state of bicyclo[2.1.0]pentene **(306)** and the transition structure **(307)** of the allowed [1_a,3_s] methylene shift, were calculated using *MOPAC*, as above. The MINDO/3 transition structure did not achieve the desired level of convergence. MNDO, AM1 and PM3 Activation energies of 49.5, 42.9 and 46.9 kcal mol^{-1} respectively, are within 15 kcal mol^{-1} of the estimated barrier of 35 kcal mol^{-1}.

Table 10.5 Transition structure geometries for the degenerate walk rearrangement of bicyclo[2.1.0]pentene.



(307)

	l_1 / Å	l_2 / Å	l_3 / Å	α / °	β / °	γ / °
PM3	1.41	1.53	2.30	84.2	91.1	93.6
AM1	1.41	1.54	2.34	83.9	91.2	93.8
MNDO	1.41	1.53	2.40	83.7	91.6	93.2

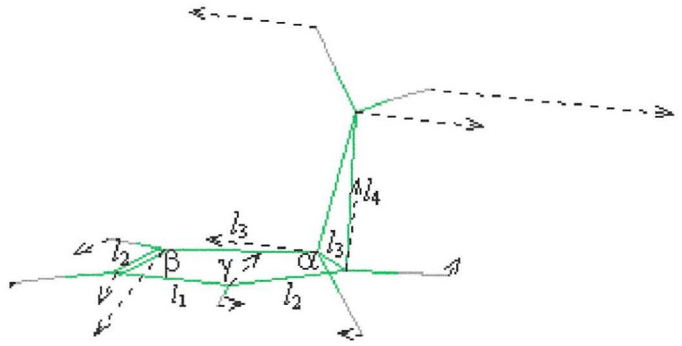
Table 10.6 Thermodynamic data for the degenerate walk rearrangement of bicyclo[2.1.0]pentene.

	MINDO/ 3		MNDO		AM1		PM3	
	(306)	**(307)**	**(306)**	**(307)**	**(306)**	**(307)**	**(306)**	**(307)**
ΔH_f^{298} kcal mol^{-1}	78.1	DNC	77.9	127.4	98.4	141.3	85.7	132.6
ν_1 / cm^{-1}	38	DNC	434	-323	447	-352	433	-448

10.5.3 [1,4] Sigmatropic shifts

RHF/3-21G Transition structures have been located for the walk rearrangement of the bicyclo[3.1.0]hexenyl cation. The forbidden [1_s,4_s] process has a distorted geometry and an activation energy of 33 kcal mol^{-1}, double that of the experimental value for the allowed [1_a,4_s] methylene shift.[903] The ground state of the bicyclo[3.1.0]hexenyl cation **(308)** and the transition structure **(309)** of the allowed [1_a,4_s] methylene shift were calculated using *MOPAC*, as above. MINDO/3, MNDO, AM1 and PM3 Activation energies of 26.3, 19.4. 15.9 and 18.7 kcal mol^{-1} respectively, are mostly within 5 kcal mol^{-1} of the experimental value of 15 kcal mol^{-1}.[903]

Table 10.7 Transition structure geometries for the degenerate walk rearrangement of the bicyclo[3.1.0]hexenyl cation.



(309)

	l_1 / Å	l_2 / Å	l_3 / Å	l_4 / Å	α / °	β / °	γ / °
PM3	1.46	1.36	1.52	2.35	102.5	109.0	109.7
AM1	1.46	1.36	1.53	2.39	102.2	109.0	109.8
MNDO	1.48	1.36	1.54	2.47	101.3	109.8	109.5

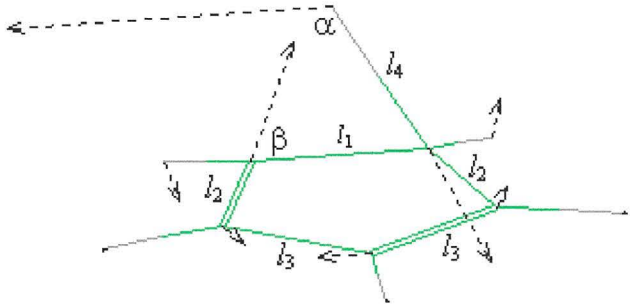
Table 10.8 Thermodynamic data for the degenerate walk rearrangement of the bicyclo[3.1.0]hexenyl cation.

	MINDO/ 3		MNDO		AM1		PM3	
	(308)	**(309)**	**(308)**	**(309)**	**(308)**	**(309)**	**(308)**	**(309)**
ΔH_f^{298} kcal mol^{-1}	231.4	257.7	240.6	260.0	249.5	265.4	249.1	267.8
ν_1 / cm^{-1}	168	-245	314	-161	304	-263	297	-324

10.5.4 [1,5] Sigmatropic shifts

The [1,5] hydrogen shift in cyclopentadiene has a low activation energy of 21 kcal mol^{-1},[846] the close proximity of the migration origin and terminus permitting good overlap of the hydrogen orbitals with orbitals on the cyclopentadiene ring. Structures of the ground state of cyclopentadiene **(310)** and the transition structure **(311)** of the allowed [1_s,5_s] hydrogen shift were calculated using *MOPAC*, as described above. MINDO/3, MNDO, AM1 and PM3 Activation energies were found to be 28.3, 48.9, 39.5 and 37.6 kcal mol^{-1} respectively, with AM1 and PM3 activation energies within 20 kcal mol^{-1} of the experimental value of 21 kcal mol^{-1}.[846]

Figure 10.3 PM3 Transition structure geometry for the degenerate [1_s,5_s] hydrogen shift in cyclopentadiene.



(311)

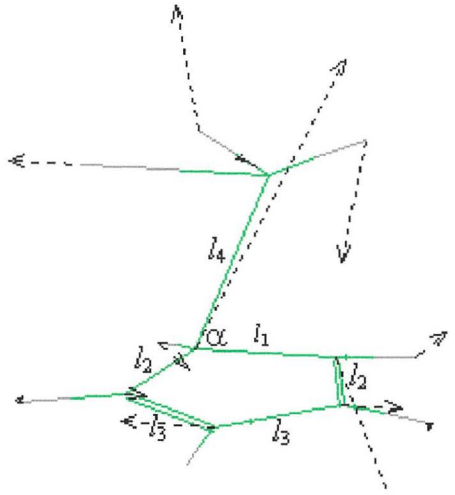
l_1 = 1.49 Å, l_2 = 1.41 Å, l_3 = 1.40 Å, l_4 = 1.42 Å, α = 63.1 °, β = 112.0 °

Table 10.9 Thermodynamic data for the degenerate [1_s,5_s] hydrogen shift in cyclopentadiene,

	MINDO/ 3		MNDO		AM1		PM3	
	(310)	**(311)**	**(310)**	**(311)**	**(310)**	**(311)**	**(310)**	**(311)**
ΔH_f^{298} kcal mol^{-1}	41.1	69.4	32.1	81.0	37.1	76.6	31.8	69.4
ν_1 / cm^{-1}	309.0	-1297	335	-2019	344	-1747	344	-1803

Many examples of [1,5] methyl shifts in cyclopentadiene systems exist, with experimental activation energies in the range of 40-44 kcal mol^{-1} for simple methyl-substituted cyclopentadienes.[904] The calculated RHF/3-21G barrier for the allowed [1_s,5_s] methyl shift of 5-methylcyclopentadiene is 53 kcal mol^{-1}, corresponding to an experimental barrier of 33 kcal mol^{-1}. Structures of the ground state of 5-methylcyclopentadiene **(312)** and the transition structure **(313)** of the allowed [1_s,5_s] methyl shift were calculated using *MOPAC*, as above. MINDO/3, MNDO, AM1 and PM3 Activation energies of 47.4, 62.5, 60.4 and 56.4 kcal mol^{-1} respectively are within 20 to 30 kcal mol^{-1} of the experimental value of 33 kcal mol^{-1},[904] and within 5 kcal mol^{-1} of the RHF/3-21G activation energy.

Figure 10.4 PM3 Transition structure geometry for the degenerate [1_s,5_s] methyl shift of 5-methylcyclopentadiene.



(313)

l_1 = 1.47 Å, l_2 = 1.41 Å, l_3 = 1.40 Å, l_4 = 1.98 Å, α = 108.8 °

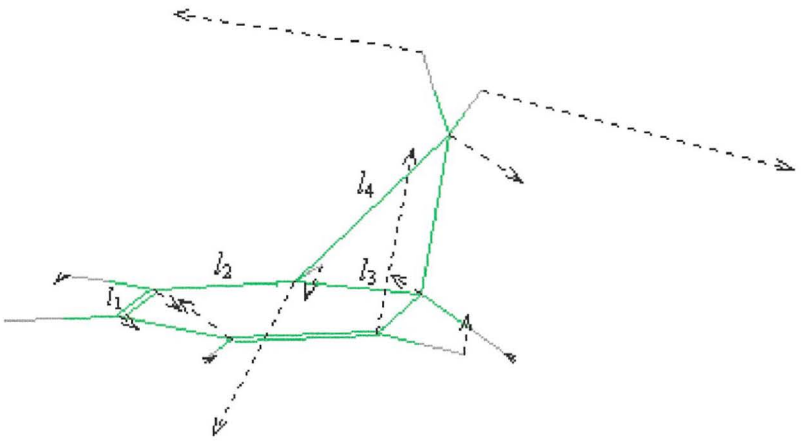
Table 10.10 Thermodynamic data for the degenerate [1_s,5_s] methyl shift of 5-methylcyclopentadiene.

	MINDO/ 3		MNDO		AM1		PM3	
	(312)	**(313)**	**(312)**	**(313)**	**(312)**	**(313)**	**(312)**	**(313)**
ΔH_f^{298} kcal mol^{-1}	38.5	85.9	27.7	90.2	32.6	93.0	27.0	83.4
ν_1 / cm^{-1}	146	-818	159	-1157	144	-992	153	-1069

The activation energy of the [1,5] sigmatropic methylene shift in bicyclo[4.1.0]heptadiene, norcaradiene, is not known experimentally. The reaction is

believed to proceed with inversion of the migrating alkyl group in substituted cases.[905] Structures of the ground state of norcaradiene **(314)** and the transition structure **(315)** of the allowed [1_s,5_s] methylene shift were calculated using *MOPAC*. AM1 and PM3 Activation energies of 53.7 and 57.2 kcal mol^{-1} respectively are in good agreement with the MP2/6-31G* single point value of 53 kcal mol^{-1}.

Figure 10.5 PM3 Transition structure for the [1_s,5_s] methylene shift of norcaradiene.



(315)

l_1 = 1.40 Å, l_2 = 1.37 Å, l_3 = 1.49 Å, l_4 = 2.24 Å

Table 10.11 Thermodynamic data for the degenerate [1_s,5_s] methylene shift of norcaradiene.

	MINDO/ 3		MNDO		AM1		PM3	
	(314)	**(315)**	**(314)**	**(315)**	**(314)**	**(315)**	**(314)**	**(315)**
ΔH_f^{298} kcal mol^{-1}	46.6	108.4	41.8	103.2	51.0	104.7	49.8	107.0
ν_1 / cm^{-1}	237	-405	237	-543	222	-492	222	-587

Other reactions such as the cyclopropane analogue of the homodienyl [1,5] shift[906] were not attempted, although further studies are envisaged.

10.5.5 [3,3] Sigmatropic shifts

[3,3] Sigmatropic shifts and other higher order sigmatropic shifts have the possibility of two stepwise mechanisms: fragmentation to two allyl or other conjugated radical fragments and recombination, or formation of a diradical by addition of the termini of the π bonds to each other, followed by cleavage. Although the concerted pathway

may involve transition structures resembling either one of these extremes, there may even be multiple competing concerted pathways for these processes.[907] *meso*-3,4-Dimethyl-1,5-hexadiene gives the *cis-trans* product, while the *dl*-compound gives the *trans-trans* diene.[908] The nature of the products demonstrate that, at least in these cases, the transition state is a chair and not a boat.[909] The degenerate Cope rearrangement of 1,5-hexadiene is a case where the precise mechanism is still a matter of uncertainty. A variety of computational techniques have been applied.[910-915] A closed-shell, delocalised, aromatic transition structure is predicted from *ab initio* CASSCF/3-21G calculations.[910] The chair transition structure is preferred, in agreement with experimental observations.[914] The 1,4-diyl is estimated to have a significantly higher energy than the observed transition state.[915]

PM3 structures of the ground state of 1,5-hexadiene and the transition structure of the degenerate Cope rearrangement were calculated using *MOPAC*. The PM3 activation energy for the degenerate Cope rearrangement of 1,5-hexadiene *via* the chair transition structure was found to be 31.1 kcal mol^{-1}, in agreement with the experimental value of 33.3 kcal mol^{-1}; while that *via* the boat transition structure was found to be 49.6 kcal mol^{-1}, 18.5 kcal mol^{-1} higher than that *via* the chair. [881]

Table 10.12 PM3 Thermodynamic data for the degenerate Cope rearrangement of 1,5-hexadiene.

	1,5-Hexadiene	Chair TS	Boat TS
ΔH_f^{298} / kcal mol^{-1}	21.7	52.8	71.3
ν_1 / cm^{-1}	40.3	-552	-1012

The MINDO/3[916] and MNDO[881] methods predict a two-step reaction with a common symmetrical intermediate consisting of equilibrating biradicaloids related to the 1,5-cyclohexadiyl biradical. Significant biradical character for both the transition state, and the intermediate, is indicated by the fact that inclusion of CI causes a substantial decrease in the energies of both species. Inclusion of CI for biradical species leads to energies that are too negative by ~ 15 kcal mol^{-1}, since MNDO *via* parameterisation allows for electron correlation, and inclusion of CI leads to an overcompensation for correlation and hence to energies which are too negative.

Table 10.13 MNDO[881] and MINDO/3[916] Thermodynamic data for the degenerate Cope rearrangement of 1,5-hexadiene.

Method	ΔH_f^{298} / kcal mol^{-1}			
	MNDO	MNDO +CI	MINDO/3	Observed
1,5-Hexadiene	19.0		19.0	20.2
Chair intermediate			52.3	
Chair TS	61.0	38.0	55.0	53.5
Boat intermediate			59.7	
Boat TS	59.7	20.8	61.3	64.9

The Claisen rearrangement of allyl vinyl ether to hex-4-en-1-al proceeds in a concerted but non-synchronous manner. The energy of the transition state is decreased by < 15 kcal mol^{-1} with the inclusion of CI, indicating very little biradical character. PM3 activation energies *via* the chair and boat transition structures were found to be 35.0 and 40.2 kcal mol^{-1} respectively, in agreement with the experimental value of 30.6 kcal mol^{-1}.[881]

Table 10.14 PM3 Thermodynamic data for the Claisen rearrangement of allyl vinyl ether to hex-4-en-1-al.

	Allyl vinyl ether	Chair TS	Boat TS	Hex-4-en-1-al
ΔH_f^{298} / kcal mol^{-1}	-4.1	30.9	36.1	-29.2
ν_1 / cm^{-1}	40.2	-927	-1014	49

Table 10.15 MNDO Thermodynamic data for the Claisen rearrangement of allyl vinyl ether to hex-4-en-1-al.[881]

Method	ΔH_f^{298} / kcal mol^{-1}		
	RHF	+CI	Observed
1,5-Hexadiene	-7.2		-8.7
Chair TS	32.3	19.8	21.9
Hex-4-en-1-al	-27.9	20.8	-25.8

For the [3,3] sigmatropic rearrangement of 1,5-hexadiyne to form 1,2,4,5-hexatetraene, which undergoes a very rapid electrocyclic reaction to produce 3,4-dimethylenecyclobutene,[917] the RHF/3-21G transition structure has forming and breaking C---C bond lengths of 1.96 Å and 2.03 Å respectively. The H=C–CH_2- fragments are bent from linear to a bond angle of 147°. The PM3 activation energy for the sigmatropic rearrangement of 1,5-hexadiyne was found to be 56.6 kcal mol^{-1},

while the MP2/6-31G* calculated activation energy is 37 kcal mol^{-1}, close to the experimental value of 34 kcal mol^{-1}.[918]

Table 10.16 PM3 Thermodynamic data for the [3,3] sigmatropic rearrangement of 1,5-hexadiyne.

	1,5-Hexadiyne	TS	1,2,4,5-Hexatetraene
ΔH_f^{298} / kcal mol^{-1}	90.4	147.0	96.4
ν_1 / cm^{-1}	56	-773	55

Transition structures for the degenerate rearrangement of hexa-1,2-dien-5-yne and the rearrangement of hex-1-en-5-yne to 1,2,5-hexatriene have also been calculated. RHF calculations favour concerted mechanisms, although diradical mechanisms are expected to be not to dissimilar in energy.[919] Activation energies for the rearrangements of hexa-1,2-dien-5-yne and hex-1-en-5-yne were found to be 52.4 kcal mol^{-1} and 49.4 kcal mol^{-1} respectively using the PM3 method.

Table 10.17 PM3 Thermodynamic data for the degenerate [3,3] sigmatropic rearrangement of hexa-1,2-dien-5-yne.

	Hexa-1,2-dien-5-yne	TS
ΔH_f^{298} / kcal mol^{-1}	93.3	145.7
ν_1 / cm^{-1}	44	-668

Table 10.18 PM3 Thermodynamic data for the [3,3] sigmatropic rearrangement of hex-1-en-5-yne to 1,2,5-hexatriene.

	Hex-1-en-5-yne	TS	1,2,5-Hexatriene
ΔH_f^{298} / kcal mol^{-1}	55.8	105.2	58.8
ν_1 / cm^{-1}	55	-584	39

The Cope rearrangement has also been applied to divinylcyclopropanes and vinylcyclobutanes.[920] *cis*-1,2-Divinylcyclopropanes undergo the rearrangement so rapidly that they cannot be isolated at room temperature,[921] although some exceptions are known.[922] 1,2-Divinylcyclopropane can only react in the boat form, demonstrating that such transition structures are not impossible. *cis*-1,2-Divinylcyclobutane rearranges to give 1,5-cyclooctadiene, while the main product for the *trans*-isomer is 4-vinylcyclohexene, formed by a diradical mechanism.[923-925] At room temperature the ^{1}H-NMR spectrum of bicyclo[5.1.0]octa-2,5-diene shows one compound.[926] At

180 °C, the NMR spectrum shows an equilibrium mixture of two rapidly interconverting valence tautomers.[927] Calculated PM3 activation energies for the rearrangements of divinylcyclopropane and bicyclo[5.1.0]octa-2,5-diene were found to be 49.2 kcal mol^{-1} and 30.3 kcal mol^{-1} respectively.

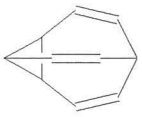
Table 10.19 PM3 Thermodynamic data for the [3,3] sigmatropic rearrangement of 1,2-divinylcyclopropane to 1,4-cycloheptadiene.

	1,2-Divinylcyclopropane	TS	1,4-Cycloheptadiene
ΔH_f^{298} / kcal mol^{-1}	52.7	101.9	20.3
ν_1 / cm^{-1}	39	-880	103

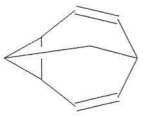
Table 10.20 PM3 Thermodynamic data for the degenerate [3,3] sigmatropic rearrangement of bicyclo[5.1.0]octa-2,5-diene.

	Bicyclo[5.1.0]octa-2,5-diene	Transition structure
ΔH_f^{298} / kcal mol^{-1}	51.4	81.7
ν_1 / cm^{-1}	232	-598

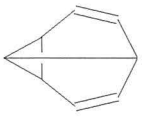
Bullvalene, tricyclo[3.3.2.$0^{2,8}$]deca-3,6,9-triene **(316)**, undergoes several Cope rearrangements, to give 10!/3 or 1.2 million tautomeric forms. At -25 °C, there are two separate vinylic and allylic resonances in its ^{1}H-NMR spectrum.[928] At 100 °C, only one resonance occurs in its ^{1}H-NMR[929] and ^{13}C-NMR[930] spectra due to the rapidly interconverting fluxional structures. Other valence tautomers, such as barbaralane **(317)** and semibullvalene **(318)** are known.[931-934] PM3 Activation energies for bullvalene and semibullvalene were found to be 31.3 and 22 kcal mol^{-1} respectively. The transition structure for barbaralane was found to be a multiple maximum.



(316)



(317)



(318)

Table 10.21 PM3 Thermodynamic data for the [3,3] sigmatropic rearrangement of bullvalene.

	Bullvalene	Transition structure
ΔH_f^{298} / kcal mol^{-1}	76.8	108.1
ν_1 / cm^{-1}	259	-622

Table 10.22 PM3 Thermodynamic data for the [3,3] sigmatropic rearrangement of semibullvalene.

	Semibullvalene	Transition structure
ΔH_f^{298} / kcal mol^{-1}	74.4	96.4
ν_1 / cm^{-1}	300	-547

10.5.6 Summary

Calculated activation energies for a selection of the above reactions have been collated together in *Table 10.23*, and are overestimated by 5 to 25 kcal mol^{-1}. *Ab initio* activation energies differ by 10 to 30 kcal mol^{-1} from semi-empirical values, while semi-empirical activation energies differ from method to method by ~ 10 kcal mol^{-1}. Semi-empirical activation energies resembling *ab initio* values can be obtained at the fraction of cost. As the physical properties of cyclopropanes and cyclopropenes are best described by the PM3 method, which also reproduces intermolecular hydrogen bonds, it will be the chosen semi-empirical method for the remainder of this chapter.

Table 10.23 Summarised activation energies / kcal mol^{-1}.

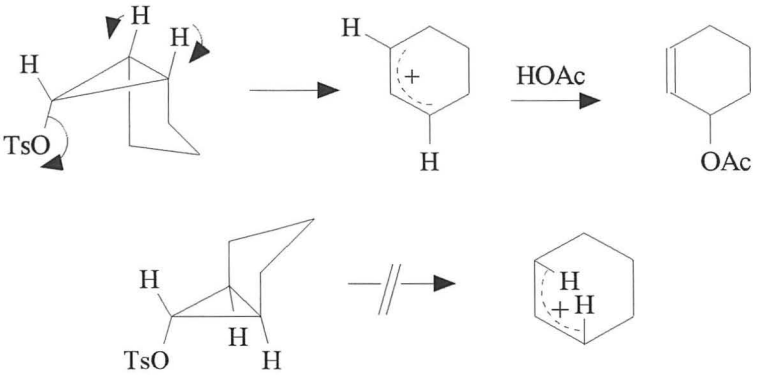
Reaction	Observed	PM3	AM1	MNDO	MINDO/3	*Ab initio*
(305)		111.3	106.6	123.4	108.4	96[†894]
(307)	35[902]	46.9	42.9	49.5	DNC	
(309)	15[903]	18.7	15.9	19.4	26.3	
(311)	21[846]	37.6	39.5	48.9	28.3	
(313)	33[904]	56.4	60.4	62.5	47.4	53[‡]
(315)		57.2	53.7	61.4	61.8	53[¶]
Cope	33.3[881]	31.1		42.0		
Claisen	30.6[881]	35.0		39.5		

† MP2/6-31G*//3-21G ‡ RHF/3-21G ¶ MP2/6-31G*

10.6 ELECTROCYCLIC REACTIONS

Electrocyclic reactions involve the cyclisation or ring-opening of conjugated π systems, and generally obey the Woodward-Hoffmann rules, unless geometrical constraints such as incorporation into a small ring prevents the allowed process.[935]

The disrotatory ring opening of cyclopropyl cations is an electrocyclic reaction governed by the orbital symmetry rules.[936] Cyclopropyl anions however undergo conrotatory ring opening.[937] Cleavage of the σ bond is concerted with departure of the leaving group in the original cyclopropyl substrate. The σ bond provides anchimeric assistance to the removal of the leaving group. Two disrotatory pathways are available, but as the σ orbital provides assistance, the two substituents which are trans to the leaving group must move outward, since the former puts the electrons of the σ bond on the side opposite that of the leaving group.[938] Strong confirmation of this picture[939] comes from acetolysis of *endo*- and *exo*-bicyclo[3.1.0]hexyl-6-tosylate, shown below, in which the groups *trans* to the tosylate must move forward. The *endo*-/*exo*- solvolysis rate ratio was found to be greater than 2.5×10^6.[940] NMR investigations using super acids, prove that the predicted cation is formed.[941] When the nucleophile is the same as the leaving group, it re-enters from the same side from which it departed.[942]



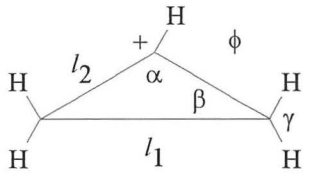
The disrotatory symmetry-allowed cyclopropyl cation opening to form the allyl cation is exothermic with a vanishingly small barrier according to MP2 calculations.[943] The breaking C---C bond length is only 1.56 Å in the transition structure, nearly the same as in the cyclopropyl cation. Experimental data indicate that the ring opening occurs in concert with ionisation in the reactions of cyclopropyl halides and tosylates.[942] The cyclopropyl radical opens with a significant barrier of ~ 20 kcal mol^{-1}.[944] Similarly, the exothermic (-43 kcal mol^{-1}) conrotatory ring opening of the cyclopropyl anion has an the MP2/6-31+G* activation barrier of 20 kcal mol^{-1}. The breaking C---C bond is stretched to 1.94 Å in the transition structure.

The PM3 transition structure for the disrotatory ring opening of the cyclopropyl cation was obtained using the same method as before. As the PM3 calculated activation energy is low, the transition structure for the ring opening is expected to be "early", *i.e.* resembling that of the cyclopropyl cation, in agreement with the experimental data. This is confirmed by analysis of the molecular geometries. Attempts to model the conrotatory ring opening of the cyclopropyl anion using the same method failed, probably due to the distorted geometry of the transition structure.

Table 10.24 PM3 Thermodynamic data for the disrotatory ring opening of the cyclopropyl cation.

	Cyclopropyl cation	TS	Allyl cation
ΔH_f^{298} / kcal mol^{-1}	261.8	268.1	232.7
ν_1 / cm^{-1}	488	-606	425

Table 10.25 PM3 Calculated molecular geometries.



Species	Bond lengths / Å		Bond angles / °			Dihedral / °
	l_1	l_2	α	β	γ	ϕ
Δ cation	1.52	1.43	64.2	57.2	105.3	69.6
TS	1.75	1.41	76.8	51.6	109.4	27.9
Allyl cation	2.41	1.38	122.2	28.9	115.5	0

3,3-Disubstituted-1,2-dichlorocyclopropenes undergo electrocyclic ring opening to produce vinyl carbenes. PM3 activation energies identical to those of Rzepa and Thoss,[945] were obtained for tetrachloro- and 1,2,3-trichlorocyclopropene. In our initial study, other substituted cyclopropenes were not studied.

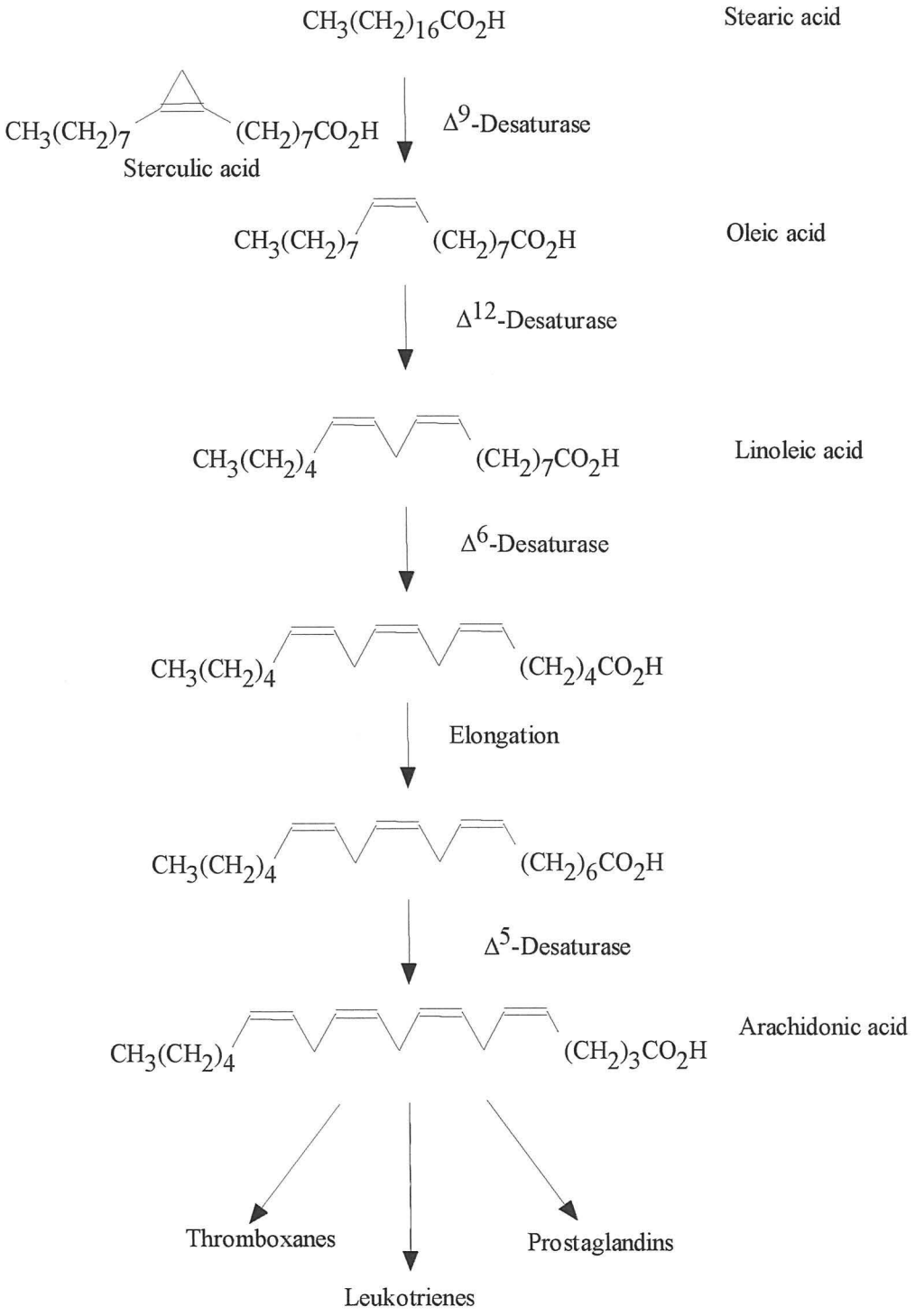
10.7 HYDROGEN ADDITIONS AND EXTRUSIONS

10.7.1 Desaturase enzymes in biological processes

Fatty acids necessary to meet the physical and biochemical requirements of membrane synthesis and lipid storage cannot be supplied by diet and chain elongation alone, and are biosynthesised by a specific oxygen dependent desaturase system. The first *cis* double bond is generally introduced into a saturated acyl chain at the Δ^9

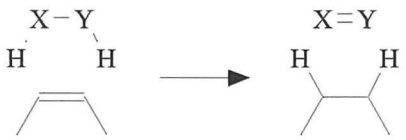
position.[946] After formation of the monounsaturated fatty acid, there is a division between the classes of organisms. In higher plants and fungi, further desaturation occurs at the Δ^{12} and Δ^{15} positions, whereas animals can only desaturase below the Δ^{9} position - initially at Δ^{6}, with further desaturation at Δ^{5} and Δ^{4}, preceded by an elongation step in order to maintain the skip conjugated alkene system.

Figure 10.6 The biosynthesis of arachidonic acid from stearic acid.



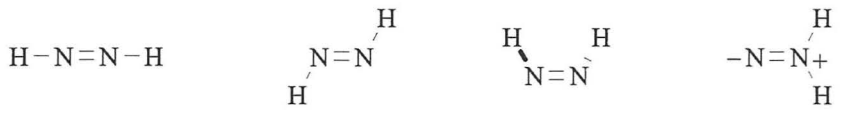
Linoleic acid is the precursor in the above biosynthesis of arachidonic acid, which gives rise to a range of metabolites known as eicosanoids, including the pharmacologically important thromboxanes, leukotrienes and prostaglandins.[947] Sterculic acid is a potent inhibitor of the Δ^9 desaturase enzyme which converts stearic acid into oleic acid.[948] It is assumed that sterculic acid, with a high degree of molecular similarity to oleic acid, can irreversibly bind to the enzyme active site, possibly through thio-groups present in the enzyme,[949] due to the high reactivity of the strained cyclopropene ring. A linoleic analogue with a cyclopropene at the Δ^{12} position is also an inhibitor of the Δ^{12} desaturase enzyme. Research at Bangor has focussed on the development of novel Δ^{12}, Δ^9, Δ^6, Δ^5 and arachidonic inhibitors.

As the desaturase enzyme removes deactivated hydrogen from the saturated backbone in a concerted manner, it was hoped to probe the desaturation mechanism by modelling the concerted addition of hydrogen – the diimide reduction. Although the removal of deactivated hydrogen cannot be carried out under synthetic conditions, analogues of the diimide reduction, where X and Y are heteroatoms or metals, could be designed and tested under experimental conditions, eventually enabling analogues for desaturation to be designed and tested experimentally.



10.7.2 Diimide reductions

Diimide[950-952] exists as a transient intermediate, which cannot be isolated under normal conditions. Diimide is prepared *in situ* by the oxidation of hydrazine,[953] acid decomposition of azodicarboxylate salts,[954] action of alkali on certain sulphonyl and acyl hydrazides,[955] alkaline decomposition of hydroxylamine-*O*-sulphonic acid,[956] the thermal decomposition of the diimide-anthracene adduct, anthracene-9,10-biimine,[957] and the alkaline decomposition of chloramine.[958] A number of possible structures have been considered for diimide:[950]



linear *planar* *non-planar* *ionic*

Diimide reduces alkenes, alkynes, and azo compounds to alkanes, alkenes and hydrazo compounds respectively, the yield of product decreasing with increasing substitution and steric crowding about the multiple bond.[959,960] Polar functional groups are usually inert towards diimide under conditions in which nonpolar multiple bonds are reduced.[959-964] Diimide attacks *trans* double bonds such as in fumaric acid more rapidly than *cis* double bonds such as in maleic acid.[960,965]

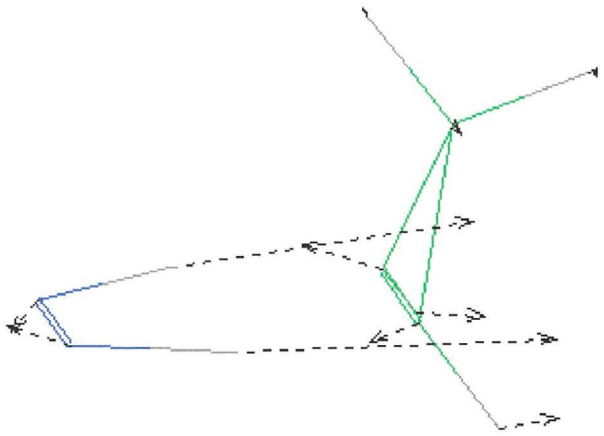
Isotopic labelling experiments established that diimide reduction proceeds with *cis*-addition across the multiple bond - symmetrical *cis*-tetrasubstituted alkenes are hydrogenated to *meso* derivatives, while alkynes are converted to *cis*-alkenes without any *trans*- alkenes being produced.[956,964,966] The mechanism of diimide reduction is believed to involve a synchronous transport of hydrogen through a cyclic transition state.[961,964,966] Diimide must however have the *syn*-conformation in the six-membered cyclic transition state.

10.7.5 Modelling studies

Rzepa has studied dihydrogen transfer reactions using the MNDO and AM1 semi-empirical methods,[967] together with a study of dihydrogen elimination reactions.[968] Previous calculations in this chapter have involved the use of single-molecule systems. Minimisation of supermolecular assemblies such as cyclopropene and diimide produce local rotational minima with small negative eigenvalues. The heat of formation for the supermolecular assembly of cyclopropene and diimide was calculated to be 109.8 kcal mol^{-1}, while summation of their individual heats of formation produced a value ~ 1 kcal mol^{-1} higher. Although the summation approach neglects intermolecular interactions between the individual molecules in the assembly, they are sufficiently nonpolar for this term to be small. For the remainder of the chapter, heats of formation of supermolecular assemblies will be obtained by summation.

The PM3 activation energy for the diimide reduction of cyclopropene was calculated to be 20.5 kcal mol^{-1}, within 4 kcal mol^{-1} of the AM1 value of 25.0 kcal mol^{-1} obtained by Rzepa.[967] Animation of the imaginary vibration using *HyperChem* verifies that two hydrogens are transferred in a synchronous manner. The highly exothermic nature of the reaction is also noted. Steric interactions between the methyl groups and diimide cause the energy of the transition structure for 3,3-dimethylcyclopropene to be approximately 5 kcal mol^{-1} higher than that of 1,2-dimethylcyclopropene. Reduction of both (*S*)- and (*R*)-1,3-dimethylcyclopropene occurs *via* the least hindered face (*syn* to H addition is 1.5 kcal mol^{-1} lower than *syn* to Me addition). The thermodynamic product (*E*)-1,3-dimethylcyclopropane is formed *via* the higher energy route.

Figure 10.7 Eigenvectors for the symmetrical PM3 transition structure of the diimide reduction of cyclopropene.



N=N = 1.16 Å, N–H = 1.19 Å, C---H = 1.51 Å, C=C = 1.39 Å.

Table 10.26 PM3 Thermodynamic data for the diimide reduction of cyclopropenes.

	ΔH_f^{298} / kcal mol^{-1}		
	Starting materials†	Transition structure	Products†
Cyclopropene	110.9	131.4	33.7
1,2-Me	89.6	112.5	18.8
3,3-Me	94.9	117.1	18.1
1,3-Me *syn* to H	92.2	114.2	18.8
1,3-Me *syn* to Me	92.2	115.6	18.2

† Calculated by summation.

Several interesting results were obtained for simple alkenes using the PM3 method. While the AM1, MNDO and MINDO/3 methods predict a synchronous transfer of hydrogen, the PM3 transition structure indicates a step-wise transfer of hydrogen, verified by animation. The N–H bond lengths in the asymmetric PM3 transition structure (1.36 Å and 1.12 Å) are markedly different from those found in the symmetric AM1 transition structure (1.25 Å). Similarly, the partial H---C bond lengths are 1.44 Å and 1.60 Å in the asymmetric PM3 transition structure, and 1.48 Å in the symmetric AM1 transition structure. The PM3 calculated activation energy for the diimide reduction of ethene was found to be 26.2 kcal mol^{-1}, while the AM1 and MNDO activation energies are 32.3 and 54.7 kcal mol^{-1}.[967] PM3 Activation energies for (*Z*)- and (*E*)-but-2-enes were calculated to be 30.1 and 29.3 kcal mol^{-1} respectively, reflecting the greater steric hindrance in the (*Z*)-transition structure. Similar heats of formation for (*Z*)- and (*E*)-but-2-enes of 36.2 kcal mol^{-1} and 36.6 kcal mol^{-1} were obtained using the AM1 method, which wrongly predicted the rate of reduction for (*Z*)-but-2-ene to be faster than that of (*E*)-but-2-ene.

Figure 10.8 Asymmetric eigenvectors for the asymmetric PM3 transition structure of the diimide reduction of ethene.

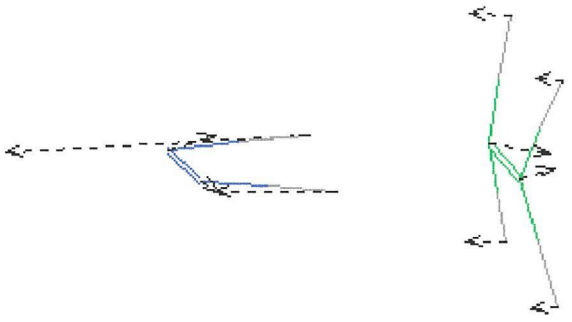


Table 10.27 PM3 Thermodynamic data for the diimide reduction of various alkenes.

	ΔH_f^{298} / kcal mol^{-1}		
	Starting materials†	Transition structure	Products†
Ethene	59.4	85.6	-0.7
(Z)-But-2-ene	40.2	70.3	-11.7
(E)-But-2-ene	40.2	69.5	-11.7

† Calculated by summation.

Coxon and McDonald[969] investigated the Diels-Alder cycloaddition of (*Z*)-diimide to butadiene. The PES for the *exo* approach reveals one valley connecting reactants and products corresponding to the forming σ bonds, each being the same length as the reaction proceeds. The transition structure is symmetrical and corresponds to a synchronous reaction process. In contrast, the PES for the *endo* approach reveals two valleys for concerted addition, related by the symmetry of the system. The bond length of each forming σ bond is markedly different (0.82 Å) at the transition structure, *i.e.* formation of the two C–N bonds is not simultaneous.

