Spartan '08

for Windows, Macintosh and Linux

Tutorial
and
User's Guide

Wavefunction, Inc.
18401 Von Karman Avenue, Suite 370
Irvine, CA 92612 U.S.A.
www.wavefun.com

Wavefunction, Inc., Japan Branch Office
3-5-2, Kouji-machi, Suite 608
Chiyoda-ku, Tokyo, Japan 102-0083
japan@wavefun.com • www.wavefun.com/japan

Spartan'08 is a collaboration with Q-Chem, Inc.

Q-CHEM, INC.
Four Triangle Drive, Suite 160, Export, PA 15632

ISBN978-1-890661-38-4

Printed in the United States of America

Acknowledgements

The ***Spartan'08*** Tutorial and User's Guide was prepared by Warren Hehre and Sean Ohlinger, who take both credit for its utility and blame for its limitations and inaccuracies. Several members of Wavefuncton and Q-Chem contributed significantly, with Philip Klunzinger, Bernard Deppmeier, Andy Driessen and Jeffrey Johnson warranting special mention. As with past manuals, sincere thanks goes to Pamela Ohsan for turning a sloppy manuscript into a finished book.

Scope of this Guide

This guide provides a general reference for ***Spartan'08*** for Windows, Macintosh and Linux and the Essential Edition of **Spartan'08** for Windows and Macintosh. It is divided into 19 chapters grouped into three sections, along with several appendices.

Section I (**Introduction**, **Chapter 1**) introduces ***Spartan*** as a tool for exploring organic, bioorganic, inorganic and organometallic chemistry by way of molecular mechanics and quantum chemical calculations, together with an array of graphical models for conveying the results of these calculations. The section ends with a brief discussion of new capabilities in ***Spartan'08*** as well as improvements over previous versions.

Section II (**Getting Started**, **Chapters 2** to **9**) describes the overall operating environment of ***Spartan***, and then provides an extensive set of "hands-on" tutorials. Most of the tutorials can be completed with the Essential Edition of ***Spartan***; sections which can not be completed are marked as such. This section is the place to start for new users of the program, but should also be examined by users of previous versions of ***Spartan***.

Section III (**Features and Functions**, **Chapters 10** to **19**) describes in detail the functions available from the menus and dialogs incorporated into the graphical user interface for ***Spartan***. The focus is on graphical input and manipulation of structure, input of other required information and text, spectral and graphical output resulting from molecular mechanics and quantum chemical calculations. This section is intended as a general reference to ***Spartan'08***.

What this guide ***does not do*** is document the performance and cost (in computation time) of the different molecular mechanics and quantum chemical models available in ***Spartan'08***, or recommend specific models or combinations of models for use on chemical problems. Nor does it show the utility of graphical models in presenting and interpreting the results of the calculations. These topics are covered in

depth in ***A Guide to Molecular Mechanics and Quantum Chemical Calculations*** included as a PDF under the **Help** menu and available from Wavefunction as a hardbound volume. The guide also provides a collection of illustrative examples.

Appendices provide an overview of the program's overall architecture as well as its present capabilities and limitations (**A**), a directory of functions under its menus (**B**), a listing of commonly-used options (**C**), a listing of units (**D**), the proper citation for the program (**E**), instructions for installing the Cambridge Structural Database (**F**), directions for making databases from ***Spartan*** calculations (**G**), examples of pharmacophore input (**H**), input of experimental infrared, UV/visible and NMR spectra (**I**) and directions for installing a network HASP (**J**). Additional materials relating to several of these appendices may be found as PDFs under the **Help** menu.

An up-to-date version of this ***Tutorial and User's Guide*** is available on Wavefunction's website (directly accessible from the **Help** menu).

Table of Contents

Section I

Introduction

Molecular mechanics calculations and quantum chemical calculations play an ever-increasing role in modern chemistry. Primary functions are to supply information about structures, relative stabilities and other properties of isolated molecules. Because of their inherent simplicity, molecular mechanics calculations on complex molecules are widespread throughout the chemical community. Quantum chemical calculations, including Hartree-Fock molecular orbital calculations, but especially calculations that take account of electron correlation, are much more time demanding. Only recently, have fast enough computers become widely available to make their application routine among mainstream chemists.

Quantum chemical calculations may also be called upon to furnish information about the mechanisms and product distributions of chemical reactions, either directly by calculations on transition states, or based on the Hammond Postulate, by modeling the steric and electronic demands of the reactants. Quantitative calculations, leading directly to information about the geometries of transition states, and about reaction mechanisms in general, are becoming more and more common, while qualitative models are still needed for systems that are too large to be subjected to the more rigorous treatments. Finally, quantum chemical calculations can supply information to complement existing experimental data or to replace it altogether, for example, atomic charges for QSAR analyses, and intermolecular potentials for molecular mechanics and molecular dynamics calculations.

Spartan'08 ("***Spartan***") has been designed to address the ever increasing role that calculations play in chemistry and related fields. It represents a continued collaboration between Wavefunction, Inc., and Q-Chem, Inc. Q-Chem codes supplement and extend the traditional strengths of ***Spartan*** as an easy to learn and use tool for molecular mechanics, semi-empirical and Hartree-Fock molecular orbital

calculations, as well as a wide variety of graphical models, with a full range of density functional models and a selection of wavefunction based important post-Hartree-Fock models. All models have been implemented using what we believe are the most robust algorithms currently available, and have been tuned for high performance on Intel and AMD processors including multi-core processors.

Spartan is intended to be utilized by chemists, not only computational chemists who are already familiar with the capabilities of molecular mechanics and quantum chemical methods, but also experimental chemists who may have little or no prior experience, but who want to use calculations much in the same way as experimental techniques such as NMR spectroscopy. This ambitious goal is directly reflected in the program's overall design criteria: "convenient access to a full range of modern molecular mechanics and quantum chemical models", and clearly distinguishes ***Spartan*** from other molecular modeling packages.

Spartan is supported under Windows (2000, XP, XP 64 and Vista), Macintosh (OS 10.4.x and higher), Linux (Linux Kernel 2.6 or later). The Linux version (only) may be used in a server mode with Windows, Macintosh or Linux acting as a front end, and can also perform calculations locally.

New features in ***Spartan'08*** include the following:

1. **Cross-Platform.** ***Spartan*** has been implemented under Qt and is now a fully cross-platform (Windows, Macintosh and Linux) application. The subtle differences that do exist are due to different management software of the three operating systems.

2. **Multi-Core Parallel.** The Hartree-Fock and density functional codes have been adopted for use of multi-core (shared-memory) machines. Speed improvements of approximately 1.9 for a two-core machine and 3.2 for a four-core machine are typically realized for energy calculation and equilibrium and transition state geometry optimization, for both Hartree-Fock and density functional methods. Vibrational frequency calculations show a lesser speed enhancement and NMR chemical shift

calculations have not been parallelized. Molecular mechanics, semi-empirical and wavefunction-based correlated calculations have not been parallelized. ***Spartan's*** parallel code is available as an option to Windows and Macintosh users. All ***Spartan'08*** for Linux licenses include the parallel option.

3. **Infrared Spectra Matching.** A tool for matching calculated and experimental infrared spectra has been provided and a database of ~40,000 spectra obtained from the EDF2/6-31G* model developed. (A subset of this database is provided with ***Spartan'08***, the full database may be licensed.) The calculated spectra are adjusted "on-the-fly" to account for systematic errors in frequencies and for line width.

4. **Percentage Enclosure of Electron Density.** Electron density surfaces can now be based on the percentage of total electrons in addition to a constant isovalue.

5. **Discrete Property Maps.** Property maps can now be displayed in terms of discrete property values in addition to a continuous range.

