

THEORETICAL MODELS

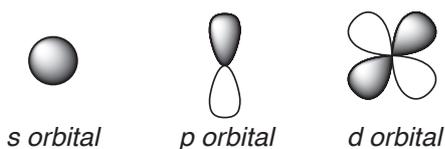
A variety of different procedures based on quantum mechanics (so-called quantum chemical models) have been developed to calculate molecular structure and properties as well as infrared, NMR and UV/visible spectra. All follow from a deceptively simple looking *Schrödinger equation* first written down in 1927.

$$\hat{H}\Psi = \epsilon\Psi$$

\hat{H} (the *Hamiltonian* or more precisely *Hamiltonian operator*) is the only known. It describes the kinetic energies of the particles that make up a molecule and the Coulombic interactions between the individual particles. Positively-charged nuclei repel other nuclei, and negatively-charged electrons repel other electrons, but nuclei attract electrons. Ψ (the *wave function*) is a function of the Cartesian coordinates, and ϵ (the *energy*) is a number. The goal in solving the Schrödinger equation is to find a function that when operated on by the Hamiltonian yields the same function multiplied by a number. Note that there are many (actually an infinite number of) solutions to the Schrödinger equation. These correspond to the ground and numerous excited states of an atomic or molecular system.

The energy of a molecule can be measured. On the other hand, the wave function has no physical meaning, and is not subject to experimental measurement, although the square of the wave function times a small volume element gives the probability of finding an electron inside that volume. This exactly corresponds to what is actually measured in an X-ray diffraction experiment.

The Schrödinger equation has been solved exactly for the hydrogen atom (a one-electron system), where the wave functions are familiar to chemists as the s, p, d, ... atomic orbitals. The lowest-energy s orbital corresponds to the ground state of the hydrogen atom, whereas the higher-energy solutions correspond to excited states.



The Schrödinger equation may easily be written down for many-electron atoms as well as for molecules, although it cannot be solved. Even something as seemingly simple as the helium atom with only two-electrons presents an insurmountable problem. Approximations must be made.

Hartree-Fock Molecular Orbital Models

Hartree-Fock molecular orbital models or *molecular orbital models*, as they are commonly referred to, were the first practical quantum chemical models to be formulated. They result from making three approximations to the Schrödinger equation:

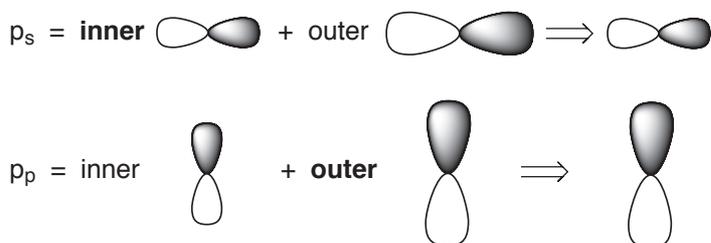
1. Separate nuclear and electron motions. The *Born-Oppenheimer approximation* says that “from the point of view of the electrons”, the nuclei are stationary. This eliminates nuclear motion and leads to an *electronic Schrödinger equation* which can be solved for the H_2^+ molecule, but cannot be solved for molecules with more than one-electron. The Born-Oppenheimer approximation is of little consequence for the description of molecular properties, for example, equilibrium geometries and reaction energies, and may be used without concern.
2. Separate electron motions. The *Hartree-Fock approximation* eliminates the need of having to simultaneously account for the motions of several electrons. It leads to a much simpler set of equations in which the motion of each electron in an environment made up of the nuclei and all the other electrons is sought.
3. Represent each one-electron solution or *molecular orbital* by a linear combination of atom-centered functions or *atomic orbitals*. The *LCAO (Linear Combinations of Atomic Orbitals) approximation* reduces the problem of finding the best functional form for the molecular orbitals to the much simpler problem of finding the best set of linear coefficients. As the number and complexity of the atomic orbitals increases, the energy and other properties approach limiting values. However, computational cost also increases. The goal is to provide as few functions as possible to yield a value for the property of interest that adequately

reflects its limit. Note, that the limiting values of properties are not expected to be the same as experimental values, but rather reflect the behavior of the Hartree-Fock model.

Basis Sets

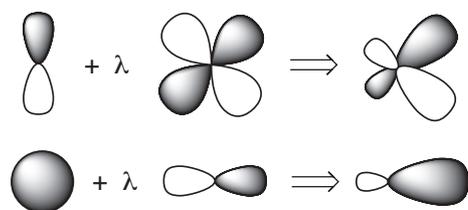
Gaussian functions are polynomials in the Cartesian coordinates times an exponential in the *square* of the distance from the origin. They are distinct although very closely related to the exact solutions of the hydrogen atom (exponential functions in the distance), and are labeled 1s, 2s, 2p, ... , the same nomenclature used to describe hydrogen atom solutions.

A *minimal basis set* includes only sufficient functions to hold all the electrons on an atom and to maintain its spherical shape. This involves a single 1s orbital for each hydrogen atom, and a set of five orbitals (1s, 2s, 2p_x, 2p_y and 2p_z) for each carbon atom. Because a minimal basis set incorporates only one set valence p functions, the components of which are the same size, atoms in nearly spherical environments will be better described than atoms in aspherical environments. A *split-valence basis set* addresses this problem by providing two different sets of valence p functions, one compact set and one loose set. This allows different linear combinations for different directions. For example, the compact p orbital can be emphasized to construct a σ bond while the loose p orbital can be emphasized to construct a π bond.



Because the functions in a minimal or split valence basis set are centered on the atoms, they may have difficulty describing electron distributions that fall in between atoms (that is, bonds). *Polarization basis sets* address this problem by providing a set of d-type functions (*polarization functions*) on main-group elements, and (optionally) a

set of p-type functions on hydrogen. The resulting combinations can be thought of as hybrid orbitals, for example, the pd and sp hybrids shown below.



The so-called 6-31G* basis set will be used for the infrared and NMR calculations described in future topics. The number “6” to the left of the “-” in the name indicates that 6 functions are used to describe each inner-shell (core) atomic orbital. The numbers, “31” to the right of the “-” indicate that groups of 3 and 1 functions are used to describe each valence-shell atomic orbital. “*” designates that polarization functions are supplied for non-hydrogen atoms. Were two stars to be present (as in 6-31G**) this would indicate that p-type polarization functions would also be placed on hydrogen atoms.

The valence basis functions can be further split and additional polarization functions can be added including f-type functions. A commonly used basis set is designated 6-311+G(2df,2p). “311” indicates a triply-split valence, “2df” indicates that two sets of d-type functions and a set of f-type functions are added to the valence of heavy atoms, and “2p” indicates that two sets of p-type functions are added to the valence of hydrogen atoms.

Taken together, these three approximations lead to a set of equations known as the *Roothaan-Hall equations*. They increase in computational cost as the cube of the size (number of basis functions), and can easily be applied to molecules incorporating up to 100 heavy (non- hydrogen) atoms.

Beyond Hartree-Fock Models

Of the three approximations we have made to reach Hartree-Fock models, the second is to be taken most seriously. According to the Hartree-Fock approximation, electrons “move independently”,

which means that both the electron-electron repulsion energy and the total energy will be too large. The limiting Hartree-Fock energy is, therefore, necessarily higher (less negative) than the experimental energy. *Electron correlation* is the term given to describe the coupling or correlation of electron motions. The *correlation energy* is defined as the difference between the Hartree-Fock energy and the experimental energy and is necessarily a negative quantity.

There are two conceptually different approaches for calculating the correlation energy, and numerous specific models arising from each of these approaches. Wave function based models start from the Hartree-Fock wave function combining it with wave functions resulting from electron excitations from filled to empty molecular orbitals. Density functional models supplement the Hartree-Fock Hamiltonian.

Configuration Interaction Models

Configuration interaction models are archetypal of wave function based correlated models. In the unachievable limit, so-called *full configuration interaction*, all possible single and multiple electron promotions from occupied to unoccupied Hartree-Fock molecular orbitals and assuming a complete basis set, the energy is same as would be achieved by solution of the electronic Schrödinger equation. More practical *limited configuration interaction* schemes have been formulated by limiting the number of electron simultaneously promoted (1-electron, 2-electron, ...) and the number of filled and unfilled molecular orbitals involved in the promotions.

Møller Plesset Models

An alternative and more commonly used wave function based model is *Møller-Plesset theory*. This assumes that the Hartree-Fock energy E_0 and wave function Ψ_0 are solutions to an equation involving a Hamiltonian, \hat{H}_0 , that is very close to the exact Schrödinger Hamiltonian, \hat{H} . This being the case, \hat{H} can be written as a sum of \hat{H}_0 and a small correction, V . λ is a dimensionless parameter.

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}$$

Expanding the exact energy in terms of a power series of the Hartree-Fock energy yields:

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \lambda^3 E^{(3)} + \dots$$

Substituting this expansion into the Schrödinger equation and collecting terms in powers of λ leads to an explicit expression for the energy correction. The sum of $E^{(0)}$ and $E^{(1)}$ is the Hartree-Fock energy. Including the next term gives rise to the so-called **MP2** (second-order Møller-Plesset) *model*.

Both the simplest configuration interaction (limited to single and double electron excitations only) and the MP2 models increase in computational cost as the power of the total number of basis 5th functions. In practice, this limits them to molecules incorporating 25-50 heavy (non-hydrogen) atoms at most.

Density functional theory or simply **DFT**, is based on two theorems elaborated by Hohenberg and Kohn, which taken together, prove that the energy and other properties of a many-electron system in its ground state may be correctly and uniquely described in terms of a function of the electron density. The term “functional” or a function of a function arises because the electron density is itself a function of the three spatial coordinates. What the two Hohenberg-Kohn theorems imply is that the Schrödinger equation can actually be solved; that is, the completely “intractable” problem involving the coupled motions of n electrons in a static field due to the nuclei (a “molecule”) may be replaced by an eminently “solvable” problem that treats the electrons as independent (that is, non-interacting) particles. Because the electron density is a function of only three coordinates, in effect a 3 dimensional problem is substituted for a $3n$ dimensional problem. Unfortunately, the Hohenberg-Kohn theorems do not tell us anything about the functional itself.

In the density functional formalism, the electronic energy, E^{el} , is written as a sum of the kinetic energy, E_{T} , the electron-nuclear interaction energy, E_{v} , the Coulomb energy, E_{J} , and a term combining the exchange and correlation energies, E_{XC} .

$$E^{\text{el}} = E_{\text{T}} + E_{\text{v}} + E_{\text{J}} + E_{\text{XC}}$$

What is “the” exchange/correlation functional? The quest has gone on for several decades and hundreds of functionals have actually been proposed and divided into several distinct classes.

Local Density Approximation (LDA)

The functional first proposed stems from a purely hypothetical problem in which a uniform gas of non-interacting electrons moves in a positively charged field. An analytical solution for the exchange energy is available and takes the form of the density raised to the 4/3's power. The correlation energy may be arrived at through numerical simulation, and is also only dependent on only the local density at each point.

$$E_{xc} = E(\rho)$$

Functionals that depend only on the electron density (ρ) are referred to as LDA functionals. They are not very successful in describing the properties of molecules and have largely been replaced.

Generalized Gradient Approximation (GGA)

GGA functionals depend on the gradient of the electron density, $\nabla\rho$, in addition to the density itself.

$$E_{xc} = E(\rho, \nabla\rho)$$

They have been around since the mid 1980's and were the first to provide a reasonable account of molecular properties and in particular the energies of chemical reactions. In this sense, they were instrumental in drawing attention to density functional theory as a viable “low-cost” alternative to wave function based correlation techniques, something that LDA functionals had failed to do. The BLYP functional is representative.

Global Hybrid Generalized Gradient Approximation (GH-GGA)

GH-GGA functionals (often referred to as hybrid functionals) replace a fixed fraction of the exchange by the “exact” Hartree-Fock exchange, the fraction being an adjustable parameter. It is likely that it was the introduction of hybrid functionals that caused the community

to recognize that density functional theory was “semi-empirical” in nature. Adding the Hartree-Fock exchange is “costly”, but led to significant improvements in the description of reaction energies. While GH-GGA functionals were introduced in the early 1990’s, they remain a mainstay in the application of density functional theory to chemistry. The B3LYP functional, in particular, is perhaps still more widely used than any other functional, even though there are now much better choices.

Range Separated Hybrid Generalized Gradient Approximation (RSH-GGA)

The idea behind range separated GGA hybrid functionals is that the “optimum” amount of Hartree-Fock exchange varies with electron-electron distance, from a small percentage in the long range limit to a large percentage in the short range limit, in the extreme from 0% to 100%. Both ω B97X-D and ω B97X-V functionals are range-separated GGA hybrids. Both incorporate so-called local corrections to account for dispersive interactions (see discussion following).

***meta* Generalized Gradient Approximation (mGGA)**

A *meta* GGA functional not only depends on the electron density and its gradient (as does a GGA functional) but also on the Laplacian of the electron density. As such, it can be construed as the next logical step beyond GGA in constructing a Taylor series expansion of the electron density.

$$E = E(\rho, \nabla\rho, \nabla^2\rho)$$

More commonly, *meta* GGA functionals are viewed as adding the so-called kinetic energy density to GGA.

$$E = E(\rho, \nabla\rho, \nabla^2\psi)$$

In either case, the addition can be construed as second-order term in a Taylor series expansion of the electron density. The B97M-V functional is an example of a “pure” *meta* functional.

Global Hybrid *meta* Generalized Gradient Approximation (GH-mGGA)

These are strictly akin to global hybrid GGA (GH-GGA) functionals in that a fixed percentage of the Hartree-Fock exchange is introduced. The “only” difference is that a *meta* GGA functional including a second-order term replaces a GGA functional. The M06-2X functional, widely accepted as an excellent choice for thermochemical comparisons, is an example of a GH-*meta* GGA functional.

Range Separated Hybrid *meta* Generalized Gradient Approximation (RSH-mGGA)

A range-separated *meta* GGA (RSH-mGGA) functional is identical to a range-separated GGA (RSH-GGA) functional except that a *meta* GGA functional has replaced the underlying GGA functional. M11 and ω B97M-V are examples of range-separated *meta* GGA functionals.

Double Hybrid *meta* Generalized Gradient Approximation (DH-RSH-mGGA)

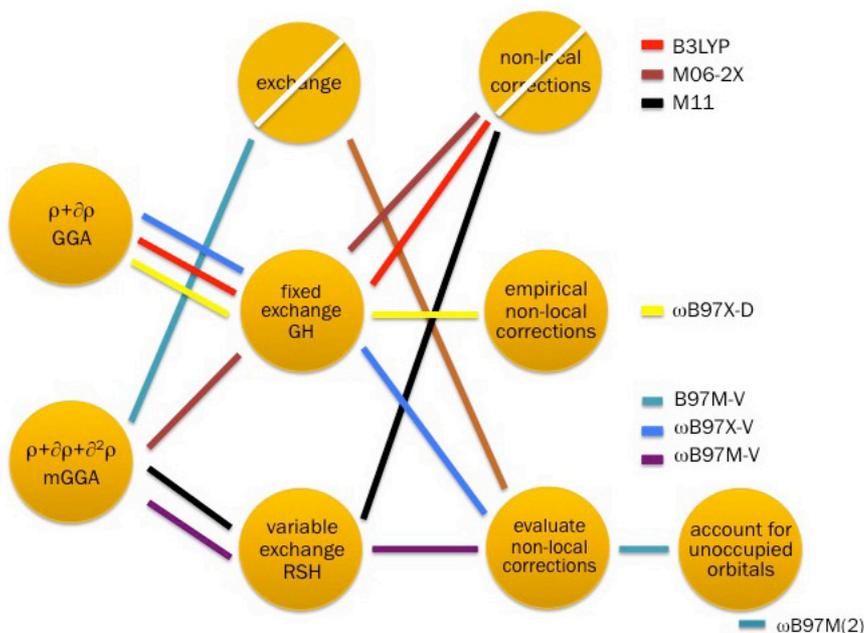
A double hybrid *meta* GGA (DH-RSH-mGGA) functional accounts for contributions of unoccupied molecular orbitals as well as occupied orbitals, by way of a MP2 like approach. The ω B97M(2) functional is now supported for energy calculations only.

Non Local Corrections

The previous functional classes are all considered “local” in that they are described in terms of a single integral over the three spatial coordinates. In order to capture dispersive van der Waals interactions, so-called non local correlation functionals are needed. These involve a double integral over two sets of coordinates, these may add significant cost to the calculations. For example, the range-separated GGA hybrid ω B97X-V functional used throughout this text, incorporates the VV10 non-local correlation functional and is 3-5 times more costly than the parent ω B97X functional. An alternative and less costly way to account for dispersive interactions is to add an empirical correction to the functional. So-called Grimme corrections are designated by

appending “-D” as in ω B97X-D, or “D3” as in B3LYP-D3, to the end of the functional.

A graphical summary of a number of commonly-used functionals stemming from different classes is provided in the figure below.



Which classes of functionals are “best” and which functionals within each class are “best”? A broad base of experience allows some general remarks to be made.

Numerical Integration

Unlike both Hartree-Fock and wave function based correlated models, density functional models cannot be evaluated wholly analytically. Some components require numerical integration, which introduces another variable into the mix, specifically the form and size of the integration grid. An oversimplified description places the grid points along a set number of angular directions and at a set number of distances from the origin. This is referred to as a *Lebedev grid*.

“Cost” of Density Functional Models

The computation “cost” of density functional models depends on the class, the number of basis functions, η , and the number of points in the numerical integration grid, κ , where $\kappa \gg \eta$. GGA and pure meta-GGA functionals such as B97M-V that do not require the Hartree-Fock exchange formally scale as $O(\eta^2\kappa)$. The other functionals combine this dependence with the cost of the Hartree-Fock exchange, which formally scales as $O(\eta^4)$ but in practice is $O(\eta^3)$ or lower. Finally, functionals such as ω B97X-V that directly account for dispersion have a step that scales as $O(\eta^2\kappa^2)$. This typically dominates the calculation. Times (in minutes) for calculation of the energy together with its gradient (“one step” in the optimization of molecular geometry) for morphine (C₁₇H₁₉NO₃) with the B3LYP (GH-GGA), ω B97X-D (RSH-GGA), ω B97X-V (RSH-GGA+dispersion), B97M-V (mGGA), M06-2X (GH-mGGA) and M11 and ω B97M-V (RSH-mGGA) functionals, as well as with RI-MP2, which formally scales as $O(\eta^5)$, with the 6-31G* basis set are provided in **Table A-1**. These have been obtained using a single core of a 3.3 GHz Intel i7 5820K processor. **Table A-2** provides times for 6-311+G(2df,2p) calculation, a dual basis set 6-311+G(2df,2p) calculation and a dual basis set cc-pVQZ-g (cc-pVQZ minus the set of g functions) energy calculation, relative to the corresponding 6-31G* energy/gradient calculations. For example, the value of 26 for the dual basis set cc-pVQZ-g ω B97X-D energy calculation given in **Table A-2**, means that this calculation is 26 times more costly than the ω B97X-D energy/gradient calculation. Put another way, a structure optimization requiring 26 cycles would cost the same as an energy obtained using the larger basis set.

The manner of reporting is deliberate. The 6-31G* basis set is usually deemed satisfactory for calculation of equilibrium geometry, whereas larger basis sets (for example, 6-311+G(2df,2p) and cc-pVQZ-g) are known to be required for accurate descriptions of the energies of most chemical reactions. Determination of geometry for molecules of the complexity of morphine typically requires upwards of 10-20 energy/gradient cycles, that is, an order of magnitude or more greater than the time reported in the first column of the table. Both parts must be considered in order to correctly assess cost and practicality.

Table A-1: Times for energy and gradient calculations on morphine with B3LYP, ω B97X-D, ω B97X-V, B97M-V, M06-2X, M11 and ω B97M-V density functional models and RI-MP2 model with the 6-31G* basis set (need ω B97M-V times)

	6-31G* + gradient
B3LYP	7
ω B97X-D	15
ω B97X-V	41
B97M-V	50
M06-2X	19
M11	22
RI-MP2	10

Table A-2: Times for B3LYP, ω B97X-D, ω B97X-V, B97M-V, M06-2X, M11 and ω B97M-V density functional and RI-MP2 energy calculations with 6-311+G(2df,2p), dual 6-311+G(2df,2p) and dual cc-pVQZ-g basis sets relative to the corresponding times for energy and gradient calculations with the 6-31G* basis set give in Table A-1 (need ω B97M-V times)

	6-311G(2df,2p)	6-311+G(2df,2p) [6-311G*]	cc-pVQZ-g [rcc-pVQZ]
B3LYP	26	5	48
ω B97X-D	14	3	26
ω B97X-V	6	2	10
B97M-V	5	2	6
M06-2X	13	3	21
M11	11	3	20
RI-MP2	18	3	33

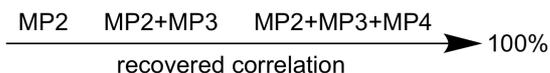
Several conclusions may be drawn:

1. Times for calculation of the energy together with its gradient vary by close to an order of magnitude among the seven models. The B3LYP model is the least costly and the ω B97X-V, B97M-V and ω B97M-V models most costly. The latter result suggests that calculation of the gradient of dispersion functional common to both models is to blame. In terms of cost, the ω B97X-D model is closest to B3LYP. The surprising result is perhaps the good performance of the RI-MP2 model, surpassed only by B3LYP.

2. 6-311+G(2df,2p) energy calculations are roughly an order of magnitude more costly than the corresponding energy/gradient calculations. At the high end and low ends is the factor of 26 for the B3LYP model and a factor of 5 for the B97M-V model. Structure optimization typically requires upwards of 20 steps (energy/gradient), and will likely dominate the overall task.
3. The dual basis set approximation applied to 6-311+G(2df,2p) basis set leads to significant cost savings. If this is the chosen energy method, the costly step will most certainly be determination of geometry.
4. Except for the ω B97X-V and B97M-V functionals, dual basis set cc-pVQZ-g calculations are roughly an order of magnitude more costly than the corresponding 6-311+G(2df,2p) calculations. This means that geometry and energy steps are likely to contribute equally to overall cost. The much smaller increase in cost for B97M-V is presumably due to its lack of the Hartree-Fock exchange.
5. The RI-MP2 model is clearly competitive, greater in overall cost only to the B3LYP models. Because of its inherent $O(\eta^5)$ scaling, RI-MP2 will eventually fall significantly behind density functional models in performance as molecular size is increased.

Disadvantages of Density Functional Models

Low computation cost relative to wave function based electron correlation models is the principal attraction of density functional models. Aside from the RI-MP2 model, density functional models are the only procedures that are both reliable and routinely applicable to molecules of moderate size. However, density functional models come with some disadvantages not shared by wave function based procedures. By far the most serious of these is that they do not offer a clear pathway to improvement. Whereas wave function based procedures return a higher and higher percentage of the correlation energy with increasing degree of complexity, for example, in successively adding orders in the Møller-Plesset models,

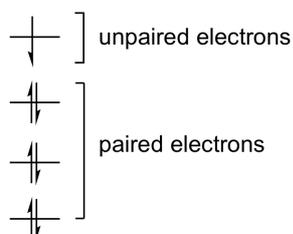


there is no guarantee that increasing the flexibility (and complexity) of a particular functional will actually lead to improvement. It is clear that present generation functionals are “better” than those of previous generations, for example, ω B97X-D and M06-2X are significantly better than B3LYP, in particular for calculation of reaction energies. What is less clear is whether the improvement is to large extent due to careful parameterization to high quality reference data (and to an increased number of parameters) rather than from insight into the “proper” functional form. Functional selection, even while guided by clear physical models, certainly has an uncertain component, and at least at present it is not possible to say from first principles which of several functionals is likely to provide the “best” results. This can only be done by way of thorough comparisons with “known” data.

In practice, density functional models increase in computational cost as the cube of the total number of basis functions (the same dependence seen for Hartree-Fock models). Because most functionals require calculation of the Hartree-Fock exchange energy, they are necessarily more costly than Hartree-Fock models, but can easily be applied to molecules incorporating up to 100 heavy (non-hydrogen) atoms.

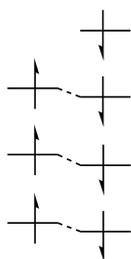
Models for Open-Shell Molecules

Thus far, discussion has been limited to molecules with closed-shell electron configurations, that is, with all electrons being paired. This covers the vast majority of organic molecules as well as most organometallic molecules. There are two ways to think about molecules with unpaired electrons. The obvious way is to insist that electrons are either paired or are unpaired.



This is referred to as *restricted* and the individual models as restricted models, for example, the restricted Hartree-Fock (or RHF) model.

The restricted procedure does not necessarily yield the lowest possible energy, simply because it forces the “paired” electron into the same spatial orbitals. Removing this constraint provides greater flexibility and generally lead to a lower energy. This is termed *unrestricted* and the resulting models are termed unrestricted models, for example, the unrestricted Hartree-Fock (or UHF) model.



Aside from yielding lower energy, unrestricted models are generally less costly than restricted models, and are much more widely used and is the default procedure in *Spartan*.

Semi-Empirical Molecular Orbital Models

The principal disadvantage of Hartree-Fock, density functional and MP2 models is their computational cost. It is possible to introduce further approximations together with empirical parameters in order to significantly reduce cost while still retaining the underlying quantum mechanical formalism. So-called *semi-empirical molecular orbital models* follow in a straightforward way from Hartree-Fock models:

1. Insist that basis functions on different atoms do not overlap (“see each other”). The so-called *NDDO approximation* is rather drastic but reduces the computation effort by more than an order of magnitude over Hartree-Fock models.
2. Restrict to a *minimal valence basis set* of atomic functions. This means that there are no inner-shell (core) functions in the basis set. As a consequence, the cost of doing a calculation involving a second-row element such as silicon, is no more than that incurred for the corresponding first-row element such as carbon.

3. Introduce adjustable parameters to reproduce specific experimental data. This is what distinguishes the various semi-empirical models currently available. Choice of parameters, more than anything else, appears to be the key to formulating successful semi-empirical models.

Molecular Mechanics Models

The alternative to quantum chemical models are molecular mechanics models. These do not start from the Schrödinger equation, but rather from a simple but chemically reasonable picture of molecular structure, a so-called *force field*. In this picture, just as with a Lewis structure, molecules are made up of atoms (as opposed to nuclei and electrons), some of which are connected (bonded). Both crowding (van der Waals) and charge-charge (Coulombic) interactions between atoms are then considered, and atom positions are adjusted to best match known structural data (bond lengths and angles).

Molecular mechanics is much simpler than attempts at solving an “approximate” Schrödinger equation, but requires an explicit description of chemical bonding, as well as a large amount of information about the structures of molecules. This biases results and seriously limits the predictive value of molecular mechanics models. Nevertheless, molecular mechanics has found an important role in molecular modeling as a tool to establish equilibrium geometries of proteins and other large molecules. It has also been widely used to establish the preferred conformation of molecules with multiple degrees of conformational freedom and hundreds or even thousands of accessible conformers. With regard to the latter, comparisons with results of high-level quantum chemical calculations have, however, clearly shown that present-generation molecular mechanics models are often not satisfactory in identifying the “best” (lowest-energy) conformer and are not to be trusted for obtaining Boltzmann weighed averages. However, they have proven to be useful for identifying highly unfavorable conformers allowing some shortening of the initial list of possible structures.

Choosing a Theoretical Model

No single method of calculation is likely to be ideal for all applications. A great deal of effort has been expended to define the limits of different molecular mechanics and quantum chemical models, and to judge the degree of success of different models. The latter follows from the ability of a model to consistently reproduce known (experimental) data. Molecular mechanics models are restricted to determination of geometries and conformations of stable molecules. Quantum chemical models also provide energy data, which may in turn be directly compared with experimental thermochemical data, as well as infrared, Raman, UV/visible and NMR spectra and properties such as dipole moments, which may be compared directly with the corresponding experimental quantities. Quantum chemical models may also be applied to transition states. While there are no experimental structures with which to compare (see the topic *Potential Energy Surfaces*), experimental kinetic data may be interpreted to provide information about activation energies (see the topic *Total Energies and Thermodynamic and Kinetic Data*).

Success is not an absolute. Different properties, and certainly different problems may require different levels of confidence to be placed in the calculation to actually be of value. Neither is success sufficient. A model also needs to be practical for the task at hand. Were this not the case, there would be no reason to look further than the Schrödinger equation itself. Models that may be practical for small to medium size organic molecules cannot be expected to be applied to proteins. Models that are successful and practical for organic molecules may not necessarily meet either criterion for inorganic molecules or transition-metal organometallics. Not only does the nature and size of the system needs to be taken into account, with due attention to the available computational resources and the experience (and patience) of the practitioner. Specifics aside, practical models usually share one common feature in that they are not likely to be the best possible treatments which have been formulated. Compromise is almost always an essential component of model selection. Continued advances in both digital computers and computer software will continue to raise the bar, but it will be some time before fully reliable models will be

routinely applicable to all chemical systems of interest.

The MMFF molecular mechanics model generally provides a satisfactory description of equilibrium geometry where, at least for organic molecules. It has also proven to be suitable for removing high-energy conformers for molecules with multiple degrees of freedom in preference to quantum chemical calculations of conformer energy differences.

Semi-empirical models are appropriate for:

- i) Equilibrium geometry determinations for large molecules.
- ii) Transition-state geometry determinations.
- iii) Equilibrium and transition-state geometry determinations involving transition metals.

Semi-empirical models are unsuitable for:

- i) Calculation of reaction energies.
- ii) Calculation of conformer energy differences.

Small basis set Hartree-Fock models such as HF/3-21G and HF/6-31G* are appropriate for:

- i) Equilibrium and transition-state structure determinations of organic molecules.
- ii) Calculation of energies of *isodesmic* reactions, including comparisons of regio and stereoisomers.

They are unsuitable for:

- i) Calculation of reaction energies that involve bond making or breaking and calculation of absolute activation energies.
- ii) Comparison of isomer energies of molecules with different bond types.
- iii) Equilibrium and transition-state structure determinations for transition-metal organometallic molecules.
- iv) Calculation of conformer energy differences.