Rzepa[970] used the AM1 method to examine pericyclic reactions where nitrogen atoms are involved in bonds which are formed or broken during the course of the reaction, which lead to the suggestion that the M_2 parameter reported as 2.1 Å for nitrogen in the core-core repulsion function was incorrect. The effect of this parameter was proposed to over-estimate the internuclear repulsion between nitrogen and other atoms when separated by distances > 1.9 Å, such as those found in transition states, favouring asynchronous reaction pathways where the internuclear interactions can be minimised relative to those of synchronous ones. Agrafiotis and Rzepa[967] have suggested that reducing the value of the M_2 parameter to about 1.6 Å will result in the AM1 method producing results more consistent with *ab initio* calculations. In PM3, the c_2 parameter (the equivalent parameter to M_2) has a value of 1.716 Å, and the results in terms of reported average errors in calculated heats of formation for nitrogen containing compounds, show a significant improvement over those for AM1.

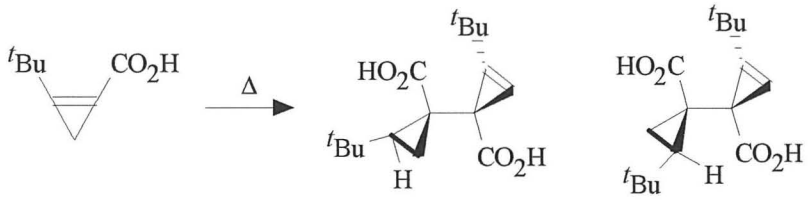
Initial studies of the diimide reduction have proved promising, although the PM3 method predicts a synchronous addition of hydrogen for cyclopropenes and a stepwise addition of hydrogen for alkenes. Further studies will prove invaluable in probing the desaturation enzyme, whose active site contains Fe–S and Fe–O complexes.

10.8 ENE REACTIONS

Cyclopropene is known to polymerise at room temperature,[971] while substituted cyclopropenes are generally more stable and dimerise at higher temperatures.[972] The products of some of these reactions are consistent with that of an Alder-ene reaction, which have been reviewed by Baird.[973] Radical mechanisms were dismissed after

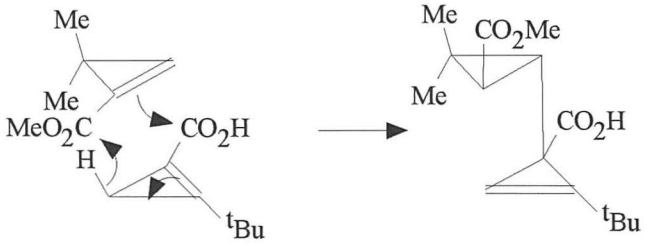
observing that the addition of radical traps such as di-*t*-butyl nitroxide and peroxide failed to reduce dimerisation rates. 1,3-Diphenylcyclopropene undergoes a highly regioselective ene reaction[972d] with a low activation energy (40.7 kJ mol^{-1} at 25 °C), large negative $\Delta S^{\ddagger}$ (-137.3 J mol^{-1} K^{-1} at 25 °C), a large k_H/k_D value for migrating hydrogen (3.1 at -30 and -40 °C), similar to k_H/k_D values in other ene reactions.[972e]

Baird *et al* studied the dimerisation of 2-*t*-butylcyclopropenecarboxylic acid **(319)**[973] at 20 °C.[972f] NMR spectroscopy and TLC showed complete reaction of the starting material and the formation of two products in the ratio 3:2. The major compound was shown by X-ray crystallography to be **(320)**, derived through an *endo* transition structure, while the minor compound **(321)** was derived *via* an *exo* transition structure. Baird noted that the *t*-butyl groups would interact unfavourably in the *exo* transition structure, causing the *endo* transition structure to be preferred.[972f] Baird *et al* also studied the ene reaction of **(319)** with the methyl ester of 3,3-dimethylcyclopropenecarboxylic acid **(322)**,[974] which reacted as the enophile. The *exo* product **(323)** is formed exclusively, perhaps due to the steric repulsion between a methyl group of the enophile and the ene in the *endo* transition structure.



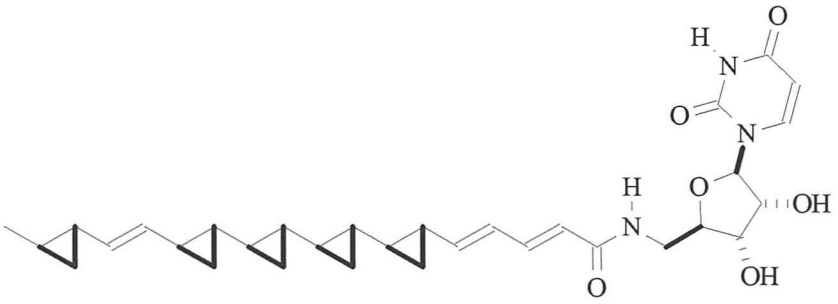
(319) **(320)** **(321)**

(322)

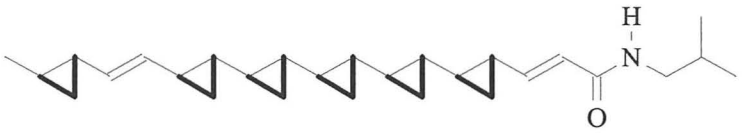


(319) **(323)**

10.8.1 Possible synthetic uses of the cyclopropene ene dimerisation



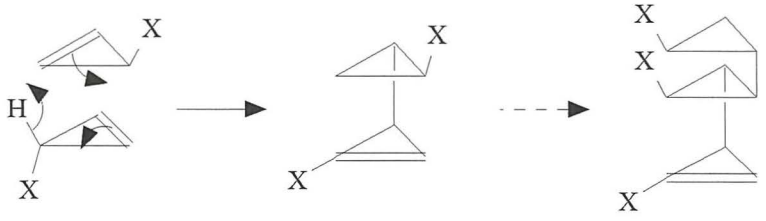
(–)-FR-900848 **(324)**



(–)-U-106305 **(325)**

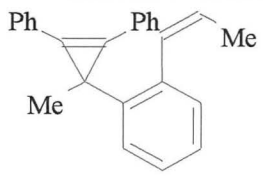
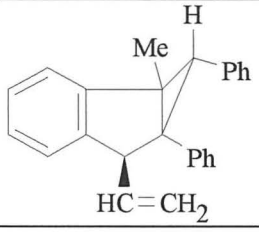
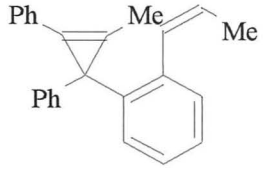
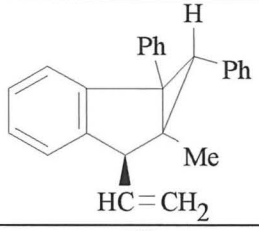
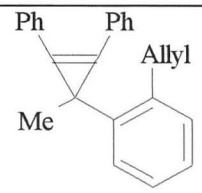
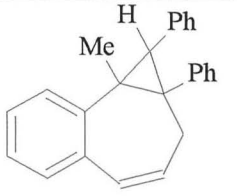
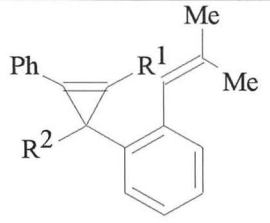
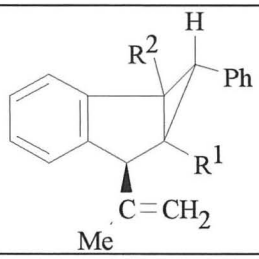
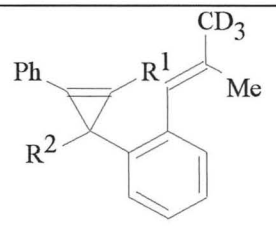
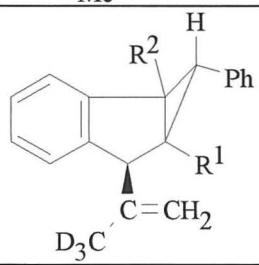
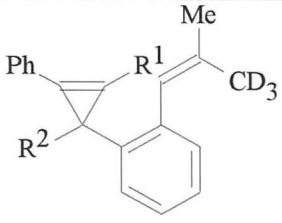
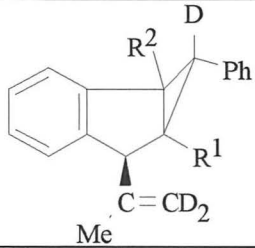
(–)-FR-900848[975] **(324)** is a nucleoside isolated from the fermentation broth of *Streptoverticillium fervens* and shows potent selective activity against filamentous fungi such as *Aspergillus niger*, but is essentially inactive against non-filamentous fungi such as *Candida albicans* and Gram-positive and Gram-negative bacteria. (–)-U-106305[976] **(325)** isolated from the fermentation broth of *Streptomyces sp. UC 11136*, is a potent *in vitro* inhibitor of the cholesteryl ester transfer protein (CETP) reaction, preventing arteriosclerosis.[977] Contiguous oligo-cyclopropanes have been synthesised using Simmons-Smith methodologies,[978,979] *via* the trapping of homo-allyl cationic intermediates,[980] and by palladium catalysed Suzuki-type cross-coupling.[981]

The ene reaction could be used to prepare contiguous oligo-cyclopropane compounds, as intermolecular examples are routinely used in synthesis.[982] As the stacking of the cyclopropene-cyclopropane oligomers is controlled in the transition structure, a particular mode of stacking, *i.e.* oligomer stereochemistry, could be forced by the use of bulky substituents such as -^{t}Bu or -$SiMe_3$.



Intramolecular ene reactions of 3-(*o*-alkenylphenyl)-substituted cyclopropenes are regioselective, with hydrogen transfer to the carbon bearing the phenyl group, in accordance with frontier orbital theory.

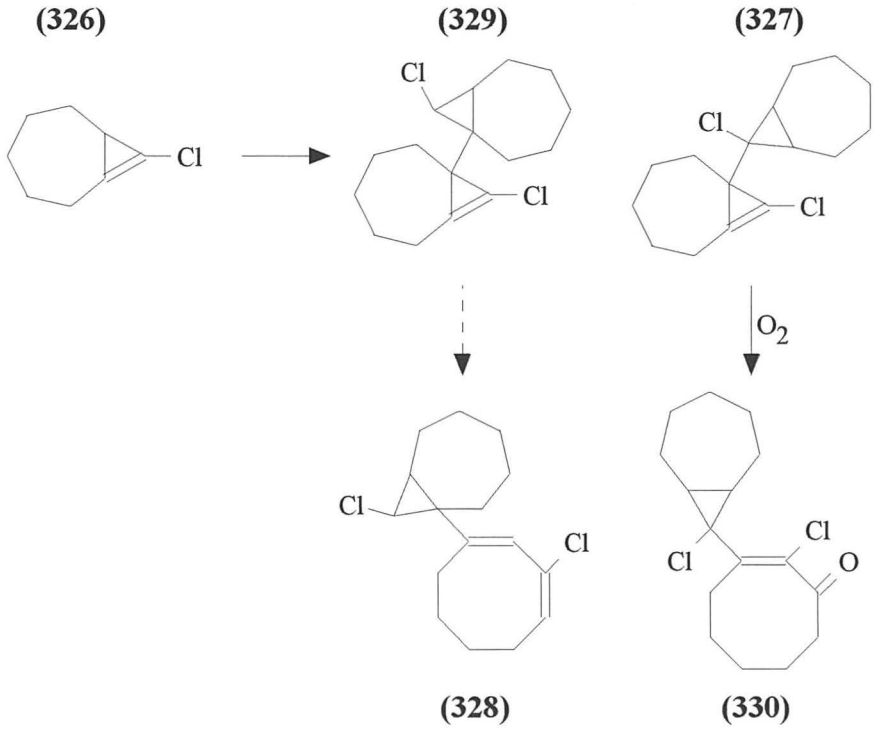
Table 10.28 Cyclopropene thermolyses.

Starting material	Product	Reference
		983
		983
		984
		984
		984
		984

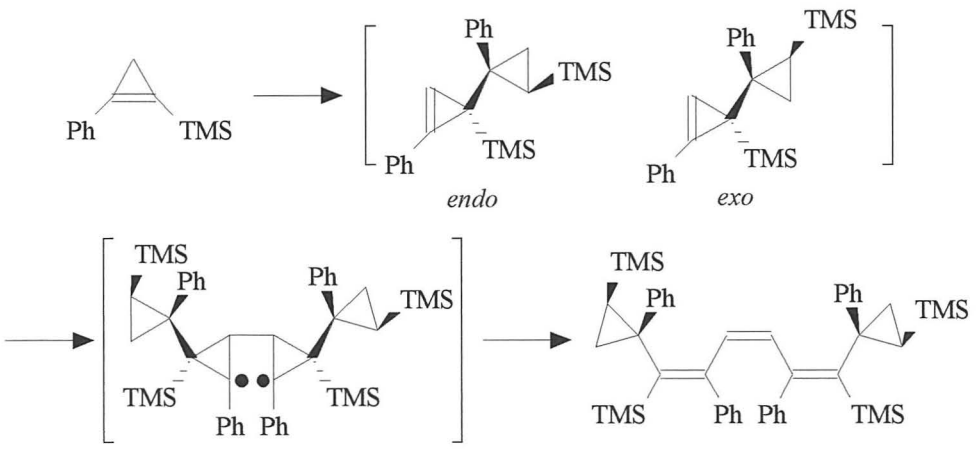
† Thermolysis of the (*E*)-isomer produced a 4.6:1 mixture of *exo* and *endo*-benzotricycloheptenes *via* a [2+2] cycloaddition.[983]

‡ Thermolysis of the (*E*)-isomer produced a 3:1 mixture of *exo* and *endo*-benzotricycloheptenes *via* a [2+2] cycloaddition.[984]

Although cyclopropane-cyclopropene dimers are the main products of cyclopropene ene-dimerisations, other competing processes, such as an initial [2+2] cycloaddition, ring opening of the ene dimer, or further dimerisation can occur, leading to the formation of several by-products. Under vacuum, 8-chlorobicyclo[5.1.0]oct-1(8)-ene **(326)** produces **(327)** and **(328)** in 65 and 2.1% yields respectively.[985] The minor compound **(328)** is formed from the ene dimer **(329)** by ring opening to give a vinyl carbene, followed by intramolecular carbene insertion to generate a cyclic allene, which then isomerises to afford **(328)**.[985] Cyclopropene-vinyl carbene rearrangement of **(327)** followed by oxidation with molecular oxygen yields α,β-unsaturated compound **(330)**.[985]



Padwa reported that two readily interconvertible 1,3,5-hexatrienes were isolated by heating 1-phenyl-2-carbomethoxy-3,3-dimethylcyclopropane.[986] Billups and Wiberg both claimed that only one 1,3,5-hexatriene was formed when 6-(bicyclo[4.1.0]hept-1-yl)bicyclo[4.1.0]hept-1(7)-ene underwent a coupling dimerisation reaction, although the configuration of the central double bond of the triene dimer was not known.[987] The *endo* and *exo* ene dimers of 1-trimethylsilyl-2-phenylcyclopropene couple to yield a tetramer **(331)** whose structure was verified by X-ray crystallography.[988]



(331)

The principal aim of this section is to predict the stereochemistry and regiochemistry of cyclopropane-cyclopropene dimers from the relative energies of the *endo* and *exo* transition structures. Although transition structures of cyclopropene and halocyclopropenes can be easily calculated, multiple maxima are obtained when larger or asymmetric non-spherical substituents are used due to incorrect alignment of the substituents. This difficulty will need to be addressed before a more detailed study of substituent effects in the transition structures can take place. As literature *endo*/*exo* ratios or activation energies are not available for the below reactions, quantitative comparison with the semi-empirical results cannot be made.

10.8.2 Computational results

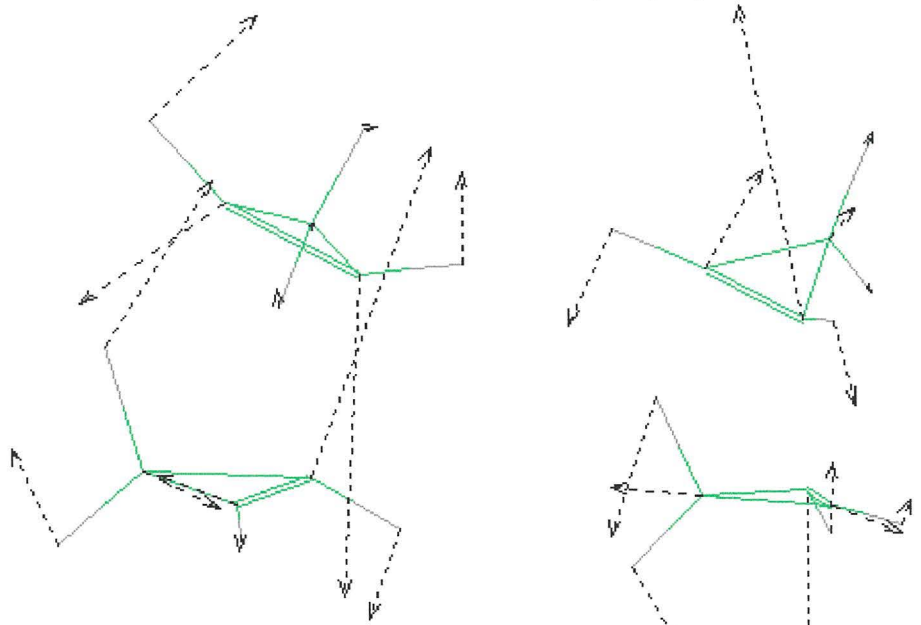
Transition structures of simple ene reactions involving alkenes have been investigated.[989,990] Houk *et al* found that the *endo* transition structure for the cyclopropene dimerisation is 2.7 kcal mol^{-1} lower in energy than the *exo* transition structure at the Becke3LYP/6-31G* + ZPE level of theory.[991,992] Secondary orbital overlap of a C–H bond of the enophile with the π-system at the central carbon of the ene is proposed to account for the preference of the *endo* transition structure. Barely stable diradical intermediates were found for both *endo* and *exo* cyclopropene dimerisations. Both transition structures are very asynchronous with forming CC bond lengths of 1.926 and 1.939 Å in the *exo* and *endo* transition structures respectively. The breaking CH bond lengths are 1.133 and 1.118 Å, while the forming CH bond lengths are 2.035 and 2.201 Å respectively. Becke3LYP/6-31G* + ZPE activation energies of 15.4 and 18.1 kcal mol^{-1} were obtained for the *endo* and *exo* transition structures.

PM3 Transition structures for the *endo* and *exo* ene reactions of cyclopropene were obtained using *MOPAC* as described earlier, employing frequency analysis to verify the existence of true transition structures and minima. Minimisation of supermolecular assemblies such as two cyclopropene molecules produce local rotational minima with small negative eigenvalues. Heats of formation were obtained by summation of the individual molecules (136.2 kcal mol^{-1}) rather than by explicit evaluation (135.4 kcal mol^{-1}), as the intermolecular interactions are quite small. *Endo* and *exo* activation energies were calculated to be 28.7 kcal mol^{-1} and 29.4 kcal mol^{-1} respectively, verifying that dimerisation of cyclopropene is more likely to proceed *via* the *endo* transition structure, although the relative *endo* specificity is predicted to be smaller than that predicted by Houk. As the activation energies are larger than those obtained by Houk, the transition structures should not be as asynchronous. The PM3 *exo* and *endo* transition structures are markedly different from those obtained by Houk, with forming CC bond lengths of 1.91 and 1.92 Å, breaking CH bond lengths of 1.19 and 1.18 Å, and forming CH bond lengths of 1.69 and 1.72 Å.

Table 10.29 PM3 Thermodynamic data for the ene reaction of cyclopropene.

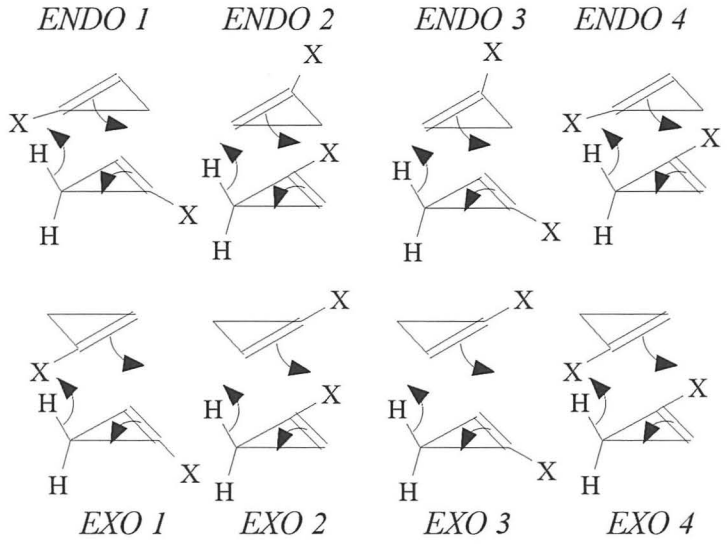
	Endo SM	*Exo* SM	*Endo* TS	*Exo* TS	Dimer
ΔH_f^{298} / kcal mol^{-1}	136.2	136.2	164.9	165.6	87.5

Figure 10.9 PM3 Transition structures for the cyclopropene ene dimerisation.



exo *endo*

Figure 10.10 Coding scheme for the dimerisation of 1-halocyclopropenes.



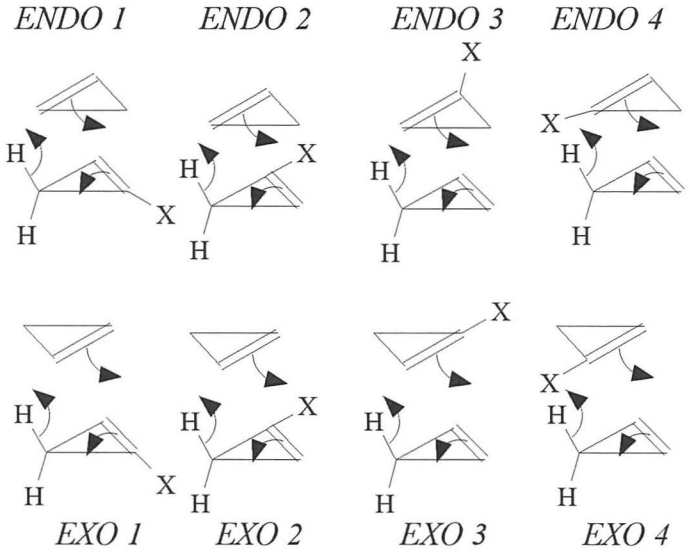
Endo and *exo* transition structures for the dimerisation of 1-halocyclopropenes, chosen due to their spherical substituents, were calculated using the PM3 method, based on the above coding scheme. Transition structures *ENDO2* (25.9 kcal mol^{-1}), *EXO1* (26.1 kcal mol^{-1}), *ENDO2* (20.1 kcal mol^{-1}), and *EXO1* (28.4 kcal mol^{-1}) had the smallest barriers for X = F, Cl, Br and I respectively. Greater specificity between the various transition states is obtained for the larger halide substituents – *EXO4* (σ = 2.3 kcal mol^{-1}), *EXO2* (σ = 7.7 kcal mol^{-1}), *ENDO1* (σ = 7.7 kcal mol^{-1}), and *EXO2* (σ = 7.4 kcal mol^{-1}) for X = F, Cl, Br and I respectively.

Table 10.30 PM3 Thermodynamic data for the ene reaction of 1-halocyclopropenes.

	ΔH_f^{298} / kcal mol^{-1}			
	F	Cl	Br	I
SM*	58.0	124.3	167.4	193.7
ENDO 1	85.0	150.5	195.2	224.6
ENDO 2	**83.9**	157.2	**187.5**	227.1
ENDO 3	85.1	155.0	192.3	223.7
ENDO 4	84.0	153.3	192.9	224.0
EXO 1	85.2	**150.4**	193.7	**222.1**
EXO 2	85.6	158.1	188.4	229.5
EXO 3	84.4	153.6	193.2	224.3
EXO 4	86.2	156.2	193.1	225.6

Transition states for the ene reactions of cyclopropene with 1-halocyclopropene were also calculated according to:

Figure 10.11 Coding scheme for the reaction of cyclopropene with various 1-halocyclopropenes.



The transition structures with the lowest heats of formation were found to be *ENDO2* (26.6 kcal mol^{-1}), *EXO1* together with *ENDO4* (27.9 kcal mol^{-1}), *ENDO2* (25.7 kcal mol^{-1}) and *ENDO1* (28.4 kcal mol^{-1}) for X = F, Cl, Br, and I respectively. All of the above lowest energy transition structures have cyclopropene acting as the enophile and the halocyclopropene component acting as the ene. The spread in activation energies are 2.4 kcal mol^{-1}, 4.0 kcal mol^{-1}, 2.8 kcal mol^{-1} and 3.7 kcal mol^{-1} for X = F, Cl, Br, and I respectively, with *EXO2*, *EXO2*, *EXO1*, and *EXO2* having the highest energies.

Table 10.31 PM3 Transition states for the ene reaction of cyclopropene with various 1-halocyclopropenes.

	ΔH_f^{298} / kcal mol^{-1}			
	F	Cl	Br	I
SM*	97.1	130.2	151.8	164.9
ENDO 1	124.8	159.1	178.4	**193.3**
ENDO 2	**123.7**	160.1	**177.5**	194.3
ENDO 3	124.4	160.6	178.0	194.5
ENDO 4	125.0	**158.1**	180.1	194.7
EXO 1	125.3	**158.1**	180.3	193.9
EXO 2	126.1	162.1	179.5	197.0
EXO 3	124.4	161.0	178.3	195.8
EXO 4	125.2	159.7	178.6	193.6

Transition structures were also calculated for other 1-substituted cyclopropenes (see *Figure 10.10*), where X = -Me, -OMe, and -CN. The lowest energy transition structures were found to be *ENDO1* (28.8 kcal mol^{-1}), *EXO1* (26.7 kcal mol^{-1}) and *ENDO1* (25.4 kcal mol^{-1}) for X = Me, OMe, and CN respectively, although *EXO1* may be a lower energy transition structure for X = CN. As the above substituents are not completely spherical, multiple maxima were produced due to incorrect substituent orientation, necessitating the use of other starting conformations. Other systems, *e.g.* X = -C≡CH, $-NO_2$, -^{t}Bu, $-SiMe_3$, $-CH_2OH$, and $-CH{=}CH_2$, (*R*)- and (*S*)-1,3-dimethylcyclopropene either did not converge, or produced multiple maxima. This serious problem will need to be addressed if transition structures are to be calculated for 3-cyclopropylcyclopropene or 1-trimethylsilyl-2-(hydroxymethyl)cyclopropene.[993]

Table 10.32 PM3 Transition states for the ene reaction of 1-substituted cyclopropenes.

	ΔH_f^{298} / kcal mol^{-1}		
	Me	OMe	CN
SM*	115.0	64.3	208.1
ENDO 1	**143.8**	91.3	**233.5**
ENDO 2	151.1	97.6	242.2
ENDO 3	147.4	94.0	237.4
ENDO 4	146.0	91.7	237.1
EXO 1	144.0	**91.0**	†
EXO 2	150.7	97.6	241.3
EXO 3	146.1	91.6	237.9
EXO 4	148.3	95.1	238.2

† Multiple maxima

10.9 CYCLOADDITION REACTIONS

10.9.1 2+2 cycloadditions

The $[2_s+2_s]$ cycloaddition is not a thermally allowed process, while the allowed $[2_a+2_s]$ process requires a geometrically difficult transition state.[994] Multiple maxima were encountered for the PM3 transition structure of the [2+2] dimerisation of cyclopropene, a process known to compete with the thermal ene reaction.

10.9.2 4+2 Diels-Alder cycloadditions

The parent Diels-Alder reaction of 1,3-butadiene with ethene to form cyclohexene has been studied extensively by both semiempirical[854] and *ab initio* methods.[855-857] *Ab initio* RHF and MCSCF calculations predict a symmetric transition structure for a route with no intermediates, while semi-empirical or *ab initio* UHF calculations predict a two-step mechanism involving a diradical intermediate. While the concerted, synchronous nature of the parent is now agreed upon, mechanisms of substituted cases are still debated.[995]

Cyclic alkenes and alkynes with pronounced angular strain are powerful dienophiles. With the exception[996] of molecules containing bulky non-electron-withdrawing substituents at C(3), cyclopropenes undergo Diels-Alder cycloadditions to give bicyclo[4.1.0]hept-3-enes or other bridged derivatives.[997] Strained alkenes possess a higher energy HOMO and a lower energy LUMO, relative to ethene, and possess the potential for accelerated participation in both normal (HOMO diene controlled) and inverse electron demand (LUMO diene controlled) Diels-Alder reactions.[998] Almost all additions to cyclopentadiene give[999] *endo* adducts exclusively, whereas addition to furan may be stereospecifically *exo*.[1000] 3,3-Difluorobis(trifluoromethyl)cyclopropene however gives[1001] *exo* and *endo* adducts in a ratio of 31:69. 3-Chlorocyclopropene with 1,3-diphenylisobenzofuran (DPIBF) produces[1002] *exo* and *endo* adducts in a ratio of 4.5:1, whereas 3,3-dichlorocyclopropene with a more electropositive C(3) atom, gives the *exo* adduct stereospecifically. An *exo* adduct is also obtained from tetracyclone and 3-chlorocyclopropene.[1002] Many reactive cyclopropenes can been trapped with DPIBF.[1003]

The initial *endo* adduct of simple cyclopropenes with thiopene-1-dioxides or cyclopentadieneones, loses[1004] SO_2 or CO cheletropically with concomitant opening of the three-membered ring,[1005] to produce cycloheptatrienes. Several adducts derived from the addition of cyclopropenes to cyclopentadienes cycloadd intramolecularly upon heating to afford homoquadricyclane derivatives.[1006] Aromatisation of the bridged adduct is also known.[1007] Various tris- and tetra-azines afford products[1008] which readily lose N_2 to give azanorcaradienes, which produce bis adducts upon further reaction with cyclopropenes.

10.9.3 Computational results

PM3 Transition structures for the Diels-Alder reactions of various cyclopropenes with dienes and furans were calculated as described above. The *exo* and *endo* transition structures have forming C---C bond lengths of 2.19 and 2.20 Å. The stability of the *exo* over *endo* transition structures gradually decreases with increasing halogen size - the *endo* transition structure for 1-iodocyclopropene being more stable than the *exo* transition structure. The stability of the *exo* transition structure over the *endo* transition structure is marked (16.2 kcal mol^{-1}) for 3,3-dimethylcyclopropene, due to steric interactions between the methyl groups of 3,3-dimethylcyclopropene and C(2) and C(3) of the butadiene molecule in the *endo* transition structure.

Figure 10.12 PM3 Transition structures for the Diels-Alder reaction of butadiene with cyclopropene.

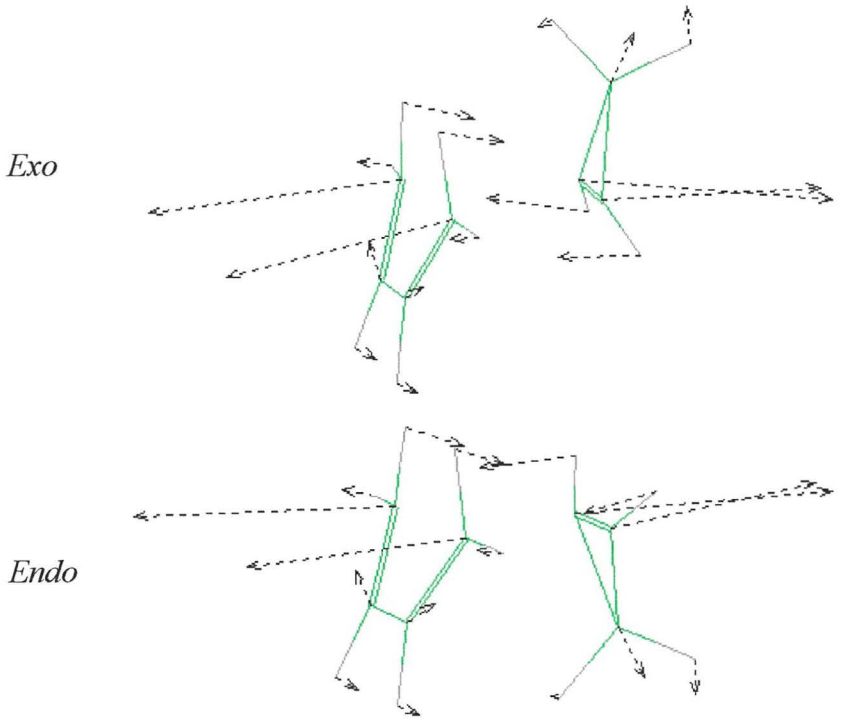


Table 10.33 Thermodynamic data for the PM3 transition structures of the Diels-Alder reactions of butadiene with various cyclopropenes.

	ΔH_f^{298} / kcal mol^{-1}							
	1-H	1-F	1-Cl	1-Br	1-I	1-Me	1,2-Me	3,3-Me
SM†	99.7	60.6	93.8	115.3	128.5	89.1	78.4	83.7
ENDO	122.4	82.6	118.4	136.7	153.7	114.8	107.1	122.9
EXO	122.1	82.2	118.2	136.8	155.0	114.4	106.7	106.7

† Obtained by summation.

Starting material heats of formation identical to those for butadiene are obtained for cyclopentadiene. The steric effect of the methylene group is seen in activation energies which are mostly 7 kcal mol^{-1} higher than their butadiene counterparts. *Endo* transition structures are ~ 1 kcal mol^{-1} less stable than *exo* transition structures. Similar activation energies to AM1 values are obtained, although PM3 *endo*/*exo* specificities are smaller.[1009] The energies of *endo* and *exo* transition structures for 3,3-dimethylcyclopropene and cyclopentadiene are similar, due to steric interactions between 3,3-dimethylcyclopropene and cyclopentadiene.

Figure 10.13 PM3 Transition structures for the Diels-Alder reaction of cyclopentadiene with cyclopropene.

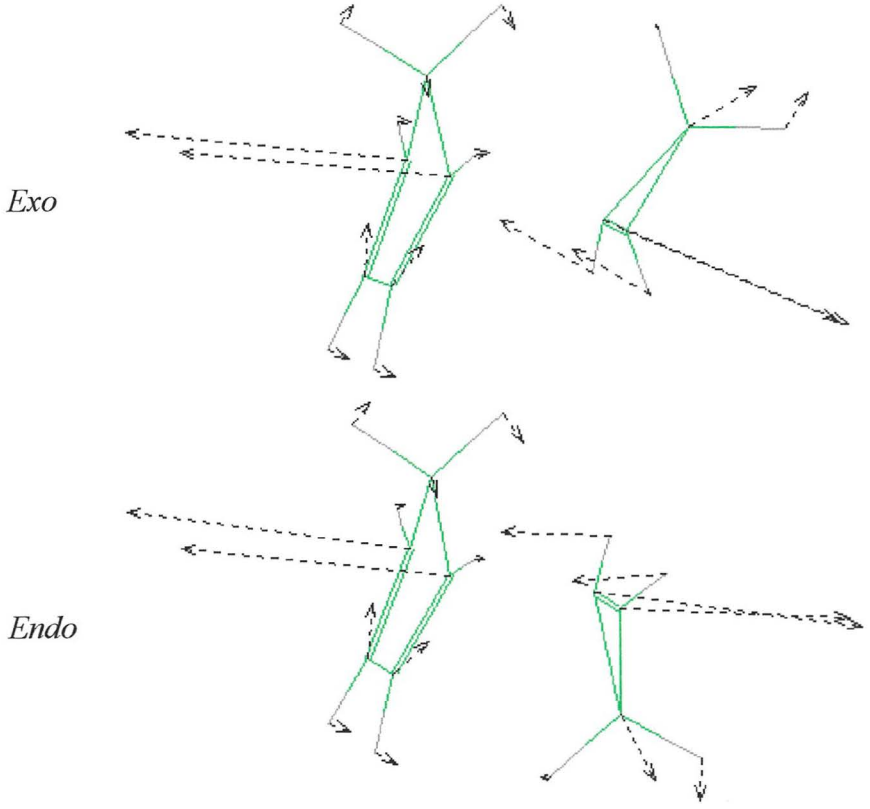


Table 10.34 Thermodynamic data for the PM3 transition structures of cyclopentadiene with various cyclopropenes.

	ΔH_f^{298} / kcal mol^{-1}							
	1-H	1-F	1-Cl	1-Br	1-I	1-Me	1,2-Me	3,3-Me
SM†	99.7	60.6	93.8	115.3	128.5	89.1	78.4	83.7
ENDO	129.4	88.9	124.9	143.6	162.3	121.4	113.3	120.9
EXO	130.4	89.3	125.8	144.9	164.1	122.3	114.4	121.0

† Obtained by summation.

While *endo* adducts are formed by reaction with cyclopentadiene,[999] *exo* adducts can be formed by addition to furan.[1000] The *endo* transition structures are favoured for cyclopropene and 1-halocyclopropenes, with greater specificity for larger halogen substituents. The *exo* transition structure for 3,3-dimethylcyclopropene and furan is ~ 2 kcal mol^{-1} more stable than that for the *endo* transition structure. The *exo* transition structure for 3,3-dimethylcyclopropene was ~ 16, 0, and 2 kcal mol^{-1} more stable than the *endo* transition structure for additions to butadiene, cyclopentadiene and furan respectively, reflecting steric interactions in the transition structures. Again, higher activation energies are obtained compared with those of butadiene.

Figure 10.14 PM3 Transition structures for the Diels-Alder reaction of furan with cyclopropene.

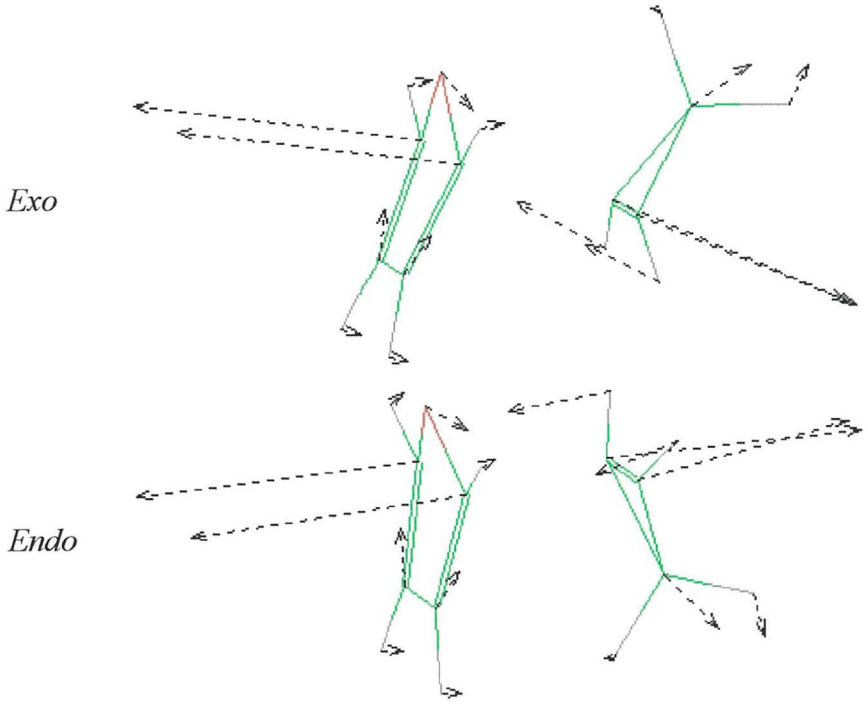


Table 10.35 Thermodynamic data for the PM3 transition structures of the Diels-Alder reactions of furan with various cyclopropenes.

	ΔH_f^{298} / kcal mol^{-1}							
	1-H	1-F	1-Cl	1-Br	1-I	1-Me	1,2-Me	3,3-Me
SM†	63.9	24.8	58.0	79.5	92.7	53.3	42.6	47.9
ENDO	94.0	53.0	89.2	108.0	126.7	85.8	77.6	85.3
EXO	94.3	53.2	89.6	108.9	128.1	86.1	77.9	83.4

† Obtained by summation.

Table 10.36 PM3 Thermodynamic data for Diels-Alder reactions of cyclopropene with various dienes and furans.

	ΔH_f^{298} / kcal mol^{-1}			
	Isoprene	Chloroprene	Dichloroprene	1,2-Dimethylfuran
SM†	90.3	92.5	87.0	46.4
ENDO	113.1	115.4	109.7	77.7
EXO	112.6	114.6	108.7	78.1

† Obtained by summation.

The Diels-Alder reactions of cyclopropene and various dienes and furans are shown above. The *exo* transition structures are more stable for the substituted butadienes, with greater *exo*/*endo* specificity with increasing substitution. The *endo* transition structure is more stable for the addition of cyclopropene to 1,2-dimethylfuran, in accordance with results obtained for the parent.

10.10 CONCLUSIONS

Transition structures for various pericyclic reactions of cyclopropenes have been calculated with the PM3 method, which correctly reproduces physical properties of cyclopropenes and cyclopropanes. Activation energies similar to either experimental or *ab initio* values are obtained for many sigmatropic shifts, with substantial reductions in computational times. Several problems have been encountered in this initial study, and will need to be resolved in further studies.

As literature *endo*/*exo* ratios or activation energies were not available for the majority of the above reactions, quantitative comparison with our results could not be made. Other important thermodynamic quantities such as entropies and Gibbs free energies of activation have not been discussed, although further studies of these thermodynamic quantities are envisaged. Heats of formation of supermolecular assemblies were obtained by summation of the individual components, rather than by explicit evaluation, due to the presence of several small negative force constants. This approximation is only valid if intermolecular interactions between the molecules in the assembly are small, and seems to be valid for ene reactions, with only a 1 kcal mol^{-1} difference between the two approaches. Problems were also encountered for non-spherical groups, *e.g.* -$SiMe_3$, -CH_2OH *etc.*, where incorrect substituent alignment caused the transition structures to be multiple maxima, *i.e.* they possessed

several negative force constants instead of one. Several projects, *e.g.* the stacking of cyclopropene oligomers could therefore not be carried out to their conclusion.

Some studies of diimide reductions of cyclopropenes have been presented as initial models for the desaturase enzyme. Heats of formation using the PM3 method were obtained which were similar to the AM1 values of Rzepa. The PM3 method predicts a synchronous symmetrical transition structure for the reduction of cyclopropene, while an asynchronous transition structure was obtained for the reduction of ethene. This may be due to incorrect nitrogen parameterisation. A further study with other cyclopropenes and reducing agents is envisaged, together with a study of concerted hydrogen elimination.

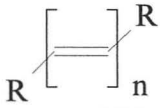
Although *ab initio* calculations of the transition structure of the ene dimerisation of cyclopropene have been published, the semi-empirical approach in this chapter has allowed the study of substituted cyclopropenes. The dimerisation of cyclopropene is more likely to proceed *via* the *endo* transition structure, although the relative *endo* specificity was smaller than that predicted by Houk. For ene reactions of 1-halocyclopropenes with cyclopropene, cyclopropene acts as the enophile, while 1-halocyclopropene acts as the ene component.

Initial studies of Diels-Alder reactions of cyclopropenes with butadiene, cyclopropene and furan have also been presented. *Exo* and *endo* transition structures of cyclopropenes with butadiene and furan are generally identical, although greater *exo* specificities are seen for 3,3-dimethylcyclopropene. Endo transition structures are preferred for the Diels-Alder reactions with cyclopentadiene, due to steric repulsions of the cyclopropene with the methylene groups of cyclopentadiene.

Secondary orbital overlap of a CH bond of the enophile with the π-system at the central carbon of the ene has been proposed to account for the *endo* transition structure.[991,992] Secondary orbital interactions have been investigated using Mulliken overlap populations and Mulliken bond orders,[1010] and could be carried out using the PM3 method. Further studies employing *ab initio* (HF, DFT, HF/DFT) calculations are also envisaged, together with a study of the dipolar reaction of cyclopropenes.

APPENDIX 1 – TRAINING SET COMPOUNDS

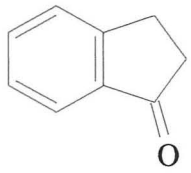
4)	Indigo	11)	Methylene blue
12)	Acid fuchsin	13)	Light green SF yellowish
25)	Orange G	31)	Methyl green
44)	Sudan III	45)	Sudan IV
54)	Bismarck brown Y	55)	Crystal violet
56)	Safranin O	68)	Methyl orange
70)	Biebrich scarlet	71)	Ponceau de xylidene
81)	Fast green FCF	85)	Congo red
95)	Brilliant green	138)	*p*-Aminobenzoic acid
139)	Octyl *p*-MeOcinnamate	140)	Oxybenzone
141)	Lawsone	142)	Padimate O
143)	Methyl anthranilate	144)	Phenyl salicylate
145)	Octyl salicylate	146)	Uranine (Eosin Y X = H)
147)	Eosin Y X = F	148)	Eosin Y X = Cl
149)	Small light green model	157)	Acridine yellow G
158)	Alizarin red S	159)	Auromine O
160)	Azure A	161)	Azure B
162)	Azure C	163)	Brilliant cresyl blue
164)	Cresyl violet acetate	165)	Guinea green B
166)	Malachite green oxalate	167)	Neutral red
168)	Orange II	169)	Pyronin Y
170)	Sudan black B	171)	Thionin
172)	toluidine blue O		



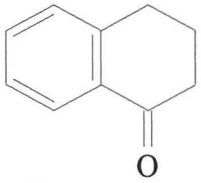
175)	n = 1, R = Me	176)	n = 2, R = H
177)	n = 2, R = Me	178)	n = 3, R = H
179)	n = 3, R = Me	180)	n = 4, R = H
181)	n = 4, R = Me	182)	n = 5, R = H
183)	n = 5, R = Me	184)	n = 6, R = Me
185)	n = 7, R = Me	186)	n = 8, R = Me

187)	Anthraquinone (An)	188)	1-OH-An
189)	1-NH_2-An	190)	1-NHMe-An
191)	1-HO,4-NH_2-An	192)	1,4-diNH_2-An
193)	1,5-diNH_2-An	194)	1,4,5,8-Tetraaminoanthraquinone
195)	Cyclopentadiene	196)	Cyclohexadiene
185)	Furan	198)	Pyrrole
199)	=COMe	200)	=CHO (acrolein)
202)	Me_2=MeCHO	203)	cyclopentenone
204)	=CO_2H	205)	*E*-Me_2=CO_2H
206)	*Z*-Me_2=CO_2H	207)	Me2=CO_2H
208)	Me_3=CO_2H	211)	*E*-Me=$CONH_2$
212)	Z-Me=$CONH_2$	213)	benzene
214)	phenol	215)	styrene
216)	styrene-*E*-Me	219)	*p*-xylene
220)	PhEt	221)	PhMe
222)	PhiPr	223)	PhF
224)	PhCl	225)	PhBr
226)	$PhSO_3H$	227)	PhOMe
228)	Piperylene	229)	PhCN
231)	$PhNH_2$	232)	$PhCO_2H$

234) Ph-alkyne
235) PhCHO
236) Ph-COMe
237) $PhNMe_2$
238) PhOPh
239) $PhNO_2$
240) Acridone
241) LG core
242) Haematein
339) Cyclohexenone
342) Cyclopentenone-αCl
344) Furan-β-CO_2H
345) Cyclohexeneone-αCl



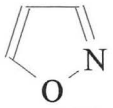
346)



351)

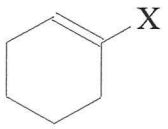
349) furan-α-CO_2Me
350) γ-pyrone
354) α-pyrone
360) pyrazole
361) imidazole
362) triazole
363) tetrazole

364)

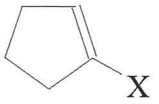


365)

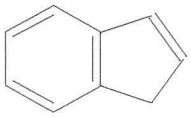
366) pyridine
367) quinoline
368) pyridazine
369) pyrazine
370) pyrimidine
371) pyridine-2Me
372) pyridine-3Me
373) pyridine-4Me
374) pyridine-2Et
375) pyridine-2^{i}Pr
376) pyridine-2F
377) pyridine-2Cl
378) pyridine-2Br
379) purine
380) *p*-CN-toluene
381) *m*-CN-toluene
382) *o*-CN-toluene
383) *p*-Cl-benzoic acid
384) *m*-Cl-benzoic acid
385) *o*-Cl-benzoic acid
386) *p*-Cl-nitrobenzene
387) *m*-Cl-nitrobenzene
388) *o*-Cl-nitrobenzene
389) *p*-nitrotoluene
390) *m*-nitrotoluene
391) *o*-nitrotoluene
392) *p*-nitrophenol
393) *m*-nitrophenol
394) *o*-nitrophenol
395) *p*-nitroaniline
396) *m*-nitroaniline
397) *o*-nitroaniline
398) *p*-nitrobenzoic acid
399) *p*-dinitrobenzene
400) *m*-dinitrobenzne
401) *m*-aminobenzoic acid
402) *o*-aminobenzoic acid
403) *p*-COMe-aniline
404) p-CN-aniline
405) *m*-CN-aniline
406) p-OH-benzaldehyde
407) *m*-OH-benzaldehyde
408) o-OH-benzaldehyde
409) *p*-OH-benzoic acid
410) m-OH-benzoic acid
411) *o*-OH-benzoic acid
412) *p*-OH-acetophenone
413) *m*-OH-acetophenone
414) *o*-OH-acetophenone



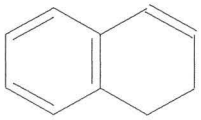
201) X = COMe
210) X = CO_2H
332) X = $CONH_2$



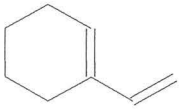
209) X = CO_2H
334) X = $CONH_2$



229)

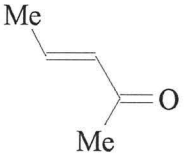


218)



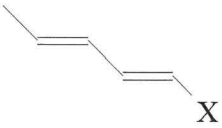
230) het
431) dichloro-het

233) het-Cl



333) ketone
336) ketone-αOCOMe
338) ketone-βOCOMe
341) ketone-βCl
347) ketone-βOH
352) ketone-αOH
355) ketone-βSH
357) ketone-βNH_2
359) ketone-βNMe_2

335) ketone-αOCOH
337) ketone-βOCOH
340) ketone-βMe
343) ketone-βCl
348) ketone-βOH
353) ketone-αOMe
356) ketone-βSMe
358) ketone-βNHMe
415) ketone-conjugated =



416) ene-conjugated =
418) eneOMe
420) eneOCOMe
422) eneSH
424) $eneNH_2$
426) $eneNMe_2$

417) eneCl
419) eneOH
421) eneOCOH
423) eneSMe
425) eneNHMe

427) homo
429) homo-α-Cl

428) homo-conj=
430) homo-β-Cl

APPENDIX 2 – SEMI-EMPIRICAL WAVELENGTHS / NM

	λ / nm	CNDO	INDO	MINDO/3	MNDO	AM1	PM3	ZINDO/S
(4)	602[528]	126.5	131.5	184.4	188.6	191.1	209.4	197.2
(11)	661[528]	141.6	148.1	204.2	208.4	209.3	217.5	237.5
(12)	545[528]			202.0	219.8	211.5	241	227
(13)	630(422)[528]			219.9	211.4	203.9		
(25)	475[528]			201.9	198.4	194.4	210.4	216.7
(31)	629(423)[528]					253.5	300.4	231.5
(33)	544[528]			187.5	192.6	193.4	205.4	194.2
(44)	507(354)[528]	144.1	145.0	203.8	223.4	216.8	215.7	245.8
(45)	520(357)[528]	140.4	146.1	204.3	224.0	217.4	216.4	246.9
(54)	457[528]	123.4	125.9	183.3	195.5	191.8	196.2	213.2
(55)	588[528]	130.1	142.3	202.1	196.0	213.9	244.3	201.5
(56)	530[528]			182.3	183.7	184.7	190.9	206.6
(68)	505[528]	140.7	151.5	190.9	198.1	190.9	217.0	211.2
(85)	497[528]			200.1	212.7	209.7	215.3	242.0
(95)	625(428)[528]	138.2	147.4	210.1	200.7	214.2	238.8	209.4
(105)	291[528]	86.3	87.7	136.1	143.2	140.7	140.8	140.8
(106)	489[528]	118.2	119.5	172.5	177.8	178.3	184.5	186.7
(115)	275[1017]	95.1	95.1	149.4	157.2	153.7		
(138)	284[526]	90.1	91.7	139.1	144.6	143.1	149.8	149.7
(139)	311[526]	102.8	103.4	150.0	157.4	155.5		
(140)	325[526]	94.9	96.9	141.2	149.5	147.7		
(141)	275,335[526]	110.3	114.9	154.9	153.2	151.9		
(142)	310[526]	90.5	93.8	142.5	137.1	144.3		
(143)	248,341[526]	92.2	94.6	145.7	151.4	150.2		
(144)	300[526]	87.6	89.5	144.0	147.1	144.8		
(145)	310[526]	96.0	97.2	134.0	143.8	140.4		
(146)	517[528]			184.4	203.4	200.7		
(147)	517[528]			193.1	204.4	202.4		
(148)	517[528]			197.6	208.0	208.4		
(157)	442[528]	118.0	119.5	172.3	178.7	177.2	183.5	185.3
(158)	556(596)[528]		195.9	194.9	185.3	181.8	204.4	211.4
(159)	434[528]	105.2	106.2	156.1	157.8	158.7	166.8	166.8
(160)	633[528]	136.7	145.8	200.7	206.1	205.9	214.4	232.6
(161)	648[528]	138.4	148.0	202.4	207.3	207.6		
(162)	616[528]	133.5	143.0	197.4	203.7	201.5		
(163)	622[528]	137.2	147.0	198.2	196.6	206.7	224.1	230.7
(164)	596[528]	132.3	137.5	189.1	198.7	196.6	202.2	230.7
(165)	618[528]			255.6	234.8	226.4		
(166)	614(425)[528]	133.9	144.5	204.7	197.5	216.4	246.7	204.5
(167)	540[528]	128.4	133.1	188.1	186.9	189.7	203.7	226.4
(168)	483[528]		254.2	204.2	204.4	198.8	226.4	220.2
(169)	548[528]	130.3	133.3	185.9	189.3	194.7	206.2	205.2
(170)	598(415)[528]					242.5	210.3	277.9
(171)	598[528]	128.7	141.7	193.8	201.4	197.5	206.9	223.9
(172)	626[528]	138.1	143.1	202.3	206.4	206.9	215.2	235.0
(175)	174[590]	71.7	71.6	112.1	115.2	114.3		
(177)	217[337]	80.6	80.6	129.5	136.8	133.8	136.1	140.1
(177)	227[590]	91.7	91.6	137.8	142.0	141.0		
(178)	267.5[352]	100.0	110.0	152.0	158.2	156.2		
(179)	274.5[348]	109.3	109.3	158.1	162.3	161.8		
(180)	304[352]	116.9	116.9	170.5	174.8	173.9		
(181)	310[348]	125.1	125.0	174.3	178.0	178.2		
(182)	334[353]	131.8	131.8	183.4	188.0	187.8		
(183)	342[348]	139.0	139.0	187.6	190.6	191.5		
(184)	380[348]	151.6	151.6	198.3	200.5	202.4		
(185)	401[348]	162.8	162.8	207.3	208.4	211.4		
(186)	411[348]	173.1	173.0	214.6	215.1	218.9		

	λ / nm	CNDO	INDO	MINDO/3	MNDO	AM1	PM3	ZINDO/S
(187)	327[591]	112.4	117.2	159.9	156.9	141.3	149.1	162.2
(188)	405[591]	113.4	118.2	161.2	164.7	156.9	158.9	170.5
(189)	465[591]	114.8	119.2	172.3	180.5	174.6	183.7	187.0
(190)	508[591]	115.4	119.8	175.2	181.8	177.8	186.7	191.5
(191)	520[591]	118.5	123.7	179.3	186.8	181.2	189.5	194.3
(192)	550[591]	127.9		190.0	200.2	195.2	207.0	207.1
(193)	480[591]	116.7	121.1	176.6	183.0	177.7	187.1	192.0
(194)	610[591]			203.0	210.3	206.4	220.8	220.1
(195)	239[337]	86.2	86.9	133.4	137.1	134.1	134.9	140.8
(197)	200[592]	74.1	73.9	122.2	117.4	115.0	116.3	128.7
(198)	210[593]	71.2	73.6	118.0	117.4	114.8	113.0	124.9
(200)	207[358]	79.6	82.6	121.1	119.5	120.5	122.5	130.3
(203)	218[595]	82.7	86.6	127.7	122.3	120.7	123.2	132.4
(204)	200[359]	77.3	80.4	116.5	118.0	114.6	117.6	125.4
(207)	216[359]	84.6	84.7	120.7	124.4	122.9	125.9	131.4
(209)	222[359]	84.9	84.5	120.8	124.1	124.7	126.1	133.0
(210)	217[359]	83.9	84.3	121.7	124.0	123.1	125.6	131.8
(211)	213[596]	82.4	85.5	122.1	121	119.4		
(213)	184[363]	69.3	69.3	117.8	125.9	121.6		
(214)	211[361]	76.2	77.4	126.5	134	130.5	131.2	129.3
(215)	248[597]	87.5	87.5	139.4	147	143.5	145.3	149.1
(217)	262[598]	94.5	94.3	144.0	149.6	147.4	148.9	153.7
(218)	249[598]	91.4	91.9	142.4	148.0	144.5	145.7	150.2
(219)	193[363]	79.0	79.0	125.7	131.1	128.7	129.6	129.8
(223)	210[653]	73.6	74.1	129.4	129.9	127.3		
(224)	210[653]	76.0	78.0	129.8	129.8	127.1		
(226)	213[653]	90.8	109.8	156.7	136.1	130.5		
(227)	217[361]	76.5	77.8	126.5	133.7	130.7		
(228)	219[590]	86.3	86.3	133.8	139.4	137.5	139.7	144.5
(229)	224[361]	81.5	81.5	130	134.1	131.1		
(230)	229[590]	88.4	88.2	135.3	140.2	138.7	132.4	139.2
(231)	230[361]	81.7	83.4	135.2	141.1	139.0		
(233)	234[590]	92.4	90.6	139.5	143.6	143.3		
(234)	236[653]	82.5	82.5	134	141.0	137.1		
(235)	242[650]	84.4	87	127.8	134.8	130.9		
(235)	445[528]	149.6	152.6	195.9	201.2	200.7	205.6	219.8
(237)	251[651]	83.7	85.1	142.7	140.8	140.6		
(238)	255[652]	81.3	83.0	138.0	144.1	140.5		
(239)	269[361]	88.7	89.7	128	137.2	132.7		
(240)	380[528]	102.8	104.5	154.2	161.7	156.2	163.1	163.8
(241)	630(422)[528]			236.6	257.7	251.9		
(325)	219[598]	81.3	84.9	125.6	119.9	119.5	121.0	133.1
(332)	213[1011]	83.8	86.9	124.8	123.3	123.5		
(334)	220[1011]	84.7	86.3	123.6	123.4	124.0		
(335)	221[590]	84.1	87.7	128.0	130.7	128.5	130.6	137.5
(336)	221[590]	83.8	86.6	125.8	129.9	127.9	130.7	136.9
(337)	221[590]	83.8	85.0	125.3	128.9	126.8	128.7	134.8
(338)	221[590]	83.7	85.2	124.9	128.6	126.4	129.1	134.7
(339)	227[590]	85.8	89.0	129.3	123.1	122.3	123.6	133.7
(340)	237[590]	83.1	87.1	127.9	123.3	121.7	124.1	133.1
(341)	227[590]	83.4	86.4	128.2	125.2	124.5	128.1	
(342)	229[590]	86.1	86.7	133.1	126.7	127.6		
(343)	230[590]	82.3	85.3	129	124.5	124.6	133.2	
(344)	238[1012]	81.3	831.3	122.8	123.5	121.5		
(345)	242[590]	86.8	88.2	132.2	127.5	128.4	135.6	
(346)	243[1013]	87.3	90.7	134.3	136.1	133.6	135.0	138.8
(347)	245[590]	82.2	84.6	123.1	128.0	125.5	127.4	132.8
(348)	245[590]	83.1	84.7	123.2	128.7	126.8	128.6	134.1
(349)	245[1014]	87.3	87.0	128	125.9	124.1		

	λ / nm	CNDO	INDO	MINDO/3	MNDO	AM1	PM3	ZINDO/S
(350)	245[1015]	89.6	90.1	130.8	131.2	128.5		
(351)	248[1013]	89.8	92.5	135.5	136.9	134.1	135.4	140.1
(352)	250[590]	81.4	84.3	125.3	132.2	129.8	131.8	135.0
(353)	250[590]	83.2	86.1	127.4	132.8	131.8	133.4	137.6
(354)	300[1016]	98.4	98.4	136.1	143.8	141.9		
(355)	300[590]	87.6	90.1	134.3	135.4	142.5	145.9	134.8
(356)	300[590]	89.4	95.0	139.5	136.9	147.2	149.0	136.9
(357)	310[590]	85.8	87.6	129.4	133.0	131.8	139.0	138.7
(358)	310[590]	88.1	90.0	132.5	134.2	134.2	141.2	141.8
(359)	310[590]	89.4	91.2	137.8	134.8	135.7	142.4	143.3
(360)	210[1017]	75.0	74.8	118.6	117.1	114.9		
(361)	207[1017]	74.4	74.2	122.0	119.6	116.9		
(362)	210[1017]	74.3	74.5	118.2	117.4	114.8		
(363)	260[1017]	71.0	73.3	122.3	107.0	106.1		
(364)	210[1017]	85.6	89.6	129.9	131.9	134.4		
(365)	210[1017]	75.4	75.1	115.7	116.6	114.4		
(366)	256[1017]	75.4	79.4	128.2	126.7	122.5		
(367)	311[1017]	92.0	95.1	145.5	150.4	147.1		
(368)	250[1017]	78.9	86.9	135.3	118.6	119		
(369)	258[1017]	78.5	84.6	138.3	127.5	128.1		
(370)	244[1017]	76.0	80.5	132.1	120.3	121.8		
(371)	262[1017]	77.4	81.8	130.6	128.7	125.4		
(372)	263[1017]	77.4	81.4	130.6	128.8	125.6		
(373)	255[1017]	77.2	82.1	129.7	126.6	122.4		
(374)	261.5[1017]	77.9	82.2	131.1	128.8	125.7		
(375)	262[1017]	77.8	81.8	130.9	128.5	124.8		
(376)	257[1017]	75.8	79.1	133.4	129.9	127.1		
(377)	263[1017]	77.4	83.4	137.9	130.1	127.4		
(379)	263[1018]	84.5	89.4	141.3	129.8	128.1		
(380)	237[361,362]	85.5	85.4	133.3	136.5	134.6		
(381)	229.5[361,362]	82.7	82.6	130.8	134.9	133.0		
(382)	228.5[361,362]	83.6	83.5	131.4	135.3	133.2		
(383)	241[361,362]	86.9	85.6	130.5	136.2	134.2		
(384)	231.5[361,362]	85.3	85.6	132.9	136.9	135.1		
(385)	229[361,362]	85.8	85.3	131.8	137.6	135.4		
(386)	280[361,362]	86.9	87.4	131.5	137.4	135.8		
(387)	264[361,362]	87.1	88.5	134.8	138.8	137.2		
(388)	260[361,362]	87.2	92.1	133.5	139.2	137.3		
(389)	285[361,362]	86.7	87.3	129.4	137.4	134.0		
(390)	273[361,362]	86.5	86.9	130.7	138.2	135.9		
(391)	266[361,362]	86.5	87.4	131.1	138.7	135.8		
(392)	317.5[361,362]	85.7	86.1	131.8	141.3	138.0		
(393)	273.5[361,362]	86.8	87.9	138.1	146.4	142.9		
(394)	278.5[361,362]	85.5	87.5	135.8	146.7	143.1		
(395)	381[361,362]	88.9	92.6	141.7	148.0	146.4		
(396)	280[361,362]	92.2	96.2	149.9	156.7	154.9		
(397)	282.5[361,362]	91.3	95.0	146.6	154.1	152.6		
(398)	264.5[361,362]	91.8	92.6	133.2	142.5	137.8		
(399)	266[361,362]	92.3	92.9	134.2	142.4	137.6		
(400)	241.5[361,362]	89.3	90.2	130.2	135.7	131.5		
(401)	250[361,362]	92.0	94.3	147.0	153.1	151.5		
(402)	248[361,362]	91.3	93.5	144.3	150.5	149.3		
(403)	311.5[361,362]	92.0	94.2	147.2	146.1	144.6		
(404)	270[361,362]	89.3	90.8	141.2	145.4	144.2		
(405)	236.5[361,362]	88.6	90.9	142.9	150.5	149.2		
(406)	283.5[361,362]	87.7	88.4	132.5	140.1	137.4		
(407)	254.4[361,362]	86.8	88.5	137.8	143.8	141.1		
(408)	256[361,362]	87.0	88.9	135.7	144.2	141.5		
(409)	255[361,362]	86.6	87.3	130.3	139.1	136.0		

	λ / nm	CNDO	INDO	MINDO/3	MNDO	AM1	PM3	ZINDO/S
(410)	236.5[361,362]	86.5	87.7	136.8	144.8	141.4		
(411)	237[361,362]	85.9	87.1	134.1	143.8	140.6		
(412)	275[361,362]	88.0	89.2	133.2	140.4	137.4		
(413)	250.5[361,362]	87.2	89.6	138.4	144.1	140.9		
(414)	252.5[361,362]	88.1	91.2	136.7	144.9	141.8		
(415)	245[590]	95.8	95.9	140.3	146.0	143.4	146.5	156.3
(416)	244[590]	104.8	104.7	155.6	160.2	159		
(417)	219[590]	91.0	88.8	138.3	143	142.4	146.0	
(418)	220[590]	91.7	92.6	138.8	145.6	144.5	146.1	151.4
(419)	220[590]	91.3	92.0	138.1	145.3	143.9	145.6	151.0
(420)	214[590]	92.9	93.6	139.7	146.8	145.8	147.3	152.9
(421)	214[590]	92.4	93.1	139.0	146.4	145.5	147.0	152.8
(422)	244[590]	93.6	96.6	146.4	147.0	152.1	154.2	150.4
(423)	244[590]	88.7	92.8	142.4	140.5	148.5	156.3	151.1
(424)	274[590]	94.6	96.0	145	149.1	148.7	154.7	155.6
(425)	274[590]	96.3	97.6	146.9	149.8	150.1	156.0	157.1
(426)	274[590]	97.0	98.3	150.0	149.9	150.8	156.5	157.9
(427)	253[590]	91.9	91.7	136.7	140.3	139.5	141.1	147.7
(428)	288[590]	100.1	99.9	147.1	151.1	150.1		
(429)	258[590]	95	93.6	141.0	143.6	143.6		
(430)	258[590]	93.8	93.2	139.2	141.6	141.3		
(431)	234[590]	92.4	90.6	139.5	143.6	143.3		

APPENDIX 3 – DFT WAVELENGTH MAXIMA / NM

Compound	λ / nm	MIN	DN	DND	DNP
(115)	275[1017]	380.7	374.1	383.9	387.8
(138)	284[526]	368.1	366.9	359.7	363.4
(176)	217[337]	328.1	326.7	330.9	336.8
(187)	327[591]	652.9	585.0	570.3	568.9
(188E)	405[591]	735.0	657.6	608.4	610.0
(188Z)	405[591]	666.2	588.0	573.9	574.6
(189)	465[591]	716.5	705.4	702.1	701.0
(190E)	508[591]	785.7	776.3	750.7	750.7
(190Z)	508[591]	847.0	809.4	791.1	788.4
(191E)	520[591]	886.6	781.6	792.5	787.0
(191Z)	520[591]	1012.0	868.0	858.2	868.0
(192)	550[591]	1087.6	977.9	988.5	980.0
(193)	480[591]	672.1	695.7	690.4	691.5
(194)	610[591]	1079.8	1077.3	1082.4	1072.2
(195)	239[337]	333.8	343.4	351.3	355.7
(197)	200[592]	257.2	265.6	273.2	278.0
(198)	210[593]	242.8	250.9	254.6	261.0
(198)	210[593]	242.8	250.9	254.6	261.0
(200)	207[358]	515.5	445.4	433.2	434.8
(203)	214[595]	459.8	412.8	409.1	411.6
(204)	200[359]	377.5	311.3	315.8	318.4
(207)	216[359]	347.3	304.4	306.7	308.3
(209)	222[359]	350.5	313.2	308.5	310.8
(210)	217[359]	339.1	306.0	304.2	307.1
(211)	213[596]	443.7	363.1	356.6	360.2
(213)	184[363]	237.2	235.1	240	243.9
(214)	211[361]	307.4	283.9	289.1	294.6
(215)	248[597]	311.1	316.5	324.1	329.0
(217)	262[598]	323.2	330.5	339.1	342.9
(218)	249[598]	311.9	316.2	323.2	326.7
(219)	193[363]	254.4	263.3	270.1	276.3
(220)	200[653]	247.9	252.0	257.9	262.5
(221)	207[361]	245.9	249.3	255	260.1
(222)	210[653]	243.8	247.1	252.9	257.5
(223)	210[653]	273.9	254.6	260.7	264.9
(224)	210[653]	305.0	263.7	268.4	272.9
(225)	210[653]	333.4	271.2	274.3	279.1
(226)	213[653]	830.0	366.3	383.6	385.5
(227)	217[361]	313.8	288.2	292.3	298.2
(228)	219[590]	337.5	342.1	346.8	352.2
(229)	224[361]	282.7	278.5	287.1	289.1
(230)	229[590]	337.8	343.1	347.3	353.5
(231)	230[361]	332.1	326.9	334.3	341.9
(232)	228[382]	415.0	317.8	323.0	325.5
(233)	234[590]	387.5	369.0	376.9	381.7
(234)	236[653]	300.8	301.0	309.2	329.0
(235)	242[650]	521.4	436.9	429.1	430.3
(235)	243[368]	511.4	436.1	429.1	431.1
(237)	251[651]	367.8	353.5	356.0	364.8
(238)	255[652]	282.3	266.5	269.2	274.0
(239)	269[361]	617.5	447.6	429.1	430.3
(240)	380[528]	509.2	462.6	474.7	477.7
(332)	213[1011]	349.2	297.6	298.6	302.8
(334)	220[1011]	405.1	335.3	329.5	332.6
(335)	221[590]	429.1	383.3	337.5	421.9
(336)	221[590]	457.1	416.2	410.5	412.8
(337)	221[590]	529.9	456.1	431.5	432.3

Compound	λ / nm	MIN	DN	DND	DNP
(338)	221[590]	517.2	450.3	429.5	430.7
(339)	227[590]	476.7	433.2	427.5	429.5
(340)	227[590]	483.7	431.9	423.1	424.7
(341)	227[590]	485.3	438.6	429.5	431.5
(342)	229[590]	443.3	406.1	403.6	406.9
(343)	230[590]	456.6	425.5	418.4	421.9
(344)	215[598]	495.9	441.1	431.5	434.4
(344)	238[1012]	345.7	331.9	339.3	342.1
(345)	242[590]	450.7	423.1	419.2	421.5
(346)	243[1012]	498.6	433.2	436.5	438.6
(347)	245[590]	450.7	405.1	388.8	391.8
(348)	245[590]	455.7	409.1	393.5	396.9
(349)	245[1014]	346.5	331.9	343.1	342.1
(350)	245[1015]	447.6	399.0	389.1	391.5
(351)	248[1013]	499.7	438.2	435.2	436.9
(352E)	250[590]	428.3	389.1	383.6	386.5
(352Z)	250[590]	464.5	406.5	392.2	392.2
(353)	250[590]	519.6	442.8	434.0	436.9
(354)	300[1016]	409.4	393.9	400.1	404.0
(355)	300[590]	463.1	425.9	416.5	419.2
(356)	300[590]	463.6	425.9	419.2	421.2
(357)	310[590]	431.1	387.2	376.6	381.7
(358)	310[590]	433.6	387.8	380.4	385.2
(360)	210[1017]	270.4	239.3	242.4	245.7
(361)	207[1017]	262.5	255.1	260.1	264.5
(362)	210[1017]	312.5	283.2	255.9	257.7
(363)	260[1017]	313.0	285.3	257.0	258.6
(364)	210[1017]	335.8	299.8	277.2	278.7
(365)	210[1017]	311.7	280.1	267.0	269.2
(366)	256[1017]	418.4	355.5	324.3	328.3
(367)	311[1017]	572.5	449.8	405.4	409.1
(368)	250[1017]	736.2	578.3	478.7	483.2
(369)	258[1017]	573.9	465.0	409.8	411.6
(370)	244[1017]	475.7	393.9	355.5	358.2
(371)	262[1017]	405.4	346.0	317.6	322.3
(372)	263[1017]	406.9	356.8	327.8	332.1
(373)	255[1017]	420.0	350.8	320.7	324.6
(374)	261.5[1017]	397.3	350.0	322.0	326.0
(375)	262[1017]	407.6	347.6	320.2	324.6
(376)	257[1017]	357.7	314.3	290.4	293.2
(377)	263[1017]	371.7	323.9	301.8	304.4
(378)	265[1017]	385.5	327.4	304.8	307.5
(379)	263[1017]	543.1	439.9	391.8	395.2
(380)	237[1017]	289.7	288	296.1	297.8
(381)	229.5[1017]	294.0	287.7	295.5	297.8
(382)	228.5[1017]	295.5	289.5	298.0	299.6
(383)	241[1017]	414.3	326.4	329.0	330.5
(384)	231.5[1017]	417.7	340.1	335.5	338.3
(385)	229[1017]	473.7	362.8	364.3	366.6
(386)	280[1017]	612.5	451.6	432.8	433.2
(387)	264[1017]	618.3	453.0	434.8	436.1
(388)	260[1017]	682.2	497.5	472.7	473.7
(389)	285[1017]	604.4	422.0	423.5	424.3
(390)	273[1017]	609.2	445.0	427.9	428.7
(391)	266[1017]	617.5	451.2	432.8	433.6
(392)	317.5[1017]	581.2	434.8	413.9	414.6
(393)	273.5[1017]	604.4	456.1	442.0	443.3
(394)	278.5[1017]	493.7	483.2	461.7	462.6
(395)	381[1017]	552.4	449.4	431.9	434.0

Compound	λ / nm	MIN	DN	DND	DNP
(396)	280[1017]	588.7	606.0	574.6	573.2
(397)	282.5[1017]	509.7	540.6	524.4	524.4
(398)	264.5[1017]	701.1	491.6	472.7	472.7
(399)	266[1017]	809.4	550.4	526.8	526.8
(400)	241.5[1017]	696.8	481.2	460.3	461.2
(401)	250[1017]	425.1	463.1	451.6	454.8
(402)	248[1017]	414.6	436.5	427.5	429.1
(403)	311.5[1017]	458.0	399.7	391.5	395.2
(404)	270[1017]	343.1	330.0	338.3	344.4
(405)	236.5[1017]	427.1	393.9	402.2	404.3
(406)	283.5[1017]	495.3	421.2	410.5	411.3
(407)	254.4[1017]	518.4	441.6	431.9	433.2
(408)	256[1017]	463.6	428.3	425.1	426.3
(409)	255[1017]	389.8	331.2	326.7	329.0
(410)	236.5[1017]	412.0	372.3	367.2	370.2
(411)	237[1017]	403.6	389.7	381.3	383.6
(412)	275[1017]	482.2	419.2	408.7	410.5
(413)	250.5[1017]	506.9	440.3	431.1	434.4
(414)	252.5[1017]	446.8	413.1	410.2	410.9

BIBLIOGRAPHIC REFERENCES AND FOOTNOTES

1) J. McLeod and J. Osborn in *Natural Automata and Useful Simulations*, H.H. Pattee, E.A. Edelsack, L. Fein and A.B. Callahan, editors, Macmillan, London, 1966, p.128.

2a) C.G. Broyden, *Journal of the Institute for Mathematics and Applications*, 1970, **6**, 222.

2b) R. Fletcher, *Computer Journal*, 1970, **13**, 317.

2c) D. Goldfarb, *Mathematics of Computation*, 1970, **24**, 23.

2d) D.F. Shanno, *Mathematics of Computation*, 1970, **24**, 647.

2e) D.F. Shanno, *Journal of Optimisation Theory and Applications*, 1985, **46**, 87.

3a) N.L. Allinger, *J. Am. Chem. Soc.*, 1977, **99**, 8127.

3b) N.L. Allinger and Y.H. Yuh, QCPE #365a.

3c) U. Berkert and N.L. Allinger, *Molecular Mechanics*, ACS Monograph 177, 1982.

4a) MM+ is derived from the public domain codes MM2(1977) and MM2(1991), disributed by the QCPE.

4b) For applications of MM+ to peptides, see:
J.H. Lii, S. Gallion, C. Bender, H. Wikstrom, N.L. Allinger, K.M. Flurchick and M.M. Teeter, *J. Comput. Chem.*, 1989, **10**, 503.

4c) For applications of MM+ to other systems see:
K.B. Lipkowitz, *QCPE Bulletin*, Indiana University, February 1992, **12**, 1.

5a) S.J. Weiner, P.A. Kollman, D.A. Case, U.C. Singh, C. Ghio, G. Alagona, S. Profeta, Jr. and P. Weiner, *J. Am. Chem. Soc.*, 1984, **106**, 765.

5b) S.J. Weiner, P.A. Kollman, D.T. Nguyen and D.A. Case, *J. Comput. Chem.*, 1986, **7**, 230.

6a) B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187.

6b) W. Reiher, *Ph.D. Thesis*, Harvard, 1985.

6c) J.C. Smith and M. Karplus, *J. Am. Chem. Soc.*, 1992, **114**, 805.

7) BIO+ is an implementation of the CHARMM force field.

8a) For applications of OPLS to proteins, see:
W.L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1988, **110**, 1657.

8b) For applications of OPLS to nucleotide bases, see:
J. Pranata, S. Wierschke and W.L. Jorgensen, *J. Am. Chem. Soc.*, 1991, **113**, 2810.

9) S. Lifson, A.T. Hagler and P. Dauber, *J. Chem. Soc.*, 1979, **101**, 5111.

10) N.L. Allinger, Y.H. Yuh and J.H. Lii, *J. Am. Chem. Soc.*, 1989, **111**, 8551.

11a) S. Lifson and A. Warshel, *J. Chem. Phys.*, 1968, **49**, 5116.

11b) A. T. Hagler, E. Huler and S. Lifson, *J. Am. Chem. Soc.*, 1977, **96**, 5319.

11c) A.T. Hagler and S. Lifson, *J. Am. Chem. Soc.*, 1974, **96**, 5327.

11d) J.R. Maple, U. Dinur and A.T. Hagler, *Proc. Natl. Acad. Sci. U.S.A.*, 1988, **85**, 5350.

12) J. Pranata and W.L. Jorgensen, *J. Am. Chem. Soc.*, 1991, **113**, 9483.

13) E. Schrödinger, *Ann. Physik*, 1926, **79**, 361, 489; **80**, 437; **81**, 109.

14) W. Pauli, *Z. Physik*, 1925, **31**, 765.

15) J.C. Slater, *Phys. Rev.*, 1929, **34**, 1293; 1930, **35**, 509.

16) J.C. Slater, *Phys. Rev.*, 1930, **36**, 57.

17) I.N. Levine, *Quantum Chemistry*, 4th edition, Prentice-Hall, Englewood Cliffs, NJ, 1991, *Chapter 8*.

18) C.C.J. Roothan, *Rev. Mod. Phys.*, 1951, **23**, 69.

19) G.G. Hall, *Proc. R. Soc. London, Ser. A*, 1951, **205**, 541.

20) S.M. Blinder, *Am. J. Phys.*, 1965, **33**, 431.

21) M. Born and J.R. Oppenheimer, *Ann. Physik*, 1927, **84**, 457.

22) J. Goodisman, *Diatomic Interaction Potential Theory*, Academic Press, New York, NY, 1973, **Volume 1**, *Chapter 1*.

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

24) I. Shavitt in *Methods in Computational Physics*, Wiley, New York, NY, 1962, **Volume 2**, p. 1.
25) S. Huzinaga, *J. Chem. Phys.*, 1965, **42**, 1293.
26) T.H. Dunning, *J. Chem. Phys.*, 1970, **53**, 2823.
27) T.H. Dunning and P.J. Hay in *Methods of Electronic Structure Theory*, H.F. Schaefer, editor, **Volume 3** of *Modern Theoretical Chemistry*, W. Miller *et al*, editors, Plenum, New York, NY, 1977, pp. 1-27.
28) W.J. Hehre, L. Radom, P. von R. Schleyer and J.A. Pople, *Ab Initio Molecular Theory*, Wiley, New York, NY, 1986, *Section 4.3*.
29a) M. Dupuis, D. Spangler and J.J. Wendoloski, *GAMESS*, National Resource for Computations in Chemistry, Software Catalogue, University of California, Berkeley, CA, 1980, *Program QG01*.
29b) M.W. Schmidt, K.K. Baldridge, J.A. Boatz, J.H. Jensen, S. Koseki, M.S. Gordon, K.A. Nguyen, T.L. Windus and S.T. Elbert, *QCPE Bulletin*, 1990, **10**, 52.
29c) M.F. Guest, *GAMESS-UK*, Compuational Science Group, SERC, Daresbury Laboratory, Warrington.
30) R.D. Amos and J.E. Rice, *CADPAC 4.0*, University of Cambridge, 1988.
31a) W. Hehre, R. Ditchfield, L. Radom and J.A. Pople, *J. Am. Chem. Soc.*, 1970, **92**, 4796.
31b) J.S. Binkley, R.A. Whiteside, K. Raghavachari, R. Seeger, D.J. DeFrees, H.B. Schlegel, M.J. Frisch, J.A. Pople and L.R. Kahn, *GAUSSIAN 82*, Carnegie-Mellon University, Pittsburgh, PA, 1982.
31c) M.J. Frisch, J.S. Binkley, H.B. Schlegel, K. Raghavachari, C.F. Melius, R.L. Martin, J.J.P. Stewart, F.W. Bobrowicz, C.M. Rohlfing, L.R. Kahn, D.J. DeFrees, R. Seeger, R.A. Whiteside, D.J. Fox, E.M. Fleuder and J.A. Pople, *GAUSSIAN 86*, Carnegie-Mellon Publishing Unit, Pittsburgh, PA, 1986.
31d) M.J. Frisch, M. Head-Gordon, H.B. Schlegel, K. Raghavachari, J.S. Binkley, C. Gonzalez, D.J. DeFrees, D.J. Fox, R.A. Whiteside, R. Seeger, C.F. Melius, J. Baker, R.L. Martin, L.R. Kahn, J.J.P. Stewart, E.M. Fluder, S. Topiol and J.A. Pople, *GAUSSIAN 88*, Gaussian Inc., Pittsburgh, PA, 1988.
31e) The Carnegie-Mellon Quantum Chemistry Archive, Carnegie-Mellon University, Pittsburgh, PA 15213.
31f) M.J. Frisch, M. Head-Gordon, G.W. Trucks, J.B. Foresman, H.B. Schlegel, K. Raghavachari, M.A. Robb., J.S. Binkley, C. Gonzalez, D.J. DeFrees, D.J. Fox, R.A. Whiteside, R. Seeger, C.F. Melius, J. Baker, R.L. Martin, L.R. Kahn, J.J.P. Stewart, S. Topio and J.A. Pople, *GAUSSIAN 90*, Gaussian Inc., Pittsburgh, PA, 1990.
31g) Latest version *GAUSSIAN 94* available from Gaussian, Inc., of Pittsburgh, PA.
32) H. Kuhn, *J. Chem. Phys.*, 1948, **16**, 840; 1949, **17**, 1198.
33) N.S. Bayliss, *Quart. Rev.*, 1952, **6**, 319.
34) J. R. Platt, K. Ruedenberg, C.W. Sherr, N.S. Ham, H. Labhart and W. Lichten, *Free-Electron Theory of Conjugated Molecules*, Wiley, New York, NY, 1964.
35) W.T. Simpson, *J. Chem. Phys.*, 1948, **16**, 1124.
36) C.A. Coulson and A. Streitwieser Jr., *Dictionary of π-Electron Calculations*, W.H. Freeman, San Fransisco, CA, 1965.
37) A. Streitwieser Jr., and J.I. Brauman, *Supplemental Tables of Molecular Orbital Calculations*, Pergamon, Elmsford, NY, 1965.
38) E. Heilbronner and H. Straub, *Hückel Molecular Orbitals*, Springer-Verlag, New York, NY, 1966.
39) R.G. Parr, *Quantum Theory of Molecular Electronic Structure*, Benjamin, New York, NY, 1963, *Chapter III*.
40) L. Salem, *The Molecular Orbital Theory of Conjugated Systems*, Benjamin, New York, NY, 1966.
41) F.L. Pilar, *Elementary Quantum Chemistry*, 2nd edition, McGraw-Hill, New York, NY, 1990.

42) J.N. Murrell and A.J. Harget, *Semi-empirical Self-Consistent-Field Molecular Orbital Theories of Molecules*, Wiley-Interscience, New York, NY, 1971, *Chapter 2*.
43) P. O'D. Offenhartz, *Atomic and Molecular Orbital Theory*, McGraw-Hill, New York, NY, 1970, *Chapter 11*.
44) R. Hoffmann, *J. Chem. Phys.*, 1963, **39**, 1397; 1964, **40**, 2745, 2474, 2480.
45) R. Hoffmann, *Tetrahedron*, 1966, **22**, 521, 539.
46) M. Wolfsberg and L. Helmholz, *J. Chem. Phys.*, 1952, **20**, 837.
47) J. Howell *et al*, QCPE #344a.
48a) J.A. Pople, D.P. Santry and G.A. Segal, *J. Chem. Phys.*, 1964, **42**, 5129.
48b) J.A. Pople and G.A. Segal, *J. Chem. Phys.*, 1965, **43**, 5136; 1966, **44**, 3289.
49) Details on the CNDO and INDO methods appear in:
a) J.A. Pople and D.L. Beveridge, *Approximate Molecular Orbital Theory*, McGraw-Hill, New York, NY, 1970.
b) G. Klopman and R.C. Evans in *Semiempirical Methods of Electronic Structure Calculations*, *Parts A* and *B*, G.A. Segal, editor, **Volumes 7 and 8** of *Modern Theoretical Chemistry*, W. Miller *et al*, editors, Plenum, New York, NY, 1977, *Part A*, p. 29.
c) J.N. Murrell and A.J. Harget, *Semi-empirical Self-Consistent-Field Molecular Orbital Theories of Molecules*, Wiley-Interscience, New York, NY, 1971, *Chapter 3*.
50) R.L. Ellis and H.H. Jaffé in *Semiempirical Methods of Electronic Structure Calculation*, *Parts A* and *B*, G.A. Segal, editor, **Volumes 7 and 8** of *Modern Theoretical Chemistry*, W. Miller et al, editors, Plenum, New York, NY, 1977, *Part B*, p. 49.
51) J. Michl in *Semiempirical Methods of Electronic Structure Calculation*, *Parts A* and *B*, G.A. Segal, editor, **Volumes 7 and 8** of *Modern Theoretical Chemistry*, W. Miller *et al*, editors, Plenum, New York, NY, 1977, *Part B*, p. 99.
52a) A.D. Bacon and M.C. Zerner, *Theor. Chim. Acta.*, 1979, **53**, 21.
52b) W.P. Anderson, W.D. Edwards and M.C. Zerner, *Inorg. Chem.*, 1986, **25**, 2728.
53a) J. Ridley and M.C. Zerner, *Theor. Chim. Acta.*, 1973, **32**, 111; 1976, **42**, 223.
53b) M.C. Zerner, G.H. Loew, R.F. Kirchner and U.T. Mueller-Westerhoff, *J. Am. Chem. Soc.*, 1980, **102**, 589.
54) G. Klopman and R.C. Evans in *Semiempirical Methods of Electronic Structure Calculation*, *Parts A* and *B*, G.A. Segal, editor, **Volumes 7 and 8** of *Modern Theoretical Chemistry*, W. Miller *et al*, editors, Plenum, New York, NY, 1977, *Part A, Chapter 2*.
55) T.A. Halgren and W.N. Lipscomb, *J. Chem. Phys.*, 1973, **58**, 1569.
56) D.S. Marynick and W.N. Lipscomb, *Proc. Natl. Acad. Sci. U.S.A.*, 1982, **79**, 1341.
57) R.G. Parr, *Quantum Theory of Molecular Electronic Structure*, Benjamin, New York, NY, 1963, pp. 51-52.
58) J.P. Nalriu in *Semiempirical Methods of Electronic Structure Calculation*, *Parts A* and *B*, G.A. Segal, editor, **Volumes 7 and 8** of *Modern Theoretical Chemistry*, W. Miller *et al*, editors, Plenum, New York, NY, 1977, *Part A, Chapter 3*.
59) T.A. Halgren, D.A. Kleier, J.H. Hall Jr., L.D. Brown and W.N. Lipscomb, *J. Am. Chem. Soc.*, 1978, **100**, 6595.
60) N.C. Baird and M.J.S. Dewar, *J. Chem. Phys.*, 1969, **50**, 1262.
61) M.J.S. Dewar and E. Haselbach, *J. Am. Chem. Soc.*, 1970, **92**, 590.
62) M.J.S. Dewar and D.H. Lo, *J. Am. Chem. Soc.*, 1972, **94**, 5296.
63) R.C. Bingham, M.J.S. Dewar and D.H. Lo, *J. Am. Chem. Soc.*, 1975, **97**, 1285, 1294, 1302, 1307.
64) M.J.S. Dewar, D.H. Lo and C.A. Ramsden, *J. Am. Chem. Soc.*, 1975, **97**, 1311.
65) M.J.S. Dewar, *Science*, 1975, **187**, 1037.
66) M.J.S. Dewar *et al*, QCPE # 506a.
67) Strengths and weaknesses of MINDO/3 and MNDO are reviewed in:
T. Clark, *A Handbook of Computational Chemistry*, Wiley, New York, NY, 1985, pp. 144-151.

68) D.N. Nanda and K. Jug, *Theor. Chim. Acta.*, 1980, **57**, 95.
69) M.J.S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, 1977, **99**, 4899, 4907.
70) M.J.S. Dewar and H.S. Rzepa, *J. Am. Chem. Soc.*, 1978, **100**, 58, 777, 784.
71) M.J.S. Dewar and M.L. McKee, *J. Am. Chem. Soc.*, 1977, **99**, 5231.
72) L.P. Davis, R.M. Guidry, J.R. Williams, M.J.S. Dewar and H.S. Rzepa, *J. Comput. Chem.*, 1981, **2**, 433.
73) M.J.S. Dewar *et al*, QCPE # 506a.
74) M.J.S. Dewar, E.G. Zoebisch, E.F. Healy and J.J.P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902.
75) M.J.S. Dewar and E.G. Zoebisch, *THEOCHEM.*, 1988, **49**, 1.
76) M.J.S. Dewar and K.M. Dieter, *J. Am. Chem. Soc.*, 1986, **108**, 8075.
77) J.J.P. Stewart, *J. Comp. Aided Mol. Design*, 1990, **4**, 1.
78) M.J.S. Dewar *et al*, QCPE # 506a.
79) J.J.P. Stewart, *J. Comput. Chem.*, 1989, **10**, 209, 221; 1990, **11**, 543; 1991, **12**, 320.
80) J.J.P. Stewart *et al*, QCPE # 506a.
81) K. Jug and D.N. Nanda, *Theor. Chim. Acta.*, 1980, **57**, 101, 131.
82) K. Jug, R. Iffert and J. Schultz, *Int. J. Quantum. Chem.*, 1987, **32**, 265.
83) F.L. Pilar, *Elementary Quantum Chemistry*, 2nd edition, McGraw-Hill, New York, NY, 1990, *Section 14-8.*
84) C.C.J. Roothaan, *Rev. Mod. Phys.*, 1960, **32**, 179.
85) J.S. Binkley, J.A. Pople and P.A. Dobosh, *Mol. Phys.*, 1974, **28**, 1423.
86) J.A. Pople and R.K. Nesbet, *J. Chem. Phys.*, 1954, **22**, 571.
87) T. Koopmans, *Physica*, 1933, **1**, 104.
88) M.S. Khan, M.S. Islam and D.R. Bates, *J. Phys. Chem. B*, 1998, **102**, 3099.
89) J.D. Gale, *J. Chem. Soc., Faraday Trans.*, 1997, **93**, 629.
90) M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias and J.D. Joannopoulos, *Rev. Mod. Phys.*, 1992, **64**, 1045.
91) F. Shimojo, K. Hoshino and H. Okazaki, *J. Phys. Soc. Japan*, 1988, **67**, 2008.
92) S. Fischer, P.D.J. Groothenuis, L.C. Groenen, W.P. van Hoorn, F.C.J.M. van Geggel, D.N. Reinhoudt and M. Karplus, *J. Am. Chem. Soc.*, 1995, **117**, 1611.
93) K.N. Houk, J. González and Y. Li, *Acc. Chem. Res.*, 1995, **28**, 81.
94) R.F. Cracknell, D. Nicholson and N. Quirke, *Molecular Simulation*, 1994, **13**, 161.
95) R.A. Dammkoehler, S.F. Karasek, E.F.B. Shands and G.R. Marshall, *J. Comp. Aided Mol. Design*, 1989, **3**, 3.
96) I.D. Kuntz, J.M. Blaney, S.J. Oatley, R. Langridge and T.E. Ferrin, *J. Mol. Biol.*, 1982, **161**, 269.
97) P.J. Goodford, *J. Med. Chem.*, 1985, **28**, 849.
98) H.J. Böhm , *J. Comp. Aided Mol. Design*, 1992, **6**, 593.
99) G. Lauri and P.A. Bartlett., *J. Comp. Aided Mol. Design*, 1994, **8**, 51.
100) P.Y.S. Lam, P.K. Jadhav, C.E. Eyermann, C.N. Hodge, Y. Ru, L.T. Bachelor, J.L. Meek, M.J. Otto, M.M. Rayner, Y.N. Wong, C-H Chang, P.C. Weber, D.A. Jackson, T.R. Sharpe and S. Erickson-Viitanen, *Science*, 1994, **263**, 380.
101) E. Gurr, *The Rational Use of Dyes in Biology*, Leonard Hill, London, 1965.
102) M. Rooseboom, *Microscopium*, Rijkmuseum voor de Gescgiedenis der Natuurwetenschappen, Leiden, Netherlands, 1956.
103) N. Grew, *The Anatomy of Vegetables*, White, London, 1682.
104) A. Leeuwenhoek, *Epistolae physiologicae super compluribus naturae arcanis*, Delphis, Beman, Netherlands, 1719.
105) Sarrabat, *Dissertation de la Sève*, Bordeaux, France, 1733.
106) J. Hill, *The Construction of Timber from its earliest Growth*, White, London, 1770.
107) A.G. Perkin and A.E. Everest, *The Natural Organic Colouring Matters*, Longmans, Green and Co., New York, NY, 1918, pp. 77, 627.
108) F. Mayer, *The Chemistry of Natural Coloring Matters*, ACS Monograph No. 89, translated and revised by A.H. Cook, Reinhold Publishing Co., New York, NY, 1943, p. 137.

109) A.G. Perkin and A.E. Everest, *The Natural Organic Colouring Matters*, Longmans, Green and Co., New York, NY, 1918, pp. 23, 37, 52, 64.
110) F. Mayer, *The Chemistry of Natural Coloring Matters*, ACS Monograph No. 89, translated and revised by A.H. Cook, Reinhold Publishing Co., New York, NY, 1943, p. 117.
111) A.G. Perkin and A.E. Everest, *The Natural Organic Colouring Matters*, Longmans, Green and Co., New York, NY, 1918, pp. 24, 42.
112) F. Mayer, *The Chemistry of Natural Coloring Matters*, ACS Monograph No. 89, translated and revised by A.H. Cook, Reinhold Publishing Co., New York, NY, 1943, p. 123.
113) A.G. Perkin and A.E. Everest, *The Natural Organic Colouring Matters*, Longmans, Green and Co., New York, NY, 1918, pp. 29, 42.
114) F. Mayer, *The Chemistry of Natural Coloring Matters*, ACS Monograph No. 89, translated and revised by A.H. Cook, Reinhold Publishing Co., New York, NY, 1943, p. 120.
115) A.G. Perkin and A.E. Everest, *The Natural Organic Colouring Matters*, Longmans, Green and Co., New York, NY, 1918, p. 27.
116) F. Mayer, *The Chemistry of Natural Coloring Matters*, ACS Monograph No. 89, translated and revised by A.H. Cook, Reinhold Publishing Co., New York, NY, 1943, p. 124.
117) A.G. Perkin and A.E. Everest, *The Natural Organic Colouring Matters*, Longmans, Green and Co., New York, NY, 1918, p. 31.
118) F. Mayer, *The Chemistry of Natural Coloring Matters*, ACS Monograph No. 89, translated and revised by A.H. Cook, Reinhold Publishing Co., New York, NY, 1943, pp. 121, 125.
119) A.G. Perkin and A.E. Everest, *The Natural Organic Colouring Matters*, Longmans, Green and Co., New York, NY, 1918, pp. 29, 32, 42.
120) P. Woulfe, *Phil. Trans.*, 1771, **61**, 114.
121) F.F. Runge, *Ann. Phys. Chem.*, 1834, **31**, 65, 70, 513.
122) H. Kolbe and Schmitt, *Ann. der Chemie*, 1861, **119**, 169.
123) B. Beneke, *Korresp. Bl. Ver. gemeinsch. Arbeiten*, 1862, **59**, 980.
124) E. Schwartz, *Sitzber. Akad. Wissensch., Wien, Math.-Naturw. Classe*, 1867, **55**, Abt. I, 671.
125) L. Ranvier, *Arch. d. Physiol.*, 1868, **1**, No.2, 319.
126) P. Ehrlich, *Arch. Anat. & Physiol. Phys. Abt.*, 1879, 571.
127) P. Ehrlich, *Zts. f. klin. Med.*, 1879, **1**, 553.
128a) A. Bolles Lee, *The Microscopist's Vade-Mecum*, 1st edition, Churchill, London, 1885.
128b) A. Bolles Lee, *The Microscopist's Vade-Mecum*, 2nd edition, Churchill, London, 1890.
128c) A. Bolles Lee, *The Microscopist's Vade-Mecum*, 3rd edition, Churchill, London, 1893.
129) M. Lavdowsky, *Arch. Mik. Anat.*, 1884, **23**, 506.
130) L. Jenner, *Lancet, (London)*, 1899, **I**, 370.
131) E. Gurr, *Staining Animal Tissues, Practical and Theoretical*, Leonard Hill (Books), London, 1962.
132) M.A. MacConaill and E. Gurr, *Irish J. Med. Sci.*, 1960, **Ser. 7**, No. 412, 182.
133) E. Gurr, *Ph.D. Thesis*, National University of Ireland, 1962.
134) J. Bodman, *Ivor Smith's Chromatographic and Electrophoretic Techniques*, William Heinemann Medical Books, London, 1960, **Volume 2**.
135) E. Gurr, *Encyclopaedia of Microscopic Stains*, Leonard Hill (Books), London; Williams & Wilkins, Baltimore, 1960.
136) E. Gurr and M.A. MacConaill, *J. Roy. Micr. Soc.*, 1961, **79**(4), 353.
137) E. Gurr, *J. Soc. Dyers & Col.*, 1963, **79**, 461.
138) M.A. MacConaill, *J. Anat. Soc. of India*, 1962, **XI**(2), 49.
139) M.A. MacConaill, *Nature*, 1957, **180**, 603.

140) F.M. Rowe, *Colour Index*, 1st edition, Society of Dyers and Colourists, Bradford, 1924.
141) M. Phillips, *Stain Technol.*, 1927, **2**, 17.
142) M.A. MacConaill and E. Gurr, *Irish J. Med. Sci.*, 1962, 1.
143) M.A. MacConaill and E. Gurr, *Irish J. Med. Sci.*, 1963, 481.
144) W.B. Leishman, *British Med. J.*, 1901, *Pt. 2*, 757.
145) E. Kardos, *Folia haematol. Archiv.*, 1911, **12**, 39.
146) F.W. Twort, *Brit. J. Exp. Path.*, 1924, **5**, 350.
147) D.L. Romanowsky, *St. Petersburger med. Woch.*, 1891, **16**, 297, 307.
148) G. Giemsa, *Centlb. Bakt. Parasit.*, 1902, *I Abt.*, **31**, 429; *I Abt.*, **32**, 307; 1904, *I Abt.*, **37**, 308.
149) E. Gurr, *Methods of Analytical Histology and Histochemistry*, Leonard Hill (Books), London, Williams and Williams, Baltimore, 1958.
150) E. Gurr and M.A. MacConaill, *Proc. Anat. Soc.*, 1959, November 27th Meeting, London.
151) T. Vickerstaff, *The Physical Chemistry of Dyeing*, Oliver & Boyd, London & Edinburgh, 1954.
152) I.M. Lewis, *J. Bacteriol.*, 1938, **35**, 573.
153) C. Gram, *Fortsch. Med.*, 1884, **2**, 185.
154) F.A. Denz, Quart. *J. Micr. Sci.*, 1949, **90**, 317.
155) J.W. Orr, *J. Path. & Bact.*, 1937, **44**, 19.
156) A.G.E. Pearse, *Acta Histochem.*, 1957, **Bd. 4** (8), 95.
157) M.A. MacConaill, *Proc. Roy. Irish Acad.*, 1949, **53**, *Sect. B, No. 1*, 1-14; 1951, **54**, *Sect. B, No. 2*, 15.
158) J.R. Baker, *Principles of Biological Microtechnique*, Methuen and Co., London, John Wiley & Sons, New York, 1958.
159) J.R. Baker, *Quart. J. Micr. Sci.*, 1960, **101**(*3*), 255; 1962, **103**(*4*), 493.
160) F.B. Mallory, *Amer. J. Path.*, 1936, **12**, 569.
161) R.G. Fargher and F.L. Pyman, *J. Chem. Soc.*, 1920, **47**, 370.
162) W.G. Christiansen, *J. Am. Chem. Soc.*, 1923, **45**, 1807.
163) Kirchoff, Korsina, Sirkin, DRP 592,870 (*Chem. Zentr.*, 1934, **I**, 2791)
164) A. Burger, *Chem. Eng. News*, 1954, **32**, 4172.
165) G. Domagk, *Angew. Chem.*, 1935, **48**, 660.
166) P. Goissedet *et al*, *Compt. rend. soc. biol.*, 1936, **121**, 1082.
167) J. Tréfouel, J. Tréfouel, F. Nitti and D. Bovet, *Compt. rend. soc. biol.*, 1935, **121**, 756.
168) U.S. Pat. 2,085,037 (June 29, 1937) F. Mietzsch and J. Klarer (to Winthrop Chemical Co.).
169) K. Imhauser, *Med. Klin.*, 1935, **31**, 282.
170) G. Domagk, *Deut. med. Wochschr.*, 1935, **61**, 250.
171) F.L. Rose, *Chem. Ind.*, 1964, 858.
172) E.R. Kennedy and J.F. Barbaro, *J. Bacteriol.*, 1953, **65**, 678.
173) J.W. Bartholomew and H. Finkelstein, *J. Bacteriol.*, 1954, **67**, 689.
174) A. Albert, *Lancet*, 1942, 278.
175) T.M. McCalla, *Stain Technol.*, 1941, **16**, 27.
176) A. Albert *et al*, *Br. J. Exp. Pathol.*, 1945, **26**, (1)60.
177) A. Albert, S.D. Rubbo and M. Burvill, *Br. J. Exp. Pathol.*, 1949, **30**, 159.
178) J.E. Weiss and L.F. Rettger, *J. Bacteriol.*, 1934, **28**, 501.
179) C.P.G. Wakeley, *Practitioner*, 1941, **146**, 27.
180) B.E. Greenberg and M.L. Brodnet, *New Engl. J. Med.*, 1933, **209**, 1153.
181) A. Albert and R.J. Goldacre, *Nature*, 1948, **101**, 95.
182) G.R. Goetchius and C.A. Lawrence, *J. Lab. Clin. Med.*, 1944, **29**, 134.
183) R. Wagner and A. Pohlner, *Z. Immunitätsforsch.*, 1938, **94**, 171.
184) A. Goerner and F.L. Haley, *J. Lab. Clin. Med.*, 1931, **16**, 957.

185) U.S. Pat. 2,396,145 (March 5 1946), E.E.A. Askelöf, N. Svartz and H.C. Willstaedt (to Akticbolaget Pharmacia).
186) G. Proske, *Arch. Hyg. Bacteriol.*, 1952, **136**, 74.
187) D.M. Parkin, E. Läärä and C.S. Muir, *Int. J. Cancer*, 1988, **41**, 184.
188) F.A. Langley and A.C. Crompton, *Recent Results in Cancer Research*, 1973, **40**, 20, 34.
189) R.M. Richart, *Cancer*, 1966, **19**, 395.
190) L.G. Koss *et al*, *Cancer*, 1963, **16**, 1160.
191) R.M. Richart, *Amer. J. Obstet. Gynec.*, 1963, **86**, 925.
192) H.F. Schellhas, *Amer. J. Obstet. Gynec.*, 1969, **104**, 617.
193) J.E. Dunn, *Proc. R. Soc. Med.*, 1966, **59**, 1198.
194) B.A. Barron and R.M. Richart, *J. Natn. Cancer Inst.*, 1968, **41**, 1343.
195) B.A. Barron and R.M. Richart, *J. Natn. Cancer Inst.*, 1970, **45**, 1025.
196) O. Petersen, *Amer. J. Obstet. Gynec.*, 1956, **72**, 1063.
197) A.I. Spriggs, *Lancet*, 1971, **ii**, 599.
198) R. Bangle, M. Berger and M. Levin, *Cancer*, 1963, **16**, 1151.
199) J. Clemmesen, *Acta Path. Microbiol. Scand. Suppl.*, 1965, **v174**.
200) I.I. Kessler, *Cancer*, 1977, **39**, 1912.
201) L. Aurelian, B. Schumann, R.L. Marcus and H.J. Davis, *Science*, 1973, **181**, 161.
202) E. Adam, D. Sharma, O. Zeigler, K. Iwamoto, J.L. Melnick, A.H. Levy and W.E. Rawls, *J. Natn. Cancer Inst.*, 1972, **48**, 65.
203) C. Muir, J. Waterhouse, T. Mack, J. Powell and Sh. Whelan, (editors), *Cancer Incidence in Five Continents*, **Volume V**, IARC Scientific Publication No. 88. Lyon: International Agency for Research on Cancer, 1987.
204) N. Muñoz and F.X. Bosch, *Epidemiology of cervical cancer* in *Human Papillomavirus and Cervical Cancer*, IARC Scientific Publications No. 94 Lyon: International Agency for Research on Cancer, editors N. Muñoz, F.X. Bosch and O.M. Jensen, 1989, pp. 9-39.
205) D.J. Buckley, R. Doll, R.W.C. Harris, M..P Vessey and P.T. Williams, *Lancet*, 1981, **ii**, 1010.
206) J. Li, F.P. Li, W.J. Blot, R.W. Miller and J.F. Fraumeni, Jr., *J. Natl. Cancer Inst.*, 1982, **69**, 1063.
207) P.G. Smith, L.J. Kinlen, G.C. White, A.M. Adelstein and A.J. Fox, *Br. J. Cancer*, 1980, **41**, 422.
208) L.A. Brinton, W.C. Reeves, M.M. Brenes, R. Herrero, E. Gaitan, F. Tenorio, R.C. de Britton, M. Garcia and W.E. Rawls, *Int. J. Cancer*, 1989, **44**, 199.
209) R.W.C. Harris, L.A. Brinton, R.H. Cowdell, D.C.G. Skegg, P.G. Smith, M.P. Vessey and R. Doll, *Br. J. Cancer*, 1980, **42**, 359.
210) C. La Vecchia, S. Franceschi, A. Decarli, M. Fasoli, A. Gentile, F. Parazzini and M. Regallo, *Cancer*, 1986, **58**, 935.
211) L.A. Brinton, R.F. Hamman, G.R. Huggins, H.F. Lehman, R.S. Levine, K. Mallin and J.F. Fraumeni, *J. Natl. Cancer Inst.*, 1987, **79**, 23.
212) N. Muñoz, X. Bosch and J.M. Kaldor, *Br. J. Cancer*, 1988, **57**, 1.
213) T.J. Allerding *et al*, Acta Cytologica, 1985, **29**, 655.
214) V. Vonka, J. Kafka, I. Hirsch, H. Zavadova, M. Hrcmar, A. Suchankova, D. Rezacova, J. Broucek, M. Press, E. Domorazkova, B. Svoboda, A. Havrankova and J. Jelinek, *Int. J. Cancer*, 1984, **33**, 61.
215) B.K. Armstrong, O.V. Allen, B.A. Brennan, I.A. Fruzynski, N.H. de Klerk, E.D. Waters, J. Machin and M.M. Gollow, *Br. J. Cancer*, 1986, **54**, 669.
216) S.K. Kjaer, E.M. De Villiers, B.J. Haugaard, R.B. Christensen, C. Reisen, K.A. Moller, P. Poll, H. Jensen, B.F. Vestergaard, E. Lynge and O.M. Jensen, *Int. J. Cancer*, 1988, **41**, 518.
217) V. Vonka, J. Kafka and Z. Roth, *Adv. Cancer Res.*, 1987, **48**, 149.
218) D. A. Galloway and J.K. McDougall, *Nature*, 1983, **302**, 21.
219) H. zur Hausen, *Lancet*, 1982, **ii**, 1370.

220) H. zur Hausen, *Lancet*, 1986, **ii**, 489.
221) D. di Luca, A. Rotola, S. Pilotti, P. Monini, E. Caselli, F. Rilke and E. Cassai, *Int. J. Cancer*, 1987, **40**, 763.
222) H. zur Hausen, *Cancer Res.*, 1989, **49**, 4677.
223) W.C. Reeves, L.A. Brinton, M. Garcia, M.M. Berens, R. Herrero, E. Gaitan, F. Tenorio, R.C. de Britton and W.E. Rawls, *N. Engl. J. Med.*, 1989, **320**, 1437.
224) M.J. Campion, D.J. McCance, J. Cuzick and A. Singer, *Lancet*, 1986, **ii**, 237.
225) A. Schneider, E. Sawada, L. Gissmann and K. Shah, *Obst. Gynecol.*, 1987, **69**, 554.
226) K. Syrjanen, R. Mantyjarvi, S. Saarikoski, M. Vayrynen, S. Syrjanen, S. Pakkinen, M. Yliskoski, J. Saasttamoinen and O. Castren, *Br. J. Obstet. Gynecol.*, 1988, **95**, 1096.
227) S.K. Kjaer, E.M. De Villiers, B.J. Haugaard, R.B. Christensen, C. Reisen, K.A. Moller, P. Poll, H. Jensen, B.F. Vestergaard, E. Lynge and O.M. Jensen, *Int. J. Cancer*, 1988, **41**, 518.
228) L.L. Villa and M. Franco, *J. Natl. Cancer Inst.*, 1989, **81**, 332.
229) N. Muñoz, G. de Thé, N. Aristizabal, C. Yee, A. Rabson and G. Pearson, *Antibodies to herpesviruses in patients with cervical cancer and controls* in *Oncogenesis and Herpesviruses II*, IARC Scientific Publication No. 11, G. de Thé, M.A. Epstein and H. zur Hausen, editors, Lyon: International Agency for Research on Cancer, 1975, pp. 45-51.
230) J.K. McDougall, A.M. Beckmann and N.B. Kiviat, *Methods for diagnosing papillomavirus infection* in *Papillomaviruses*, Ciba Foundation Symposium No. 120, 1984, p. 86.
231) L.A. Brinton, K.T. Tashima, H.F. Lehman, R.S. Levine, K. Mallin, D.A. Savitz, P.D. Stolley and J.F. Fraumeni, Jr., *Cancer Res.*, 1987, **47**, 1706.
232) R.K. Peters, D. Thomas, D.G. Hagan, T.M. Mack and B.E. Henderson, *J. Nat. Cancer Inst.*, 1986, **77**, 1063.
233) I. Parazzini, C. La Vecchia, E. Negri, M. Fasoli and G. Ceccheti, *Br. J. Cancer*, 1988, **41**, 184.
234) V. Beral, P. Hannaford and C. Kay, *Lancet*, 1988, **ii**, 1331.
235) K. Jayant, P.N. Notani, V.V. Gadre, S.S. Gulati and P.R. Shah, *Ind. J. Cancer*, 1987, **24**, 47.
236) L.A. Brinton, C. Schairer, W. Haenszel, P. Stolley, H.F. Lehman, R. Levine and D.A. Savitz, *JAMA*, 1986, **255**, 3265.
237) IARC, *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Tobacco Smoking*, **Volume 38**, Lyon: International Agency for Research on Cancer, 1986.
238) S.E. Barton, P.H. Maddox, D. Jenkins, R.E. Edwards, J. Cuzick and A. Singer, *Lancet*, 1988, **ii**, 652.
239) A.L. Herbst, D.C. Poskanzer, S.J. Robboy, L. Friedlander and R.E. Scully, *New Engl. J. Med.*, 1975, **292**, 334.
240) A.L. Herbst, R.E. Scully and S.J. Robboy, *Clin. Obstet. Gynecol.*, 1975, **18**, 185.
241) IARC, *IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Sex Hormones (II)*, **Volume 21**, Lyon: International Agency for Research on Cancer, 1979.
242) A.M. Lilienfeld, M.L. Levin and I.I. Kessler, *Cancer in the United States*, Harvard University Press, Cambridge, MA, 1972.
243) J. Chomet and J. Chomet, *Smear Tests*, Thorsons, London, 1991, pp. 82-83.
244) T. Langerhans, *Virch. Arch.*, 1890, **120**, 28.
245) Lugol's iodine is prepared by shaking KI (2 g) with water (5 cm^3) until dissolved. Iodine (1 g) is also shaken in until dissolved and further water (95 cm^3) added.
246) H.C. Mack, *Harper Hosp. Bull.*, 1942, **1**, 54 proposed the use of Lugol's iodine in a "new and rapid method of staining vaginal smears, based upon a specific reaction for glycogen".
247) G.N. Papanicolaou, *Proc. of the Third Race Betterment Conference*, 1928, p. 538.

248) G.N. Papanicolaou, *Am. J. Anat.*, 1933, **52**, 519.

249) G.N. Papanicolaou and H.F. Traut, *Am. J. Obst. and Gynec.*, 1941, **42**, 193.
P.L. Johnson and M.N. Klein, *Stain Technol.*, 1956, **31**, 223 used the EA36 stain to study cyclic changes in vaginal mucosa.

250) G.N. Papanicolaou, *J. Lab. and Clin. Med.*, 1941, **26**, 1200.

251) P. Ehrlich, *Z. wiss. Mikr.*, 1886, **3**, 150.
Dissolve hæmatoxylin (2 g) in absolute alcohol (100 cm^3) and then add a 2.5% aqueous solution of potash alum (100 cm^3), glycerine (100 cm^3), glacial acetic acid (10 cm^3). Ripen for 3 months.

252) Eosin (1.25 g), water blue (0.5 g) and phosphomolybdic acid (2 g) in distilled water (350 cm^3).

253) A corresponding mixture can be prepared as follows:- a 0.5% aqueous solution of aniline blue water soluble (60 cm^3), a 0.5% aqueous solution of yellow eosin (300 cm^3), and phosphomolybdic acid (1.65 g). For intensifying the staining of the cornified cells, the relative amount of eosin should be raised.

254) Stain DA 15 is prepared as follows:- a 0.5% aqueous solution of light green F S (20 cm^3), a 0.5% aqueous solution of orange G (20 cm^3), a 0.5% aqueous solution of acid fuchsin (20 cm^3), a 0.5% aqueous solution of eosin yellowish (40 cm^3) and phosphomolybdic acid (0.45 g).

255) Stain CY 18 is prepared as follows:- a 0.5% aqueous solution of Wasserblau 6B extra (10 cm^3), a 0.5% aqueous solution of orange G (25 cm^3), a 0.5% aqueous solution of acid fuchsin (25 cm^3) , a 0.5% aqueous solution of eosin yellowish (60 cm^3) and phosphomolybdic acid (0.75 g).

256) Stain DF 11 is prepared as follows:- a 0.5% aqueous solution of aniline blue water soluble (12 cm^3), a 0.5% aqueous solution of orange G (25 cm^3), a 0.5% aqueous solution of acid fuchsin (25 cm^3), a 0.5% aqueous solution of eosin yellowish (60 cm^3) and phosphomolybdic acid (0.6 g).

257) Stain DA 19 is prepared as follows:- a 0.5% aqueous solution of light green SF yellowish (12 cm^3), a 0.5% aqueous solution of orange G (24 cm^3), a 0.5% aqueous solution of acid fuchsin (20 cm^3), a 0.5% aqueous solution of eosin yellowish (40 cm^3) and phosphomolybdic acid (0.45 g).

258) Stain DF 20 is prepared as follows:- a 0.5% aqueous solution of aniline blue water soluble (16.5 cm^3), a 0.5% aqueous solution of orange G (25 cm^3), a 0.5% aqueous solution of acid fuchsin (25 cm^3), a 0.5% aqueous solution of eosin yellowish (60 cm^3) and phosphomolybdic acid (0.7 g).

259) Stain DF 32 is prepared as follows:- a 0.5% aqueous solution of aniline blue water soluble (12 cm^3), a 0.5% aqueous solution of orange G (25 cm^3), a 0.5% aqueous solution of acid fuchsin (21 cm^3), a 0.5% aqueous solution of eosin yellowish (42 cm^3), phosphotungstic acid (0.1125 g) and phosphomolybdic acid (0.225 g).

260) Good fixation may be obtained within one to two minutes, although a fixation of five to fifteen minutes may be needed if the smear is thick. Slides can be kept in the fixative for several hours or days without being harmed. Fixation for more than a week or two can change the cellular staining reaction cells.

261) C.R. Stockard and G.N. Papanicolaou, *Am. J. Anat.*, 1917, **22**, 225.

262) H.F. Harris, *Micr. Bull.*, 1900, **15**, 47.
Mix a 10% solution of hæmatoxylin in absolute alcohol (5 cm^3) with a 10% aqueous solution of potash alum (100 cm^3), heating to boiling point. Mercuric oxide (0.25 g) is then added. Once a purple colour appears, cool and add glacial acetic acid (4 cm^3).

263) P. Mayer, *Zeit. wiss. Mik.*, 1903, **20**, 409.

264) R.D. Lillie, *Stain Technol.*, 1942, **17**, 89.

265) F. Delafield, *Z. wiss. Mikr.*, 1885, **3**, 150.

266) G. Gomori, *Amer. J. Path.*, 1941, **17**, 395.

267) M. Heidenhain, *Festsch. f. Kolliker, Englemann*, Leipzig, 1892.

268) M.A. MacConaill, *Proc. Roy. Irish. Acad.*, 1949, **53**, Sect B., No. 1, 1.
M.A. MacConaill, *Proc. Roy. Irish. Acad.*, 1951, **54**, Sect B., No. 2.

269) C. Dobell, *J. Parasitology*, 1942, **34**, 101.
270) F.B. Mallory, *J. Exp. Med.*, 1900, **5**, 15.
271) Stain DF 50 is prepared as follows:- a 0.5% aqueous solution of aniline blue water soluble (16 cm^3), a 0.5% aqueous solution of orange G (17 cm^3), a 0.5% aqueous solution of acid fuchsin (23 cm^3), a 0.5% aqueous solution of eosin yellowish (44 cm^3), phosphomolybdic acid (0.2 g) and phosphotungstic acid (0.2 g).
272) P.F. Fletcher, *Am. J. Obst. and Gynec.*, 1940, **39**, 562 - Discussion by P.F. Schneider.
273) E. Shorr, *Science*, 1940, **91**, 321.
274) E. Shorr, *Science*, 1940, **91**, 579.
275) E. Shorr, *Science*, 1941, **94**, 545.
276) G.N. Papanicolaou, *Jour. Lab. and Clin. Med.*, 1941, **26**, 1200.
277) Stain EA36 is prepared using:- a 0.5% solution of light green SF yellowish in 95% alcohol (45 cm^3), a 0.5% solution of bismarck brown Y in 95% alcohol (10 cm^3), a 0.5% solution of eosin yellowish in 95% alcohol (45 cm^3), phosphotungstic acid (0.2 g), and saturated aqueous lithium carbonate solution (1 drop).
278) Satin EA25 is prepared using:- a 0.5% solution of light green SF yellowish in 95% alcohol (44 cm^3), a 0.5% solution of bismarck brown Y in 95% alcohol (12 cm^3), a 0.5% solution of eosin yellowish in 95% alcohol (44 cm^3), phosphotungstic acid (0.170 g), and saturated aqueous lithium carbonate solution (1 drop).
279) The staining procedure (a-g) is:
a) Smears are fixed immediately before drying in equal parts of 95% alcohol and ether for 5 to 15 minutes. Pronlonged fixation of a week or more affects the staining reaction of the cells.
b) Stain in Harris hæmatoxylin for 5 to 10 minutes. Rinse in distilled water. Rinse 3 to 4 times in 0.5% aqueous solution of hydrochloric acid. Rinse thoroughly in water. Leave for 1 minute in a weak solution of lithium carbonate (3 drops of a saturated aqueous solution per 100 cm^3 of water). Rinse thoroughly in water.
c) Rinse in distilled water, then in 50%, 70%, 80% and 95% alcohols.
d) Stain for 1 minute in OG6 solution (acidified 0.5 % orange G solution in 95% alcohol).
e) Rinse 5 to 10 times in each of two jars containing 95% alcohol to remove excess stain.
f) Stain in EA 36 or EA 25 for 2 minutes.
g) Rinse 5 to 10 times in each of three jars containing 95% alcohol - not the same alcohol used after orange G. Rinse in absolute alcohol and xylol. Mount in clarite, Canada Balsam or Gum Dammar.
280) The addition of phosphotungstic acid to the orange G solution intensifies the orange colour. For normal slides a slight acidification of 0.010 g per 100 cm^3 (OG 8) or 0.015 g per 100cm^3 (OG 6) is suggested. For cancer diagnosis a higher acidification of 0.025 g per 100 cm^3 (OG 5) is preferable, giving sharper contrast of the abnormal cell types.
281) Some EA stains such as the Gill Modified EA stain do not contain Bismarck brown Y.
282a) G.F. Bahr, P.H. Bartels, M. Bibbo, M. deNicholas and G.L. Wied, *Acta Cytol.*, 1973, **17**, 106.
282b) P.H. Bartels, G.F. Bahr, M. Bibbo, D.L. Richards, M.G. Sonek and G.L. Wied, *Acta Cytol.*, 1974, **18**, 522.
283) H.C. Cook, *Manual of Histological Demonstration Techniques*, Butterworths and Co., London, 1974, pp. 182.
284) N.C. Foot, *Stain Technol.*, 1933, **8**, 101.
285) N.C. Foot, *Am. J. Path.*, 1938, **14**, 245.
286) The exact procedure (a-h) is as follows:
a) Fix slide while wet in 95% alcohol:ether (1:1) and carry through alcohols to water.
b) Stain with Harris hæmatoxylin for 2 minutes.
c) Rinse 3 to 4 times in water and stand for 5 minutes in running water.

d) Stain with ponceau de xylidene-acid fuchsin-orange G solution for 5 minutes. Rinse 3 to 4 times in water.
e) Stain with 3% phosphomolybdic solution for 10 minutes. Rinse 3 to 4 times in water.
f) Stain with light green for 8 minutes.
g) Stain with 0.25% acetic acid for 3 minutes.
h) Dehydrate, clear in xylol and mount in damar.

287) R.D. Lillie, *Stain Technol.*, 1940, **15**, 17.

288) Shorr's revised staining procedure (a-e) is as follows:
a) After fixing, carry through alcohols to water, stain with Harris Hæmatoxylin for 2 minutes, and wash under running water for 5 minutes.
b) Instead of the Ponceau de Xylidene-Acid Fuchsin-Orange G solution use 1% Biebrich Scarlet, water soluble, and 0.4% Orange G in 1% acetic acid. Stain for 1 minute and rinse in water.
c) Instead of the 3% phosphotungstic acid mordant, use a mixture of equal parts of 5% phosphomolybdic and phosphotungstic acids. Mordant for 1 minute and rinse.
d) Instead of 0.3% Light Green, use a 0.25% solution of Fast Green FCF in 0.3 % acetic acid. Stain for 2 minutes. Do not rinse.
e) Differentiate in 1% acetic acid for 1 minute, carry through alcohols to xylol and mount in damar.

289) The S3 differential stain is prepared as follows: 50% ethyl alcohol (100 cm^3), Biebrich Scarlet, water soluble (0.5 g), Orange G (0.25 g), Fast Green FCF (0.075 g), phosphotungstic acid (0.5 g), phosphomolybdic acid (0.5 g) and glacial acetic acid (1 cm^3).
The staining procedure (a-d) is as follows:
a) Fix, while wet in equal parts of ether and 95% alcohol for 1 to 2 minutes.
b) Stain for 1 minute in Solution S3.
c) Carry through 70%, 95% and absolute alcohols, dipping slide 10 times in each.
d) Clear in xylol and mount in damar.

290) Dissolve powdered May-Grunwald dye (0.3 g) in methanol (100 cm^3). Dilute 20 parts of above solution with 30 parts of pH 6.8 buffer. Dilute 10 parts Giemsa stock solution[291] with 40 parts pH 6.8 buffer. Stain fixed smears in diluted May-Grunwald solution for 10 minutes. Rinse in tap water. Stain in diluted Giemsa soltuion for 30 minutes. Wash and differentiate in buffer for 5-20 minutes until desired colour balance is achieved. Dry and mount in a DPX-type mountant.

291) Dissolve Giemsa dry stain (7.36 g) in glycerol (500 cm^3) heated to 50 °C in a water bath. Leave for 30 mins at 50 °C with periodic mixing. Cool, add methanol (500 cm^3), mix and filter.

292) L. von Bertalanffy, F. Masin and M. Masin, *Science, N.Y.*, 1956, **124**, 1024.

293) L. von Bertalanffy, M. Masin and F. Masin, *Cancer, N.Y.*, 1958, **11**, 873.

294) F.D. Bertalanffy, *Canadian Assn. J.*, 1960, **83**, 211.

295) F.D. Bertalanffy, *Modern Med. of Canada*, 1960, **15**, 55.

296) L. von Bertalanffy, *Protoplasma*, 1963, **57**, Heft 1-4, 51.

297) E. Gurr, *Staining Animal Tissues, Practical and Theoretical*, Leonard Hill (Books), London, Williams & Wilkins, Baltimore, 1962.

298) The procedure (a-f) is as follows: Prepare a 0.1% aqueous solution of acridine orange. Dilute 1 part stain with 10 parts of pH 6.0 M/15 phosphate buffer to give a 0.01% solution. Prepare M/10 calcium chloride differentiator using calcium chloride (11.099 g) in distilled water (100 cm^3).
a) Take alcohol-fixed smears to distilled water.
b) Immerse in 1% acetic acid for about 10 seconds (prevents rapid fading of fluorescence) and rinse in two changes of distilled water over 1 minute.
c) Stain in diluted acridine orange solution at pH 6.0 for 3.0 minutes.
d) Rinse in pH 6.0 buffer for 1 minute to remove excess dye.
e) Differentiate in M/10 calcium chloride solution for 1 minute.
f) Wash in phosphate buffer and mount in same.

299) J. Brachet, *Chemical Embryology*, translated by L.G. Barth, Interscience, New York, NY, 1950.
300) E. Chargaff and J.N. Davidson, *The Nucleic Acids; Chemistry and Biology*, 2 volumes, Acadmeic Press, New York, NY, 1955.
301) W.J. Pirozynski and L. von Bertalanffy, *Exp. Med. Surg.*, 1955, **13**, 261.
302) L. von Bertalanffy and F.D. Bertalanffy, *Ann. New York Acad. Sci.*, 1960, **84**, 225.
303) A pure basic dye in aqueous solutions absorbs light in a distinct manner. When the dye reacts with tissue elements to produce a colour similar to that of the dye, it stains **orthochromatically**. However, certain dyes may stain some tissue components orthochromatically but will produce a different colour with other component. Toluidine blue O **(155)** stains cells blue, but collagen red.[303d] Sylvén[303b] showed that toluidine blue stains **metachromatically**, probably due to dye micelle formation (dimers, trimers etc.),[303e] only when the acidic groups are present in a polymer, and when these groups are closely packed (<5 Å apart) together in the molecule. Similarly, acridine orange **(106)** produces a different coloured fluorescence when it is bound to low polymer nucleic acids than when it is linked to "high-polymer" DNA.[303f-j] When excited with 365 nm ultraviolet light, DNA fluoresces greenish yellow and RNA crimson red. It also shows its high polymer colour when linked to other polymers, such as keratin, which fluoresces green.
303a) J.R. Baker, *Principles of Biological Microtechnique*, Methuen, London, 1958, *Chapter 13*.
303b) B. Sylvén, *Quart. J. micr. Sci.*, 1954, **95**, 327.
303c) R. Chayen and E.R. Roberts, *Sci. J. Roy. Coll. Sci.*, 1955, **25**, 50.
303d) J. Ball and D.S. Jackson, *Stain Technol.*, 1954, **28**, 33.
303e) L. Michaelis, *Cold Spring Harbour Symp. Quant. Biol.*, 1947, **12**, 131.
303f) J.A. Armstrong, *Exptl. Cell Res.*, 1956, **11**, 640.
303g) P.P.H. de Bruyn, R.S. Farr, H. Banks and F.W. Morthland, *Exptl. Cell Res.*, 1953, **4**, 174.
303h) L. Massart, G. Peeters, J. De Ley, R. Vercauteren and A. Von Honcke, *Experientia*, 1947, **3**, 288.
303i) N. Schümmelfeder, *J. Histochem. Cytochem.*, 1958, **6**, 392.
303j) C.R. Austin and E.C. Amoroso, *Endeavour*, 1959, **18**, 130.
304) L. von Bertalanffy and I. Bickis, *J. Cytochem. & Histochem.*, 1956, **4**, 481.
305) T. Lowhagen, *Acta Cytol.*, 1966, **10**, 194.
306) E. Gurr, *J. Roy. Naval Med. Ser.*, 1951, **37**, 3, 133.
307) E. Gurr, *A Practical Manual of Medical and Biological Staining Techniques*, 1st edition, Leonard Hill (Books), London, Interscience Publishers, New York, 1953, 2nd edition, 1956.
308) W.K. Cuyler, *J. Lab. & Clin. Med.*, 1932, **18**, 314.
309) The procedure (a-k) is as follows:
a) Prepare solution A - Harris Haematoxylin.
b) Prepare solution B - absolute alcohol (70 cm^3), distilled water (28.8 cm^3), strong ammonia solution (1.2 cm^3).
c) Prepare solution C - indigocarmine (0.25 g), eosin yellowish (0.75 g), distilled water (99 cm^3), chloroform (5 drops).
d) Allow smears to dry on slides.
e) Immerse in 95% alcohol for 3 minutes.
f) Immerse in distilled water for 1 minute.
g) Immerse in Harris hæmatoxylin for 20 minutes.
h) Wash with running tap water for 2 minutes.
i) Immerse in B for a few seconds.
j) Stain in C overnight. Rinse quickly in tap water.
k) Dehydrate with 70%, 90% and absolute alcohol, clear in xylol and mount.
310) E. Gurr, *Methods of Analytical Histology and Histochemistry*, Leonard Hill (Books), London, Williams & Wilkins, Baltimore, 1958, pp. 37.

311) The procedure (a-g) is as follows:
a) Fix wet smear in equal volumes of ether and absolute alcohol for one minute.
b) Rinse in two changes of 80% alcohol.
c) Rinse in two changes of 70% alcohol.
d) Rinse by dipping seven times in distilled water.
e) Immerse in Michrome KH1, KH8 or KH9 staining solution for two minutes.
f) Rinse in 70% and 90% alcohols.
g) Dehydrate with absolute alcohol, clear in xylol and mount.
312) A.W. Asscher, C.J. Turner and C.H. De Boer, *J. Anat.*, 1956, **90**, 547.
313) The procedure (a-f) is as follows:
a) Fix wet smears in equal parts of ether and absolute alcohol for one minute.
b) Rinse with two changes of 80% alcohol, followed by two changes of 70% alcohol.
c) Rinse by dipping slide seven times in distilled water.
d) Immerse in MF4 staining solution for two minutes.
e) Rinse with 70% and 90% alcohols.
f) Dehydrate with absolute alcohol, clear in xylol and mount.
314) The procedure (a-f) is as follows:
a) Fix wet smears in equal parts of ether and absolute alcohol for two minutes.
b) Rinse 70% and 50% alcohols.
c) Rinse in water.
d) Immerse in PX staining solution for three minutes.
e) Immerse in dioxane for a few seconds.
f) Wash rapidly with absolute alcohol, rinse well in xylol and mount.
315) C. Van Duijn, Jr., *Mikroskopie*, 1956, **10**, 396.
316) The procedure (a-p) is as follows:
a) Fix wet smears in a mixture of equal volumes of absolute alcohol and ether for five to fifteen minutes.
b) Rinse successively in 90%, 70% and 50% alcohols.
c) Rinse in distilled water.
d) Stain for 20 minutes in Ehrlich's hæmatoxylin.
e) Rinse in distilled water.
f) Differentiate in 0.5% hydrochloric acid.
g) Wash in running tap water for a few seconds.
h) Blue in saturated aqueous lithium carbonate.
i) Wash in tap water.
j) Immerse in 0.5% aqueous erythrosin solution for 15 minutes.
k) Wash with tap water.
l) Mordant in 2% phosphotungstic acid for five minutes.
m) Wash well with tap water and rinse in distilled water.
n) Stain with 2% aqueous water blue for 20 minutes.
o) Wash well with tap water.
p) Dry thoroughly in air and mount.
317) J.S. Greenstein, *Stain Technol.*, 1961, **36**, 87.
318) The procedure (a-o) is as follows:
a) Prepare solution A - hæmatoxylin (1 g), 95% alcohol (50 cm^3), glycerin (50 cm^3), iron alum (15 g), ferrous sulphate (15 g), distilled water (100 cm^3). Dissolve hæmatoxylin in alcohol. Add glycerin and hæmatoxylin to stock bottle. Dissolve iron alum and ferrous sulphate in water by shaking. Pour this solution into the stock bottle. Shake well.
b) Prepare solution B - ponceau 2R (1 g), distilled water (100 cm^3), glacial acetic acid (1 cm^3).
c) Prepare solution C - acid fuchsin (0.5 g), distilled water (100 cm^3).
d) Prepare solution D - 1% aqueous solution of phosphomolybdic acid.
e) Prepare solution E - fast green FCF (2 g), glacial acetic acid (1 cm^3), distilled water (99 cm^3).

f) Prepare solution F - a saturated aqueous solution of picric acid.
g) Prepare solution G - 8 parts A, 3 parts B, 1 part C - filter.
h) Prepare solution H - 1 part D, 1 part E, 1part F - filter.
i) Carry sections through to distilled water.
j) Immerse in G for 3 to 5 minutes, stirring occasionally.
k) Wash thoroughly in tap water.
l) Immerse in H to 2 to 3 minutes with stirring.
m) Rinse in distilled water until water runs clean.
n) Treat with two changes of 95% alcohol.
o) Dehydrate in absolute alcohol, clear in xylol and mount.

319) K. Herzberg, *Zbl. Bakt. Parasit.*, 1934, **131**, 358.

320) The procedure (a-p) is as follows:

a) Prepare solution A - Kernechtrot (Herzberg) (0.1 g), aluminium sulphate (5 g), distilled water (100 cm^3). Dissolve by boiling, allow to cool, filter then add 10% aqueous acetic acid (0.5 cm^3).
b) Prepare solution B - a 1% aqueous solution of light green (100 cm^3) and 10% aqueous acetic acid (0.5 cm^3)
c) Prepare solution C - a 1% aqueous solution of Victoria blue 4R (10 cm^3), distilled water (90 cm^3) and 10% aqueous acetic acid (5 cm^3).
d) Make smears of infected tissue on clean slides and dry in air. Fix by flaming gently.
e) Wash in distilled water for ten minutes.
f) Stain in Kernechtrot solution for an hour.
g) Rinse for thirty seconds in each of two changes of distilled water.
h) Drain and dry in air for half an hour.
i) Stain in light green solution for two minutes.
k) Rinse for five seconds in each of two changes of distilled water.
l) Immerse in 10% aqueous tartaric acid solution for two minutes.
m) Drain off excess tartaric acid solution, then wash for 5 to 10 seconds in each of two changes of distilled water.
n) Stain with victoria blue solution for two minutes.
o) Drain off excess stain, and wash for five seconds in two changes of distilled water.
p) Drain and leave to dry, protected from dust in air or in an incubator.

321) G.I. Roskin and M.F. Struve, *Stain Technol.*, 1947, **22**, 83.

322) P.G. Unna, *Arch. Mikr. Anat.*, 1911, **78**, 1.

323) The procedure (a-g) is as follows:

a) Prepare solution A - Hyposulphite *leuco* base of methylene blue (HLM) - 0.1 % aqueous solution of methylene blue (10 cm^3), sodium thiosulphate ($Na_2S_2O_3.5H_2O$) A.R (0.8 g), pure conc. hydrochloric acid diluted 1:3 with distilled water (4 drops). Add thiosulphate to methylene blue solution, shake to dissolve and add acid. Heat carefully, agitating solution over a bunsen flame for 20-30 minutes, until light grey sediment thrown down. Allow solution to cool for ten minutes. Filter through filter paper. If *leuco* base is prepared it will be colourless or slightly yellowish, otherwise it is useless. If filtered *leuco* base assumes a blue colour, stand for a while in dark.
b) Prepare solution B - a 0.05% aqueous solution of acid fuchsin.
c) Dried sections on slide covered with *leuco* methylene blue solution applied by pipette.
d) Allow staining solution to act until normal tissue stains blue, while malignant tissue remains colourless - usually 2-3 minutes.
e) Pour off excess liquid and rinse with a large volume of fresh, double-distilled water.
f) Stain in B for about 2 minutes.
g) Pour off excess stain, dry and mount.

324) S.C. Kasdon, W.H. Fishman and F. Homburger, *J. Am. Med. Ass.*, 1950, **144**, 892; S.C. Kasdon, F. Homburger, E. Yorshis and W.H. Fishman, *Surg. Gynec. Obstet.*, 1953, **97**, 579; S.C. Kasdon, J. McGowan, W.H. Fishman and F. Homburger, *Am. J. Obstet. Gynec.*, 1951, **61**, 647.

325) L.D. Odell and J.C. Burt, *Canc. Res.*, 1949, **9**, 362; L.D. Odell, J.C. Burt and R. Bethea, *Science*, 1949, **109**, 564.
326) See:
W.H. Fishman, *Bull. New Eng. Med. Centr.*, 1951, **13**, 12.
J.G. Lawson and D.K. Watkins, *J. Obstet. Gynaec. Brit. Comm.*, 1965, **72**, 1.
327) D.G. Bonham and D.F. Gibbs, *Brit. Med. J.*, 1962, **ii**, 823; D.G. Bonham, *Triangle*, 1964, **6**, 157.
See also:
C.B. Cameron and O.A.N. Husain, *Brit. Med. J.*, 1965, **i**, 1529.
O.A.N. Husain and C.B. Cameron, *Proc. Roy. Soc. Med.*, 1966, **59**, 982.
328) J. Chayen, L. Bitensky, E.K. Aves, G.R.N. Jones, A.A. Silcox and G.J. Cunningham, *Nature*, 1962, **195**, 714.
329) S. Cohen and S. Way, *Brit. Med. J.*, 1966, **i**, 88.
330) J.W. Sidman, *Chem. Rev.*, 1958, **58**, 689.
331) A.I. Scott, *Interpretation of the Ultraviolet Spectra of Natural Products*, Pergamon Press, Oxford, 1964, pp. 8-9.
332) P.S. Ellington and G.D. Meakins, *J. Chem. Soc.*, 1960, 697.
333) P. Bladon, H.B. Henbest and G.W. Wood, *J. Chem. Soc.*, 1952, 2727.
334) J.R. Platt, H.B. Klevens and W.C. Price, *J. Chem. Phys.*, 1949, **17**, 466.
335) E.A. Braude, *Ann. Rep. Chem. Soc.*, 1945, **42**, 105.
336) K. Bowden, E.A. Braude, E.R.H. Jones and B.C.L. Weedon, *J. Chem. Soc.*, 1946, 45.
337) H. Booker, L.K. Evans and A.E. Gillam, *J. Chem. Soc.*, 1940, 1453.
338) R.S. Mulliken, *J. Chem. Phys.*, 1939, **7**, 121.
339) L.F. Fieser, *Natural Products Related to Phenanthrene*, 2nd edition, Reinhold, New York, 1937, p. 398.
340) H. Dannenberg, *Abhandl. preuß. Akad. Wiss.*, 1939, **21**, 3.
341) R.B. Woodward, *J. Am. Chem. Soc.*, 1942, **64**, 72.
342) L.F. Fieser and M. Fieser, *Steroids*, Reinhold, New York, 1959, pp. 15-24.
343) L. Dorfman, *Chem. Rev.*, 1953, **53**, 47.
344) G.N. Lewis and M. Calvin, *Chem. Rev.*, 1939, **25**, 273.
345) W. Kuhn, *Helv. Chim. Acta.*, 1948, **31**, 1780.
346) H. Kuhn, *J. Chem. Phys.*, 1949, **17**, 1198.
347) M.J.S. Dewar, *J. Chem. Soc.*, 1952, 3544.
348) P. Nayler and M.C. Whiting, *J. Chem. Soc.*, 1955, 3037.
349) E.A. Braude, *J. Chem. Soc.*, 1949, 1902.
350) E.A. Braude, *J. Chem. Soc.*, 1949, 1890.
351) E.A. Braude, *J. Chem. Soc.*, 1950, 370.
352) G.F. Woods and L.H Schwartzman, *J. Am. Chem. Soc.*, 1948, **70**, 3394.
353) A.D. Mebane, *J. Am. Chem. Soc.*, 1952, **74**, 5227.
354) L.K. Evans and A.E. Gillam, *J. Chem. Soc.*, 1941, 815.
355) Calculated from the λ values of Δ^4-cholesten-3-one in the appropriate solvent.
356) U. Westphal and B.A. Ashley, *J. Biol. Chem.*, 1958, **233**, 57.
357) A.E. Gillam and E.S. Stern, *Electronic Absorption Spectroscopy*, 2nd edition, Arnold, London, 1957, p. 232.
358) W.F. Forbes and R. Shilton, *J. Am. Chem. Soc.*, 1959, **81**, 786.
359) A.T. Nielsen, *J. Org. Chem.*, 1957, **22**, 1539.
360a) K. Dimroth and C. Reichardt, *Z. Anal. Chem.*, 1966, **215**, 344.
360b) B.P. Johnson, M.G. Khaledi and J.G. Dorsey, *Anal. Chem.*, 1986, **58**, 2354.
361) B. Doub and J.M. Vandenbelt, *J. Am. Chem. Soc.*, 1947, **69**, 2714.
362) B. Doub and J.M. Vandenbelt, *J. Am. Chem. Soc.*, 1949, **71**, 2414.
363) H. Conrad-Billroth, *Z. Phys. Chem.*, 1935, **B29**, 170.
364) W.R. Moore, E. Marcus, S.E. Fenton and R.T. Arnold, *Tetrahedron*, 1959, **5**, 179.
365) H.C. Longuet-Higgins and J.N. Murrell, *Proc. Phys. Soc.*, 1955, **68A**, 601 and references cited therein.
366) J.N. Murrell, *J. Chem. Soc.*, 1956, 3779.

367) R.A. Morton and A.L. Stubbs, *J. Chem. Soc.*, 1940, 1347.
368) R.A .Morton and Z. Sawires, *J. Chem. Soc.*, 1940, 1052.
369) B.E. Cross, J.F. Grove, J. Macmillan and T.P.C. Mulholland, *J. Chem. Soc.*, 1958, 2520.
370) J.C. Dearden and W.F. Forbes, *Canad. J. Chem.*, 1958, **36**, 1362.
371) E. Heilbronner and R. Gerdil, *Helv. Chim. Acta.*, 1956, **39**, 1996.
372) J. Tanaka and S. Nagakura, *J. Chem. Phys.*, 1956, **24**, 1274.
373) M.T. O'Shaugnessy and W.H. Rodebush, *J. Am. Chem. Soc.*, 1940, **62**, 2906.
374) E.A. Braude and F. Sondheimer, *J. Chem. Soc.*, 1940, 1347.
375) D.J. Cram and F.W. Cranz, *J. Am. Chem. Soc.*, 1940, **62**, 2906.
376) T.W. Campbell and G.M. Coppinger, *J. Am. Chem. Soc.*, 1951, **73**, 2708.
377) P. Grammaticakis, *Bull. Soc. Chim.*, 1953, **20**, 93, 821, 865.
378) L. Doub and J.M. Vandenbelt, *J. Am. Chem. Soc.*, 1955, **77**, 4535.
379) W.M. Schubert, J. Robins and J.L. Haun, *J. Am. Chem. Soc.*, 1957, **79**, 910.
380) E.A. Braude, *J. Chem. Soc.*, 1949, 1902.
381) L.H. Conover, *Chem. Soc. Special Publications*, 1956, **5**, 48.
382) C.M. Moser and A.I. Kohlenberg, *J. Chem. Soc.*, 1951, 804.
383) There is a point of inflexion at 317 nm in the spectrum of 2,4-diaminobenzoic acid when measured at pH 4.4. The calculated λ for 2,4-diaminobenzoic acid is 301 nm.
384) Spectra of aminobenzoic acids[385-387] return to that of benzoic acid[388] in acidic medium because the ammonium ion is unable to transfer its *p*-electrons to the ring, as these are now involved in sp^3-bond hybridisation.
385) The UV spectrum of 2,4-diaminobenzoic acid in 2N HCl has a benzenoid band at 270 nm (ε 970), and an E.T. band at 226.5 nm (ε 12,300) due to cationic form.[361]
386) The UV spectrum of 2,4-diaminobenzoic acid at pH 3.75 has a benzenoid band at 219.5 nm (ε 9900), and an E.T. band at 284 nm (ε 14,000) due to free amino acid.[361]
387) The UV spectrum of 2,4-diaminobenzoic acid in 0.1N NaOH has an E.T. band at 265 nm (ε 14,900) due to anionic form.[361]
388) The UV spectrum of benzoic acid in ethanol has benzenoid bands at 202 nm (ε 8000) and 279 nm (ε 550), and an E.T. band at 228 nm (ε 10,000).[376]
389) S.E. Hunt, J.I. Jones and A.S. Lindsey, *J. Chem. Soc.*, 1956, 3099.
390) E.A. Braude, *Ann. Reports Chem. Soc.*, 1945, **42**, 105.
391) D. Marshall and M.C. Whiting, *J. Chem. Soc.*, 1957, 537.
392) A.L. Sklar, *J. Chem. Phys.*, 1937, **7**, 339, 353.
393) A.L. Sklar, *J. Chem. Phys.*, 1942, **10**, 135.
394a) F.A. Matsen, *J. Am. Chem. Soc.*, 1950, **72**, 5243.
394b) W.W. Robertson and F.A. Matsen, *J. Am. Chem. Soc.*, 1950, **72**, 5248, 5250, 5252.
394c) J.F. Music and F.A. Matsen, *J. Am. Chem. Soc.*, 1950, **72**, 5256.
394d) W.W. Robertson, J.F. Music and F.A. Matsen, *J. Am. Chem. Soc.*, 1950, **72**, 5260.
395) J.R. Platt, *J. Chem. Phys.*, 1951, **19**, 263.
396) J. Petruska, *J. Chem. Phys.*, 1961, **34**, 1120.
397) C.N.R. Rao, *J. Sci. Ind. Research (India)*, 1958, **17B**, 56.
398) C.N.R. Rao, *Curr. Sci. (India)*, 1957, **26**, 276.
399) C.N.R. Rao, *Chem. & Ind.*, 1956, 666.
400) C.N.R. Rao, *Chem. & Ind.*, 1957, 1239.
401) C.N.R. Rao and co-workers, *Proc. 10th Intern. Conf. Spectroscopy*, Maryland, 1962, Spartan Press, Washington, D.C., 1963.
402) C.N.R. Rao, J. Ramachandran and A. Balasubramanian, *Canad. J. Chem.*, 1961, **39**, 171. See also H.H. Jaffe, *J. Chem. Phys.*, 1954, **22**, 1430.
403) C.N.R. Rao, *J. Sci. Industr. Research (India)*, 1958, **17B**, 89.
404) E. Lieber, C.N.R. Rao, J. Ramachandran and C.N. Pillai, *Canad. J. Chem.*, 1959, **37**, 563.
405) J.E. Bloor and D.B. Copley, *Chem. & Ind.*, 1960, 526.
406) O. Exner, M. Horak and J. Pliva, *Chem. & Ind.*, 1958, 1174.

407) W.F. Forbes and W.A. Mueller, *Canad. J. Chem.*, 1955, **33**, 1145.
W.F. Forbes and M.B. Sheratte, *Canad. J. Chem.*, 1955, **33**, 1829.
408) W.F. Forbes and W.A. Mueller, *Canad. J. Chem.*, 1956, **34**, 1340.
W.F. Forbes and A.S. Ralph, *Canad. J. Chem.*, 1956, **34**, 1447.
409) W.F. Forbes and W.A. Mueller, *Canad. J. Chem.*, 1957, **35**, 488.
W.F. Forbes, W.A. Mueller, A.S. Ralph and J.F. Templeton, *Canad. J. Chem.*, 1957, **35**, 1049.
410) W.F. Forbes, *Canad. J. Chem.*, 1958, **36**, 1350.
J.C. Dearden and W.F. Forbes, *Canad. J. Chem.*, 1958, **36**, 1362.
W.F. Forbes and I.R. Leckie, *Canad. J. Chem.*, 1958, **36**, 1371.
411) W.F. Forbes *et al*, *Canad. J. Chem.*, 1959, **37**, 1294, 1305.
W.F. Forbes, *Canad. J. Chem.*, 1959, **37**, 1977.
412) W.F. Forbes, *Canad. J. Chem.*, 1960, **38**, 1104.
413) W.M. Schubert, J. Robins and J.L. Haun, *J. Am. Chem. Soc.*, 1957, **79**, 910.
414) W.M. Schubert, J.M. Craven and H. Steadly, *J. Am. Chem. Soc.*, 1959, **81**, 2695.
415) W.M. Schubert, H. Steadly and J.M. Craven, *J. Am. Chem. Soc.*, 1960, **82**, 1353.
W.M. Schubert and J.M. Craven, *J. Am. Chem. Soc.*, 1960, **82**, 1357.
416) V. Baliah and M. Uma, *Naturwiss.*, 1958, **45**, 212.
417) W.M. Schubert and W.A. Sweeney, *J. Org. Chem.*, 1956, **21**, 119 and references cited therein.
418) W.M. Schubert and J. Robins, *J. Am. Chem. Soc.*, 1958, **80**, 559.
419) L. Doub and J.M. Vandenbelt, *J. Am. Chem. Soc.*, 1955, **77**, 4435.
420) D.P. Stevenson and H.M. McConnell, *Spectrochim. Acta*, 1958, **12**, 262.
421) L. Goodman and H. Shull., *J. Chem. Phys.*, 1957, **27**, 1388.
422) G.M. Badger, *The Stucture and Reactions of Aromatic Compounds*, Cambridge University Press, 1954.
423) Naphthalene, anthracene, naphthacene and pentacene absorb at 314 nm (log ε 2.50), 380 nm (log ε 3.90), 480 nm (log ε 4.05), and 580 nm (log ε 4.10) respectively.[424]
424) R.A. Friedel and M. Orchin, *Ultraviolet Spectra of Aromatic Compounds*, Wiley, New York, NY, 1951.
425) R.N. Jones, *Chem. Rev.*, 1947, **41**, 353.
426) E. Heilbronner and J.N. Murrell, *J. Chem. Soc.*, 1962, 2611.
427) S.F. Mason, *Quart. Revs. (London)*, 1961, **15**, 287.
428) M. Gordon, *Chem. Rev.*, 1952, **50**, 185 and references cited therein.
429) E. Heilbronner in *Non-Benzenoid Aromatic Compounds*, G. Ginsburg, editor, Interscience, New York, NY, 1959, *Chapter V.*
430) E. Heilbronner and R. Gerdil, *Helv. Chim. Acta*, 1956, **39**, 997.
431) J.E. Rice, B. Liu, T.J. Lee and C.M. Rohlfing, *Chem. Phys. Lett.*, 1989, **161**, 277.
432) I.L. Alberts and H.F. Schaefer, III, *Chem. Phys. Lett.*, 1989, **161**, 375.
433) H. Suzuki, *Electronic Absorption Spectra*, Academic Press, New York, N.Y., 1967, p. 204.
434) H. Suzuki, *Electronic Absorption Spectra*, Academic Press, New York, N.Y., 1967, p. 365.
435) J.E. Lennard-Jones, *Proc. R. Soc. London, Ser. A*, 1937, **158**, 280.
436a) D. Désilets, P.M. Kazmaier and R.A. Burt, *Canad. J. Chem.*, 1995, **73**, 319.
436b) D. Désilets, P.M. Kazmaier, R.A. Burt and G.K. Hamer, *Canad. J. Chem.*, 1995, **73**, 325.
437a) J. Fabian, H. Nakazumi and M. Matsuoka, *Chem. Rev.*, 1992, **92**, 1197.
437b) P.F. Gordon in *The chemistry and application of dyes*, D.R. Waring and G. Hallas, editors, *Topics in applied chemistry*, A.R. Katrizky and G.J. Sabongi, editors, Plenum Press, New York, NY, 1990, *Chapter 8.*
437c) M. Matsuoka in *Colour chemistry, the design and synthesis of organic dyes and pigments*, A.T. Peters and H.S. Freeman, editors, *Advances in colour chemistry*, Elsevier Applied Science, London, 1991, *Chapter 3.*

437d) P. Gregory in *Colour chemistry, the design and synthesis of organic dyes and pigments*, A.T. Peters and H.S. Freeman, editors, *Advances in colour chemistry*, Elsevier Applied Science, London, 1991, *Chapter 9*.
438a) R.O. Loutfy, A.M. Hor, P. Kazmaier and M. Tam, *J. Imaging Sci.*, 1989, **33**, 151.
438b) R.O. Loutfy, A.M. Hor, C.K. Hsaio, G. Baranyi and P. Kazmaier, *Pure Appl. Chem.*, 1988, **60**, 1047.
438c) K.C. Nguyen and D.S. Weiss, *Denshi Shashin Gakkaishi*, 1988, **27**, 2.
438d) K.C. Nguyen and D.S. Weiss, *J. Imaging Sci.*, 1989, **15**, 158.
439a) A. Kral and H. Laatsch, *Z. Naturforsch, Teil B*, 1993, **48**, 1401.
439b) C. Göltner and H. Laatsch, *Liebigs Ann. Chem.*, 1991, 1085.
440) I. Baraldi, A. Carnevali, F. Momicchioli and G. Ponterini, *Spectrochim. Acta*, 1993, **49A**, 471.
441) L. Antonov and S. Stoyanov, *Dyes and Pigments*, 1995, **28**, 31.
442) G. Hallas and A.D. Towns, *Dyes and Pigments*, 1997, **33**, 205.
443) K.A. Bello, *J. Soc. Dyers & Colourists*, 1997, **113**, 361.
444) R. Naef, *Dyes and Pigments*, 1994, **24**, 83.
445) M.G. Hutchings, *Dyes and Pigments*, 1995, **29**, 95.
446) J. Lipinski and W. Bartkowiak, *J. Phys. Chem.*, 1997, **101A**, 2159.
447) K.Y. Burshtein, A.A. Bagaturyants and M.V. Alfimov, *Russian Chemical Bulletin*, 1997, **46**, 62.
448) K. Ando, *J. Chem. Phys.*, 1997, **107**, 4585.
449) S.F. Alberti and J. Echave, *Chem. Phys.*, 1997, **223**, 183.
450) S. Marguet, J.C. Mialocq, P. Millie, G. Berthier and F. Momicchioli, *Chem. Phys.*, 1992, **160**, 265.
451) D. Markovitsi, H. Sigal, C. Ecoffet, P. Millie, F. Charra and C. Fiorini, *Chem. Phys.*, 1994, **182**, 69.
452) S.T. Abdelhalim and M.K. Awad, *J. Phys. Chem.*, 1993, **97**, 3160.
453) G. Grimme, S. Grimme, P.G. Jones and P. Boldt, *Chem. Ber.*, 1993, **126**, 1015.
454) L. Serrano Andres and B.O. Roos, *Chemistry – A European Journal*, 1997, **3**, 717.
455) S. Dähne, *Proc. Indian Acad. Sci. (Chem. Sci.)*, 1992, **104**, 311. Dyes having desired light absorption can be designed on the basis of five simple, generally valid colour rules within the framework of triad theory:
a) Molecules in the ideal polymethine state have the longest wavelength absorption and (apart from exceptional cases of forbidden transitions), the most intense absorption meaning that they are deepest in colour in comparison to aromatic or polyenic molecules of comparable size.
b) If the characteristic π-electron density alternation in polymethines is maintained on structural modifications, a slight blue shift occurs.
c) If the characteristic π-electron density alternation in polymethines is perturbed on structural modifications, a strong shift of the light absorption occurs.
d) Aromatisation within molecules originates a strong blue shift.
e) Polyene-formation in consequence of the nonsymmetric localisation of the π-electron system causes a moderate blue shift.
All deviations from these five basic colour rules can be explained by second order effects such as *cis-trans* isomerisation, steric hindrances, intermolecular interactions, *etc.*
456) O.N. Witt, *Ber. Dtsch. Chem. Ges.*, 1876, **9**, 522; 1888, **21**, 325.
457) R. Wizinger, *Ber. Dtsch. Chem. Ges.*, 1927, **60**, 1377.
458) R. Wizinger, *Symposium over Kleur en Structur van organische Verbindingen*, Antwerpen, 1956; Vlaamse Chemische Vereniging, Brussels, 1956.
459) W. König, *J. Prakt. Chem.*, 1926, **112**, 1.
460) The replacement of the X, X' atoms by heteroatoms while maintaining the polymethine electronic structure **(123)** is known as isologous substitution. Known examples with (p+1) π electrons to p chain atoms are: cyanines – N...N, oxonols –

O...O, merocyanines – O...N (generally meropolymethines for **(123)**, X≠X'), carbon polymethines – C...C (particularly carbeniate polymethines with $(p + 1)$ π electrons / p atoms and carbonium polymethines with $(p - 1)$ π electrons / p atoms).

461) H. Kuhn, *Chimia*, 1995, **9**, 237; *Angew. Chem.*, 1959, **71**, 93.
462) D. Leupold, *Z. Physik. Chem.*, 1963, **223**, 405.
463) G. Scheibe, W. Seiffert, H. Wengenmayr and C. Jutz, *Ber. Bunsenges. Physik. Chem.*, 1963, **67**, 650.
464a) S. Dähne and J. Ranft, *Z. Physik. Chem.*, 1963, **224**, 65; 1966, **232**, 259; *Angew. Chem.*, 1963, **75**, 1175; *Angew. Chem. Int. Ed. Engl.*, 1963, **2**, 740.
464b) S. Dähne, D. Leupold, H.-E. Nikolajewski and R. Radeglia, *Z. Naturforsch.*, 1965, **20b**, 1006.
464c) R. Radeglia and S. Dähne, *Ber. Bunsenges. Physik. Chem.*, 1966, **70**, 745.
465) A. von Bayer, *Liebigs Ann. Chem.*, 1907, **354**, 152.
466) R. Willstatter and J. Piccard, *Ber. Dtsch. Chem. Ges.*, 1908, **41**, 1467.
467) S. Dähne, *Z. Chem.*, 1965, **5**, 444.
468) S. Dähne and D. Leupold, *Ber. Bunsenges. Physik. Chem.*, 1966, **70**, 618.
469) L.G.S. Brooker, *Rev. Mod. Physics*, 1942, **14**, 289 and in C.E.K. Mees, *The Theory of the Photographic Process*, MacMillan, 2nd edition, New York, NY, 1954, p. 371-429.
470a) S. Dähne and H. Paul, *Contribution at the 2nd International Color Symposium*, Schloß Elmau (Germany), April 1964.
470b) S. Dähne and H. Paul, *Chem. Ber.*, 1964, 97, 1625.
470c) S. Dähne, J. Ranft and H. Paul, *Tetrahedron Lett.*, 1964, 3355.
470d) D. Leupold and S. Dähne, *Theor. Chim. Acta*, 1965, **3**, 1; *Mber. Dtsch. Akad. Wiss. Berlin*, 1965, **7**, 46.
470e) S. Kulpe, D. Leupold and S. Dähne, *Angew. Chem.*, 1966, **78**, 639; *Angew. Chem. Int. Ed. Engl.*, 1966, **5**, 599.
471) A. Streitweiser, *Molecular Orbital Theory for Organic Chemists*, Wiley, London, 1961.
472) S. Dähne and D. Leupold, *Z. Naturforsch.*, 1965, **20b**, 76; D. Leupold and S. Dähne, *J. Mol. Spectrosc.*, 1965, **17**, 325.
473) S. Dähne and D. Leupold, *Angew. Chem. Int. Ed. Engl.*, 1966, **5**, 984.
474) D. Buckley, H.B. Henbest and P. Slade, *J. Chem. Soc.*, 1957, 4891.
475) R.A. Morton and W.T. Earlam, *J. Chem. Soc.*, 1941, 159.
476) W. Lüttke and M. Klessinger, *Chem. Ber.*, 1964, **97**, 2342; 1966, **99**, 2136.
477) S. Dähne, *Z. Chem.*, 1965, **5**, 441.
478) S. Dähne and D. Leupold, *Angew. Chem.*, 1966, **78**, 1029.
479) S. Dähne, *Ber. Bunsenges. Phys. Chem.*, 1966, **70**, 618.
480) J. Fabian and H. Hartmann, *J. Signalaufzeichnungsmaterialien*, 1974, **2**, 457; *Theor. Chim. Acta*, 1975, **36**, 351.
481) J. Fabian and H. Hartmann, *J. Mol. Struct.*, 1975, **27**, 67.
482) S. Dähne, *Wiss. Z. Tech. Univ. Dresden*, 1971, **20**, 671.
483) R. Zahradník, *Fortschr. Chem. Forsch.*, 1968, **10**, 1.
484) H.-E. Nikolajewski, S. Dähne, B. Hirsch and E. Jauer, *Angew. Chem.*, 1966, **78**, 1063; J. Kucera and Z. Arnold, *Collect. Czech. Chem. Commun.*, 1967, **32**, 1704; D. Lloyd and H. Mc Nab, *Angew. Chem.*, 1976, **88**, 495.
485) That is, a maximum absorption intensity; exceptions may occur in special cases if transitions are more or less forbidden for reasons of symmetry [*e.g*, see D. Leupold and S. Dähne, *J. Mol. Spectrosc.*, 1965, **17**, 325].
486) D. Meuche, H. Strauss and E. Heilbronner, Helv. Chim. Acta, 1958, **41**, 57; G. Naville, H. Strauss and E. Heilbronner, *Helv. Chim. Acta*, 1960, **43**, 1221, 1243.
487) J. Koutecký, P. Hochman and J. Michl, *J. Chem. Phys.*, 1964, **40**, 2439.
488) J. Koutecký and J. Paldus, *Collect. Czech. Chem. Commun.*, 1963, **28**, 1483.
489) J. Koutecký, J. Paldus and R. Zahradník, *J. Chem. Phys.*, 1962, **36**, 3129.

490a) P. Hochmann, J. Koutecký, É Szabo and J. Fišer, *Collect. Czech. Chem. Commun.*, 1966, **31**, 2287.
490b) W.T. Simpson, *J. Am. Chem. Soc.*, 1951, **73**, 5363; 1955, **77**, 6164.
490c) L.S. Forster, *Theor. Chim. Acta*, 1966, **5**, 81.
491) J. Michl and E.W. Thulstrup, *Tetrahedron*, 1978, **32**, 205.
492) P.A. Clark, F. Brogli and E. Heilbronner, *Helv. Chim. Acta*, 1972, **55**, 1415.
493) R.S. Becker and E. Chen, *J. Chem. Phys.*, 1966, **45**, 2403.
494) J.B. Birks, *Photophysics of Aromatic Molecules*, Wiley-Interscience, New York, NY, 1970.
495) R.H. Huebner, W.F. Frey and R.N. Compton, *Chem. Phys. Lett.*, 1973, **23**, 587.
496) P.M. Rentzepis, *Chem. Phys. Lett.*, 1969, **3**, 717.
497) R. Zahradník, I. Tesarová and J. Pancir, *Collect. Czech. Chem. Commun.*, 1971, **36**, 2867.
498a) A. Klamt, *J. Phys. Chem.*, 1996, **100**, 3349. Geometry changes of the solute in different solvents were taken into account for the first time using the Conductor-like screening solvation model (COSMO), and appear to be of considerable importance.
498b) A.V. Yatsenko, V.A. Tafeenko, K.A. Pasechnichenko, T.V. Stalnaya and S.L. Solodar, *Zeitschrift fur Kristallographie*, 1998, **213**, 466. Changes in molecular geometries due to intermolecular interactions were used to interpret spectral properties (AM1 and INDO/S) of 9-amino-1H-phenalen-1-one, *N*-(1-oxo-1H-phenalen-9-yl)-acetamide and 9-(methylamino)-1H-phenalen-1-one-,1,4-dioxan-2-yl hydroperoxide solvate, in particular the decreasing Stokes shift of 6-aminophenalenone with increasing solvent polarity.
498c) V. Buss and K. Kolster, *Chem. Phys.*, 1996, **203**, 309. CNDO/S calculations including configuration interaction of up to 300 singly excited states have been performed on MMP2 and PM3 optimised geometries of [*n*]-helicenes (n = 3 to 9). Oscillator and rotatory strengths were calculated for all states down to 200 nm, in generally good agreement for long wavelength regions.
498d) J. Dods, D. Gruner, P. Brumer, *Chem. Phys. Lett.*, 1996, **261**, 612. Discrepancies between experimental and theoretical fluroescence/absorption spectra are minimised using a genetic algorithm to shift the calculated excited electronic surface position. The method depends on the extreme sensitivity of polyatomic Franck-Condon factors to the geometrical shift that molecules undergo upon electronic excitation and on the power of genetic algorithms to rapidly locate the required origin shift.
498e) J. Cornil, A.J. Heeger and J.L. Bredas, *Chem. Phys. Lett.*, 1997, **272**, 463. Intermolecular interactions strongly affect absorption and emission properties characteristic of isolated conjugated chains. Emission properties of conjugated systems for electrooptic devices are significantly modified when geometry relaxation phenomena occur in the lowest excited state.
498f) P.L. Muino and P.R. Callis, *J. Chem. Phys.*, 1994, **100**, 4093. Solute-solvent interactions and dynamics were simulated with a fully molecular hybrid method consisting of a semi-empirical method with singly excited configurations for the solute and classical MD for the solvent.
499) K. Pierloot, J.O.A. DeKerpel, U. Ryde and B.O. Roos, *J. Am. Chem. Soc.*, 1997, **119**, 218.
500) D.W. Turner, *Adv. Phys. Org. Chem.*, 1966, **4**, 31 and references cited therein.
501) P.A. Cox, S. Evans, A.F. Orchard, N.V. Richardson and P.J. Roberts, *Faraday Discussions of the Chemical Society*, 1972, **No 54**, 26.
502) K.D. Jordan and P.D. Burrow, *Acc. Chem. Res.*, 1978, **11**, 341.
503) R.S. Ruoff, K.M. Kadish, P. Boulas and E.C.M. Chen, *J. Phys. Chem.*, 1995, **99**, 8843.
504) J.R. Wiley, E.C.M. Chen and W.E. Wentworth, *J. Phys. Chem.*, 1993, **97**, 1256.
505) E.J. Chen and R.S. Becker, *J. Chem. Phys.*, 1966, **45**, 240.
506) For recent tabulations of ionisation potentials and electron affinities, see:

S.G. Lias, J.E. Bartmess, J.F. Liebman, J.L. Holmes, R.D. Levin and W.G. Mallard, *J. Phys. Chem. Ref. Data*, 1988, **17** (S1), 1.
507) W.E. Wentworth and R.S. Becker, *J. Am. Chem. Soc.*, 1962, **84**, 4263.
508) R.S. Becker and W.E. Wentworth, *J. Am. Chem. Soc.*, 1963, **85**, 2210.
509) W.E. Wentworth, E.C.M. Chen and J.E. Lovelock, *J. Phys. Chem.*, 1966, **70**, 445.
510) E.C.M. Chen, J.R. Wiley, C.F. Batten and W.E. Wentworth, *J. Phys. Chem.*, 1994, **98**, 88.
511) E.K. Fukada and R.T. McIver, *J. Am. Chem. Soc.*, 1985, **107**, 2291.
512) E.P. Grimsud, G. Caldwell, S. Chowdhury and P. Kebarle, *J. Am. Chem. Soc.*, 1985, **107**, 4627.
513) P. Kebarle and S. Chowdhury, *Chem. Rev.*, 1987, **87**, 513.
514) L. Crocker, T. Wang and P. Kebarle, *J. Am. Chem. Soc.*, 1993, **115**, 7818.
515) J.T. Sprague, J.C. Tai, Y. Yuh and N.L. Allinger, *J. Comput. Chem.*, 1987, **8**, 581.
516) C. Baird, *Environmental Chemistry*, Freeman, New York, NY, 1995.
517) S. Madronich and F.R. de Gruijl, *Nature*, 1993, **366**, 23.
518) J.-S. Taylor, *J. Chem. Educ.*, 1990, **67**, 835.
519) A.C. Guyton and J.E. Hall, *Textbook of Medical Physiology*, 9th edition, Saunders, Philadelphia, 1996.
520) J.M. Wilmott, M.C. Duggan and A.P. Znaidin in *Physician's Guide to Sunscreens*, N.J. Lowe, editor, Dekker, New York, NY, 1991, *Chapter 6*.
521) A.I. Scott, *Interpretation of the Ultraviolet Spectra of Natural Products*, Macmillan, New York, NY, 1964.
522) G.H. Pittet and S.A. Givaudan, *Drug Cosmetic Ind.*, 1988, **143**, 24.
523) S.H. Dromgoole and H.I. Maibach in *Physician's Guide to Sunscreens*, N.J. Lowe, editor, Dekker, New York, NY, 1991, *Chapter 10*.
524) N.J. Lowe in *Physician's Guide to Sunscreens*, N.J. Lowe, editor, Dekker, New York, NY, 1991, *Chapter 11*.
525a) N.S. Isaacs, *Physical Organic Chemistry*, Wiley, New York, NY, 1987.
525b) T.H. Lowry and K.S. Richardson, *Mechanism and Theory in Organic Chemistry*, 3rd edition, Harper & Row, New York, NY, 1987.
526) M.R. Davis and M.N. Quigley, *J. Chem. Educ.*, 1995, **72**, 279.
527) C. Walters, A. Keeney, C.T. Wigal, C.R. Johnston and R.D. Cornelius, *J. Chem Educ.*, 1997, **74**, 99.
528) *Aldrich Catalogue / Handbook of Fine Chemicals*, Aldrich Chemical Co. Ltd., Gillingham, 1992-93. Stocks of dyes, *e.g.* light green SF yellowish can become limited.
529) The experimental wavelength is for eosin Y, X = Br.
530) Approximate value only as ponceau de xylidene is an obsolete dye.
531) F.H. Allen and O. Kennard, *Chemical Design Automation News*, 1993, **8**, 130.
532) D.A. Fletcher, R.F. Meeking and D. Parkin, *J. Chem. Int. Comput. Sci.*, 1996, **36**, 746.
533) I. Willner, Y. Eichen, M. Rabinovitz, R. Hoffman and S. Cohen, *J. Am. Chem. Soc.*, 1992, **114**, 637.
534) C. de Martinis and M.F. Mackay, *Acta Crystallogr.*, 1980, **36B**, 1606.
535) W.H. Ojala, L.K. Lu, K.E. Albers, W.B. Gleason, T.I. Richardson, R.E. Lovrien and E.A. Sudbeck, *Acta Crystallogr.*, 1994, **50B**, 684.
536) W.H. Ojala, E.A. Sudbeck, LK. Lu, T.I. Richardson, R.E. Lovrien and W.B. Gleason, *J. Am. Chem. Soc.*, 1996, **118**, 2131.
537) Available *via* anonymous ftp from *joplin.biosci.arizona.edu* in *pub/Babel*
538) Acridine orange is used:
a) As a flurorescent dye for cytological staining – J.F. Golden and S.S. West, *J. Histochem. Cytochem.*, 1974, **22**, 495.
b) As a DNA intercalcating dye – A.D. Wolfe, R.G. Allison and F.E. Hahn, *Biochemistry*, 1972, **11**, 1569.
c) For differentiating known tumour cells from normal cells and sputum.

i) J.F. Golden, S.S. West, H.M. Shingleton, T.M. Murad and C.K. Echols, *J. Histochem. Cytochem.*, 1979, **27**, 522.

ii) J.K. Frost, H.W. Tyrer, N.J. Pressman, C.D. Albright, M.H. Vansickel and G.W. Gill, *J. Histochem. Cytochem.*, 1979, **27**, 545.

iii) H.W. Tyrer, J.F. Golden, M.H. Vansickel, C.K. Echols, J.K. Frost, S.S. West, N.J. Pressman, C.D. Albright, C.A. Adams and G.W. Gill, *J. Histochem. Cytochem.*, 1979, **27**, 552.

d) For staining semi-thin sections of plant material – R.N. Paul, *Stain Technol.*, 1980, **55**, 195.

539) Alizarin red S is used:

a) For gross staining of skeletons in small invertebrates (especially embryos).

b) For differentiating bone from cartilage in mammalian embryos.

c) As an acid-base indicator and in the analysis of metals, *e.g.* for hafnium (IV), gallium (III) and indium (III)

540) Auramine O is used for fluorescent staining of acid-fast bacteria in sputum or in paraffin sections of infected tissue.

541) Azure A is used as a constituent of MacNeal's Tetrachrome stain for differentiating leucocytes.

542) Azure B is used for differentiating cellular RNA and DNA in plant tissues – M.H. Flax and M.H. Hines, *Physiol. Zool.*, 1952, **25**, 297.

543) Brilliant cresyl blue ALD is used in the reticulum of immature red blood cells and platelets using the Wright stain as a counterstain.

544) Brilliant green is used in the determination of:

i) the presence of colon organisms in drinking water (also a bacteriostatic agent),

ii) fatty acids by ion-pair chromatography – J. Di Nunzio, H. Freiser, *Talanta*, 1979, **26**, 587.

545) Congo red is used:

a) For staining cytoplasm, erythrocytes, spindle fibres and collagen.

b) For showing amyloid in pathological human tissue.

c) As a useful probe for the conformation of nucleotide-binding enzymes – R.A. Edwards and R.W. Woody, *Biochemistry*, 1979, **18**, 5197.

546) Cresyl violet acetate is used:

a) For staining Nissl substances in nerve cells – M.M. Powers and G. Clark, *Stain Technol.*, 1955, **30**, 83.

b) For staining fresh tumour tissue for biopsy.

547) Malachite green oxalate is used:

a) With methyl green or toluidine blue for differentiating between *diptheria* and other bacteria.

b) As a reagent for bacterial polysaccharides.

c) As a contrast stain for *gonococci* and *meningococci* in smears.

548) Methylene blue is used:

a) As a stain for elastic fibres and connective tissue – M.A. Darrow, *Stain Technol.*, 1952, **27**, 329.

b) As a counterstain after Basic Fuchsin for staining *tubercle* and *lepra bacilli* in mammalian tissue.

c) As a component of MacNeal's Tetrachrome stain for differentiation of blood corpuscles.

549) Methyl green is used:

a) With pyronin B and Y for staining *gonococcus* and mast cells – R. Virchows, *Arch. Pathol. Anat.*, 1899, **157**, 19.

b) As a stain for differentiating between *diptheria* and other bacteria.

c) As a tissue stain for mitochondria and nucleoli in animal tissue.

550) Neutral red is used with Janus Green B for embryological tissues.

551) Pararosaniline chloride is used for staining *bacilli* (acid-fast using methylene blue as a counterstain), especially *influenza* and *tubercle* in tissue.

552) Pyronin Y is used:

a) With methyl green for staining basophil elements such as plasma cells – E.B. Taft, *Stain Technol.*, 1951, **26**, 205.

b) For staining bacteria in mammalian tissue – *Pathological Technique*, W.B. Saunders Co., Philadelphia, PA, 1938, p. 271.

553) Safranin O is used:

a) As a cationic lipophilic probe – P.X. Petit, J.E. O'Connor, D. Grunwald and S.C. Brown, *Eur. J. Biochem.*, 1990, **194**, 389.

b) In the detection of glycosaminoglycans and proteoglycans – D.A. Carrino, J.L. Arias and A.I. Caplan, *Biochem. Int.*, 1991, **24**, 485.

c) As a general stain for showing nuclei, cellulose walls and chromosomes of eristematic tissue, and with picric acid and aniline blue for differentiating between lignified and unlignified cellulose walls in woody tissue.

554) Sudan III is used as a stain for fats in animals, and for differentiating suberised and cutinised tissue in plants.

555) Sudan black B is used:

a) As a fat stain for animal tissues, and for bacteria and other micro-organisms.

b) As a stain for chromosomes, Golgi apparatus and leucocyte granules – H.N. Subramaniam and K.A. Chaubal, *J. Biochem. Biophys. Methods*, 1990, **21**, 9.

556) Thionin is used:

a) As a general nuclear stain for chromatin and mucin – R.D. Lillie, *Histopathologic Technic and Practical Histochemistry*, The Blakiston Co., New York, NY, 1954, p. 284.

b) For staining Nissl granules by endpoint buffering.

c) For showing malignant cells in frozen biopsy sections – *Pathological Technique*, W.B. Saunders Co., Philadelphia, PA, 1938, p. 35.

d) For staining semi-thin sections of plant material – R.N. Paul, *Stain Technol.*, 1980, **55**, 195.

557) Toluidine blue O is a metachromatic nuclear stain which is a component of a tribasic stain along with malachite green and basic fuchsin – P.M. Grimley, *Stain Technol.*, 1964, **39**, 229.

558) W. Bradley and F.W. Pexton, *J. Chem. Soc.*, 1954, 4436.

559) See, among others:

a) F.A. Matsen, *J. Chem. Phys.*, 1956, **24**, 602.

b) G.W. Wheland, *J. Am. Chem. Soc.*, 1941, **63**, 2025.

560) G.W. Wheland and D.E. Mann, *J. Chem. Phys.*, 1949, **17**, 264 assume that the increase in α, expressed in units of β, is proportional to the positive charge $(1-q)\beta$ on the carbon atom, where q is the relative negative charge on carbon.

561) S.E. Ehrenson, *J. Phys. Chem.*, 1961, **66**, 706.

562a) A. Streitwieser and P. Nair, *Tetrahedron*, 1959, **5**, 149.

562b) A. Streitweiser, *J. Am. Chem. Soc.*, 1960, **82**, 4123.

562c) Hoijtink's calculations,[562d-f] based on the one-electron half-wave reduction potential of aromatic hydrocarbons, in combination with Hückel's equation (*Equation 5.3*) and Wheland's f' values,[560] gave $\beta = -2.26$ eV.
See also:

562d) P. Balk, S. de Bruyn and G.J. Hoijtink, *Recl. Trav. Chim. Pays-Bas*, 1957, **76**, 860.

562e) G.J. Hoijtink, *Recl. Trav. Chim. Pays-Bas*, 1955, **74**, 1525.

562f) G.J. Hoijtink, E. de Boer, P.H. van der Meij and W.P. Weijland, *Recl. Trav. Chim. Pays-Bas*, 1956, **75**, 487.

563) *Equation 5.8* can be fitted to experimental data only with limited accuracy as ionisation energies are known to within an accuracy of 0.1 to 0.2 eV for only a few aromatic hydrocarbons. Moreover, often the first excitation energy of aromatic

hydrocarbons cannot be determined with great accuracy. This energy is determinined by taking the maximum of the longest wavelength absorption band as ν_0, or by basing the calculation on the 0,0-frequency, ν_{00}, provided the latter can be identified.[566]

564) E. Clar, *Aromatische Kolenwasserstoffe*, Springer, Berlin, 1941.

565) R.S. Becker and W.E. Wentworth, *J. Am. Chem. Soc.*, 1963, **85**, 4263.

566) The 0,0-frequency is the excitation energy required to raise an electron from its zero vibrational energy level in the ground state to the zero vibrational level in the excited state.

567) G. Briegleb, *Angew. Chem. Int. Ed. Engl.*, 1964, **3**, 617-632.

568a) J.A. Pople, *Trans. Faraday Soc.*, 1953, **49**, 375.

568b) A. Brickstock and J.A. Pople, *Trans. Faraday Soc.*, 1954, **50**, 901.

568c) J.A. Pople, *J. Phys. Chem.*, 1957, **61**, 6.

568d) C.J. Roothaan, *Rev. Mod. Phys.*, 1951, **23**, 61.

569) $$F = 2 \sum_{\mu < \nu} \chi_{N\mu}\, \chi_{N\nu}\, (\beta_{\mu\nu} - \gamma_{\mu\nu} \sum_{i=1}^{N} \chi_{i\mu}\, \chi_{i\nu})$$

where $\gamma_{\mu\nu}$ is the Coulomb interaction between one electron at atom μ and one electron at atom ν, $\beta_{\mu\nu}$ is the resonance integral for two adjacent atoms μ and ν, N is the number of occupied orbitals in the aromatic hydrocarbon, $\chi_{i\mu}$ and $\chi_{i\nu}$ are the Hückel coefficients of the μ^{th} and ν^{th} atomic orbitals in the i^{th} molecular orbital. The χ-values can be calculated with good approximation using simple MO theory.

570) Antisymmetrised Molecular Orbital Self-Consistent.

571a) R.M. Hedges and F.A. Matsen, *J. Chem. Phys.*, 1958, **28**, 950.

571b) A. Streitweiser and P.M. Nair, *Tetrahedron*, 1959, **5**, 149.

571c) A. Streitweiser, *J. Am. Chem. Soc.*, 1960, **82**, 4123.

571d) W.F. Wolf, *Abstracts J. Am. Chem. Soc. Meeting*, New York, 1960.

572) J.R. Hoyland and L. Goodman, *J. Chem. Phys.*, 1962, **36**, 12, 21.

573) E.W. Wentworth and R.S. Becker, *J. Am. Chem. Soc.*, 1962, **84**, 4263 using electron absorption coefficients, χ,[573a] for the capture of thermal electrons e^- by a few organic molecules, A, determined experimentally by Lovelock.[573b] Under experimental conditions for which $[e^-] \gg [A]$, the absolute electron absorption coefficient, χ, becomes identical with the equilibrium constant K_A.[573c,d] The electron affinity can be calculated by setting $g_A^-/g_A = 2$, and hence $g_A^-/g_A.g_{e^-} = 1$.

573a) $[e^-] = [e^-]_0 \exp(-\chi[A])$

573b) J.E. Lovelock and R.S. Lipsky, *J. Am. Chem. Soc.*, 1960, **82**, 431.
J.E. Lovelock, *Nature (London)*, 1960, **187**, 49; **188**, 401; 1961, **189**, 729.

573c) $K_A = (g_A^-/g_A.g_{e^-}) \exp(E_A/RT) = [A^-]/[A][e^-]$

573d) $[e^-]$ is the stationary electron concentration at equilibrium, $[e^-]_0$ is the electron concentration before approach of acceptor A to electron gas, and g is the number of realisable quantum states. For e^-, $g = 2$ ($s = \pm\frac{1}{2}$) - T.L. Hill, *Introduction to Statistical Thermodynamics*, Addison-Wesley Publishing Co., Reading, M.A., 1960. Expression 573a) can be approximated by:
$[e^-] = [e^-]_0 - [A^-] = [e^-]_0 (1 - \chi[A])$.
So for the case where $[e^-] \gg [A]$ or $[A^-] \ll [e^-]_0$:
$\chi = [A^-]/[e^-]_0[A] \approx [A^-]/[e^-][A]$.

574) S.E. Ehrenson, *J. Phys. Chem.*, 1961, **66**, 706, 712.

575) S.E. Ehrenson[574] uses $\omega = 3.5$.

576) D.R. Scott and R.S. Becker, *J. Phys. Chem.*, 1962, **66**, 2713.

577) H.C. Longuet-Higgins and J.A. Pople, *Proc. Phys. Soc.*, 1955, **68**, 591.

578a) M.J.S. Dewar and S. Olivella, *J. Chem. Soc., Faraday Trans. II*, 1979, **75**, 829.

578b) Used in the MOPAC and AMPAC series of programs

578c) J.J.P. Stewart, *QCPE Bull.*, 1990, **10**(*4*), 86.

578d) Dewar Research Group, *QCPE Bull.*, 1987, **7**(*2*), 85.

579) N.J. Turro, *Modern Molecular Photochemistry*, University Science Books, Mill Valley, CA, 1991, p.292 – 293.
580) J.B. Birks in *Organic Molecular Photophysics*, Volume 1, J.B. Birks, editor, John Wiley & Sons, London, 1973, Chapter 1. Values are for zero-point excitation energies.
581) Personal communication with H.S. Rzepa.
582) L. Brillouin, *Actualities Sci. Ind.*, 1934, **71**, 59.
583) *HyperChem for Windows Reference Manual*, Hypercube Inc., Canada, April 1994.
584) H.S. Rzepa and G.A. Suñer, *J. Chem. Soc., Perkin Trans. 2*, 1994, 1397.
585) A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans 2*, 1993, 799; J. Andzelm, C. Kolmel and A. Klamt, *J. Chem. Phys.*, 1995, **103**, 9312; A. Klamt, *J. Phys. Chem.*, 1996, **100**, 3349; A. Klamt, V. Jonas, T. Burger and J.C.W. Lohrenz, *J. Phys. Chem. A*, 1998, **102**, 5074.
586) O. Tapia and O. Goscinski, *Mol. Phys.*, 1975, **29**, 1653; M.M. Karelson and M.C. Zerner, *J. Phys. Chem.*, 1992, **96**, 6949.
587) Overlap weighting factors modify the resoance integrals for the off-diagonal elements in the Fock matrix. Resoance integrals account for σ- and π-bonding patterns in molecules. The magnitude of the resonance integral is proportional to the overlap integral $S_{\mu\nu}$. σ-π Overlap is usually absent due to orthogonality conditions. Adjusting the weights provides a mechanism to adjust the relative contributions of σ- *versus* π-bonding.
588) The default value for Sigma-Sigma for ZINDO/1 is 1.0.
589) The default value for Pi-Pi for ZINDO/1 is 1.0.
590) Calculated using Woodward's rules.
591) *Kirk-Othmer Encylopedia of Chemical Technology*, M. Grayson and D. Eckroth, editors, John Wiley & Sons Inc., New York, NY, 1979, **Volume 8**, p. 212.
592) W.C. Price and A.D. Walsh, *Proc. R. Soc. London, Ser. A.*, 1941, **179**, 201.
L.W. Pickett, N.J. Hoeflich and T.-C. Liu, *J. Am. Chem. Soc.*, 1951, **73**, 4865.
593) K. Bowden and E.A. Braude, *J. Chem. Soc.*, 1946, 948.
594) E.A. Braude and E.S. Waight, *Progress in Stereochemistry*, W. Klyne, editor, Butterworth, London, 1954, **Volume 1**, *Chapter 4*.
E.R.H. Jones, G.H. Mansfield and M.C. Whiting, *J. Chem. Soc.*, 1956, 4073.
595) W.M. Schubert and W.A. Sweeney, *J. Am. Chem. Soc.*, 1955, **77**, 2297.
596) R.H. Mazur, *J. Am. Chem. Soc.*, 1959, **81**, 1454.
597) C.G. Overberger and D. Tanner, *J. Am. Chem. Soc.*, 1955, **77**, 369.
H.P. Koch, *J. Chem. Soc.*, 1948, 1123.
598) R.A. Morton and A.J.A. deGouveia, *J. Chem. Soc.*, 1934, 196.
599) J. Griffiths, *Dyes & Pigments*, 1982, **3**, 211.
600) P. Hohenberg and W. Kohn, *Phys. Rev. B*, 1964, **136**, 864.
601) M. Levy, *Proc. Natl. Acad. Sci. U.S.A.*, 1979, **76**, 6062.
602) C.C.J. Roothaan, *Rev. Mod. Phys.*, 1951, **23**, 69.
603) J.C. Slater, *Phys. Rev.*, 1951, **81**, 385, 538.
604) M.J.S. Dewar, *J. Mol. Struct.*, 1983, **100**, 41.
605) J.A. Pople and R.K. Nesbet, *J. Chem. Phys.*, 1954, **22**, 571.
606) L. Hedin and B.I. Lundqvist, *J. Phys. C*, 1971, **4**, 2064.
607) D.M. Ceperley and B.J. Alder, *Phys. Rev. Lett.*, 1980, **45**, 566.
608) S. Lundqvist and N. March, *Theory of the Inhomogenous Electron Gas*, Plenum, New York, NY, 1983.
609) U. von Barth and L. Hedin, *J. Phys. C*, 1972, **5**, 1629.
610) S.J. Vosko, L. Wilk and M. Nusair, *Canad. J. Phys.*, 1980, **58**, 1200.
611) J.P. Perdew and Y. Wang, *Phys. Rev.*, 1992, **B45**, 13244.
612) T. Ziegler, *Chem. Rev.*, 1991, **91**, 651.
613) K. Labanowski and J. Andzelm, *Density Functional Methods in Chemistry*, Springer-Verlag, New York, NY, 1991 and references cited therein.

614) P. Politzer and J.M. Seminario, *Density Functional Theory: A Tool for Chemistry*, Elsevier, 1995 and references cited therein.
615) A.D. Becke, *J. Chem. Phys.*, 1988, **88**, 2547.
616) C. Lee, W. Yang and E.G. Parr, *Phys. Rev. B*, 1988, **37**, 786.
617) W. Kohn and L.J. Sham, *Phys. Rev. A*, 1965, **140**, 1133.
618) J. Andzelm, E. Wimmer and D.R. Salahub, *Spin density functional approach to the chemistry of transition metal clusters: Gaussian-type orbital implementation in The Challenge of d- and f-Electrons: Theory and Computation*, D.R. Sulahub and M.C. Zerner, editors, ACS Symposium Ser. 394, 1989.
619) L. Versluis and T. Ziegler, *J. Chem. Phys.*, 1988, **88**, 3322.
620) B. Delley, *J. Chem. Phys.*, 1990, **92**, 508.
621) D. Ellis and G. Painter, *Phys. Rev. B*, 1968, **2**, 2887.
622) P.M. Boerrigter, G. te Velde and E.J. Baerends, *Int. J. Quant. Chem.*, 1988, **37**, 87.
623) A.H. Stoud, *Approximate Calculation of Multiple Integrals*, Prentice-Hall, Englewood Cliffs, 1971.
624) V.I. Lebedev, *Zh. Vychisl. Mat. Fiz.*, 1975, **15**, 48; 1978, **16**, 293.
625) S.I. Konyaev, *Mat. Zametki*, 1979, **25**, 629.
626) F.L. Hirshfeld, *Theor. Chim. Acta B*, 1977, **44**, 129.
627) An alternative approach is to fit the charge density to an analytic multipolar expansion *via* a least-squares fitting procedure.
E.J. Baerends, D.E. Ellis and P. Ros, *Chem. Phys.*, 1973, **2**, 41.
628) B. Delley, D.E. Ellis, A.J. Freeman, E.J. Baerends and D. Post, *Phys. Rev. B*, 1983, **27**, 2132.
629) B. Delley, *J. Chem. Phys.*, 1991, **94**, 7245.
630) J. Harris, *Phys. Rev. B*, 1985, **31**, 1770.
631) P. Pulay, *Mol. Phys.*, 1969, **17**, 197.
632) B.R. Brooks, W.D. Laidig, P. Saxe, J.D. Goddard, Y. Yamaguchi and H.F. Schaefer, *J. Chem. Phys.*, 1980, **72**, 4652.
633) J.A. Pople, R. Krishnan, H.B. Schlegel and J.S. Binkley, *Int. J. Quant. Chem. Symp.*, 1979, **13**, 225.
634) C. Satako, *Chem. Phys. Lett.*, 1981, **83**, 111.
635) J. Baker, *J. Comput. Chem.*, 1986, **7**, 385; 1992, **13**, 240.
636) J. Baker and W.J. Hehre, *J. Comput. Chem.*, 1991, **12**, 606.
637) P. Pulay, *J. Comput. Chem.*, 1982, **3**, 556.
638) λ is chosen from an eigenvalue equation.
A. Bannerjee, N. Adams, J. Simons and R. Shepard, *J. Phys. Chem.*, 1985, **89**, 52.
639) C.J.Miller and W.H. Miller, *J. Chem. Phys.*, 1981, **75**, 2800.
640) R.G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, NY, 1989.
641) JMW = Hedin-Lundqvist/Janak-Morruzi-Williams local correlation functional.
642) VWN = Vosko-Wilk-Nusair local correlation functional - parameterisation of Ceperley and Alder's Monte Carlo data.
643) BP = Perdew-Wang gradient-corrected correlation functional, and Becke's B88 gradient-corrected exchange functional using a value of energy for the Nonlocal keyword causing the gradient-corrected "nonlocal" exchange and correlation energies to be calculated as a pertubation on a converged LDA density at the end of an LDA calculation.
644) BPS = Perdew-Wang gradient-corrected correlation functional, and Becke's B88 gradient-corrected exchange functional using a value of potential for the Nonlocal keyword causing the gradient-corrected functional to be applied self-consistently.
645) BLYP = Lee-Yang-Parr gradient-corrected correlation functional, and Becke's B88 gradient-corrected exchange functional using a value of energy for the Nonlocal keyword causing the gradient-corrected "nonlocal" exchange and correlation energies to be calculated as a pertubation on a converged LDA density at the end of an LDA calculation.

646) BLS = Lee-Yang-Parr gradient-corrected correlation functional, and Becke's B88 gradient-corrected exchange functional using a value of potential for the Nonlocal keyword causing the gradient-corrected functional to be applied self-consistently.

647) BVWN = Vosko-Wilk-Nusair local correlation functional, and Becke's B88 gradient-corrected exchange functional using a value of energy for the Nonlocal keyword causing the gradient-corrected "nonlocal" exchange and correlation energies to be calculated as a pertubation on a converged LDA density at the end of an LDA calculation.

648) BVS = Vosko-Wilk-Nusair local correlation functional, and Becke's B88 gradient-corrected exchange functional using a value of potential for the Nonlocal keyword causing the gradient-corrected functional to be applied self-consistently.

649) E.A. Braude and F. Sondheimer, *J. Chem. Soc.*, 1955, 3754.

650) H.E. Ungnade, *J. Am. Chem. Soc.*, 1953, **75**, 432.

651) H.E. Ungnade, E.E. Pickett, L. Rubin and E. Youse, *J. Org. Chem.*, 1951, **16**, 1318.

652) Calculated using empirical methods for aromatic systems.

653) J.E. Baerends, O.V. Gritsenko and R. vanLeeuwen, *ACS Symp. Ser.*, 1996, **629**, 20.
See also:
R.G. Pearson and W.E. Palke, *Int. J. Quant. Chem.*, 1990, **37**, 103.
Y.F. Guo and M.A. Whitehead, *J. Comput. Chem.*, 1991, **12**, 803.
J.B. Krieger, J. Chen and G.J. Iafrate, *Int. J. Quant. Chem.*, 1998, **69**, 255.
P. Politzer, F. AbuAwwad and J.S. Murray, *Int. J. Quant. Chem.*, 1998, **69**, 607.

654) Well stained chromatin, differential counterstaining and cytoplasmic transparency.
G.W. Gill and K.A. Miller, *Staining in Compendium on Cytopreparatory Techniques*, C.M. Keebler, J.W. Reagan and G.L. Wied, editors, *Tutorials of Cytology*, Chicago, IL, 1974, pp. 9-25.

655) D.P. Penney and J.M. Powers, *Biotechnic & Histochemistry*, 1995, **70**, 217.

656) G.L. Wied, G.F. Bahr, D.G. Oldfield and P.H. Bartels, *Acta Cytol.*, 1969, **13**, 21.

657) J. Holmquist, Y. Imasato, E. Bengtsson, B. Olsen and B. Stenkvist, *J. Histochem. Cytochem.*, 1976, **24**, 1218.

658) W. Galbraith, P.N. Marshall, E.S. Lee and J.W. Bacus, *Anal. Quant. Cytol.*, 1979, **1**, 160.

659) G.L. Wied, G.F. Bahr, D.G. Oldfield and P.H. Bartels, *Acta. Cytol.*, 1968, **12**, 357.

660) R.K. Aggarwal and J.W. Bacus, *J. Histochem. Cytochem.*, 1977, **25**, 668.

661) P.N. Marshall, W. Galbraith and J.W. Bacus, *Anal. Quant. Cytol.*, 1979, **1**, 169.

a) Smears stained with an alcoholic solution of bismarck brown Y (2 g dm^{-3}) for 4 minutes, did not show any incorporation of bismarck brown Y.

b) However, smears stained with a saturated alcoholic solution of Bismarck brown Y for 15 minutes, showed slight staining of intermediate nuclei.
W. Galbraith and P.N. Marshall, *Stain Technol.*, 1984, **59**, 133.

662) The spectra of the dyes changed shape on binding to different cellular structures due to dimer formation. Spectra of monomer and dimer forms of the dyes were then used as components.

663) P.N. Marshall and R.W. Horobin, *Histochem J.*, 1972, **4**, 493.

664) R.D. Lillie and R.W. Palmer, *Histochemie*, 1965, **5**, 44.

665) S.E. Sheppard and A.L. Geddes, *J. Am. Chem. Soc.*, 1944, **66**, 1995.

666) B. Soderborg, *Ann. Phys.*, 1913, **40**, 381.

667) W.E. Speas, *Phys. Rev.*, 1928, **31**, 569.

668) E. Coates, *J. Soc. Dyers Colour*, 1969, **85**, 355.

669) B.L. Dunton, *Acta Cytol.*, 1972, **16**, 361.

670) D.S. Meyers, *Am. J. Med. Technol.*, 1968, **34**, 240.

671) Stockard and Papanicolaou's original method was applied to guinea pigs, which have more mucus cells in the cervix.

672) L. Lilienfeld, *Arch. Anat. Physiol., Lpz. Physiol. Abt.*, 1893, p. 554.

673) H. Hoyer, *Arch. mikr. Anat.*, 1890, **36**, 310.

674) J.H. List, *Z. wiss. Mikr.*, 1885, **2**, 145.

675) W.B. Hardy and F.F. Wesbrook, *J. Physiol.*, 1895, **18**, 490.
676) Anonymous, *Gill's modified OG-6 and EA for Papanicolaou staining*, Polysciences Inc. Data Sheet 196, Polysciences Inc., Warrington, PA.
677) G.W. Gill, *Scanner*, 1975, **14**, 2.
678) S.L. Pharr, D.A. Wood and H.F. Traut, *Am. J. Clin. Pathol.*, 1954, **24**, 239.
679) R.W. Payne, *J. Soc. Dyers Colour*, 1934, **50**, 320.
680) S.R. Trotman and T.B. Frearson, *J. Soc. Dyers Colour*, 1931, **47**, 344.
681) P. Gray, *The Microtomist's Formulary and Guide*, Constable and Company Ltd., London, 1954.
682) E. Gurr, *A Practical Manual of Medical and Biological Staining Techniques*, Leonard Hill (Books) Ltd., London, Interscience Publishers Inc., New York, NY, 2nd edition, 1956.
683) E. Gurr, *Encyclopædia of Microscopic Stains*, Leonard Hill (Books) Ltd., London, 1960.
684) H. Puchtler, S.N. Meloan and F.S. Waltrop, *Histochemisty*, 1986, **85**, 353.
685a) F.M. Arshid, J.N. Desai, D.J. Duff, C.H. Giles, S.K. Jain and I.R. McNeal, *J. Soc. Dyers Colour*, 1954, **70**, 392.
685b) F.M. Arshid, R.F. Connelly, J.N. Desai, R.G. Fulton, C.H. Giles and J.G. Kefalas, *J. Soc. Dyers Colour*, 1954, **70**, 402.
686a) G.C. Lalor and S.L. Martin, *J. Soc. Dyers Colour*, 1959, **75**, 513.
686b) G.C. Lalor and S.L. Martin, *J. Soc. Dyers Colour*, 1959, **75**, 517.
687) N.K. Imes, D.C. Sanders, C.R. Crane and G. Clark, *Stain Technol.*, 1969, **44**, 167.
688) P.N. Marshall and R.W. Horobin, *Stain Technol.*, 1974, **49**, 137.
689) P. Feather, E. Inman and A.T. Peters, *J. Soc. Dyers Colour*, 1962, **78**, 321.
690) E. Asmus, H.J. Altmann and E. Thomasz, *Z. Anal. Chem.*, 1966, **216**, 3.
691a) M. Chevreul, *Ann. Chim.*, 1812, **81**, 128.
691b) M. Chevreul, *Ann. Chim.*, 1812, **82**, 53, 126.
692) O.L. Erdmann, *Liebigs Ann. Chem.*, 1842, **44**, 292.
693) O. Hesse, *Liebigs Ann. Chem.*, 1859, **109**, 332.
694) W.H. Perkin and R. Robinson, *J. Chem. Soc.*, 1907, **91**, 1073.
695) O. Dann and H. Hofmann, *Chem. Ber.*, 1965, **98**, 1498.
696) F. Morsingh and R. Robinson, *Tetrahedron*, 1970, **26**, 281.
697) J. Cymerman Craig, A.R. Naik, R. Pratt, E. Johnson and N.S. Bhacca, *J. Org. Chem.*, 1965, **30**, 1573.
698) Ch. Bettinger and H.W. Zimmermann, *Histochemistry*, 1991, **95**, 279.
699) The solvent and concentration of the solution was not reported.
700) G. Clark, *Stain Technol.*, 1974, **49**, 225.
701) Attempts to prepare 2,4-dinitrophenylhydrazone or semicarbazone derivatives have failed.[698]
702) Structures **(242)**, **(248)** and **(249)** are covalent and ionic valence structures in the sense of the valence bond approximation. Only the quantum mechanical combination of these structures describes the actual ground state of the molecule.
703) d_6-Ethylene glycol exhibits a transparent window in the 1400-1800 cm^{-1} region and is suitable for investigating the C=O stretching vibration of dissolved molecules.
704) G.C. Lalor, *J. Soc. Dyers Colour*, 1962, **78**, 549.
705) C.J. Pouchert, *The Aldrich library of FT-IR-spectra*, **Volumes I and II**, 1st edition, Aldrich Chemical Company, Milwaukee, WI.
706) Two equivalent valence structures favours a shift to longer wavelength and an increase in the intensity of the absorption bands.
707) W.R. Orndorff, R.C. Gibbs and S.A. McNulty, *J. Am. Chem. Soc.*, 1925, **47**, 2767.
708) B. Romeis, *Mikroskopische Technik*, Oldenburg Verlag, München, 1968.
709) H. Harms, *Handbuch der Farbstoffe für die Mikroscopie, Teil II*, 2. Lieferung, Staufen-Verlag, Kamp-Lintfort, 1957.
710) J.R. Baker, *Quart. J. Microsc. Sci.*, 1962, **103**, 493.
711) P.N. Marshall and R.W. Horobin, *Histochemie*, 1973, **35**, 361.

712) Barbiturate buffer: M. Michaelis, *Biochem. Z.*, 1931, **234**, 139.
713) Ch. Bettinger and H.W. Zimmermann, *Histochemistry*, 1991, **96**, 215.
714) It is not known how many water molecules are coordinatively bound as ligands.
715) The species are in equilibrium with each other.
716) Although both dye complexes absorb at the same wavelength 556 nm, $\varepsilon_{(258)} = 22\ 000$ and $\varepsilon_{(259)} = 41\ 000$.
717) HeLa cells are easily cultured and also grow as monolayers in culture vessels. Monolayers are highly suited for microspectrophotometric investigations. The cells were fixed with methanol for 5 minutes before staining.
718) Micro- and solution-spectra correspond to each other, but the former are somewhat shifted and broadened, due to the binding of the dye complexes to substrates in the cells and to changes in the medium.
719) R.D. Lillie, *Histopathologic technic and practical histochemistry*, 3rd edition, McGraw-Hill, New York, 1965.
720) J. Pauluhn, A. Naujok and H.W. Zimmerman, *Z. Naturforsch*, 1980, **35C**, 585.
721) E. Zipfel, J.R. Grezes, A. Naujok, W. Seifert, D.H. Wittekind and H.W. Zimmerman, *Histochemistry*, 1984, **81**, 337.
722) Cover slips with HeLa cells Bristol S3 were briefly rinsed with Earle BBS solution, dried in air for 30 minutes, incubated with the enzyme solution for 30 minutes at 37 °C with shaking, fixed for 5 minutes with methanol, dried for 5 minutes in air, and stained as normal.
Enzyme solution for DNA digestion: DNAse I (2.4 mg) in 0.05 M acetate buffer (2.4 cm^3) with 1% $MgCl_2$, pH 7.0.
723) Enzyme solution for RNA digestion: RNAse A (1 mg) in *Tris*-buffer (1ml - 0.05M *Tris* + 0.01 M Na_2EDTA + 0.01 M NaCl), pH 7.7.
724) Cover slips with HeLa cells were briefly rinsed with Earle BBS solution, dried for 30 minutes in air, incubated for 45 minutes with the enzyme solution for DNA digestion at 37 °C, rinsed with *Tris*-buffer, fixed for 5 minutes in methanol, incubated for 1 hour with the enzyme solution for RNA digestion, fixed for 5 minutes in methanol, dried for 10 minutes in air, and stained as normal.
725) R. Robinson, *Bull. Soc. Chim. Fr.*, 1958, 125.
726) D. Johansen, *Plant Microtechnique*, McGraw-Hill Book Co. Inc. New York and London, 1940.
727) J.B. Gatenby and H.W. Beams, *The Microtomist's Vade-Mecum*, J. & A. Churchill, London, 1950.
728) F. Morsingh and R. Robinson, *Tetrahedron*, 1970, **25**, 281.
729) R.H. Jaeger, P.M.E Lewis and R. Robinson, *Tetrahedron*, 1974, **30**, 1295.
730) J.N. Chatterjea, R. Robinson and M.L. Tomlinson, *Tetrahedron*, 1974, **30**, 507.
731) P.M. Chakrabarti, *Tetrahedron Lett.*, 1963, 1907.
732) J.N. Chatterjea, S.C. Shaw and N.D. Sinha, *J. Indian Chem. Soc.*, 1974, **51**, 752.
733) J. Xu and J.C. Yadan, *Tetrahedron Lett.*, 1996, **37**, 2421.
734) For a review, see: S.W. McCombie, *Reduction of Saturated Alcohols and Amines to Alkanes* in *Comprehensive Organic Synthesis*, B.M. Trost and I. Fleming editors, Pergamon Press, Oxford, 1991, pp. 811-833.
735) H. Deshayes and J.P. Pete, *Canad. J. Chem.*, 1984, **62**, 2063.
736) C. Portella, H. Deshayes, J.P. Pete and D. Scholler, *Tetrahedron*, 1984, 3635.
737) For a review, see: W. Harting, *Tetrahedron*, 1983, **39**, 2609.
738) O. Dann and H. Hofmann, *Justus Liebigs Ann. Chem.*, 1963, **667**, 116.
739) M.J. Robins, J.S. Wilson and F. Hansske, *J. Am. Chem. Soc.*, 1983, **105**, 4059.
740) D.H.R. Barton and S.W. McCombie, *J. Chem. Soc., Perkin Trans. 1*, 1975, 1574.
741) D.H.R. Barton and C. Jaszberenyi, *Tetrahedron Lett.*, 1989, **30**, 2619.
742) D.H.R. Barton and D. Crich, *J. Chem. Soc., Perkin Trans. 1*, 1986, 1603.
743) S.C. Dolan, J. MacMillan, *J. Chem. Soc., Chem. Commun.*, 1985, 1588.
744) F. Proescher and A.S. Arkush, *Stain Technol.*, 1928, **3**, 28.
745) A.C. Lendrum, *J. Path. Bact.*, 1935, **40**, 415.

746) A.C. Lendrum and D. McFarlane, *J. Path. Bact.*, 1940, **50**, 381.
747) P. Gray, E. Bereczy, M.D. Maser and C. Nevsimal, *Stain Technol.*, 1958, **33**, 215.
748) S. Becher, *Untersuchengen ueber Echtfaerbung der Zellkerne*, Borntraeger, Berlin, 1921.
749) D. Wittekind, M. Hilgarth and V. Kretschmer, *Geburtsh Frauenheilkd*, 1979, **39**, 969.
750) D. Wittekind, E.R. Reinhardt, V. Kretschmer and E. Zipfel, *Anal. Quant. Cytol.*, 1983, **5**, 55.
751) D.H. Wittekind, M. Hilgarth, V. Kretschmer, W. Seiffert and E. Zipfel, *Anal. Quant. Cytol.*, 1982, **4**, 286.
752) E. Schulte and D. Wittekind, *Anal. Quant. Cytol.*, 1990, **12**, 149.
753) E. Schulte and D. Wittekind, *Histochemistry*, 1989, **91**, 321.
754) E. Schulte, D. Wittekind and V. Kretschmer, *Histochemistry*, 1988, **88**, 427.
755) P.J. Eason and J.H. Tucker, *J. Histochem. Cytochem.*, 1979, **27**, 25.
756) O.A.N. Husain and K.C. Watts, *J. Clin. Pathol.*, 1984, **37**, 99.
757) E. Schulte and D. Wittekind, *Histochemistry*, 1989, **91**, 321.
758) E. Schulte, *Diagn. Cytopathol.*, 1986, **2**, 324.
759) H.J. Soost, E.W. Falter and K. Otto, *Anal. Quant. Cytol.*, 1979, **1**, 37.
760) E. Schulte, *Hematoxylin and the Feulgen reagent in nuclear staining in Standardization and Quantitation of Diagnostic Staining in Cytology*, M.E. Boon and L.P. Lok, editors, Coulomb Press, Leyden, 1986, pp. 9-15.
761) G. Clark, *Stain Technol.*, 1969, **44**, 15.
762) W. Sandritter, *Acta Cytol.*, 1958, **2**, 324.
763) E. Schulte, *Acta Histochem.*, 1988, **36**, 341.
764) P.S. Oud, D.J. Zahniser, M.C.T. Raaijmakers, P.G. Vooijs and R.T. Van de Walle, *Anal. Quant. Cytol.*, 1981, **3**, 289.
765) L. Sachs, *Angewandte Statistik*, Springer Verlag, Berlin, 1984.
766) P. Haour and C. Conti, *Acta Cytol.*, 1958, **2**, 326.
767) W.R. Orndorff, R.C. Gibbs and C.V. Shapiro, *J. Am. Chem. Soc.*, 1926, **48**, 1327.
W.R. Orndorff, R.C. Gibbs and S.A. McNulty, *J. Am. Chem. Soc.*, 1926, **48**, 1994.
W.R. Orndorff, R.C. Gibbs, S.A. McNulty and C.V. Shapiro, *J. Am. Chem. Soc.*, 1927, **49**, 1545.
W.R. Orndorff and W.R. Barrett, *J. Am. Chem. Soc.*, 1924, **46**, 2483.
768) D. Fuller, *J. Lab. & Clin. Med.*, 1943, **28**, 1475.
769) Glycosaminoglycans are large, usually heterogeneously sulphated polysaccharides implicated in embryonic development, blood coagulation and wound healing. See:
a) T.E. Hardingham and A.J. Fosang, *FASEB J.*, 1992, **6**, 861.
b) N. Sugiura, K. Sakurai, Y. Hori, K. Karasawa, S. Suzuki and K. Kimata, *J. Biol. Chem.*, 1993, **268**, 15779.
770a) P. Mohan, *Pharm. Res.*, 1992, **9**, 703.
770b) M. Baba, D. Schols, P. Mohan, E. De Clercq and S. Shigeta, *Antiviral Chem. Chemother.*, 1993, **4**, 229.
770c) J.L. Weaver, P.S. Pine, R. Anand, S. Bell and A. Aszalos, *Antiviral Chem. Chemother.*, 1992, **3**, 147.
770d) P. Mohan, R. Singh and M. Baba, *J. Med. Chem.*, 1991, **34**, 212.
770e) D. García-Villalón and C. Gil-Fernández, *Antiviral Chem. Chemother.*, 1992, **3**, 9.
771a) R.I. Brinkworth and D.P. Farlie, *Biochem. Biophys. Res. Commun.*, 1992, **118**, 624.
771b) J. Balzarini, H. Mitsuya, E. De Clercq and S. Broder, *Int. J. Cancer*, 1986, **37**, 451.
772a) W.G Turnell and J.T Finch, *J. Mol. Biol.*, 1992, **227**, 1205.
772b) W.E. Klunk, J.W. Pettegrew and D.J. Abraham, *J. Histochem. Cytochem.*, 1989, **37**, 1273.
772c) S. Zhang, T. Holmes, C. Lockshin and A. Rich, *Proc. Natl. Acad. Sci. U.S.A.*, 1993, **90**, 3334.
773a) B. Caughey, D. Ernst and R.E. Race, *J. Virol.*, 1993, **67**, 6270.
773b) B. Caughey and R.E. Race, *J. Neurochem.*, 1992, **59**, 678.
774) P. Mohan, A.J. Hopfinger and M. Baba, *Antiviral Chem. Chemother.*, 1991, **2**, 215.

775) C. Sachsenmaier, A. Radler-Pohl, R. Zinck, A. Nordheim, P. Herrlich and H. Rahmsdorf, *J. Cell.*, 1994, **78**, 963.
776) A. Gagliardi, H. Hadd and D.C. Collins, *Cancer Res.*, 1992, **52**, 5073.
777) M. Jozefowicz and J. Jozefonvicz, *Pure Appl. Chem.*, 1984, **56**, 1335.
778a) M. Ito, M. Baba, S. Shigeta, S. Wada, M. Takagi, T. Kimura and T. Okuyama, *Antiviral Chem. Chemother.*, 1991, **2**, 41.
778b) M.O. McClure, D. Whitby, C. Patience, N.J. Gooderham, A. Bradshaw, R. Cheingsong-Popov, J.N. Weber, D.S. Davies, G.M.W. Cook, R.J. Keynes and R.A. Weiss, *Antiviral Chem. Chemother.*, 1991, **2**, 149.
778c) M. Witvrouw, D. Schols, G. Andrei, R. Snoeck, M. Hosoya, R. Pauwels, J. Balzarini and E. De Clercq, *Antiviral Chem. Chemother.*, 1991, **2**, 171.
778d) D. Schols, E. De Clercq, M. Witvrouw, H. Nakashima, R. Snoek, R. Pauwels, A. Van Schepdael and P. Claes, *Antiviral Chem. Chemother.*, 1991, **2**, 45.
778e) R. Anand, S. Nayyar, J. Pitha and C.R. Merril, *Antiviral Chem. Chemother.*, 1990, **1**, 41.
778f) M.R. Kozlowski and A. Watson, *Antiviral Chem. Chemother.*, 1990, **1**, 175.
778g) T. Otake, D. Schols, M. Witvrouw, L. Naesens, H. Nakashima, T. Moriya, H. Kurita, K. Matsumoto, N. Ueba and E. De Clercq, *Antiviral Chem. Chemother.*, 1994, **5**, 155.
779) W.H. Ojala, C.R. Ojala and W.B. Gleason, *Antiviral Chem. Chemother.*, 1995, **6**, 25.
780) Z.F. Kanyo and D.W. Christianson, *J. Biol. Chem.*, 1991, **266**, 4264.
781) P. Chakrabarti, *Int. J. Peptide Protein Res.*, 1994, **43**, 284.
782) Y. Nawata, K. Ochi, M. Shiba, K. Morita and Y. Iitaka, *Acta Crystallogr.*, 1981, **B37**, 246.
783) X. Zhu, B.T. Hsu and D.C. Rees, *Structure*, 1993, **1**, 27.
784a) W.H. Ojala, W.B. Gleason, T.I. Richardson and R.E. Lovrien, *Acta Crystallogr.*, 1994, **C50**, 1615.
784b) W.H. Ojala, L.K. Lu, K.E. Albers, W.B. Gleason, T.I. Richardson, R.E. Lovrien and E.A. Sudbeck, *Acta Crysallogr.*, 1994, **B50**, 684.
785) A.C. Oliveri, R.B. Wilson, I.C. Paul and D.Y. Curtin, *J. Am. Chem. Soc.*, 1989, **111**, 5525.
786) In the eclipsed sulphonate conformation, the S–O bond most nearly coplanar with the naphthalene ring must be located within a torsional andgle of 15° with respect to the ring.
787) L. Leiserowitz, *Acta Crystallogr.*, 1976, **B32**, 775.
788a) V.A. Russell, M.C. Etter and M.D. Ward, *J. Am. Chem. Soc.*, 1994, **116**, 1941.
788b) V.A. Russell, M.C. Etter and M.D. Ward, *Chem. Mater.*, 1994, **6**, 1206.
789) Resorcinol (50 g) and phthalic anhydride (34 g) are ground to a homogeneous mixture, transferred to a nickel crucible and heated to 180 °C, on an oil bath or in a thermostatically controlled oven. Anhydrous zinc chloride powder (16 g) is added in small amounts at a time with constant stirring. After addition of zinc chloride, the temperature is raised to 210 °C, and maintained for one to two hours until the melt becomes solid. The product is cooled to room temperature, chipped out of the crucible and ground to a fine powder. The powder is shaken with 5% aqueous sodium hydroxide solution (500 cm^3) until a solution of crude uranin is obtained. Concentrated hydrochloric acid is added until no further precipitate (fluorescein acid) appears. The fluorescein is washed well with water and dried at 50-55 °C. The free acid fluorescein (30 g) is shaken with absolute ethyl alcohol (70 cm^3) in a flask, which is then transferred to a fume hood. Bromine (11 cm^3) is added in very small amounts, shaking the flask after each addition. After the preparation is allowed to stand for 2 hours, the precipitated eosin (free acid) is filtered off, washed with alcohol, followed by washing with water, drained, and transferred to a large evaporating basin. After a little water has been stirred in, 10% aqueous sodium carbonate is added until effervescence stops (do not add excess of carbonate solution!). The solution is evaporated to dryness on a water bath, or a thermostatically controlled oven at about 90 °C, or less.

790) R.D. Lillie, *Arch. Path.*, 1940, **29**, 705.
791) E.V. Cowdry, *Laboratory Technique in Biology and Medicine*, Baillière, Tindall & Cox, Ltd., London, 1952.
792) S.G. Scott, *J. Path. Bact.*, 1911, **16**, 390.
793) C.F. Geschickter, *Stain Technol.*, 1930, **5**, 81.
794) H.J. Conn, *Biological Stains*, Biotech. Publications, Geneva, NY, 1953.
795) Three methods of preparation of light green SF yellowish are known:
a) Condense benzylethylaniline (2 eq) and benzaldehyde (1 eq) in the presence of oxalic acid (removes water) and a lewis acid (zinc chloride). The resulting leuco compound is trisulphonated with oleum and oxidised with lead oxide.
b) Condense sulphobenzylethylaniline (2 eq) with benzaldehyde (1 eq), followed by further sulphonation and oxidation.
c) Sulphobenzylethylaniline (2 eq) is condensed with *p*-sulphobenzaldehyde (1 eq), followed by oxidation of the leuco compound.
796a) W.S. Ollet, *J. Path. Bact.*, 1947, **59**, 351
796b) W.S. Ollet, J. Path. Bact., 1951, **63**, 166.
797) R.D. Lillie, *J. Tech. Methods*, 1945, **25**, 1.
798) J. Van Gieson, *N.Y. Med. J.*, 1889, **50**, 57.
799) H. Wales, *J. Chem. Soc.*, 1924, **46**, 2124.
800) G.H. Williams, P. Romanowski, L. Morris, M. Madine, A.D. Mills, K. Stoeber, J. Marr, R.A. Laskey and N. Coleman, *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 14932.
801) Keratins are fibrous intracellular proteins which are components of intermediate filaments in epithelial tissues. Around 30 different keratins are known with molecular weights of 40,000 to 70,000 Daltons. Intermediate filament subunits all contain a central left-handed α-helical rod superhelix of 311-314 amino acids made up of repeated heptads $(a–b–c–d–e–f–g–)_n$ of amino acids. Hydrophobic side chains occur at residues *a* and *d* of the heptad repeat, while electrostatic interactions occur primarily between residues *e* and *g*. Position *a* is most frequently Leu, Ile or Ala, and *d* is primarily Leu or Ala. Residues *e* and *g* are often Glu or Gln, with Arg and Lys prominent at *g*. See:
a) F.H.C. Crick, *Acta Cryst.*, 1953, **6**, 689.
b) L. Pauling and R.B. Corey, *Nature*, 1953, **171**, 59.
c) C. Cohen and D.A.D. Parry, *Trends Biochem. Sci.*, 1986, **11**, 245.
d) P.M. Steinert and D.R. Roop, *Ann. Rev. Biochem.*, 1988, **57**, 593.
e) C. Cohen and D.A.D. Parry, *Proteins Struct. Funct. Genet.*, 1990, **7**, 1.
802) Collagen is a microfibrillar protein whose subunit tropocollagen, is a triple helix composed of two identical polypeptide chains (α_1) and one slightly different chain (α_2), in which the amino acid sequences are essentially $(Gly–Xaa–Yaa)_n$ where Xaa is frequently Pro and Yaa is often Hyp, although Lys, hydroxylysine, His and other residues occur at Xaa and Yaa. The three polypeptide chains have a left-handed threefold helical conformation similar to poly(Pro) II and poly(Gly) II. The surface of the triple helix is defined by the side chains of Xaa and Yaa. See:
a) P. Bornstein and W. Traub, *The chemistry and biology of collagen* in *The Proteins*, H. Neurath and R.L. Hill, editors, Academic Press, New York, NY, 1979, 3rd edition, **Volume 4**, pp. 411-632.
b) P. Bornstein and H. Sage, *Ann. Rev. Biochem.*, 1980, **49**, 957.
c) D.R. Eyre, *Science*, 1980, **207**, 1315.
d) K. Okuyama, K. Okuyama, S. Arnott, M. Takayanagi, M. Kakudo, *J. Mol. Biol.*, 1981, **152**, 427.
e) R.D.B. Fraser, T.P. MacRae, A. Miller and E. Suzuki, *J. Mol .Biol.*, 1983, **167**, 497.
f) D.R. Eyre, M.A. Paz and P.M. Gallop, *Ann. Rev. Biochem.*, 1984, **53**, 717.
g) R.E. Burgeson, *Ann. Rev. Cell Biol.*, 1988, **4**, 551.
h) E.Y. Jones and A. Miller, *J. Mol. Biol.*, 1991, **218**, 209.
i) J. Engel and D.J. Prockop, *Ann. Rev. Biophys. Biophys. Chem.*, 1991, **20**, 137.

803) K. Müller, *Angew. Chem.*, 1980, **92**, 1; *Angew. Chem. Int. Ed. Engl.*, 1980, **19**, 1.
804) This definition has been put into common use by John A. Pople.
805) W.J. Hehre, L. Radom, P. von R. Schleyer and J.A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley, New York, NY, 1986.
806) For a recent review on the concepts and terminology of *ab initio* quantum chemistry see: J. Simons, *J. Phys. Chem.*, 1991, **95**, 1017.
807) S. Gladstone, K.J. Laidler and H.M. Eyring, *The Theory of Rate Processes*, McGraw-Hill, New York, NY, 1941.
808) J.I. Steinfeld, J.S. Fransisco and W.L. Hase, *Chemical Kinetics and Dynamics*, Prentice Hall, New Jersey, 1989.
809) R.D. Levine and R.B. Bernstein, *Molecular Reaction Dynamics*, Oxford University Press, New York, NY, 1987.
810) D.G. Truhlar and M.S. Gordon, *Science*, 1990, **249**, 491.
811) B.R. Gelin and M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.*, 1975, **72**, 2002.
812) H.B. Schlegel, *J. Comput. Chem.*, 1982, **3**, 214.
813) R. Czerminski and R. Elber, *Int. J. Quantum Chem.*, 1990, **S24**, 167.
W. Nowak, R. Czerminski and R. Elber, *J. Am. Chem. Soc.*, 1991, **113**, 5627.
814) R. Elber and M. Karplus, *Chem. Phys. Lett.*, 1987, **139**, 375.
815) S. Fischer and M. Karplus, *Chem. Phys. Lett.*, 1992, **194**, 252.
816) F. Bernardi and M.A. Robb in *Ab Initio Methods in Quantum Chemistry I*, K.P. Lawley, editor, Wiley, New York, NY, 1987, p.156-248.
817) H.B. Schlegel in *New Theoretical Concepts for Understanding Organic Reactions*, J. Bertran and I.G. Csizmadia, editors, Kluwer Academic Publishers, Boston, MA, 1989, p.33.
818) F. Bernardi, A. Bottoni, J.J.W. McDouall, M.A. Robb and H.B. Schlegel, *Faraday Symp. Chem. Soc.*, 1984, **19**, 137.
819) A.C. Wahl and G. Das in *Methods of Electronic Structure Theory: Modern Theoretical Chemistry*, **Volume 3**, Part 2, H.F. Schaefer, III, editor, Plenum, New York, NY, 1979, p. 51.
820) R. Shepard, *Adv. Chem. Phys.*, 1987, **69**, 63.
821) B.O. Roos, *Int. J. Quantum. Chem. Symp.*, 1980, **14**, 175; *Adv. Chem. Phys.*, 1987, **69**, 399.
822) P. Pulay and T.P. Hamilton, *J. Chem. Phys.*, 1988, **88**, 4926.
823) S.W. Benson, *Thermochemical Kinetics*, Wiley-Interscience, New York, NY, 1976.
824) J.A. Pople, H.B. Schlegel, R. Krishnan, D.J. DeFrees, J.S. Binkley, M.J. Frisch, R.A. Whiteside, R.A. Hout and W.J. Hehre, *Int. J. Quantum Chem. Symp.*, 1981, **15**, 269.
825) A.B. Hess, Jr., L.J. Schaad and P.L. Polavara, *J. Am. Chem. Soc.*, 1984, **106**, 4348.
826) J. Bigeleisen and M. Goeppert-Mayer, M., *J. Chem. Phys.*, 1947, **15**, 261.
827) J. Bigeleisen and M. Wolfsberg, *Adv. Chem. Phys.*, 1958, **1**, 15.
828) C.J. Collins and N.S. Bowman, *Isotopes Effects in Chemical Reactions*, Van Nostrand Reinhold, New York, NY, 1970.
829) J.E. Baldwin, V.P. Reddy, B.A. Hess, Jr. and L.J. Schaad, *J. Am. Chem. Soc.*, 1988, **110**, 8554.
830) J.E. Baldwin, V.P. Reddy, L.J. Schaad and B.A. Hess, Jr., *J. Am. Chem. Soc.*, 1988, **110**, 8555.
831) L. Chantranupong and T.A. Wildman, *J. Am. Chem. Soc.*, 1990, **112**, 4151.
832) *Isotopes in Organic Chemistry*, E. Buncel and E.E. Lee, editors, Elsevier, Amsterdam, 1987.
833) J.P. Foster and F. Weinhold, *J. Am. Chem. Soc.*, 1980, **102**, 7211.
834) A.E. Reed, L.A. Curtiss and F. Weinhold, *Chem. Rev.*, 1988, **88**, 899.
835) I. Mayer, *Chem. Phys. Lett.*, 1983, **97**, 270; *Int. J. Quantum. Chem.*, 1986, **29**, 477.
836) L. Pauling, *J. Am. Chem. Soc.*, 1947, **69**, 524.
837) S.J. Formosinho, *Tetrahedron*, 1986, **42**, 4557
838) L. Valko and P. Simon, *Chem. Phys.*, 1985, **99**, 447.

839) An intrinsic reaction barrier is defined for A–B + C → A + B–C as the mean of the barriers for the reactions: A–B + A → A + B–A and C–B + C → C + B–C. R.A. Marcus, *J. Phys. Chem.*, 1968, **72**, 891.
840) J.R. Murdoch, *J. Am. Chem. Soc.*, 1983, **105**, 2159, 2660, 2667.
841) Femtosecond transition-state spectroscopy studies of chemical reactions may change this; see:
a) M.O. Trulson, G.D. Dollinger and R.A. Mathies, *J. Am. Chem. Soc.*, 1987, **109**, 587.
b) A.H. Zewail, *Science*, 1988, **242**, 1645.
c) I.W.M. Smith, *Nature*, 1987, **328**, 760.
842) J.A. Pople, M. Head-Gordon, D.J. Fox, K. Raghavachari and L.A. Curtiss, *J. Chem. Phys.*, 1989, **90**, 5622.
843) H.E. O'Neal and S.W. Benson in Free Radicals, **Volume 2**, J.K. Kochi, editor, Wiley, New York, NY, 1973, p275-359.
844) W.R. Roth and J. König, *Ann. Chem.*, 1966, **699**, 24; W.R. Roth, *Chimia*, 1966, **299**, 20.
845a) G.J.M. Dormans and H.M. Buck, *J. Mol. Struct. (THEOCHEM)*, 1986, **136**, 121.
845b) G.J.M. Dormans and H.M. Buck, *J. Am. Chem. Soc.*, 1986, **108**, 3253.
845c) B.A. Hess, Jr., L.J. Schaad and J. Pancir, *J. Am. Chem. Soc.*, 1985, **107**, 149.
845d) B.A. Hess, Jr. and L.J. Schaad, *J. Am. Chem. Soc.*, 1983, **105**, 7185.
846a) F. Jensen and K.N. Houk, *J. Am. Chem. Soc.*, 1987, **109**, 3139.
846b) S.D. Kahn, W.J. Hehre, N.G. Rondan and K.N. Houk, *J. Am. Chem. Soc.*, 1985, **107**, 8291.
846c) N.G. Rondan and K.N. Houk, *Tetrahedron Lett.*, 1984, **25**, 2519.
846d) M.J.S. Dewar, E.F. Healy and J.M. Ruiz, *J. Am. Chem. Soc.*, 1988, **110**, 2666.
847) Location of a typical MP2/6-31G* transition structure of six heavy atoms requires 3 CPU hours of Cray Y-MP using the GAUSSIAN 90.
848) M.J.S. Dewar, K.M. Merz and J.J.P. Stewart, *J. Chem. Soc., Chem. Commun.*, 1985, 166.
849a) M.J.S. Dewar, E.G. Zoebisch, E.F. Healy and J.J.P. Stewart, *J. Am. Chem. Soc.*, 1985, **107**, 3902.
M.J.S. Dewar and S. Kirchner, *J. Am. Chem. Soc.*, 1971, **93**, 4290, 4291, 4292.
849b) J. Breulet and H.F. Schaefer, III, *J. Am. Chem. Soc.*, 1984, **106**, 1221.
849c) F. Bernardi, S. De, M. Olivucci and M.A. Robb, *J. Am. Chem. Soc.*, 1990, **112**, 1737 and references cited therein.
850a) K. Hsu, R.J. Buenker and S.D. Peyerimhoff, *J. Am. Chem. Soc.*, 1971, **93**, 2717; K. Hsu, R.J. Buenker and S.D. Peyerimhoff, *J. Am. Chem. Soc.*, 1972, **94**, 5639; R.J. Buenker, S.D. Peyerimhoff and K. Hsu, J. Am. Chem. Soc., 1971, **93**, 5005.
850b) W. Thiel, *J. Am. Chem. Soc.*, 1981, **103**, 1420.
850c) J.W. McIver, Jr., and A. Komornicki, *J. Am. Chem. Soc.*, 1972, **94**, 2625.
850d) G. Frenking and N. Heinrich, *Theor. Chim. Acta*, 1984, **65**, 65.
850e) A. Jensen and H. Kunz, *Theor. Chim. Acta*, 1984, **65**, 33.
850f) T.A. Halgren and W.N. Lipscomb, *Chem. Phys. Lett.*, 1977, **49**, 225.
851a) D.C. Spellmeyer and K.N. Houk, *J. Am. Chem. Soc.*, 1988, **110**, 3412.
851b) N.G. Rondan and K.N. Houk, *J. Am. Chem. Soc.*, 1984, **106**, 7989; 1985, **107**, 2099.
851c) W. Kirmse, N.G. Rondan and K.N. Houk, *J. Am. Chem. Soc.*, 1984, **106**, 1871.
851d) E.A. Kallel, Y. Wang and K.N. Houk, *J. Org. Chem.*, 1989, **54**, 6006.
852a) M. Kraus, M. Vavruska and V. Bazant, *Coll. Czech. Chem. Commun.*, 1957, **22**, 484.
852b) L. Kuchler, *Trans. Faraday Soc.*, 1939, **35**, 874.
852c) C.K. Lewis, J. Bergmann, R. Manjoney, R. Paddock, and B.L. Kalra, *J. Phys. Chem.*, 1984, **88**, 4112.
852d) D. Rowley and H. Steiner, *Discuss. Faraday Soc.*, 1951, **10**, 198.
852e) S.R. Smith and A.S. Gordon, *J. Phys. Chem.*, 1961, **65**, 1124.
852f) D.C. Tardy, R. Ireton and A.S. Gordon, *J. Am. Chem. Soc.*, 1979, **101**, 1508.
852g) W.J. Tsang, *Chem. Phys.*, 1965, **42**, 1805; W.J. Tsang, *Int. J. Chem. Kinet.*, 1970, **2**, 311.

852h) M. Uchiyama, T. Tomioka and A. Amano, *J. Phys. Chem.*, 1964, **68**, 1878.
853) W. von E. Doering, W.R. Roth, R. Breuckmann, L. Figge, H.W. Lennartz, W.D. Fessner and H. Prinzbach, *Chem. Ber.*, 1988, **121**, 1.
854a) M.J.S. Dewar, S. Olivella and J.J.P. Stewart, *J. Am. Chem. Soc.*, 1986, **108**, 5771; M.J.S. Dewar, S. Olivella and H.S. Rzepa, *J. Am. Chem. Soc.*, 1978, **100**, 5650.
854b) J.W. McIver, Jr., *Acc. Chem. Res.*, 1974, **7**, 72.
854c) M.V. Basilevsky, A.G. Shamov and V.A. Tikhomirov, *J. Am. Chem. Soc.*, 1977, **99**, 1369; M.V. Basilevsky, V.A. Tikhomirov and I.E. Chenlov, *Theor. Chim. Acta*, 1971, **23**, 75.
854d) O. Kikuchi, *Tetrahedron*, 1971, **27**, 2791.
854e) J. Pancir, *J. Am. Chem. Soc.*, 1982, **104**, 7424.
855a) K.N. Houk, Y.T. Lin, and F.K. Brown, *J. Am. Chem. Soc.*, 1986, **108**, 544; F.K. Brown and K.N. Houk, *Tetrahedron Lett.*, 1984, **25**, 4609.
855b) R.D. Bach, J.J.W. McDougall and H.B. Schlegel, *J. Org. Chem.*, 1989, **54**, 2931.
856a) R.E. Townshend, G. Ramunni, G. Segal, W.J. Hehre and L. Salem, *J. Am. Chem. Soc.*, 1976, **98**, 2190.
856b) L.A. Burke, *Theor. Chim. Acta*, 1985, **68**, 101; L.A. Burke and G. Leroy, *Theor. Chim. Acta*, 1977, **44**, 219; L.A. Burke, G. Leroy and M. Sana, *Theor. Chim. Acta*, 1975, **40**, 313.
856c) M. Ortega, A. Oliva, J.M. Lluch and J. Bertran, *Chem. Phys. Lett.*, 1983, **102**, 317.
857) F. Bernardi, A. Bottoni, M.J. Field, M.F. Guest, I.H. Hillier, M.A. Robb and A. Venturini, *J. Am. Chem. Soc.*, 1988, **110**, 3050; F. Bernardi, A. Bottoni, M.A. Robb and M.J. Field, *J. Chem. Soc., Chem. Commun.*, 1985, 1051.
858) W. Cooper, and W.D. Walters, *J. Am. Chem. Soc.*, 1958, **80**, 4220; R.W. Carr, Jr., and W.D. Walters, *J. Phys. Chem.*, 1965, **69**, 1073.
859) S. Huzinaga, *J. Chem. Phys.*, 1965, **42**, 1293; T.H. Dunning, *J. Chem. Phys.*, 1970, **53**, 2823.
860) R.B. Woodward and R. Hoffmann, *J. Am. Chem. Soc.*, 1965, **87**, 2046, 4389; R. Hoffmann and R.B. Woodward, *J. Am. Chem. Soc.*, 1965, **87**, 395, 2511.
861) R.B. Woodward and R. Hoffmann, *Angew. Chem.*, 1969, **81**, 797; *Angew. Chem. Int. Ed. Engl.*, 1969, **8**, 781; *The Conservation of Orbital Symmetry*, Verlag Chemie, Weinheim, 1970.
862a) K. Fukui, *Tetrahedron Lett.*, 1965, 2009, 2427; *Acc. Chem. Res.*, 1971, **4**, 57.
862b) *Orbital Symmetry Papers*, H.E. Simmons and J.F. Bunnett, editors, American Chemical Society, Washington, DC, 1976.
863a) H.E. Zimmerman, *Acc. Chem. Res.*, 1971, **4**, 272.
863b) M.J.S. Dewar, *Angew. Chem.*, 1971, **83**, 859; *Angew Chem. Int. Ed. Engl.*, 1971, **10**, 761.
864) M.J.S. Dewar, *J. Am. Chem. Soc.*, 1984, **106**, 209.
865) G. Desimoni, G. Tacconi, A. Barco and G.P. Pollini, *Natural Products Synthesis Through Pericyclic Reactions*, Americal Chemical Society, Washington, DC, 1983.
866) Cortisone was synthesised *via* a Diels-Alder reaction. R.B. Woodward, F. Sondheimer, D. Taub, K. Heusler and W.M. McLamore, *J. Am. Chem. Soc.*, 1952, **74**, 4223.
867) Lasalocid A was synthesised using the Ireland version of the Claisen rearrangement. R.E. Ireland, S. Thaisrivangs and C.S. Wilcox, *J. Am. Chem. Soc.*, 1980, **102**, 1155.
868) Xylopinine was prepared *via* the electocyclic openings of benzocyclobutenes and electrocyclisation of a hexatriene. T. Kametani, M. Kajiwara, T. Takahashi and K. Fujimoto, *J. Chem. Soc. Perkin Trans.*, 1975, **1**, 737
T. Kametani, K. Ogasawara and T. Takahashi, *Tetrahedron*, 1973, **29**, 73.
869) Ergosterol undergoes a series of photochemical and thermal pericyclic reactions to form vitamin D2. M.P. Rappold, E. Havinga, *Recl. Trav. Chim. Pays-Bas*, 1960, **79**, 369; E. Havinga, *Tetrahedron*, 1961, **16**, 146.
For a recent study see: J.D. Enas, G.Y. Shen and W.H. Okamura, *J. Am. Chem. Soc.*, 1991, **113**, 3873 and references cited therein.

870) The chorismate → prephenate rearrangement occurring in the biosynthesis of aromatic amino acids in plants, is a Claisen rearrangement catalysed by the enzyme, *chorismate mutase*. S.G. Sogo, T.S. Widlanski, J.H. Hoare, C.E. Grimshaw, G.A. Berchtold and J.R. Knowles, *J. Am. Chem. Soc.*, 1984, **106**, 2701; Y. Asano, J.I. Lee, T.L. Shieh, F. Spreafico, C. Kowal and H.G. Floss, *J. Am. Chem. Soc.*, 1985, **107**, 4314.

871) The electrocyclic formation of a biradical that reacts with DNA is the origin of activity of the *esperamicin* and *calicheamicin* series of anitbiotics. See: J. Golik, J. Clardy, G. Dubay, G. Groenewold, H. Kawaguchi, M. Konishi, B. Krishnan, H. Ohkuma, K. Saitoh and T.W. Doyle, *J. Am. Chem. Soc.*, 1987, **109**, 3461; J. Golik, G. Dubay, G. Groenewold, H. Kawaguchi, M. Konishi, B. Krishnan, H. Ohkuma, K. Saitoh and T.W. Doyle, *J. Am. Chem. Soc.*, 1987, **109**, 3462; M.D. Lee, T.S. Dunne, M.M. Siegel, C.C. Chang, G.O. Morton and D.B. Borders, *J. Am. Chem. Soc.*, 1987, **109**, 3464; M.D. Lee, T.S. Dunne, C.C. Chang, G.A. Ellestad, M.M. Siegel, G.O. Morton, W.J. McGahren and D.B. Borders, *J. Am. Chem. Soc.*, 1987, **109**, 3466; N. Zein, M. Poncin, R. Nilakantan and G.A. Ellestad, *Science*, 1989, **244**, 697; K.C. Nicolaou and W.M. Dai, *Angew. Chem.*, 1991, **103**, 1453; *Angew. Chem. Int. Ed. Engl.*, 1991, **30**, 1387.

872) The biradical mechanism has prompted wide interest in this reaction, and in the use of thermal rearrangements to generate reactive intermediates for biological action. K.C. Nicolaou, Y. Ogawa, G. Zuccarello and H. Kataoka, *J. Am. Chem. Soc.*, 1988, **110**, 7247; K.C. Nicolaou, G. Zuccarello, Y. Ogawa, E.J. Schwieger and T. Kumazawa, *J. Am. Chem. Soc.*, 1988, **110**, 4866; J.P. Snyder, *J. Am. Chem. Soc.*, 1989, **111**, 7630; 1990, **112**, 5367; P. Magnus, S. Fortt, T. Pitterna and J.P. Snyder, *J. Am. Chem. Soc.*, 1990, **112**, 4986; N. Koga and K. Morokuma, *J. Am. Chem. Soc.*, 1991, **113**, 1907.

873) K.N. Houk in *Pericyclic Reactions*, **Volume 2**, A.P. Marchand and R.E. Lehr, editors, Academic Press, New York, NY, 1977.

874) M.T. Reetz, *Tetrahedron*, 1973, **29**, 2189.

875) G.S. Hammond, *J. Am. Chem. Soc.*, 1955, **77**, 334.

876a) R.P. Bell, *Proc. R. Soc. London A*, 1936, **154**, 414.

876b) M.G. Evans and M. Polanyi, *Trans. Faraday Soc.*, 1938, **34**, 11.

877) R.H. Higgins, *J. Chem. Educ.*, 1995, **72**, 703.

878) *PCModel* is a product of Serena Software, Box 3076, Bloomington, IN 47402.

879) The *.mop* file is set out in Z-matrix format as follows:

Line 1				Commands are written here						
Line 2				Name of compound						
Line 3										
Line 4	Atom	0	0†	0	0	0	0	0¶	0	0	q_1
Line 5	Atom	l_{21}	1	0	0	0	0	1	0	0	q_2
Line 6	Atom	l_{31}	1	θ_{321}	1	0	0	2	1	0	q_3
Line 7	Atom	l_{43}	1	θ_{432}	1	ϕ_{4321}	1	3	2	1	q_4
Last line										

where:

	l_{21}	=	bond length atom2-atom1
	θ_{321}	=	bond angle atom3-atom2-atom1
	ϕ_{4321}	=	torsional angle atom4-atom3-atom2-atom1
	q_4	=	partial atomic charge on atom 4
¶	Connectivity	=	atoms which define the internal coordinates.

† Each internal coordinate is followed by an integer:

1 Optimise the internal co-ordinate.

0 Do not optimise the internal co-ordinate.

-1 Reaction co-ordinate, or grid index.

880) The third line of the *.mop* file is empty to signify the end of the command lines. The last line of the *.mop* file is empty to signify the end of the file.
881) M.J.S. Dewar, E.F. Healy and J.J.P. Stewart, *J. Chem. Soc., Faraday Trans. 2*, 1984, **80**, 227.
882) D.K. Agrafiotis and H.S. Rzepa, *J. Chem. Res (S)*, 1988, 100.
883) K. Fukui, *Acc. Chem. Res.*, 1981, **14**, 368.
884) K. Ishida, K. Morokuma and A. Komornicki, *J. Chem. Phys.*, 1977, **66**, 2153.
885) K. Müller and L.D. Brown, *Theor. Chemica Acta*, 1979, **53**, 75.
886) C. Gonzalez and H.B. Schlegel, *J. Chem. Phys.*, 1988, **90**, 2154.
887a) W.A. Lathan, W.J. Hehre and J.A. Pople, *J. Am. Chem. Soc.*, 1971, **93**, 808.
887b) P.C. Hariharan, W.A. Lathan and J.A. Pople, *Chem. Phys. Lett.*, 1972, **14**, 385.
887c) B. Zurawski, R. Ahlrichs and W. Kutzelnigg, *Chem. Phys. Lett.*, 1973, **21**, 309.
887d) J. Weber and A.D. McLean, *J. Am. Chem. Soc.*, 1976, **98**, 875.
887e) J. Weber, M. Yoshimine and A.D. McLean, *J. Chem. Phys.*, 1976, **64**, 4159.
887f) H. Lischka and H.J. Köhler, *J. Am. Chem. Soc.*, 1978, **100**, 5297.
887g) K. Raghavachari, R.A. Whiteside, J.A. Pople and P. von R. Schleyer, *J. Am. Chem. Soc.*, 1981, **103**, 5649.
887h) E.D. Jemmis, K.S. Sarma and P.N.V. Pavankumar, *THEOCHEM*, 1985, **22**, 305.
888a) L.B. Harling, *J. Am. Chem. Soc.*, 1981, **103**, 7469.
888b) Y. Osamura, T. Takeughi and K. Nishimoto, *Bull. Chem. Soc. Jpn.*, 1988, **61**, 3387.
888c) T.P. Hamilton and H.F. Schaefer III, *J. Phys. Chem.*, 1989, **93**, 7560.
889) W.K. Li, R.H. Nobes and L. Radom, *THEOCHEM*, 1986, **29**, 121.
890) W.R. Rodwell, W.J. Bouma and L. Radom, *Int. J. Quantum Chem.*, 1980, **18**, 107.
891a) R.A. Poirier, D. Majlessi and T.J. Zielinski, *J. Comput. Chem.*, 1986, **7**, 464.
891b) S. Schröder, and W. Thiel, *J. Am. Chem. Soc.*, 1986, **108**, 7985.
892) F. Bernardi, M.A. Robb, H.B. Schlegel and G. Tonachini, *J. Am. Chem. Soc.*, 1984, **106**, 1198.
893a) T. Clark, *J. Am. Chem. Soc.*, 1987, **109**, 6838; 1989, **111**, 761.
893b) P. Du, D.A. Hrovat and W.T. Borden, *J. Am. Chem. Soc.*, 1988, **110**, 3405.
893c) M.T. Nguyen, L. Landuyt and L.G. Vanquickenborne, *Chem. Phys. Lett.*, 1991, **182**, 225.
894) D.F. McMillen and D.M. Golden, *Annu. Rev. Phys. Chem.*, 1982, **33**, 483.
895a) J.A. Berson and L. Salem, *J. Am. Chem. Soc.*, 1972, **94**, 8917.
895b) J.A. Berson, *Acc. Chem. Res.*, 1982, **5**, 406.
896) For a review of "walk" rearrangement see: F.G.T. Klärner, *Top. Stereochem.*, 1984, **15**, 1; R.F. Childs, *Tetrahedron*, 1982, **38**, 567.
897) For reviews, see:
R.H. DeWolfe in *Comprehensive Chemical Kinetics*, C.H. Bamford and C.F.H. Tipper editors, American Elsevier Publishing Company, New York, NY, 1973, **Volume 9**, pp. 470-474.
C.D. Gutsche and D. Redmore, *Carbocylic Ring Expansion Reactions*, Academic Press, Inc., New York, NY, 1968, pp. 163-170.
For a review of expansions and contractions of three- and four-membered rings, see:
J.M. Conia and M.J. Robson, *Angew. Chem. Int. Ed. Engl.*, 1975, **14**, 473; H.M. Frey, *Adv. Phys. Org. Chem.*, 1966, **4**, 147.
898) A.D. Ketley, *Tetrahedron Lett.*, 1964, 1687; G.R. Branton and H.M. Frey, *J. Chem. Soc. A*, 1966, 1342.
899) For evidence favouring the concerted mechanism, see:
T.C. Shields, W.E. Billups and A.R. Lepley, *J. Am. Chem. Soc.*, 1968, **90**, 4749; J.A. Berson and P.B. Dervan, *J. Am. Chem. Soc.*, 1973, **95**, 267, 269; W.E. Billups, K.H. Leavell, E.S. Lewis and S. Vanderpool, *J. Am. Chem. Soc.*, 1973, **95**, 8096.
900) For evidence favouring the diradical mechanism, see:
M.R. Willcott and V.H. Cargle, *J. Am. Chem. Soc.*, 1967, **89**, 723; W. von E. Doering and E.K.G. Schmidt, *Tetrahedron*, 1971, **27**, 2005; W.R. Roth and Th. Schmidt, *Tetrahedron Lett.*, 1971, 3639; J.M. Simpson and H.G. Richey, Jr., *Tetrahedron Lett.*,

1973, 2545; J.C. Gilbert and D.P. Higley, *Tetrahedron Lett.*, 1973, 2075; P. Caramella, R. Huisgen and B. Schmolke, *J. Am. Chem. Soc.*, 1974, **96**, 2997, 2999; P.H. Mazzocchi and H.J. Tamburin, *J. Am. Chem. Soc.*, 1975, **97**, 555.
A continuous diradical transition state has also been proposed:
W. von E. Doering and K. Sachdev, *J. Am. Chem. Soc.*, 1974, **96**, 1168; 1975, **97**, 5512.

901a) J.J. Quirante, F. Enriquez and J.M. Hernando, *J. Mol. Struct. (THEOCHEM)*, 1990, **204**, 193.

901b) J.J. Gakewski and M.P. Squicciarini, *J. Am. Chem. Soc.*, 1989, **111**, 6717.

901c) M.R. Willcott III and V.H. Cargle, *J. Am. Chem. Soc.*, 1967, **89**, 723.
M.R. Willcott III and V.H. Cargle, *J. Am. Chem. Soc.*, 1969, **91**, 4310.

902a) F. Jensen, *J. Am. Chem. Soc.*, 1989, **111**, 4643.

902b) P.N. Skancke, K. Yamashita and K. Morokuma, *J. Am. Chem. Soc.*, 1987, **109**, 4157.

903) T.S. Sorensen and A. Rauk in *Pericyclic Reactions*, **Volume 2**, A.P. Marchand and R.E. Lehr, editors, Academic Press, New York, 1977.

904) For a review of thermal [1,j]-sigmatropic rearrangement: C.W. Spangler, *Chem. Rev.*, 1976, **76**, 187 and references cited therein.

905) W.T. Borden, J.G. Lee and S.D. Young, *J. Am. Chem. Soc.*, 1980, **102**, 4841.

906) R.J. Ellis and H.M. Frey, *Proc. Chem. Soc.*, 1964, 221; W.R. Roth and J. König, *Justus Liebigs Ann. Chem.*, 1965, **688**, 28; G. Ohloff, *Tetrahedron Lett.*, 1965, 3795; M.J. Jorgenson and A.F. Thacher, *Tetrahedron Lett.*, 1969, 4651; E.J. Corey, H. Yamamoto, D.K. Herron and K. Achiwa, *J. Am. Chem. Soc.*, 1970, **92**, 6635.
See also: J.K. Crandall and R.J. Watkins, *Tetrahedron Lett.*, 1967, 1717.

907) M.J.S. Dewar and E.F. Healy, *Chem. Phys. Lett.*, 1982, **141**, 521.

908) W. von E. Doering and W.R. Roth, *Tetrahedron*, 1962, **18**, 67.
See also: R.K. Hill and N.W. Gilman, *Chem. Commun.*, 1967, 619; M.J. Goldstein and M.R. DeCamp, *J. Am. Chem. Soc.*, 1974, **96**, 7356; H.-J. Hansen and H. Schmid, *Tetrahedron*, 1974, **30**, 1959.

909) Preference for the chair transition state is a consequence of orbital symmetry relationships: R. Hoffmann and R.B. Woodward, *J. Am. Chem. Soc.*, 1965, **87**, 4389; K. Fukui and H. Fujimoto, *Tetrahedron Lett.*, 1966, 251.

910) K. Morokuma, W.T. Borden and D.A. Hrovat, *J. Am. Chem. Soc.*, 1988, **110**, 4474.

911) M. Bearpark, F. Bernardi, M. Olivucci and M.A. Robb, *J. Am. Chem. Soc.*, 1990, **112**, 1732.

912a) M.J.S. Dewar and C. Jie, *J. Am. Chem. Soc.*, 1987, **109**, 5983.

912b) E.A. Halevi and R. Rom, *Isr. J. Chem.*, 1989, **29**, 311.

913) Y. Osamura, S. Kato, K. Morokuma, D. Feller, E.R. Davidson and W.T. Borden, *J. Am. Chem. Soc.*, 1984, **106**, 3362.

914) W. von E. Doering, V.G. Toscano and G.H. Beasley, *Tetrahedron*, 1971, **27**, 5299.

915a) W.T. Borden, R.J. Loncharich and K.N. Houk, *Annu. Rev. Phys. Chem.*, 1988, **39**, 213.

915b) W.R. Roth, H.W. Lennartz, W. von E. Doering, L. Birladeanu, C.A. Guyton and T. Kitagawa, *J. Am. Chem. Soc.*, 1990, **112**, 1722.

915c) Note added in proof: The best available calculations predict concerted and diradical mechanisms of similar energy: M. Dupis, C. Murray and E.R. Davidson, *J. Am. Chem. Soc.*, 1991, **113**, 9756.

916) M.J.S. Dewar, G.P. Ford, M. L. McKee, H.S. Rzepa and L.E. Wade, *J. Am. Chem.Soc.*, 1977, **99**, 5069.

917) For reviews of Cope rearrangements involving triple bonds, see:
A. Viola, J.J. Collins and N. Filipp, *Tetrahedron*, 1981, **37**, 3765; W.D. Huntsman, *Intra-Sci. Chem. Rep.*, 1972, **6**, 151-159.
F. Théron, M. Verny and R. Vessière in *The Chemistry of the Carbon-Carbon Triple Bond*, S. Patai, editor, Wiley, New York, NY, 1978, **Part 1**, pp. 381-445.

918a) W.D. Huntsman and H.J. Wristers, *J. Am. Chem. Soc.*, 1967, **89**, 342.

818b) K.A. Owens and J.A. Berson, *J. Am. Chem. Soc.*, 1990, **112**, 5973.

919) K.A. Owens and J.A. Berson, *J. Am. Chem. Soc.*, 1990, **112**, 5973.

920) von E. Vogel, K.-H. Ott and K. Gajek, *Justus Liebigs Ann. Chem.*, 1961, **644**, 172.

921) Unsubstituted *cis*-1,2-divinylcyclopropane is fairly stable at -20 °C.
J.M. Brown, B.T. Golding and J.J. Stofko, Jr., *J. Chem. Soc., Chem. Commun.*, 1973, 319; M.P. Schneider and J. Rebell, *J. Chem. Soc., Chem. Commun.*, 1975, 283.

922) J.M. Brown. *Chem. Commun.*, 1965, 226; D. Schönleber, *Chem. Ber.*, 1969, **102**, 1789; J.E. Baldwin and C. Ullenius, *J. Am. Chem. Soc.*, 1974, **96**, 1542.

923) G.S. Hammond and C.D. De Boer, *J. Am. Chem. Soc.*, 1964, **86**, 899; J.T. Trecker and J.P. Henry, *J. Am. Chem. Soc.*, 1964, **86**, 902; H. Kessler and W. Ott, *J. Am. Chem. Soc.*, 1976, **98**, 5014; W.R. Dolbier, Jr., and G.J. Mancini, *Tetrahedron Lett.*, 1975, 2141.

924) Part of the cyclooctadiene produced may come from a prior epimerisation of *trans*- to *cis*-divinylcyclobutadiene, followed by Cope rearrangement of the latter. See:
J.A. Berson and P.B. Dervan, *J. Am. Chem. Soc.*, 1972, **94**, 8949; J.E. Baldwin and K.E. Gilbert, *J. Am. Chem. Soc.*, 1976, **98**, 8283.
For a similar result in the 1,2-divinyl cyclopropane series, see:
J.E. Baldwin and C. Ullenius, *J. Am. Chem. Soc.*, 1974, **96**, 1542.

925) Another type of diradical two-step mechanism in which the 1,6 bond is formed before the 3,4 breaks may be preferred by some substrates. W. von E. Doering, V.G. Toscano and C.H. Beasley, *Tetrahedron*, 1971, **27**, 5299; M.J.S. Dewar and L.E. Wade, *J. Am. Chem. Soc.*, 1973, **95**, 290; A. Padwa and T.J. Blacklock, *J. Am. Chem. Soc.*, 1980, **102**, 2797; M. Dollinger, W. Henning and W. Kirmse, *Chem. Ber.*, 1982, **115**, 2309; D. Kaufmann and A. de Meijere, *Chem. Ber.*, 1984, **117**, 1128; M.J.S. Dewar and C. Jie, *J. Am. Chem. Soc.*, 1987, **109**, 5893; *J. Chem. Soc., Chem. Commun.*, 1989, 98.
For evidence against this view, see:
J.J. Gwajewski and N.D. Conrad, *J. Am. Chem. Soc.*, 1978, **100**, 6268, 6269; 1979, **101**, 6693; J.J. Gajewski, *Acc. Chem. Res.*, 1980, **13**, 142; K. Morokuma, W.T. Borden and D.A. Hrovat, *J. Am. Chem. Soc.*, 1988, **110**, 4474; K.A. Owens and J.A. Berson, *J. Am. Chem. Soc.*, 1990, **112**, 5973.

926) W. von E. Doering and W.R. Roth, *Tetrahedron*, 1963, **19**, 715.

927) For reviews of valence tautomerisations, see:
B. Decock-Le Révérend and P. Goudmand, *Bull. Soc. Chim. Fr.*, 1973, 389; L.A. Paquette, *Angew. Chem. Int. Ed. Engl.*, 1971, **10**, 11-20; T.V. Domareva-Mandel'shtam and I.A. D'yakonov, *Russ. Chem. Rev.*, 1966, **35**, 559; G. Schröder, J.F.M. Oth and R. Merényi, *Angew. Chem. Int. Ed. Engl.*, 1965, **4**, 752.

928) G. Schröder, *Angew. Chem. Int. Ed. Engl.*, 1963, **2**, 481; *Chem. Ber.*, 1964, **97**, 3140.
R. Merényi, J.F.M. Oth and G. Schröder, *Chem. Ber.*, 1964, **97**, 3150.
For a review of bullvalenes, see:
G. Schröder and J.F.M. Oth, *Angew. Chem. Int. Ed. Engl.*, 1967, **6**, 414.

929) A number of azabullvalenes have been synthesised, being fluxional when heated, with fewer tautomeric forms than bullvalene itself.
L.A. Paquette and T.J. Barton, *J. Am. Chem. Soc.*,1967, **89**, 5480; L.A. Paquette, T.J. Barton and E.B. Whipple, *J. Am. Chem. Soc.*, 1967, **89**, 5481; P. Wegener, *Tetrahedron Lett.*, 1967, 4985; L.A. Paquette, J.R. Malpass, G.R. Krow and T.J. Barton, *J. Am. Chem. Soc.*, 1969, **91**, 5296.

930) J.F.M. Oth, K. Müllen, J.-M. Gilles and G. Schröder, *Helv. Chim. Acta*, 1974, **57**, 1415; H. Nakanishi and O. Yamamoto, *Tetrahedron Lett.*, 1974, 1803; H. Günther and J. Ulmen, *Tetrahedron*, 1974, **30**, 3781; R. Poupko, H. Zimmermann, Z. Luz, *J. Am. Chem. Soc.*, 1984, **106**, 5391; P. Luger, J. Buschmann, R.K. McMullan, J.R. Ruble, P. Matias and G.A. Jeffrey, *J. Am. Chem. Soc.*, 1986, **108**, 7825.

931) Hypostrophene: J.S. McKennis, L. Brener, J.S. Ward and R. Pettit, *J. Am. Chem. Soc.*, 1971, **93**, 4957; L.A. Paquette, R.F. Davis and D.R. James, *Tetrahedron Lett.*, 1974, 1615.

932) Barbaralane: U. Biethan, H. Klusacek and H. Musso, *Angew. Chem. Int. Ed. Engl.*, 1967, **6**, 176; H. Tsuruta, K. Kurabayashi and T. Mukai, *Tetrahedron Lett.*, 1965, 3775; W. von E. Doering, B.M. Ferrier, E.T. Fossel, J.H. Hartenstein, M. Jones Jr., G. Klumpp, R.M. Rubin and M. Saunders, *Tetrahedron*, 1967, **23**, 3943; J.G. Henkel and J.T. Hane, *J. Org. Chem.*, 1983, **48**, 3858. ^{1}H-NMR Spectra indicate that even at room temperature a rapid interchange of both tautomers is present, though by about -100 °C this has slowed to a point where the spectrum is in accord with a single structure.

933) Semibullvalene has the lowest energy barrier of any known compound capable of undergoing the Cope rearrangement.
A. Cheng, F.A.L. Anet, J. Mioduski and J. Meinwald, *J. Am. Chem. Soc.*, 1974, **96**, 2887; D. Moskau, R. Aydin, W. Leber, H. Günther, H.. Quast, H.D. Martin, K. Hassenrück, L.S. Miller and K. Grohmann, *Chem. Ber.*, 1989, **122**, 925.
There is a rapid interconversion even at -110 °C. H.E. Zimmerman and G.L. Grunewald, *J. Am. Chem. Soc.*, 1966, **88**, 183; J. Meinwald and D. Schmidt, *J. Am. Chem. Soc.*, 1969, **91**, 5877; H.E. Zimmerman, J.D. Robbins and J. Schantl, *J. Am. Chem. Soc.*, 1969, **91**, 5878; H.E. Zimmerman, R.W. Binkley, R.S. Givens, G.L. Grunewald and M.A. Sherwin, *J. Am. Chem. Soc.*, 1969, **91**, 3316.

934) E. Ciganek, *J. Am. Chem. Soc.*, 1965, **87**, 1149.
For other examples of norcaradiene-cycloheptatriene valence tautomerisations, see:
M. Görlitz and H. Günther, *Tetrahedron*, 1969, **25**, 4467; H.J. Reich, E. Ciganek and J.D. Roberts, *J. Am. Chem. Soc.*, 1970, **92**, 5166; G.E. Hall and J.D. Roberts, *J. Am. Chem. Soc.*, 1971, **93**, 2203; E. Ciganek, *J. Am. Chem. Soc.*, 1971, **93**, 2207; H. Günther, B.D. Tunggal, M. Regitz, H. Scherer and T. Keller, *Angew. Chem. Int. Ed. Engl.*, 1971, **10**, 563; H. Dürr and H. Kober, *Chem. Ber.*, 1973, **106**, 1565; W. Betz and J. Daub, *Chem. Ber.*, 1974, **107**, 2095.

935) J.E. Baldwin in *Pericyclic Reactions*, A.P. Marchand and R.E. Lehr, editors, Academic Press, New York, NY, 1977, **Volume 2**, p.173.

936) For discussions, see:
C.H. DePuy, *Acc. Chem. Res.*, 1968, **1**, 33; U. Schöllkopf, *Angew. Chem. Int. Ed. Engl.*, 1968, **7**, 588.

937) For a review of ring opening of cyclopropyl anions and related reactions, see:
G. Boche, *Top. Curr. Chem.*, 1988, **146**, 1.
For evidence that this is so, see:
M. Newcomb and W.T. Ford, *J. Am. Chem. Soc.*, 1974, **96**, 2968; G. Boche, K. Buckl, D. Martens, D.R. Schneider and H.-U. Wagner, *Chem. Ber.*, 1979, **112**, 2961; R.M. Coates and L.A. Last, *J. Am. Chem. Soc.*, 1983, **105**, 7322.

938) The statement was first proposed by: C.H. DePuy, L.G. Schnack, J.W. Hausser and W. Wiedemann, *J. Am. Chem. Soc.*, 1965, **87**, 4006.

939) The second pathway is possible in certain cases: J.W. Hausser and M.J. Grubber, *J. Org. Chem.*, 1972, **37**, 2648; J.W. Hausser and J.T. Uchic, *J. Org. Chem.*, 1972, **37**, 4087.

940) There is more evidence, for example:
C.W. Jefford and R. Medary, *Tetrahedron Lett.*, 1966, 2069; C.W. Jefford, E.H. Yen, and R. Medary, *Tetrahedron Lett.*, 1966, 6317; C.W. Jefford and D.T. Hill, *Tetrahedron Lett.*, 1969, 1957; P. von R. Schleyer, G.W. Van Dine, U. Schöllkopf and J. Paust, *J. Am. Chem. Soc.*, 1966, **88**, 2868; P. von R. Schleyer, W.F. Sliwinski, G.W. Van Dine, U. Schöllkopf, J. Paust and K. Fellenberger, *J. Am. Chem. Soc.*, 1972, **94**, 125; W.F. Sliwinski, T.M. Su, and P. von R. Schleyer, *J. Am. Chem. Soc.*, 1972, **94**, 133; S.R. Sandler, *J. Org. Chem.*, 1967, **32**, 3876; L. Ghosez, P. Laroche and G. Slinckx, *Tetrahedron Lett.*, 1967, 2767; L. Ghosez, G. Slinckx, M. Glineur, P. Hoet and P. Laroche, *Tetrahedron Lett.*, 1967, 2773; W.E. Parham and R.J. Sperley, *J. Org. Chem.*, 1967, **32**, 924, 926; W.E. Parham and K.S. Yong, *J. Org. Chem.*, 1968, **33**, 3947; M.S. Baird and C.B. Reese, *Tetrahedron Lett.*, 1969, 2117; M.S. Baird, D.G. Lindsay and C.B. Reese, *J. Chem. Soc., C.*, 1969, 1173.

941) U. Schöllkopf, K. Fellenberger, M. Patsch, P. von R. Schleyer, T. Su and G.W. Van Dine, *Tetrahedron Lett.*, 1967, 3639.
942) P. von R. Schleyer, T.M. Su, M. Saunders and J.C. Rosenfeld, *J. Am. Chem. Soc.*, 1969, **91**, 5174.
943) P. von R. Schleyer and M. Bremer, *J. Org. Chem.*, 1988, **53**, 2362.
944) S. Olivella, A. Sole and J.M. Bofill, *J. Am. Chem. Soc.*, 1990, **112**, 2160 and references cited therein.
945) V. Thoss, *Imperial College International Diploma Thesis*, Imperial College, London, 1992; M.S. Baird, J.R. Al-Dulayymi, H.S. Rzepa and V. Thoss, *J. Chem. Soc., Chem. Commun.*, 1992, 1323.
946) D.K. Bloomfield and K. Bloch, *J. Biol. Chem.*, 1960, **235**, 337.
947) D.A. van Dorp, *Proc. Nutr. Soc.*, 1975, **34**, 279.
948) A.R. Johnson, J.A. Pearson, F.S. Shenstone and A.C. Fogerty, *Nature (London)*, 1967, **214**, 1244.
949) H.W. Kircher, *J. Am. Oil. Chem. Soc.*, 1964, **41**, 4.
950) D.A. Dows, G.C. Pimentel and E. Whittle, *J. Chem. Phys.*, 1955, **23**, 1606.
951) S.F. Foner and R.C. Hudson, *J. Chem. Phys.*, 1958, **28**, 719.
952) E.J. Blau, B.F. Hochheimer and H.J. Unger, *J. Chem. Phys.*, 1961, **34**, 1060.
953) P. Falciola and A. Mannino, *Ann. Chim. Appl.*, 1914, **2**, 351; L.F. Audrieth and B.A. Ogg, *The Chemistry of Hydrazine*, John Wiley and Sons, Inc., New York, NY, 1951, *Chapter VI*; F. Aylward and M. Sawistowska, *Chem. Ind.*, 1961, **404**, 433; D.J. Cram and J.S. Bradshaw, *J. Am. Chem. Soc.*, 1963, **85**, 1108; J. Hanus and J. Vorisek, *Coll. Czech. Chem. Commun.*, 1929, **1**, 223.
954) J. Thiele, *Ann.*, 1892, **271**, 127; C.V. King, *J. Am. Chem. Soc.*, 1940, **62**, 379; E.E. van Tamelen and S.P. Pappas, *J. Am. Chem. Soc.*, 1963, **85**, 3297; E.E. van Tamelen, R.S. Dewey and R.J. Timmons, *J. Am. Chem. Soc.*, 1961, **83**, 3725.
955) F. Raschig, *Z. Angew. Chem.*, 1910, **23**, 972; R.S. Dewey and E.E. van Tamelen, *J. Am. Chem. Soc.*, 1961, **83**, 3729; R. Buyle, A. van Overstraeten and F. Eloy, *Chem. Ind.*, 1964, 839.
956) R. Appel and W. Büchner, *Angew. Chem.*, 1961, **73**, 807.
957) E.J. Corey and W.L. Mock, *J. Am. Chem. Soc.*, 1962, **84**, 1962.
958) F. Raschig, *Schwefel- und Stickstoffstudien*, Verlag Chemie, Leipzig-Berlin, 1924, p. 74; E. Schmitz and R. Ohne, *Angew. Chem.*, 1961, **73**, 807.
959) E.J. Corey, W.L. Mock and D.J. Pasto, *Tetrahedron Lett.*, 1961, **11**, 347.
960) S. Hünig and H.R. Müller, *Angew. Chem.*, 1962, **74**, 215.
961) E.E. van Tamelen, R.S. Dewey, M.F. Lease and W.H. Pirkle, *J. Am. Chem. Soc.*, 1961, **83**, 4302.
962) F. Aylward and M. Sawistowska, *Chem. Ind.*, 1962, 484.
963) H. Fischer and H. Gibian, *Ann.*, 1941, **548**, 183; 1942, **550**, 208.
964) S. Hünig, H. Müller and W. Thier, *Tetrahedron Lett.*, 1961, **11**, 353.
965) There is little difference in the rate of reaction with *cis*-oleic and *trans*-elaidic acids.[962]
966) E.J. Corey, D.J. Pasto and W.L. Mock, *J. Am. Chem. Soc.*, 1961, **83**, 2957.
967) D.K. Agrafiotis and H.S. Rzepa, *J. Chem. Soc., Chem. Commun.*, 1987, 902; *J. Chem. Soc., Perkin Trans. 2*, 1989, 475.
968) D.K. Agrafiotis and H.S. Rzepa, *J. Chem. Soc., Perkin Trans. 2*, 1989, 367.
969) J.M. Coxon and D.Q. McDonald, *Tetrahedron Lett.*, 1992. **33**, 3673.
970) L. Grierson, M.J. Perkins and H.S. Rzepa, *J. Chem. Soc., Chem. Commun.*, 1987, 1779; H.S. Rzepa, *J. Chem. Res.*, 1988, (S), 224.
971) M.J. Schlatter, *J. Am. Chem. Soc.*, 1941, **63**, 1733; K.B. Wiberg and W.J. Bartley, *J. Am. Chem. Soc.*, 1960, **82**, 6375; P. Dowd and A. Gold, *Tetrahedron Lett.*, 1969, 85.
972a) R. Breslow and P. Dowd, *J. Am. Chem. Soc.*, 1963, **85**, 2729.
972b) F. Fischer and D.E. Applequist, *J. Org. Chem.*, 1965, **30**, 2089.
972c) F.J. Weigert, R.L. Baird and J.R. Shapley, *J. Am. Chem. Soc.*, 1972, **92**, 6630.

972d) K. Komatsu, T. Niwa, H. Akari and K. Okamato, *J. Chem. Res.*, 1985, **(S)** 252, **(M)** 2847.

972e) H.M.R. Hoffmann, *Angew. Chem. Int. Ed. Engl.*, 1969, **8**, 556.

972f) M.S. Baird, H.H. Hussain and W. Clegg, *J. Chem. Res.*, 1988, **(S)** 110, **(M)** 1101.

973) M.S. Baird and H.H. Hussain, *J. Chem. Soc., Perkin Trans. 1*, 1986, 1845.

974) M.S. Baird, H.H. Hussain and W. Clegg, *J. Chem. Soc., Perkin Trans. 1*, 1987, 1609, 2109.

975) M. Yoshida, M. Ezaki, M. Hasimoto, M. Yamashita, N. Shigematsu, M. Okuhara, M. Kohsaka and K. Horikoshi, *J. Antibiotics*, 1990, **43**, 748.

976) M.S. Kuo, R.J. Zielinski, J.I. Cialdella, C.K. Marschke, M.J. Dupuis, G.P. Li, D.A. Klooserman, C.H. Spilmann and V.P. Marshall, *J. Am. Chem. Soc.*, 1995, **117**, 10629.

977) M. Yalpani, *Chem. Ind.*, 1996, 85.

978) For total synthesis, see:
A. Charette and H. Lebel, *J. Am. Chem. Soc.*, 1996, **118**, 10327; A.G.M. Barrett and K. Kasdirf, *J. Chem. Soc. Chem. Commun.*, 1996, 325; A.G.M. Barrett, D. Hamprecht, A.J. White and D.J. Williams, *J. Am. Chem. Soc.*, 1996, **118**, 7863; J.R. Falck, B. Meckonnen, J. Yu and J.-Y. Lai, *J. Am. Chem. Soc.*, 1996, **118**, 6096.

979) C.R. Theberge and C.K. Zercher, *Tetrahedron Lett.*, 1994, **35**, 9181; C.R. Theberge and C.K. Zercher, *Tetrahedron Lett.*, 1995, **36**, 5496; R.W. Armstrong and K.W. Maurer, *Tetrahedron Lett.*, 1995, **36**, 357; R.E.J. Cebula, M.R. Hanna, C.R. Theberge, C.A. Verbicky and C.K. Zercher, *Tetrahedron Lett.*, 1996, **37**, 8341.

980) R.E. Taylor, M.K. Ameriks and M.J. LaMarche, *Tetrahedron Lett.*, 1997, **38**, 2057.

981) A.B. Charette and R.P. De Freitas-Gil, *Tetrahedron Lett.*, 1997, **38**, 2809.

982) W. Oppolzer and V. Snieckus, *Angew. Chem. Int. Ed. Engl.*, 1978, **17**, 476.

983) A. Padwa and W.F. Riecker, *J. Am. Chem. Soc.*, 1981, **103**, 1859.

984) A. Padwa, W.F. Riecker and R.J. Rosenthal, *J. Am. Chem. Soc.*, 1985, **107**, 1710.

985) G.-A. Lee, C.-S. Shiau, C.-S. Chen and J. Chen, *J. Org. Chem.*, 1995, **60**, 3565.

986) A. Padwa, Kennedy, G.R. Newkome and F.R. Fronczek, *J. Am. Chem. Soc.*, 1983, **105**, 137.

987) W.E. Billups, W. Luo, G.-A. Lee, J. Chee, B.E. Arney Jr., K.B. Wiberg and D.R. Artis, *J. Org. Chem.*, 1996, **61**, 764; W.E. Billups, G.-A. Lee, B.E. Arney Jr. and K.H. Whitmire, *J. Am. Chem. Soc.*, 1991, **113**, 7980; K.B. Wiberg, D.R. Artis and G. Bonneville, *J. Am. Chem. Soc.*, 1991, **113**, 7969.

988) G.-A. Lee and Ch.-Y. Chang, *Tetrahedron Lett.*, 1998, **39**, 3013.

989) S.H. Nahm and H.N. Cheng, *J. Org. Chem.*, 1986, **51**, 5093.

990) R. Longcharich and K.N. Houk, *J. Am. Chem. Soc.*, 1987, **109**, 6947.

991) B.E. Thomas IV, K.N. Houk and P. Dowd, *Abstracts of the Papers of the American Chemical Society*, 1992, **203**, Part 2, 264.

992) Q. Deng, B.E. Thomas IV, K.N. Houk and P. Dowd, *J. Am. Chem. Soc.*, 1997, **119**, 6902.

993) A.R. Al-Dulayymi, J.R. Al-Dulayymi, M.S. Baird, M.E. Gerrard, G. Koza, S.D. Harkins and E. Roberts, *Tetrahedron*, 1996, **52**, 3409. The stereochemistry of the dimerisation product of 2-trimethylsilyl-1-(hydroxymethyl)cyclopropene was assigned on the basis of an *endo* transition state controlled by the size of the silyl substituents.

994a) F. Bernardi, M. Olivucci, J.J.W. McDougall and M.A. Robb, *J. Am. Chem. Soc.*, 1987, **109**, 544.

994b) F. Bernardi, A. Bottoni, M.A. Robb, H.B. Schlegel, and B. Tonachini, *J. Am. Chem. Soc.*, 1985, **107**, 2260.

994c) J. Doubleday Jr., M. Page and J.W. McIver Jr., *J. Mol. Struct. (THEOCHEM)*, 1988, **163**, 331.

994d) B.A. Hess, L.J. Schaad and D.N. Reinhoudt, *Int. J. Quantum Chem.*, 1986, **29**, 345.

994e) D.J. Pasto, *J. Am. Chem. Soc.*, 1979, **101**, 39.

994f) F. Bernardi, A. Bottoni, M. Olivucci, M.A. Robb, H.B. Schlegel and G. Tonachini, *J. Am. Chem. Soc.*, 1988, **110**, 5993.
994g) F. Bernardi, A. Bottoni, M.A. Robb, and A. Venturini, *J. Am. Chem. Soc.*, 1990, **112**, 2106.
994h) X. Wang and K.N. Houk, *J. Am. Chem. Soc.*, 1990, **112**, 1754.
995) M.J.S. Dewar and A.B. Pierini, *J. Am. Chem. Soc.*, 1984, **106**, 203.
996) G.L. Closs, L.A. Closs and W.E. Boll, *J. Am. Chem. Soc.*, 1973, **85**, 3796; M.L. Deem, *Synthesis*, 1972, 675.
997) P. Müller and M. Rey, *Helv. Chim. Acta*, 1981, **64**, 354; 1982, **65**, 1157; R.M. Albert and G.B. Butler, *J. Org. Chem.*, 1977, **42**, 674; N. Galloway, B.R. Dent and B. Hallon, *Aust. J. Chem.*, 1983, **36**, 593; D.N. Butler and R.I. Gupta, *Canad. J. Chem.*, 1978, **56**, 80; V.V. Plemenkov, V.A. Tikonov and V.V. Klochov, *Zh. Org. Khim.*, 1978, **14**, 1220; H. Sakuri, Y. Kobayashi and Y. Nakadaira, *J. Organomet. Chem.*, 1976, **120**, C1.
998) D.L. Boger and C.E. Brotherton, *Tetrahedron*, 1986, 42, 2777.
999) G. McMullen, G. Soedelmeier, R. Hildebrand, H. Fritz and H. Prinzbach, *Tetrahedron Lett.*, 1979, 3847; E.A. Shapiro, G.V. Lun'kova, A.O. Nefedov, I.E. Dolgii and O.M. Nefedov, *Izv. Akad. Nauk. SSSR, Ser. Khim.*, 1981, 2535.
1000) C. Dietrich-Buchecker and M. Franck-Neumann, *Tetrahedron*, 1977, 33, 751; C. Dietrich-Buchecker, D. Martina and M. Franck-Neumann, *J. Chem. Res. (S)*, 1978, 79; V. Plemenkov and I.G. Bolesov, *Dokl. Akad. Nauk. SSSR*, 1979, **248**, 887; I.G. Bolesov, L.G. Zaitseva, V.V. Plemenkov, I.B. Avezov and L.S. Surmina, *Zh. Org. Khim.*, 1978, **14**, 71.
1001) P.B. Sargeant, *J. Am. Chem. Soc.*, 1969, **91**, 3061.
1002) R. Breslow, G. Ryan and J.J. Groves, *J. Am. Chem. Soc.*, 1970, **92**, 998.
1003) T.H. Chan and D. Massuda, *Tetrahedron Lett.*, 1975, 3383; K.B. Wiberg and G. Bonneville, *Tetrahedron Lett.*, 1982, **23**, 5385.
1004) K. Takeuchi, Y. Yokmichi, T. Kurosaki, Y. Kimura and K. Okamoto, *Tetrahedron*, 1979, **35**, 949; D.N. Reinhoudt, P. Smael, W.J.M. van Tilbourg and J.P. Visser, *Tetrahedron Lett.*, 1979, 3755; I.G. Bolesov, L.G. Zaitseva, V.V. Plemenkov and L.S. Surmina, *Zh. Org. Khim.*, 1978, **14**, 285.
1005) B. Halton, M.A. Battiste, R. Rehberg, C.L. Deyrup and M.E. Brennan, *J. Am. Chem. Soc.*, 1967, **89**, 5964.
1006) H.-D. Martin, L. Kandy and D. Stusche, *Tetrahedron Lett.*, 1977, 3561.
1007) D.C.F. Law and S.W. Tobey, *J. Am. Chem. Soc.*, 1968, **90**, 2376.
1008) F.X. Huber, J. Sauer, W.S. McDonald and H. Noeth, *Chem. Ber.*, 1982, **115**, 444; U. Göckel, Y. Hartmannsgruber, A. Steigel and J. Sauer, *Tetrahedron Lett.*, 1980, **21**, 599; H.D. Fühlhuber, G. Gousetis, T. Troll and J. Sauer, *Tetrahedron Lett.*, 1978, 3903; H.D. Fühlhuber, C. Gousetis, J. Sauer and H.J. Linder, *Tetrahedron Lett.*, 1979, 1299.
1009) B.S. Jursic, *J. Heterocyclic Chem.*, 1995, **32**, 1445.
1010) B.S. Jursic, *Tetrahedron Lett.*, 1997, **38**, 1305; Y. Apeloig and E. Matzner, *J. Am. Chem. Soc.*, 1995, **117**, 5375.
1011) O.H. Wheeler, *J. Am. Chem. Soc.*, 1956, **78**, 3216.
1012) M. Beroza, *J. Am. Chem. Soc.*, 1952, **74**, 1585.
1013) A. Hassner and N.H. Cromwell, *J. Am. Chem. Soc.*, 1958, **80**, 893.
1014) J.R. Willard and C.S. Hamilton, *J. Am. Chem. Soc.*, 1951, **73**, 4805.
1015) S. Aronoff and J.B. Harborne, *J. Org. Chem.*, 1939, **4**, 389.
1016) J. Fried and R.C. Elderfield, *J. Org. Chem.*, 1941, **6**, 566.
1017) C.N.R. Rao, *Ultraviolet and Visbile Spectroscopy - Chemical Applications*, Butterworth and Co. (Publishers) Ltd., London, 1975, 3rd edition, and references cited therein.