6. **Corrected ^{13}C Chemical Shifts.** Carbon chemical shifts obtained from the B3LYP/6-31G* model may now be empirically corrected according to the number and kinds of directly-bonded atoms. The resulting shifts are now (in the mean) within 2 ppm of experimental values compared to 4 ppm for uncorrected shifts.

7. **Solvation Models.** The empirical SM8 solvation model has been implemented for use with Hartree-Fock and density functional methods. This not only affects the energy but also the wavefunction, meaning that equilibrium and transition-state geometries may be obtained in solvent and solvent effects on calculated infrared spectra, properties and graphical displays may be evaluated. SM8 has been parameterized for water as well as for a variety of important organic solvents.

8. **Next-Generation Functionals and User-Defined Functionals.** A variety of new functionals have been provided. These represent

a next generation of functional development in that they address some of the major problems associated with commonly-used functionals such as B3LYP. Particular emphasis ("menu status") has been given to the M06 and ωB97X-D functionals. A facility allowing users to define combinations of exchange and correlation components has been provided.

9. **HH Coupling Constants.** An empirical procedure for estimating HH coupling constants has been implemented. Proton spectra that incorporate coupling constants may be displayed and a tool for "spectra magnification" has been provided. In addition, COSY spectra (based on empirical coupling constraints) may be displayed.

10. **RI-CIS(D).** The RI-CIS(D) method for excited states has been implemented. This is significantly faster than CIS(D) both for energies and geometries but provides nearly identical results.

11. **Improved T1.** The T1 thermochemical recipe has been improved and a database of ~40,000 molecules from T1 has been developed. (A subset of this database is provided with ***Spartan'08*** and the full database may be licensed.)

12. **QuickTime.** QuickTime movies may now be made based on manipulations in ***Spartan's*** main screen. (Not available with ***Spartan'08*** for Linux.)

13. **Reaction Energies.** A reaction energy calculator has been provided. This may either operate using molecules in the selected document or from entries in the Spartan Molecular Database corresponding to molecules in the selected document. The latter allows reactants and products to be "substituted" molecules, that is, they do not need to be explicitly specified.

14. **Geometries with Dual-Basis-Sets.** Dual-basis-set calculations have been extended to include calculation of equilibrium and transition-state geometries.

15. **Spartan Molecular Database.** The Spartan Molecular Database has increased in overall size by 40%. A significant number of

entries with the 6-311+G** basis set for Hartree-Fock, B3LYP and MP2 models as well as ~1,500 molecules from the G3(MP2) recipe have been added. (A subset of the full database is provided with ***Spartan'08***, the full database may be licensed separately.)

16. **Spartan Reaction Database.** The Spartan Reaction Database has been extended and now includes data for ~2,000 organic and organometallic reactions. The full reaction database is supplied with ***Spartan'08***.

17. **Input of Experimental Spectra.** Experimental infrared, UV/visible and NMR spectra may now be directly read into ***Spartan'08*** in CML and JCAMP formats.

18. **Cambridge Structural Database for Macintosh.** The Cambridge Structural Database (CSD) may now be accessed from ***Spartan*** on the Macintosh (in addition to Windows and Linux). The CSD must be licensed from the Cambridge Crystallographic Data Centre.

19. **Use of Substituents in Searches of the Spartan Molecular Database.** Substituents specified in ***Spartan's*** substituent builder may now be used in queries of the Spartan Molecular Database.

20. **Embedded Files.** External files, such as Word, Excel and PowerPoint files, may be embedded into ***Spartan*** documents either at the document level and/or at the individual molecule level.

Since the release of ***Spartan'06***, ***Spartan*** users have diligently reported program bugs via ***support@wavefun.com***. All reported bugs have been addressed. Wavefunction would like to thank our customers for their help in making ***Spartan*** a better program. As ***Spartan'08*** has been completely rewritten with the Qt-crossplatform tool kit, customer input will remain an important part of ***Spartan's*** continued development. User-feedback is greatly appreciated!

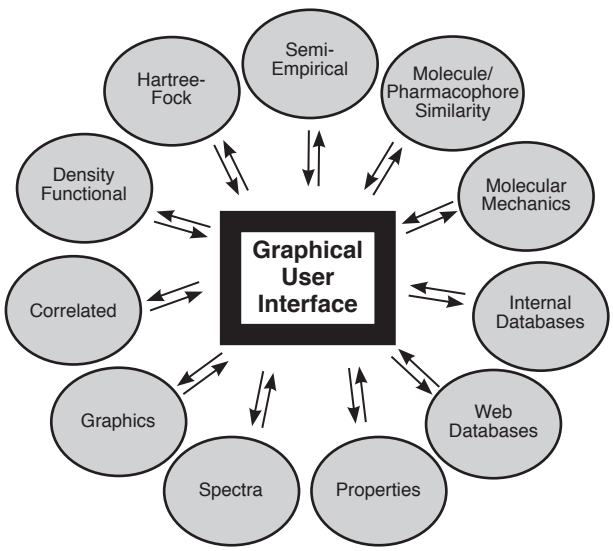
Chapter 1

Spartan'08

*This chapter describes the architecture of **Spartan'08**, focusing on the connectivity of graphics, database and computational components. Available molecular mechanics and quantum chemical methods are enumerated, and their overall role suggested.*

Spartan'08 ("***Spartan***") is a cross-platform application. With very few exceptions, ***Spartan's*** operations and features are identical between Windows, Mac and Linux versions. The following nomenclature will be used throughout this manual to indicate the few cases where features are not available: **NAL**; Not Available Linux, **NAM**; Not Available Macintosh and **NAW**; Not Available Windows.

Spartan comprises a series of independent modules tightly connected via a graphical user interface that is highly functional, yet elementary and uncluttered. It has been designed not only to greatly reduce the drudgery and possibility for human error associated with the preparation of input, but also to guide the interpretation of output. The interface is perhaps best viewed as an interactive and intuitive window into a full range of modern computational techniques.



Included in the interface are builders for organic, inorganic and organometallic molecules, polypeptides and polynucleotides, as well as for creating substituted molecules, and a procedure for guessing transition states. Access to ChemDraw[1] is provided without having to exit ***Spartan's*** interface. Also available is the Spartan Molecular Database (SMD), a collection of over 150,000 calculated structures and associated properties of molecules, each obtained from as many as ten different theoretical models (Hartree-Fock models with 3-21G*, 6-31G* and 6-311+G** basis sets, B3LYP and MP2 models with the 6-31G* and 6-311+G** basis sets, the EDF1 model with the 6-31G* basis set, the G3(MP2) model and the T1 model)[2], as well as a collection of approximately 1,200 small molecules bound to proteins and nucleotides from the PDB[3]. New to ***Spartan'08*** is the Spartan Infrared Database (SIRD), a collection of approximately 40,000 organic molecules, the infrared spectra of which have been obtained from the EDF2/6-31G* model[2] and which may be searched by matching to an input (experimental) spectrum. The Spartan Reaction Database (SRD) has been significantly enhanced and now contains transition states for approximately 2,000 organic and organometallic reactions each obtained at one or more theoretical models. It may be accessed either by substructure searching (as SMD) or automatically by using "arrows" attached to reactants.

Also provided is an interface to the Cambridge Structural Database, comprising over 450,000 experimental X-ray crystal structures[4], and automatic retrieval from the PDB database[3] of protein and related crystal structures. Finally, ***Spartan*** allows for import of a variety of data files, most important among them being PDB (protein structures) and SDF (structure drawings). Thus, it is not only possible to construct molecules from scratch, but also to access existing structures.

Spartan's interface provides the gateway to a range of modern computational methods, including molecular mechanics models, semi-empirical and Hartree-Fock molecular orbital models, and a variety of correlated models including density functional and Møller-Plesset models. None of these models is likely to be ideal for every application[5]. While the most sophisticated quantum chemical models

may yield excellent results, they will likely be too time consuming for routine application, and it will usually be necessary to contend with lesser treatments. ***Spartan's*** interface facilitates mixing and matching different molecular mechanics and quantum-chemical models. Results from one model may easily be passed on for further analysis with another (more rigorous) model.

The simplest computational methods in ***Spartan*** are molecular mechanics models using the SYBYL and MMFF force fields, with aqueous energy corrections available for the latter (MMFFaq). These are available to determine equilibrium geometries and conformations of molecules comprising upwards of several thousand atoms. These are the only computational techniques that are routinely applicable to biopolymers. MMFF has been shown to provide reasonable assignments of low-energy conformers and has been used to construct libraries of diverse conformers[6].

Quantum chemical models are required to account for the geometries of transition states as well as for reaction and activation energies. The simplest of these, semi-empirical molecular orbital models, can be routinely applied to systems with a few hundred atoms. Supported in ***Spartan*** are the MNDO model (with extensions for second-row and heavier main-group elements), the AM1 model, the PM3 model (with parameters for transition metals) and the RM1 model (a reparameterization of AM1). While semi-empirical models have proven to be quite successful for determining equilibrium geometries including the geometries of transition-metal compounds, they have not proven to be successful for the calculation of the reaction energies.

Hartree-Fock molecular orbital models are a mainstay of quantum chemical techniques, in particular, for equilibrium and transition-state structure determination, and for thermochemical comparisons. The models available in ***Spartan*** may be routinely applied to molecules with up to one hundred atoms. Hartree-Fock models are, however, ***not adequate*** for thermochemical comparisons where bonds are broken or formed, nor do they provide a proper account of the geometries of molecules incorporating transition metals. So-called correlated models are required. Several classes of correlated models are available in

Spartan[7]: density functional models, Møller-Plesset models, coupled cluster models and quadratic configuration interaction (QCI) models. ***Spartan'08*** supports a wider variety of functionals than available in previous versions of ***Spartan***, and in addition, allows users to make their own functionals by combining exchange and correlation components. Density functional models may be routinely applied to molecules of the same size as with Hartree-Fock models. On the other hand, the simplest Møller-Plesset model (MP2) is limited to molecules comprising thirty or fewer atoms. The recently introduced RI-MP2 model is nearly an order of magnitude faster than MP2 for energy calculations and 2-3 times faster for equilibrium and transition-state geometry calculations, but provides nearly identical results.

Higher order (MP3 and MP4) Møller-Plesset, coupled-cluster and QCI models are available for energy calculations only. They are limited in practice to molecules with less than 10-20 heavy (non-hydrogen) atoms.

Hartree-Fock and correlated models may be used with a variety of all-electron Gaussian basis sets. Hartree-Fock models are often quite successful with split-valence basis sets, while correlated models require polarization basis sets, and in some cases basis sets incorporating diffuse functions. Dual-basis set calculations, which approximate changes in energy and wavefunction resulting from extension from a small basis set to a large basis set, are also available. They can lead to an order of magnitude performance enhancement over calculations that directly use large basis sets. New in ***Spartan'08***, is support for gradient calculations with dual basis sets. Pseudopotentials are provided for use with both Hartree-Fock and correlated models for calculations on molecules incorporating heavy elements for which all-electron bases sets are not available.[7]

A series of recipes (combinations of different methods) are available to estimate heats of formation[7]. G2, G3 and G3 (MP2) recipes, while applicable only to very small systems, have been shown to yield heats of formation in excellent accord with experiment, and are benchmarks against which other calculations may be judged. The T1 recipe closely reproduces heats of formation from the G3 (MP2) recipe but requires 2-3 orders of magnitude less computation time.

Spartan provides singles configuration interaction (CIS) techniques[7] for energies, equilibrium and transition-state geometry and conformation calculations on molecules in electronically-excited states. In addition, energy calculations (only) on excited states with the CIS (D) and RI-CIS(D) models as well as a range of density functional models may be performed.

Spartan provides access to several common spectral quantities, in particular infrared spectra (molecular mechanics, semi-empirical, Hartree-Fock, density functional and MP2 models), NMR spectra[7,8] (Hartree-Fock and density functional models) and UV/visible spectra[7] (CIS, CIS(D) and density functional models). These are available both as numerical data (vibrational frequencies, chemical shifts, etc.) as well as spectral plots. Spectra may be directly input to ***Spartan*** and internet connections to IR, NMR and UV/visible databases are available allowing direct comparison of calculated and experimental spectra. New to ***Spartan'08*** are an "on-the-fly" procedure allowing fitting of a calculated infrared spectrum to an experimental spectrum, using both a multiplicative scale of calculated frequencies and peak width at half height as parameters, and an empirical correction to ^{13}C chemical shifts obtained from the B3LYP/6-31G* model.

New to ***Spartan'08*** is its SM8 solvation model[9] for water and a selection of important organic solvents. This alters the wavefunction in addition to the energy, meaning that solvent effects on geometries, infrared spectra and molecular properties may be examined. Also supported is the SM 5.4 model which affects only the energy.

Spartan provides tools to quantify both the similarity of molecular structures and of chemical environments characteristic of these structures. Also available is the ability to identify molecules that fit into a specific chemical environment (a so-called pharmacophore). Finally, pharmacophores may be extracted from small molecules bound to proteins and nucleotides in the PDB.

Spartan provides a variety of graphical tools to assist in interpreting the results of calculations as well as similarity analyses. These include structure models, and also molecular orbitals, electron and spin densities, local ionization potentials and electrostatic potentials

that can be displayed as surfaces, slices and property maps. ***Spartan*** provides the ability to distinguish accessible and inaccessible regions on a density surface and on property maps based on this surface. New in ***Spartan'08*** are displays where discrete bands replace continuous colors and density surfaces the size of which are based on percentage enclosure of the total electron density.

Graphical models may be manipulated in real time, and as many as desired can be displayed simultaneously. Animation files may be made, and can be used to depict conformational changes or chemical reactions. New to ***Spartan'08*** is the ability to make QuickTime movies from on-screen manipulations (**NAL**).

---

1. ChemDraw is not included with ***Spartan***, but may be obtained from CambridgeSoft (www.cambridgesoft.com). **NAL**, **NAM**. While all versions of ***Spartan*** are able to read ChemDraw files, seamless access is presently available only in the Windows version.
2. Subsets of SMD and SIRD are provided with ***Spartan***. The full databases may be licensed separately and are provided under maintenance contract.
3. PDB web reference: http://www.rcsb.org.
4. The Cambridge Structural Database is not included with ***Spartan***, but is available by subscription from the Cambridge Crystallographic Data Centre (www.ccdc.cam.ac.uk).
5. Full discussion and assessment of the specific molecular mechanics and quantum chemical models available in ***Spartan*** is provided in: W.J. Hehre, ***A Guide to Molecular Mechanics and Quantum Chemical Calculations***, Wavefunction, Irvine, 2003. This is available as a PDF under the **Help** menu. See also: W.J. Hehre, L. Radom, P.v.R. Schleyer and J.A. Pople, ***Ab Initio Molecular Orbital Theory***, Wiley, New York, 1986; Y. Shao, L.F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S.T. Brown, A.T.B. Gilbert, L.V. Slipchenko, S.V. Levchenko, D.P. O'Neill, R.A. DiStasio Jr., R.C. Lochan, T. Wang, G.J.O. Beran, N.A. Besley, J.M. Herbert, C.Y. Lin, T. Van Voorhis, S.H. Chien, A. Sodt, R.P. Steele, V.A. Rassolov, P.E. Maslen, P.P. Korambath, R.D. Adamson, B. Austin, J. Baker, E.F.C. Byrd, H. Dachsel, R.J. Doerksen, A. