CREATING COLOUR
The Chemistry of Dyes

Alex Thow
Section 1 Introduction

To truly understand the chemistry of dyes we first must consider their colour. Colour has fascinated humans for thousands of years, from Aristotle believing it was sent from heaven, to Newton separating light into its colours with a prism. Our curiosity with colour comes from a young age, asking our parents why the grass is green and the sky blue.

To start this exploration of colour we will look at some basic quantum mechanics to study the behaviour of light and particles. These ideas can be applied to give an accurate picture of how electrons behave in atoms and molecules. From this, a description of bonding and delocalisation in organic molecules can be built up and used to explain how certain molecules can absorb visible light, giving them colour.

The focus of the text will then shift to a more general look at dyes. A glance over the importance of dyes in human history and their development will be covered before moving into the organic composition of dyes. We will then look at the structure of some materials and how they can be dyed and bleached. To finish off our discussions we will consider some specific examples that display how an understanding of the chemistry of dyes can explain different physical and biological phenomena.

Section 2 Quantum Mechanics

2.1 A Brief History

2.1.1 Blackbody Radiation

In 1900, Max Planck began a scientific revolution. At the time there was a problem, known as the ultraviolet catastrophe, with something called blackbody radiation. This is the radiation of electromagnetic waves from an object that absorbs and emits all frequencies of light. A law can be derived from classical physics, Rayleigh-Jeans’ law, that links the intensity, \( B \), of radiation to the frequency, \( \nu \).

\[
B(\nu) = \frac{2\nu^2 k_b T}{c^2}
\]

In this equation, \( k_b \) is the Boltzmann constant, \( c \) the speed of light in a vacuum and \( T \) the temperature of the black body. This law worked for low frequencies but as the frequency increased, it diverged from the observed spectrum. What Planck did was to assume that the emissions of radiation did not take a continuous range of energies but had discrete energies proportional to integer multiples of their frequency.

\[
E \propto n\nu, \quad n = 1, 2, 3 \ldots
\]

He called the constant of proportionality Planck’s constant, \( h \), and used his assumption to create a new law for blackbody radiation, Planck’s law.

\[
B(\nu) = \frac{2h\nu^3}{c^2 \left(e^{\frac{h\nu}{k_b T}} - 1 \right)}
\]
Using this equation and experimental data, he calculated the value of $h$ to be $6.626 \times 10^{-34} \text{ J s}$. Planck’s law fit the observed spectrum perfectly, but he had no idea why this was the case.

\[ E = h \nu \]

The constant of proportionality is Planck’s constant, $h$. Einstein reasoned that a photon could transfer its energy, $h \nu$, to an electron. If the photon had enough energy, it could kick the electron out of the metal. Any extra energy that the photon had would become the kinetic energy of the electron, $E_k$.

\[ E_k = h \nu - \phi \]

The letter $\phi$ denotes the work function and is the minimum energy required to kick the electron out of the metal. The assumption that light was made of particles was a radical one, so was not widely accepted until 1916, when experimental results confirmed Einstein’s theory.

2.1.2 The Photoelectric Effect

Within the year 1905, Albert Einstein published four groundbreaking papers and so this year is called Einstein’s ‘Annus Mirabilis’. The first of these, for which he was awarded the Nobel prize in 1921, explained the photoelectric effect. This is a phenomenon in which light above a certain frequency, when shone onto the surface of a metal, can kick out electrons. The higher the frequency, the faster these electrons travel. No classical explanation had been found since the effect’s discovery in 1887.
2.2 The Schrödinger Equation

2.2.1 An Introduction

Over the years, quantum theory developed as more rigorous mathematical descriptions were formulated. Perhaps the most well-known is the Schrödinger equation, which describes something called the wavefunction of a system, denoted by the Greek letter, ψ. The first postulate of quantum mechanics is that the state of a quantum system is completely specified by the wavefunction. A classical parallel would be a set containing the positions and momenta of all the particles in a system. Schrödinger’s equation does have time-dependence, but for the systems we will be covering, a time-independent equation can be used.

\[ \hat{H} \psi = E \psi \]

2.2.2 Operators, Eigenfunctions and Eigenvalues

The first term of the time-independent Schrödinger equation is \( \hat{H} \). The hat symbolises that this is an operator, which is a mathematical object that turns a function into another function. They can take a wide variety of forms. For example, the operator \( \hat{A} = x \) simply multiplies the original function by \( x \), whereas the operator \( \hat{B} = \frac{d}{dx} \) differentiates it with respect to \( x \).

\[
\hat{A} \sin(kx) = x \times \sin(kx) = x \sin(kx) \\
\hat{B} \sin(kx) = \frac{d}{dx} \sin(kx) = k \cos(kx)
\]

For a given operator, \( \hat{O} \), if a function, \( f(x) \), gives back the same function multiplied by a constant, \( \lambda \), that function is called an eigenfunction of \( \hat{O} \) and \( \lambda \) is the corresponding eigenvalue.

\[ \hat{O} f(x) = \lambda f(x) \]

2.2.3 Operators in Quantum Mechanics

The second postulate of quantum mechanics is that for every classical observable, there is a linear, Hermitian operator in quantum mechanics. We will not discuss here what it means for the operator to linear or Hermitian, but all the operators we will see obey both criteria. An example of the second postulate is that, corresponding to the classical observable of linear momentum along the \( x \)-axis, there is a quantum mechanical operator, \( \hat{p}_x \).

\[
\hat{p}_x = -i \hbar \frac{\partial}{\partial x}
\]

Note that \( i = \sqrt{-1}, \ h = \frac{\hbar}{2\pi} \) (the reduced Planck’s constant) and \( \frac{\partial}{\partial x} \) is the partial derivative with respect to \( x \), meaning that any other variable is treated as a constant. A discussion of the origins of this operator is too in-depth for this text, but it can be taken as an axiom.

The operator in the Schrödinger equation, \( \hat{H} \), is the Hamiltonian operator. It is the operator corresponding classically to the total energy of a system. From classical mechanics, the total energy is the sum of the kinetic energy and potential energy.

\[ \hat{H} = \hat{T} + \hat{V} \]
Here, $\hat{T}$ is the kinetic energy operator and $\hat{V}$ is the potential energy operator. The kinetic energy operator can be written in terms of the momentum operator if the classical formulae for kinetic energy and momentum are combined. The known momentum operator can then be substituted into the resulting expression.

$$T = \frac{1}{2}mv^2, \quad p = mv$$

$$\Rightarrow T = \frac{p^2}{2m}$$

$$\therefore \hat{T} = \frac{\hat{p}^2}{2m} = \left( -i\hbar \frac{\partial}{\partial x} \right)^2 = \frac{i^2\hbar^2}{2m} \left( \frac{\partial}{\partial x} \right)^2 = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$$

Note that $\frac{\partial^2}{\partial x^2}$ is the second partial derivative. We cannot derive a general expression for the potential energy operator in the same way because it depends on the system. Therefore, it is written as a function of the position, $x$.

$$\hat{V} = V(x)$$

These expressions for the potential and the kinetic energy operators can be added to find $\hat{H}$.

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x)$$

2.2.4 Eigenfunctions and Eigenvalues in Quantum Mechanics

The third postulate of quantum mechanics is that if a classical observable is measured, the only values that the observable can take are the eigenvalues of the corresponding operator. This postulate finally makes sense of the Schrödinger equation.

$$\hat{H}\psi = E\psi$$

The equation represents the Hamiltonian operator, $\hat{H}$, acting on the wavefunction, $\psi$, and giving back the same $\psi$ multiplied by a constant, $E$. Therefore, to solve the Schrödinger equation, the eigenfunctions and eigenvalues of $\hat{H}$ must be found. The eigenvalues, from the third postulate, give the measurable values of the classical observable corresponding to $\hat{H}$, and so are effectively the allowed total energies of the system, $E$. The eigenfunctions, $\psi$, are the wavefunctions corresponding to those values of energy.

2.2.5 The Wavefunction

There is no precise definition of the wavefunction, but there are certain useful interpretations, the most common being the Born interpretation. This states that the square of the wavefunction gives a probability density. For a wavefunction associated with a particle, this can be used to calculate the probability that a particle is found in a particular place. If, in a region of space, the probability density is $p(x)$, then the probability of finding the particle there is given by $p(x)dx$, where $dx$ is the ‘size’ of that region. In three dimensions this would instead be $p(x,y,z)dV$.

$$\text{Probability of finding particle between } x \text{ and } x + dx = p(x)dx = |\psi(x)|^2 dx$$
The total probability of finding the particle anywhere in space is 1, so the sum of the probabilities at every point in space must equal 1. This is represented mathematically as an integral.

\[ \int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1 \]

A wavefunction that obeys this is said to be normalised, and any wavefunction can be normalised by multiplying it by some normalisation constant, \( N \).

\[ \int_{-\infty}^{\infty} N^2 |\psi(x)|^2 dx = 1 \]

The Born interpretation imposes some restrictions on the wavefunction:

- The wavefunction must be single-valued because it does not make sense for a particle to have two different probabilities of being somewhere.

\[ \text{Figure 3 - A multi-valued function. This could not be a wavefunction.} \]

- The wavefunction must be continuous because a break in the wavefunction would lead to a particle effectively having an undefined probability of being somewhere.

\[ \text{Figure 4 - A discontinuous function. This could not be a wavefunction.} \]

- The wavefunction must be able to be normalised, so must have a finite integral over all space.

\[ \text{Figure 5 - A function with an infinite integral. This could not be a wavefunction.} \]
2.2.6 Solving the Schrödinger Equation

The Schrödinger equation can be solved for several systems, including the harmonic oscillator and rigid rotor. The system we are going to solve for is called ‘particle in a box’ which consists of a particle travelling in one dimension between $x = 0$ and $x = L$. In this region, the potential energy is zero and outside this region, the potential energy is infinite. This means that the particle cannot leave this region so perhaps a more suitable name would be ‘particle in an endless hell from which it can never escape’.

As the particle always experiences zero potential energy, the Hamiltonian operator is equal to the kinetic energy operator.

$$\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$$

Using this, the Schrödinger equation for the system can be written.

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x) = E\psi(x)$$

We are looking for a function whose second derivative is itself multiplied by a constant. Luckily, this has been solved many times before, so we can take a very well-informed guess at the solution.

$$\psi(x) = A \cos(kx) + B \sin(kx)$$

The constants $A$, $B$, and $k$ are to be determined. We can now differentiate this function twice.

$$\frac{\partial}{\partial x} (A \cos(kx) + B \sin(kx)) = -kA \sin(kx) + kB \cos(kx)$$

$$\frac{\partial^2}{\partial x^2} (A \cos(kx) + B \sin(kx)) = \frac{\partial}{\partial x} (-kA \sin(kx) + kB \cos(kx)) = -k^2 A \cos(kx) - k^2 B \sin(kx)$$

$$\therefore \frac{\partial^2}{\partial x^2} \psi(x) = -k^2 \psi(x)$$

This can now be substituted into the Schrödinger equation.

$$-\frac{\hbar^2}{2m} (-k^2 \psi(x)) = E\psi(x) \Rightarrow \frac{\hbar^2 k^2}{2m} \psi(x) = E\psi(x) \Rightarrow E = \frac{\hbar^2 k^2}{2m}$$

We have derived an expression for the allowed energy levels, which depends on the constant, $k$. Applying boundary conditions, outside the interval $[0, L]$ there is no probability of finding the particle, so the wavefunction must be zero at $x = 0$ and $x = L$.

$$\psi(0) = 0 \Rightarrow A \cos(0) + B \sin(0) = 0 \Rightarrow A = 0$$

$$\psi(L) = 0 \Rightarrow B \sin(kL) = 0$$

The constant $B$ cannot be zero or the wavefunction would be zero everywhere so $\sin(kL)$ must equal zero. For $\sin x$ to equal zero, $x$ must be an integer multiple of $\pi$.

$$kL = n\pi \Rightarrow k = \frac{n\pi}{L}, \quad n = 0, \pm 1, \pm 2, \pm 3, \pm 4 \ldots$$
This expression for $k$ can be inserted into the expressions for the wavefunction and the allowed energy levels.

$$\psi_n(x) = B \sin \left(\frac{n\pi x}{L}\right)$$

$$E_n = \frac{\hbar^2 n^2 \pi^2}{2ml^2}$$

If $n = 0$, then the wavefunction is simply is zero everywhere so this value is disallowed. If $n$ is negative, the wavefunctions are the positive wavefunctions multiplied by $-1$. The square of the wavefunction is, therefore, the same and so these values can also be ignored.

Our final solution to the Schrödinger equation can now be written.

$$\psi_n(x) = B \sin \left(\frac{n\pi x}{L}\right), \quad E_n = \frac{\hbar^2 n^2 \pi^2}{2ml^2}, \quad n = 1, 2, 3 \ldots$$

There is still an undetermined constant, $B$, which can be worked out by normalising the wavefunction. This process will not be covered here as the important result is the allowed energy levels, but the result is that $B = \sqrt{\frac{2}{L}}$.

Section 3 Atomic and Molecular Orbitals

3.1 A History of Atomic Structure

3.1.1 The Discovery of The Electron

In the late 19th century, there was great interest in cathode rays, which were created when a large potential difference was applied across a vacuum tube. In 1897, J. J. Thompson directed a stream of these rays between two charged plates and they deflected towards the positively charged plate. He concluded that the rays were made of negatively charged particles, electrons. From the magnitude of the deflection of the rays in electric and in magnetic fields, Thompson calculated the mass of the electron as only $\frac{1}{1836}$ the mass of a proton.

Thompson then suggested a model for the atom called the plum pudding model. This was based on the knowledge that atoms contained electrons but must also contain positive charge for the overall atom to be neutral. It consists of electrons distributed within a positively charged medium.
3.1.2 The Discovery of the Nucleus

In 1909, Ernest Rutherford and two fellow scientists Geiger and Marsden conducted an experiment in which a stream of alpha particles was fired at a very thin gold foil, and the angle of deflection of each alpha particle was measured. If the plum pudding model had been true, the alpha particles would have been deflected by a very small angle. However, while the majority passed through almost undeflected, a small fraction of the alpha particles was deflected by a large angle.

The conclusion was that Thompson’s model was incorrect. In Rutherford’s new model, most of the atom was empty space apart from a tiny sphere of positive charge in the centre of the atom, the nucleus, surrounded by orbiting electrons.

3.1.3 The Bohr Model

The flaw with Rutherford’s model was that the orbiting electrons, according to classical physics, would constantly emit electromagnetic radiation, and therefore lose energy and spiral into the nucleus in 16 picoseconds. This would make atoms inherently unstable.

In 1913, Niels Bohr proposed a new model to overcome this issue. He put forward three postulates:

1. Electrons revolve around the nucleus in stable orbits without radiating energy. These orbits are at certain discrete distances from the nucleus.

2. These discrete distances are such that the angular momentum, \( l \), of an electron is an integer multiple of the reduced Planck’s constant.

\[
l = n\hbar, \quad n = 1, 2, 3 \ldots
\]

The smallest possible orbital radius, occurring when \( n = 1 \), is called the Bohr radius, \( a_0 \).

3. An electron can only change its energy by jumping between orbits, either emitting or absorbing electromagnetic radiation of a frequency given by the Planck relation.

\[
\Delta E = h\nu
\]
This model of the atom quantises the energy levels of the electrons and can accurately predict the frequency of the lines in the emission spectrum of hydrogen. But just like the other models, it eventually was replaced by the modern theory of atomic orbitals.

3.2 Atomic Orbitals

3.2.1 An Introduction to Orbitals

In Bohr’s model, we know the electron’s exact distance from the nucleus and its momentum. This violates Werner Heisenberg’s uncertainty principle. We will not derive this, but it means we can never know both the position and momentum of any particle with complete accuracy. The product of their uncertainties must be greater than or equal to half of the reduced Planck’s constant.

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}$$

To solve this issue, the electrons must be described by wavefunctions. This means only knowing the probability of an electron being somewhere, not exactly where it is, no longer violating the uncertainty principle. These wavefunctions are atomic orbitals.

3.2.2 A Mathematical Description of Orbitals

The Schrödinger equation can be used to find mathematical expressions for the atomic orbitals. Unfortunately, it can only be solved exactly for single-electron systems as with multiple electrons there are too many interactions to deal with. However, these solutions do provide a good description of atomic orbitals in general.

For atomic orbitals, it is much easier to use polar coordinates than Cartesian coordinates. Cartesian coordinates use the variables $x$, $y$ and $z$. Polar coordinates use the distance from the origin $r$, and two angles, $\theta$ and $\phi$.

![Diagram showing an electron losing energy via emission of a photon.](image)

![Diagram showing polar coordinates.](image)
When the Schrödinger equation is solved for a hydrogen atom, each wavefunction is characterised by three quantum numbers, similar to how the particle in a box wavefunction was characterised by $n$:

1. The principal quantum number, $n$, takes values 1, 2, 3 ...
2. The angular momentum quantum number, $l$, takes values 0, 1, 2 ..., $n - 1$
3. The magnetic quantum number, $m_l$, takes values $-l$, ..., $0$, $+l$.

The principal quantum number determines the shell the electron is in and can completely determine the energy of a hydrogen orbital.

$$E_n = -\frac{Z^2 R_H}{n^2}$$

The nuclear charge, $Z$, is simply 1 for hydrogen and the Rydberg constant, $R_H$, has a value of $2.180 \times 10^{-18}$ J. The energy is negative because a free electron is said to have zero energy and when bound by a nucleus it has less energy.

The principal and angular momentum quantum numbers determine the subshell an electron is in. For each value of $l$ there is a corresponding letter, $l = 0$ is $s$, $l = 1$ is $p$, $l = 2$ is $d$ and $l = 3$ is $f$. The possible values of $l$ mean that in the first shell there is only an $s$ subshell, in the second only $s$ and $p$ subshells and so on.

The principal, angular momentum and magnetic quantum numbers determine the orbital the electron is in. For the $s$ subshells, $m_l$ can only be 0 so there is one $s$ orbital per subshell. For the $p$ subshells, $m_l$ can be $-1$, 0, or 1, so there are three $p$ orbitals per subshell. In the same way, there are five $d$ and seven $f$ orbitals per subshell. The orbitals are distinguished by a subscript. For example, $2p_x$, $2p_y$ and $2p_z$ are the three $p$ orbitals in the second shell.

The wavefunctions of hydrogen can be written as the product of a radial part, $R$, and an angular part, $Y$.

$$\psi_{n,l,m_l}(r, \theta, \phi) = R_{n,l}(r) \times Y_{l,m_l}(\theta, \phi)$$

The probability of an electron being found at a certain distance, $r$, from the nucleus can be calculated by multiplying the probability density of a point at that distance from the nucleus, $p(r)$, by the surface area of a sphere of that radius.

$$\text{Surface area of sphere} = 4\pi r^2$$

$$\text{Probability density of electron at radius } r = 4\pi r^2 p(r)$$

From the Bohr interpretation, the probability density, $p(r)$, is the square of the wavefunction. This defines the radial distribution function, $P(r)$, which is effectively a measure of the electron densities at different radii.

$$P(r) = 4\pi r^2 |\psi(r)|^2$$

3.2.3 A Closer Look at the $s$ Orbitals

The simplest orbital is the $1s$ orbital, whose wavefunction we will consider now.

$$\psi_{1s} = N_{1s} e^{-\frac{r}{\alpha_0}}$$
\( N_{1s} \) is a normalisation constant whose value is not important and \( a_0 \) is the Bohr radius. The value of the wavefunction only depends on the distance from the nucleus, so the orbital has spherical symmetry. The wavefunction has a maximum in the centre of the nucleus and decays exponentially outwards. The radial distribution function has a maximum at \( r = a_0 \), so the electron density is highest at this radius.

![Figure 10 - Graph of the 1s wavefunction, blue, and radial distribution function, orange.](image)

The issue with depicting orbitals on paper is that there are not enough dimensions. There are several ways of overcoming this, but each has its flaws. The one we are going to use is created by taking a cross-section through the orbital and plotting large numbers of dots based on the value of the wavefunction. The denser the dots, the higher the value of the wavefunction.

![Figure 11 - Dot representation of the 1s orbital.](image)

The wavefunction of the \( 2s \) orbital is slightly more complex than that of the \( 1s \).

\[
\psi_{2s} = N_{2s} \left( 2 - \frac{r}{a_0} \right) \left( e^{-\frac{r}{\sqrt{2}a_0}} \right)
\]

Just like the \( 1s \), this orbital is spherically symmetric. To compare the \( 2s \) orbital to the \( 1s \) their radial distribution functions can be plotted on the same graph.

![Figure 12 - Graphs of the radial distribution functions of the 1s, blue, and 2s, orange.](image)
The main difference is that the 1s has higher electron density nearer the nucleus, and the 2s is more spread out and effectively larger. At 2a₀ there is no electron density in the 2s orbital, which is called a node. The dot representations can also be placed next to each other for comparison. Note that the blue colour represents a negative value of the wavefunction.

\[ \psi_{3s} = N_{3s} \left( 27 - 18 \left( \frac{r}{a_0} \right) + 2 \left( \frac{r}{a_0} \right)^2 \right) \left( e^{-\frac{r}{3a_0}} \right) \]

We can plot the graph of the radial distribution function compared with the 1s and 2s and show the dot representation of the 3s.

From these, we can see that the 3s is larger still than the 2s orbital and that the 3s flips sign twice at two nodes. It should now be easy to predict the shape of the s orbitals for higher values of \( n \).

3.2.4 A Closer Look at the \( p \) Orbitals

The \( p \) orbitals have more interesting shapes than the \( s \) orbitals as they no longer have spherical symmetry. We will look first at the radial part of the 2p orbitals.

\[ R_{2p} = N_{2p} \left( \frac{r}{a_0} \right) \left( e^{-\frac{r}{2a_0}} \right) \]
To understand this, we can plot the radial distribution function of the 2\textit{p} orbitals compared with the 1\textit{s} and 2\textit{s}.

It is a similar size to the 2\textit{s} orbital, which should not be surprising as it has the same energy. It has no radial nodes, but it does have another type of node not seen in the radial part. The angular part depends on the particular \textit{p} orbital.

\[ Y_{pz} = \cos(\theta), \quad Y_{px} = \sin(\theta) \cos(\phi), \quad Y_{py} = \sin(\theta) \sin(\phi) \]

Explaining what these mean mathematically can get confusing, so we will simply look at the result using the dot representation.

Each orbital has a nodal plane, shown in the diagram by a dotted line. Each side of this dotted line the wavefunction has the opposite sign. It turns out that each orbital has cylindrical symmetry, meaning it is symmetrical by rotation about a certain axis. The 2\textit{p}_z lies along the \textit{z}-axis and so on. Going into detail discussing any more orbitals is unnecessary for our future discussions so they will be left here.

3.2.5 Multi-Electron Atoms

For more complex atoms, we are adding electrons and increasing the nuclear charge. It is impossible to solve the Schrödinger equation analytically for these systems, so approximations and numerical methods are used to find the wavefunctions.

The way electrons occupy the orbitals is governed by their energy and the Pauli exclusion principle. An electron in hydrogen will usually occupy the lowest energy orbital, the 1\textit{s}, but in multi-electron atoms, the electrons start to occupy different orbitals. This is because the Pauli exclusion principle prevents two electrons from having identical quantum numbers. There is a quantum number that we have not
Electrons can either have a spin of $+\frac{1}{2}$ (spin-up) or $-\frac{1}{2}$ (spin-down). The result is that at most two electrons can occupy any orbital, as long as they have opposite spin. Therefore, electrons fill up the orbitals two at a time from lowest energy to highest energy.

The energies of the orbitals change in multi-electron atoms such that those with the same $\pi$ no longer need have the same energy. We will not go into a detailed discussion as to why this is, but we will show a graph of the energies of some of the orbitals as you increase the atomic number.

The unit of electron volts, $eV$, makes the numbers easier to deal with as $1\ eV = 1.60 \times 10^{-19}\ J$. When the 1s orbital is full, the electrons then go into the 2s as this is actually lower in energy than the 2p. Then they go into the 2p, then the 3s, and so on. The results can be summarised in an electron configuration. For fluorine, it would be $1s^22s^22p^5$. This shows that there are two electrons in both the 1s and 2s subshells, and five in the 2p subshell.

### 3.3 Diatomic Molecular Orbitals

#### 3.3.1 The Simple Case – $H_2^+$

Due to the presence of only one electron in $H_2^+$, it is possible to solve the Schrödinger equation and the result is a molecular orbital (MO). MOs are filled up by electrons in order of increasing energy, just like atomic orbitals (AOs). For molecules with multiple electrons, some rules can be used to guess the form of the MOs and their relative energies, all of which can be derived from quantum mechanics.

One simple approach is called the linear combination of atomic orbitals, LCAO. This takes the algebraic sum of the AOs at every point in space to construct the MO.

$$MO = c_1(AO1) + c_2(AO2)$$

The constants, $c_1$ and $c_2$, are orbital coefficients which can be determined through detailed quantum mechanical calculations. In the $H_2^+$ molecule, it would make sense for the lowest energy MO to be constructed from a combination of the two 1s orbitals. It turns out that there are two MOs formed from the combination of these AOs. For the first, $c_1 = c_2 = 1$ and for the second, $c_1 = 1$, $c_2 = -1$.

$$MO_1 = N_1(1s\ of\ first\ atom + 1s\ of\ second\ atom)$$

$$MO_2 = N_2(1s\ of\ first\ atom - 1s\ of\ second\ atom)$$
Again, \( N_1 \) and \( N_2 \) are normalisation constants. When the atoms are far apart, the MOs will effectively be identical in shape to the AOs. At shorter distances, MO1 will contain a region between the nuclei where the AOs are adding together and producing high electron density, called constructive overlap. For MO2, the AOs are subtracting and producing low electron density, called destructive overlap. Halfway between the nuclei, there will be a node of zero electron density.

\[
\text{Node}
\]

In \( \text{H}_2^+ \), the lowest energy MO turns out to be MO1, so the electron occupies this orbital. This creates high electron density between the nuclei and pulls them towards each other until the attraction balances the repulsion of the nuclei. This is a chemical bond and so MO1 is described as a bonding orbital. If the electron were for some reason to be MO2, then the low electron density between the nuclei means that the repulsive force is strong, and the electron density either side of the nuclei pulls them apart. For this reason, MO2 is described as an antibonding orbital.

Molecular orbitals can be classified using symmetry. In \( \text{H}_2^+ \), the orbitals are symmetric by rotation about the internuclear axis, and so are called \( \sigma \) orbitals. The only other type of orbital we will meet is one where rotation by 180° about the internuclear axis flips the sign of the wavefunction. This is called a \( \pi \) orbital. Bonds formed by these orbitals are called \( \sigma \) and \( \pi \) bonds, respectively. Antibonding orbitals are given an asterisk, for example \( \sigma^* \).

The energies of MOs can be represented on a MO diagram. The horizontal lines represent orbitals, with those at the sides being the original AOs, and those in the centre being the MOs. At large separation, the lines are the same level as the MOs are equivalent to the AOs. As the atoms get closer, the bonding MO lowers in energy and the antibonding MO increases in energy.

3.3.2 More Complex Cases – \( \text{H}_2 \), \( \text{He}_2^+ \), and \( \text{He}_2 \)

The MOs can be filled in a similar way to how AOs are filled. Two electrons are placed in each orbital with opposite spin, starting with the lowest energy orbital. For the simple case of two 1s orbitals forming a \( \sigma \) and a \( \sigma^* \) orbital, it can be seen that the first two electrons fill the \( \sigma \) and the next goes into the \( \sigma^* \). This can be represented using MO diagrams.
If we were to add another electron it would go into the $\sigma^*$ orbital. It turns out that the $\sigma^*$ orbital is slightly more raised in energy than the $\sigma$ is lowered. This would mean that going from two He atoms to an He$_2$ molecule would lead to a net increase in energy, showing why helium does not form diatomic molecules.

3.3.3 Rules of Forming Molecular Orbitals

We will not cover how these rules are derived but they are very useful for constructing MOs from AOs.

1. If $n$ AOs are combined, $n$ MOs are formed.
2. AOs can only combine if they have the correct symmetry.
3. AOs that are closer in energy have a larger interaction when forming MOs.
4. AOs which are closer in energy to an MO contribute more to it.
5. AOs can only interact strongly if their sizes are compatible.

The first of these is quite simple to understand. If we pull apart a molecule, the MOs will steadily become AOs. Since this must be a continuous process, the number of orbitals cannot change.

To understand the second, we will cover a disallowed interaction. If we try to combine a $2s$ and a $2p_x$ orbital, taking the internuclear axis to be the $z$-axis, then one lobe of the $2p_x$ orbital will constructively overlap with the $2s$ and the other will destructively overlap with it. The overlaps will be equal and opposite and so the total overlap will be zero, so these AOs do not form an MO. For a $2s$ and $2p_z$, this is not a problem and so an MO can be formed.

The third rule can be easily explained with MO diagrams.

It can be seen that as the difference in energy between the AOs increases, the MOs become closer in energy to the original AOs, so the interaction decreases. The fourth rule is self-explanatory and simply means that the higher energy AO contributes more to the antibonding MO and vice versa.
The fifth rule is also a result of the overlap as, if the two orbitals differ greatly in size, then the overlap will not be effective, and the AOs will not interact strongly.

3.3.4 Combining Different Atomic Orbitals

When 2s orbitals are combined the result is simple and we end up with another $\sigma$ and $\sigma^*$ of higher energy than with a combination of 1s orbitals. When 2$p_z$ orbitals are combined, as these lie on the internuclear axis, we get another $\sigma$ and $\sigma^*$, which are higher in energy still. They are, however, slightly different from the $\sigma$ MOs created from $s$ orbitals.

Figure 24 - Diagram showing the bonding (top) and antibonding (bottom) MOs created from two 2$p_z$ orbitals.

When 2$p_x$ and 2$p_y$ orbitals are combined, the resulting MOs are $\pi$ and $\pi^*$. These do not have cylindrical symmetry and so look more complex than the $\sigma$ orbitals.

Figure 25 - Diagram showing the bonding (top) and antibonding (bottom) MOs created from two 2$p_x$ or 2$p_y$ orbitals.

With this knowledge, we can draw an MO diagram displaying the bonding between many different atoms. As an example, we will draw one for the nitrogen molecule. For simplicity, we will assume that only interactions between identical orbitals can occur, although in reality the $s$ and $p$ orbitals can interact.

Figure 26 - MO diagram for the $N_2$ molecule.

The 1$s$ interactions have been ignored as they are much lower in energy. The most important interactions are those of the $p$ orbitals, and the order of energy of their MOs is $\sigma, \pi, \pi^*, \sigma^*$. The electrons shown fill four bonding MOs and one antibonding, giving $N_2$ an overall triple bond.
Section 4 Bonding in Organic Molecules and Colour

4.1 Hybridisation

4.1.1 An Introduction to Hybrid Orbitals

When describing bonding in organic molecules, using the molecular orbitals described in the last section would be almost impossible. For more than a few atoms there are too many interactions to deal with, and the molecular orbitals become very difficult to interpret. Hybridisation is a different approach that helps solve this issue.

Hybridisation is the mixing of AOs to creating hybrid atomic orbitals (HAOs). These HAOs often combine to form much more simple MOs. For example, with methane, the issue with using the full MO approach is that the AOs do not point directly towards the four hydrogen atoms. The $2s$ points equally in all directions and the $2p$ orbitals point at right-angles, not at the tetrahedral angle of $109.5^\circ$. The hybrid approach allows for four simple HAOs that point directly to each hydrogen, creating simple MOs.

![Diagram showing the HAOs of methane.](image)

4.1.2 Different HAOs – $sp^3$, $sp^2$, and $sp$

In organic molecules, we are usually considering bonding to carbon atoms. If the carbon atom is bonded to four other atoms, four different HAOs need to be formed just like in methane. The way this is done is by directly combining the single $2s$ orbital with the three $2p$ orbitals. The new orbitals are $sp^3$ hybridised. Each of these are identical and are separated by $109.5^\circ$.

The energy of this hybrid orbital is between the energy of the $2s$ and $2p$ orbitals and turns out to be similar in energy to the $1s$ orbital of a hydrogen atom. We can draw an MO diagram for the combination of an $sp^3$ HAO with a hydrogen $1s$ AO. There is one electron contributed from each orbital and so both go into the bonding MO, forming a C-H bond.

![MO diagrams showing the creation of $sp^3$ HAOs (left) and the creation of a C-H bond (right).](image)
When a carbon is double bonded to an atom, we must use a different hybridisation. The required HAOs should be at 120° to each other and in the same plane. This can be done by combining the 2s orbital with only two of the 2p orbitals. This creates three $sp^2$ hybridised orbitals and a leftover 2p orbital. In the case of ethene, $C_2H_4$, each carbon has two $\sigma$ bonds to hydrogens created by a combination of the hydrogen 1s and an $sp^2$ HAO from the carbon. There is also a $\sigma$ bond between the two carbons created by the head-on overlap of two $sp^2$ HAOs. Finally, the double bond is a $\pi$ bond formed by the overlap of the leftover 2p orbitals.

![Figure 29 - Diagram showing the $\sigma$ and $\pi$ interactions in ethene.](image)

Finally, when a carbon has a triple bond, it must be $sp$ hybridised. This leaves two 2p orbitals that can generate two $\pi$ bonds in addition to the $\sigma$ bond, creating a triple bond. This can be used to describe the bonding in ethyne, $C_2H_2$.

![Figure 30 - Diagrams to show the interactions in ethyne. One helps show the $\sigma$ bond (left) and one the $\pi$ bonds (right).](image)

### 4.2 Delocalisation and Conjugation

#### 4.2.1 Delocalised Bonding

So far, we have covered localised bonding, meaning that the bonding occurs between two atoms. This gives an accurate picture in most circumstances, but there are notable exceptions where this fails to explain certain properties of a molecule, giving rise to delocalised bonding.

One well-known example is benzene, a molecule with the molecular formula $C_6H_6$. In 1865, August Kekulé proposed a structure for benzene.

![Figure 31 - Kekulé structure of benzene.](image)
A double bond is stronger and therefore shorter than a single bond, so we would expect benzene to have alternating long and short bonds. However, it turns out that each bond length is identical and approximately halfway between that of a single bond and a double bond. This suggests that each bond is effectively one and a half bonds. This is because the electrons in the p orbitals of benzene are delocalised over all six carbon atoms equally, and so instead of alternating double and single bonds, a circle is drawn.

![Figure 32 - Delocalised structure of benzene.](image)

To represent delocalisation, multiple structures can be drawn with a double-headed arrow between them. This means that the molecule is somewhere between these two forms. The curly arrows here do not represent a reaction but simply that the electrons are delocalised around the ring.

![Figure 33 - Diagram showing the delocalisation in benzene.](image)

As benzene chooses to adopt this delocalised form, we can conclude that delocalisation is a stabilising effect.

### 4.2.2 Delocalisation in Carbon Chains

Delocalisation is not confined to rings but shows up in many different organic molecules. For example, trans-hexatriene.

![Figure 34 - Trans-hexatriene.](image)

There are two main issues with a localised bonding description of this molecule. One is that the molecule should rotate freely about the single bonds, but it turns out to be planar. The other is that the single bonds are shorter than they should be and the double bonds slightly longer.

If we carry over the idea from benzene that the p orbitals can delocalise over the whole molecule, then we can understand why it is planar. If the molecule were not planar, the p orbitals would not overlap effectively. It adopts a planar form to maximise delocalisation and its stabilising effect.

![Figure 35 - Diagram showing why trans-hexatriene is planar.](image)
Conjugation is the presence of alternating double and single bonds and it allows for $p$ orbital delocalisation. Note that if two single bonds separate the double bonds, the $p$ orbitals are too far away to overlap and if there are two double bonds directly next to each other, the $p$ orbitals lie perpendicular to each other and cannot overlap. A clear example of an extended conjugated system is in $\beta$-carotene, the molecule that gives carrots their orange colour.

![Figure 36 – The clear conjugated system in $\beta$-carotene.](image)

4.2.3 Other Examples of Delocalisation

Delocalisation does not just occur over carbon atoms. A good example is in propenal.

![Figure 37 - Propenal.](image)

There are two double bonds separated by a single bond, so this molecule is conjugated. The presence of an oxygen changes the energies of the $p$ orbitals, but they still overlap. The resultant MO covering the whole molecule is not symmetric, but there is still delocalisation.

Charge on a molecule can create delocalisation where otherwise it would not be seen. An example is the allyl anion.

![Figure 38 - Allyl anion.](image)

There is only one double bond, but the rightmost carbon is $sp^2$ hybridised with a leftover $p$ orbital, in which there are two electrons. This $p$ orbital can delocalise with the $p$ orbitals of the double bond. This delocalisation makes the anion symmetric.

![Figure 39 - Diagram showing how delocalisation makes the allyl anion symmetric.](image)

4.2.4 The Molecular Orbitals of Conjugated Systems

Now we can discuss the form of the MOs that the overlapping $p$ orbitals create. To do this we will consider how the $\pi$ and $\pi^*$ MOs combine. One of the simplest examples is 1,3-butadiene.

![Figure 40 - 1,3-butadiene.](image)
We are combining two \( \pi \) and two \( \pi^* \) localised MOs, so we will create a total of four new delocalised MOs. The lowest energy combination will be when the two \( \pi \) MOs combine in phase (constructively), so there are no nodes. The next highest will be when the two \( \pi \) MOs combine out of phase (destructively), creating a node through the central bond. Then we combine the two \( \pi^* \) MOs in phase, creating two nodes. The highest energy orbital will be when the two \( \pi^* \) MOs combine out of phase.

This can be represented on an MO diagram. Notation is to label the new MOs \( \psi_1, \psi_2, \psi_3, \) and \( \psi_4 \) in order of increasing energy.

The diagram has been filled with four electrons, two from each \( \pi \) orbital, and the total effect is a slight lowering of energy. The highest energy orbital with electrons in it is called the HOMO (highest energy occupied molecular orbital), and the lowest energy orbital without electrons in it is called the LUMO (lowest energy unoccupied molecular orbital).

### 4.3 Colour of Conjugated Systems

#### 4.3.1 Modelling Conjugated Systems with Particle in a Box

We can now explain the absorbance of visible light by conjugated systems. When light is shone on a molecule with a conjugated system, an electron can absorb a photon of the right frequency to jump from the HOMO of that conjugated system to the LUMO. The frequency of the light depends on the energy difference between the HOMO and LUMO via the Planck relation \( \Delta E = h\nu \).
To energies of the HOMO and LUMO can be estimated using particle in a box. An electron moving through a conjugated system can be thought of as a particle moving from one end of the conjugated system to the other as the electron is effectively trapped within the molecule. For simplicity, we will deal only with conjugated systems in simple carbon chains.

4.3.2 Calculating the Colour of Molecules

We have already derived the energy levels for particle in a box.

$$E_n = \frac{\hbar^2 n^2 \pi^2}{2mL^2}$$

We can turn this into a formula depending on the number of double bonds in a straight-chain conjugated system, $k$. The length of the system, $L$, can be approximated by taking the average bond length, $R$, which is approximately the mean of the C=C and C-C average bond lengths, and multiplying it by the total number of bonds in the chain, which is $2k - 1$.

$$L \approx (2k - 1)R$$

To work out the difference in energy between the HOMO and the LUMO, we must write $n$, the energy level, in terms of $k$. Each double bond contributes two electrons to the conjugated system, but there are two electrons per energy level and so for the HOMO, $n = k$. The energy level of the LUMO is therefore $n = k + 1$.

$$\Delta E = E_{LUMO} - E_{HOMO} = \frac{\hbar^2 (k + 1)^2 \pi^2}{2mL^2} - \frac{\hbar^2 k^2 \pi^2}{2mL^2} = \frac{\hbar^2 \pi^2}{2mL^2} ((k + 1)^2 - k^2) = \frac{\hbar^2 \pi^2}{2mL^2} (2k + 1)$$

Into this, we can substitute the Planck relation, $\Delta E = hv = \frac{hc}{\lambda}$.

$$\Delta E = \frac{hc}{\lambda} = \frac{\hbar^2 \pi^2}{2mL^2} (2k + 1) \Rightarrow \lambda = \frac{2mL^2 hc}{\hbar^2 \pi^2} (2k + 1) = \frac{8mc}{h} \left( \frac{L^2}{2k + 1} \right)$$

$$\therefore \lambda = \frac{8mc}{h} \left( \frac{R^2(2k - 1)^2}{2k + 1} \right)$$

In this equation, $m$, the mass of an electron, $h$, $c$, and $R$ are all constants. Given the number of double bonds, we have an equation that gives us the absorbed wavelength of light. The colour that the molecule appears will be the complementary colour to the absorbed colour.

![Figure 43 - A colour wheel can be used to easily work the colour of a molecule given the wavelength it absorbs.](image-url)
Let us calculate values for the absorbed wavelength for $k = 1 - 5$.

<table>
<thead>
<tr>
<th>No. Double Bonds</th>
<th>Absorbed Wavelength/nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>123</td>
</tr>
<tr>
<td>3</td>
<td>244</td>
</tr>
<tr>
<td>4</td>
<td>372</td>
</tr>
<tr>
<td>5</td>
<td>503</td>
</tr>
</tbody>
</table>

*Figure 44 - Table showing the calculated absorbed wavelengths for 1-5 double bonds.*

Past this, the calculated wavelength diverges from the actual value as this method is just an approximation. However, we are shown the very important point that a molecule must have a conjugated system of a reasonable length, at least five or six double bonds, to absorb in the visible spectrum. These extended conjugated systems turn out to be what give dyes their colour.

**Section 5 Terminology and History of Dyes**

**5.1 Some Simple Terminology**

To avoid confusion in our later discussions of dyes, there are a few terms that should be defined.

- **Dye** – A water-soluble organic substance that is coloured and can impart colour to an object, e.g., betanin (found in beetroot).

- **Pigment** – A water-insoluble, coloured, inorganic substance, e.g., hematite ($\text{Fe}_2\text{O}_3$).

- **Fastness** – The resistance of a dye to being removed from a material by washing or by exposure to light or heat.

**5.2 History of Dyes**

**5.2.1 Ancient History**

Dyes have a very long history. The first appearance of dyes is debated but prehistoric rock paintings, done using pigments, dating back to thousands of years BC are believed to depict coloured garments. One of the oldest dyes is indigo, used in denim. The oldest example of an indigo-dyed fabric is from 6000 years ago. The earliest written reference to dyeing is from China in 2600 BC.

In ancient civilisations, dyes were often seen as a luxury. Tyrian Purple was worth more than its weight in gold. It was extracted from sea snails and it is estimated that for 1g of the dye, about 8,500 snails must be used. Alexander the Great is said to have found old purple robes when he conquered Susa, Persia’s capital, which are suspected to have had a value equivalent to $6$ million.

*Figure 45 - Depiction of Justinian I wearing Tyrian Purple robes.*
5.2.2 Pre-Synthetic Dyes

Before the 19th century, dyes were found purely from natural sources. In the 15th century, insects were often used as a source. For example, the dye carmine comes from the insect cochineal. In the 1600s, dyeing ‘in the wood’ was introduced which involved extracting dyes from wood chippings. Logwood and fustic dyes are examples.

Prior to the 18th century, bleaching was done using alkaline or acid baths, but this often took months. In 1774, Carl Scheele discovered a yellow-green gas that removed colour from certain objects, chlorine. Claude Berthollet was the first to suggest its use in bleaching fabrics. He also discovered sodium hypochlorite, the first commercial bleach. Since then, many bleaching chemicals have been discovered such as hydrogen peroxide.

5.2.3 Synthetic Dyes

The turning point in dye history was the year 1856. William Perkin, 18 at the time, was attempting to synthesise quinine. Organic synthesis, however, was poorly understood at this point. One of his attempts involved oxidising aniline using potassium dichromate. The structures of aniline and quinine are shown below, and it is clear that this was not going to work.

![Aniline and quinine structures](image)

Figure 46 - Aniline (left) and quinine (right).

After this failed attempt, Perkin noticed a purple colour in his apparatus. Toluidine impurities in the aniline sample had reacted with the aniline, producing a variety of derivatives of mauveine differing only in the placement of methyl groups. These all absorbed light in the region 540-550nm due to the extended conjugated system, giving them their purple colour. A colour-coded structure of mauveine A is shown below, along with o-toluidine and p-toluidine. The green atoms are from o-toluidine, the red from aniline, and the blue from p-toluidine. The black bonds are those formed in the reaction.

![Mauveine and toluidines structures](image)

Figure 47 - The structures of p-toluidine (left), o-toluidine (centre) and mauveine A (right).
Perkin began mass-producing mauveine under the name aniline purple. From that point, discoveries in dye chemistry were being made frequently. In 1858, Peter Griess discovered the processes called diazotisation and coupling, reactions important in the synthesis of azo dyes. Ten years later, Graebe and Liebermann produced alizarin, the first natural dye to be produced synthetically. The first azo dye was successfully synthesised in 1880 by two English scientists, Thomas and Holliday. These are just a few of many discoveries made in the 19th and 20th centuries concerning the chemistry of dyes.

Section 6 Structure of Dyes

6.1 Chromophores and Auxochromes

6.1.1 Basic Structure

Graebe and Liebermann, in 1868, put forward the idea of dyes containing conjugated systems. In 1876 another German, Otto Witt, suggested a general structure of dyes consisting of a conjugated system of benzene rings with certain groups attached such as -NO₂, -C=O, or -N=N-, called the chromophore, and a polar group such as -OH or -NH₂, called the auxochrome. The chromophore imparts colour to the dye and the auxochrome deepens that colour (shifts the absorption to a longer wavelength). A couple of examples are shown below with the chromophores coloured blue and the auxochromes red.

6.1.2 Chromophores

Dyes are often built around a central chromophore, so certain groups in the chromophore can be used to classify the dye. Witt’s idea of the chromophore, while generally correct, took a slightly restricted view of the possible structures it can take.

Azo dyes contain an -N≡N- group in their chromophore. The groups attached to these nitrogens are often aryl groups. With the right substituents, they can absorb a wide range of wavelengths, but are mostly used for red, orange, and yellow dyes. Many azo dyes contain an -SO₃H group as an auxochrome. This is mostly present as -SO₃⁻ in neutral conditions so azo dyes are often produced as salts, with the organic body of the dye being the anion.

Figure 48 – The structures of Sudan I (left) and Acridine Orange (right).

Figure 49 - The structure of Trypan Blue, an azo dye.
Arylmethane dyes are divided into two main subgroups, diarylmethane dyes and triarylmethane dyes. The backbones of these two groups are shown below.

![Figure 50 - Diarylmethane (left) and triarylmethane (right) general structure.](image)

Auramine O is currently the only commonly used diarylmethane dye. Methyl violet dyes are triarylmethane dyes with -N(CH₃)₂ groups at the para position of at least two of the aryl groups. Fuchsine dyes can have primary or secondary amines at these positions. A final example is phenol dyes, where at least two aryl groups have -OH groups at their para positions. These phenol dyes often change colour with pH, for example, phenolphthalein. Many of these dyes, due to their amine groups, are basic and are often produced as salts with the organic body as the cation.

![Figure 51 - The structures of Auramine O (left) and Methyl Violet 6B (right).](image)

Anthraquinone dyes are built around an anthraquinone chromophore.

![Figure 52 - Anthraquinone.](image)

Anthraquinone is colourless and so auxochromes such as -OH are needed. Alizarin contains two -OH groups on the same aryl group. This structure allows alizarin to act as a pH indicator.

![Figure 53 - The structure of alizarin.](image)
Xanthene-based dyes contain heterocycles, which in organic chemistry are rings containing elements other than carbon. Xanthene itself is not particularly useful, but when it is slightly altered it becomes a good chromophore.

The main classes of xanthene-based dyes involve substitutions of R1 and R2.

<table>
<thead>
<tr>
<th>R1</th>
<th>R2</th>
<th>Dye Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>-CH=</td>
<td>-S-</td>
<td>Thiopyronine</td>
</tr>
<tr>
<td>-CH=</td>
<td>-O-</td>
<td>Pyronine</td>
</tr>
<tr>
<td>-CH=</td>
<td>-NH-, -NMe-</td>
<td>Acridine</td>
</tr>
<tr>
<td>-N=</td>
<td>-S-</td>
<td>Thiazine</td>
</tr>
<tr>
<td>-N=</td>
<td>-O-</td>
<td>Oxazine</td>
</tr>
<tr>
<td>-N=</td>
<td>-NH-, -NMe-</td>
<td>Azine</td>
</tr>
</tbody>
</table>

These derivatives all have different uses. Oxazines and thiazines fade quickly in light on silks but are useful for dyeing acrylics. Mauve, the first synthetic dye, was an azine. Fluorescein is a type of fluorescent xanthene-based dye.

The phthalocyanines were first developed in the 20th century, more recently than the other classes. They are the synthetic analogues of chlorophyll and haemoglobin. Phthalocyanine itself is similar in structure to the common organic molecule, porphyrin.
Some of the carbons of phthalocyanine can be substituted for nitrogen, and some of the outer hydrogens for auxochromes such as -OH or -NH₂ groups. This allows the colour of the dye to be finetuned. Phthalocyanines are often produced as metal complexes, with the most common metal ion being copper. For example, the dye Monastral Fast Blue G is simply phthalocyanine as its copper complex.

![Figure 58 - The structure of Monastral Fast Blue G.](image)

6.1.3 Auxochromes

Auxochromes usually contain a polar group with a lone pair. For example, -OH, -NH₂ and -SH. These donate electrons into the conjugated system and lengthen it, decreasing the energy difference between the HOMO and the LUMO, causing an absorbance shift to a longer wavelength known as a bathochromic shift.

There are certain examples which can cause a shift to a shorter wavelength, a hypsochromic shift. This is often due to the presence of carbonyl groups next to heteroatoms. However, we will not cover the details of how this works here.

Auxochromes can be divided into two groups, acidic and basic. This division mainly affects how a dye is applied. Nitrogen-based auxochromes are generally basic, and oxygen and sulphur-based auxochromes are generally acidic.

Section 7 Application and Bleaching

7.1 Application of Dyes

7.1.1 Material Structure

To understand how dyes are applied we must cover the chemical structure of certain materials.

Many fabrics, such as cotton, are made of plant fibres which are composed mainly of cellulose. Cellulose is a polymer of glucose and the repeat unit of cellulose is shown below. The chains are stacked on top of each other and attract each other via hydrogen bonds.

![Figure 59 - The repeat unit of cellulose.](image)
In all plant-based fibres, there are substances other than cellulose present such as lignin and pectin. This is not important, however, as the cellulose is abundant enough for a deep colour to be achieved anyway. For example, cotton is usually 94% cellulose and most plant fibres contain at least 60%. Rayon is a type of synthetic cellulose, and acetate rayon replaces some of the -OH groups with -OAc groups. Both of these are commonly used in the clothing industry.

Silk, and other animal fibres such as hair, wool and leather, are mainly composed of proteins. These are chains of amino acids joined by peptide bonds. For example, the main protein in silk is fibroin, generally consisting of the recurring sequence of amino acids Gly-Ser-Gly-Ala-Gly-Ala.

![Figure 60 - Recurring sequence of amino acids in the protein fibroin.](image)

In hair and wool, the main protein is keratin and in leather, it is collagen. Different proteins have different amino acid sequences. The amino acids often contain acidic, basic, and polar groups which serve as useful points for interacting with dyes molecules.

Synthetic fibres are usually polymers of some sort. Nylon, for example, is similar in structure to proteins, with amide groups separated by carbon chains of certain lengths. The polar amide groups serve as sites for interaction with dyes. Acrylics in their purest form are chains of hydrocarbons bearing nitrile groups. To improve dyeability, they are often polymerised in the presence of other molecules which add groups such as -OH. The structures of nylon-6 and a general acrylic polymer are shown below.

![Figure 61 - The repeat unit of nylon-6 (left) and an acrylic polymer (right).](image)

Some synthetic materials are difficult to dye. PET accounts for 50% of worldwide synthetic fibre production, and while it does contain polar groups that would serve as reasonable dyeing sites, the polymer chains are packed too close together for effective dyeing.

![Figure 62 - The repeat unit of PET.](image)

Tight chain packing leads to poor dyeability because dyeable fibres must be porous so that dye molecules can enter the fibre through pores and fix themselves to the chains. If these pores are too small, like in PET, then very few dye molecules can enter the material and the material is not easily dyed.
7.1.2 Dyeing Methods

There are multiple ways in which dyes can bond to fibres. These depend on both the material and the type of dye being used.

Direct dyes are applied simply by placing the material in a hot, aqueous solution of the dye. The temperature improves dye solubility, and direct dyes must be water-soluble. For this reason, they usually contain ionic or polar groups. Direct dyes can be applied to most fabrics, except tightly packed ones such as PET. The main bonding interactions are between ionic or polar groups of the fibre and the dye. Direct dyes are often large, resulting in strong intermolecular forces of attraction (dispersion forces) between the dye and the fibre, further strengthening its attachment.

![Diagram of an ionic interaction between an arbitrary protein and the dye Direct Blue 1.](Figure 63)

For dyeing tightly packed synthetic fibres such as PET and acetate rayon, disperse dyes are required. They are insoluble in water and the dyeing process involves the material being placed in a very fine, hot suspension of the dye with a chemical called a carrier. The carrier, an example being benzyl alcohol, transports the dye into the fibre when hot as the pores have expanded. Then, upon cooling, the dye becomes trapped. As the dye is insoluble and the fibre tightly packed, this dyeing method creates materials with excellent wash-fastness. Disperse dyes must be relatively small to enter the fibre.

![The structure of Disperse Blue 6.](Figure 64)
Vat dyeing is used to apply insoluble dyes to fibres. First, a reduced form of the dye that is itself soluble in water is applied to the material. Then, the dye is oxidised within the fibre, usually in air, into its original, insoluble form. Because of the insolubility of vat dyes, they often have very good wash-fastness. Indigo is a common example of a vat dye. Indigo itself is insoluble in water, but the reduced form, indigo white, is soluble.

![Indigo White and Indigo Dye](image)

Figure 65 - The structures of Indigo White (left) and Indigo Dye (right).

Reactive dyes react with the fibre and so attach themselves via covalent bonds. They are usually used to dye plant-based fibres as they can react with the -OH groups of cellulose. The most common reactive dyes are those based on haloheterocycles, heterocycles with attached halogens, and vinyl sulfones, sulfonyl groups next to double bonds. The haloheterocycles are reacted with cellulose in basic conditions. A deprotonated -OH group from cellulose attacks the heterocycle in a nucleophilic aromatic substitution. The vinyl sulfones are formed by elimination and go on to react with cellulose in basic conditions. A deprotonated -OH group from cellulose attacks the end carbon in a mechanism called Michael addition.

![Halo and Vinyl Reactions](image)

Figure 66 - Example reactions of a haloheterocycle (top) and vinyl sulfone (bottom) with cellulose, denoted Cell.

Azo dyes are synthesised by reacting a diazonium salt with a coupling component. To apply azo dyes, a material can be treated with a solution of the coupling component and then placed in a solution of the diazonium salt, forming the dye on the fabric, or vice versa.

![Azo Reaction](image)

Figure 67 - Example of an azo coupling reaction.
7.1.3 Mordants

It is often the case that the dye alone cannot form strong enough interactions with the fibre and so has poor fastness. This is often the case with direct dyes. To solve this issue mordants can be used. A mordant is a substance that helps fix a dye to a material, and dyes that require mordants are called mordant dyes.

Mordants are usually inorganic metal salts containing a metal ion with an oxidation state of at least +2, often aluminium or iron (III). A common example is alum, KAl(SO$_4$)$_2$·12H$_2$O. The metal ion can attach to the dye via both a covalent and a coordinate bond in a process known as chelation.

When a mordant dye and metal ion chelate together, the complex formed is called a lake. Generally, the covalent bond is to an -OH oxygen and the coordinate bond to a -C=O oxygen.

![Figure 68 - Two examples of metal-dye lakes.](image)

The metal can attach to more than one dye molecule, and this does happen. However, not all of the metal ions form complete complexes with the dye, as some of them form a complex with the fibre too. In this way the dye is attached to the fibre through the metal ion.

In a way Mordants can also alter the colours of dyes due to the presence of the metal in the conjugated system. This is because the orbitals on the metal can affect the energy difference between the HOMO and the LUMO.

Cellulose-based materials do not bond well to mordant dyes as they are unable to form chelations with the metal ions. Mordant dyes are normally used to dye protein-based substances. In proteins, there are often many hydroxyl and carboxyl groups available, with which the metal can chelate, creating a dye-metal-fibre complex.

7.2 Bleaching

7.2.1 Bleaching Mechanism

Bleaching a material is done by destroying the chromophore of the dye. Chemical bleaches react with the chromophore and either remove double bonds from it or break it apart completely. This interrupts the conjugated system, meaning the dye can no longer absorb visible light, making it colourless.

Most bleaches are strong oxidising agents which can break bonds in several different ways. Some are reducing agents which can convert double bonds into single bonds. It is also possible to cleave bonds simply using energy, so heat and light of high enough energy such as UV light, can bleach materials.

As a result of their ability to cleave bonds and destroy certain molecules, bleaches are highly effective at denaturing proteins. This makes them dangerous to humans, but also makes them very good disinfectants as they can easily kill bacteria.
7.2.2 Common Bleaches

Some of the most common bleaching agents are those based on chlorine. The first chemical bleach was chlorine gas and since then several other bleaches containing chlorine have been discovered. Chlorine itself is mostly used for disinfecting water but can be used to bleach wood pulp. Sodium hypochlorite, NaClO, is often simply called bleach and is generally used in households. A mixture of calcium hypochlorite, Ca(ClO)₂, calcium hydroxide, Ca(OH)₂, and calcium chloride, CaCl₂, creates bleaching powder, which has many of the same uses as sodium hypochlorite.

An example of sodium hypochlorite acting as a bleach is its reaction with an alkene. The chlorine atom in the ClO⁻ ion has a partial positive charge and can act as an electrophile. The double bond attacks this atom, forming a three-membered ring as an intermediate, which is then attacked by the solvent, water, forming a halohydrin, which can undergo further reactions. The result is that the double bond has been broken.

![Figure 69 - Reaction of sodium hypochlorite with an alkene.](image)

The other most common bleaches are those based on peroxides. These are characterised by an O-O peroxide bond which is easily broken, producing highly reactive species. Hydrogen peroxide, H₂O₂, is used to bleach wood pulp and hair and can be used to prepare other bleaching agents. Sodium percarbonate, Na₂CO₆, and sodium perborate, Na₂H₄B₂O₈, are other examples of peroxide bleaches and have a wide range of uses.

Section 8 Specific Examples

8.1 Important Examples in Living Organisms

8.1.1 Chlorophyll

Chlorophyll is the green pigment found in plants, algae, and any other living organism that photosynthesises. Its function is as a photoreceptor, meaning it traps light. This light then catalyses the production of glucose via photosynthesis. The ability of chlorophyll to absorb light is very closely related to the theory of dyes we have covered.

Chlorophyll, like any other dye, contains a large conjugated system. The chromophore is a substituted porphyrin ring chelated with a central metal ion of magnesium. The structures of the two main types of chlorophyll, A and B, are shown below. The only difference is that in chlorophyll A, the R group is -CH₃, and in chlorophyll B, it is -CHO.

![Figure 70 - The structure of chlorophyll.](image)
The extended conjugated system is clear, stretching over the porphyrin and some of the attached groups. The wavelength of light absorbed is further altered by the magnesium, whose orbitals can interact with those of the conjugated system.

Chlorophyll A and B work together in plants to absorb a large range of the visible spectrum, allowing sunlight to be used as a source of energy. Wavelengths that one type fails to absorb are often made up for by the other. However, there is still a gap in the absorbance spectrum from 500nm to 600nm, which corresponds to green light, causing the green colour of chlorophyll.

In photosynthesis, energy from sunlight is absorbed by the conjugated system, leading to the excitation of an electron in chlorophyll. This electron can now be easily transferred to other molecules and a long chain of electron transfers leads to the transfer of an electron to a CO₂ molecule, reducing it, and the removal of an electron from H₂O molecule. Chlorophyll is bound to the back of a large, complex protein which positions it perfectly to react with any nearby water and carbon dioxide quickly and efficiently.

As chlorophyll has such a strong absorbance in the visible spectrum, it masks the colours of other molecules present in leaves. This can be seen in autumn when the chlorophyll decays, allowing the oranges and reds of carotenoids to be seen. Also, when leaves are cooked, they become slightly more yellow as the magnesium is removed from the chlorophyll molecules.

8.1.2 Retinal

Retinal is found in the eye and is a molecule that allows humans to see. It undergoes a cycle of chemical changes, together with the protein opsin, called the visual cycle. The retinal molecule contains an extended conjugated system. The structure of this conjugated system is why we see the wavelengths of light that we do, and not, for example, ultraviolet light.

![Figure 71 - The structure of all-trans-retinal.](image1)

Retinal is very similar in structure to β-carotene. As humans cannot synthesise retinal, we must ingest β-carotene through foods such as carrots, which can then be cleaved to give vitamin A (retinol) and through oxidation, retinal.

![Figure 72 - A summary of the conversion of β-carotene to all-trans-retinal.](image2)
At a certain point in the visual cycle, a molecule of 11-cis-retinal binds to the protein opsin via a lysine residue. This new retinal-opsin molecule is called rhodopsin and has a strong absorbance in the region of 400 nm – 700 nm, depending on the type of opsin. When a photon of the right wavelength is absorbed by the rhodopsin, the energy goes converts the double bond between the 11th and 12th carbon into a single bond, meaning it is free to rotate. When it has rotated through 180°, the double bond reforms, now in a trans configuration. The retinal is now all-trans-retinal, which does not fit the opsin binding site well, so the link between the retinal and the opsin weakens, and several fast reactions occur before they detach completely. This sudden movement is transferred through the protein to the membrane it is attached to, and eventually is picked up by nerve cells in the optic nerve which send a signal to the brain. The brain detects these and so we can see. The all-trans-retinal can undergo a series of enzyme-catalysed changes to reform 11-cis-retinal and the cycle repeats.

[Diagram of the visual cycle]

There are three types of rhodopsin, each with a slightly different structure, that absorb at different wavelengths. The three are responsible for light in the red, green, and blue portions of the visible spectrum and all three working together results in our ability to see the colours that we do.

8.2 pH Indicators

8.2.1 General Mechanism

Certain chemicals appear different colours in acidic and basic conditions and can, therefore, be used as pH indicators. The pH of a solution is defined as the negative logarithm, base ten, of the concentration of H₃O⁺ ions.

\[ pH = -\log_{10}[H_3O^+] \]

The colour change is usually caused by the dissociation of an H⁺ ion from the indicator. The neutral, HIn, and deprotonated, In⁻, forms of the indicator are in equilibrium with each other.

\[ HIn + H_2O \rightleftharpoons In^- + H_3O^+ \]

The equilibrium constant for this reaction is called the acid dissociation constant, K_a. Another constant, the pK_a, is defined as the negative logarithm, base ten, of the acid dissociation constant. The pK_a values of different indicators can be found in data books.

\[ K_a = \frac{[H_3O^+][In^-]}{[HIn]}, \quad pK_a = -\log_{10}K_a \]
There is an equation linking the concentration of $HIn$ and $In^-$, the pH of the solution and the $pK_a$ of the indicator (the $pK_{In}$) called the Henderson-Hasselbalch equation. It can be derived by taking logarithms of both sides of the $K_a$ equation.

$$pH = pK_a + \log_{10} \frac{[In^-]}{[HIn]}$$

From this equation, it can be deduced that if the pH is equal to $pK_{In}$, the concentration of $In^-$ is equal to the concentration of $HIn$. If the pH is below $pK_{In}$, the concentration of $In^-$ is less than that of $HIn$. If the pH is above $pK_{In}$, the concentration of $In^-$ is greater than that of $HIn$.

For a weak acid indicator where $HIn$ has a colour, say red, and $In^-$ a different colour, say blue, then the colour changes can be estimated. At a pH equal to $pK_{In}$, the concentrations of $HIn$ and $In^-$ are equal, so the solution is purple. At a lower pH, the concentration of $In^-$ is smaller than the concentration of $HIn$. As the pH scale is logarithmic, after about one unit of pH, the concentration of $HIn$ is already about ten times larger than that of $In^-$, so the solution is red. At a higher pH than $pK_{In}$, the concentration of $HIn$ is smaller than the concentration of $In^-$, so the solution is blue.

For any pH indicator, there is a $pK_{In}$ and a range either side of this in which the solution is undergoing a colour change. The best indicators have a clear colour change and a small range over which the colour is changing.

8.2.2 Common Examples and Acid-Base Titrations

A universal indicator gives an approximate measure of pH by having a wide range of colours in different pH values. Universal indicators are composed of a mixture of several indicators, each specifically selected for their $pK_{In}$ and colour change. When the correct indicators are mixed, a solution can be created that steadily changes through multiple different colours as the pH is altered, which is very useful for qualitative analysis.

The most common universal indicator is composed of a mixture of thymol blue, methyl orange, methyl red, bromothymol blue and phenolphthalein. This mixture results in a colour change from red in strong acid to violet in strong alkali, following the visible spectrum.

Indicators are often used in acid-base titrations, and the two most common indicators for this are phenolphthalein and methyl orange. Phenolphthalein loses two protons between pH 8.2 and 10.0, going from the colourless $H_2In$ to the bright pink $In^{2-}$ in this range. Methyl orange changes between pH 3.2 and pH 4.4 from the red zwitterion, $HIn$, to the yellow $In^-$. 

Figure 74 - The structures of phenolphthalein below pH 8.2 (left) and above pH 10.0 (right).
When choosing an indicator for a titration, it must change colour at a pH near the equivalence point, the point where the two substances have been mixed in equal proportions. The shape of the titration curve allows for a large variation in pH from the equivalence point with minimal loss of accuracy. For this reason, both phenolphthalein and methyl orange change colour close enough to pH 7 that they can be used for strong acid-strong base titrations. But for weak acid-strong base titrations, the equivalence point is at a higher pH, so methyl orange is no longer a valid indicator to use. The opposite is true for strong acid-weak base titrations, so phenolphthalein is not valid. For weak acid-weak base titrations, the shape of the titration curve is such that neither of these indicators are useful.

Section 9 Concluding Remarks

This exploration of the chemistry of dyes has taken us through many branches of chemistry from quantum mechanics and atomic and molecular orbitals to biochemistry and acid-base indicators. It is clear that the study of dyes is a rich field full of interest and that there is a lot to be gained from studying them.

Dyes not only hold an important place in chemistry but also our lives. All the clothes we wear every day have most likely been dyed or bleached in some way. Then there is the fabric on the furniture, the food dyes in the cupboard, the list goes on. Even outside, away from man-made objects, the landscape is filled with the vibrant colours of plant life. It is impossible to escape the importance of dyes in our daily lives.

Section 10 Bibliography

10.1 Section 2 Quantum Mechanics


10.2 Section 3 Atomic and Molecular Orbitals


Hanson, D and Harvey, E and Sweeney, R and Zielinski, T. (2020). The Schrödinger Wave Equation for the Hydrogen Atom. [online] Available at: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Map%3A_Physical_Chemistry_for_the_Biosciences_(Chang)/11%3A_Qualum_Mechanics_and_Atomic_Structure/11.10%3A_The_Schr%C3%B6dinger_Wave_Equation_for_the_Hydrogen_Atom [Accessed 24 Apr. 2020].


10.3 Section 4 Bonding in Organic Molecules and Colour


10.4 Section 5 Terminology and History of Dyes


### 10.5 Section 6 Structure of Dyes


10.6 Section 7 Application and Bleaching


(2016). Dyes and Dyeing. [online] Available at: https://www.flinnsci.com/api/library/Download/c46ad9cbff64e9cae7caf3aed645692 [Accessed 14 May 2020].


### 10.7 Section 8 Specific Examples


*Retinal*. [online] Available at: [https://www.ch.ic.ac.uk/vchemlib/mim/bristol/retinal/retinal_text.htm](https://www.ch.ic.ac.uk/vchemlib/mim/bristol/retinal/retinal_text.htm) [Accessed 12 Jun. 2020].


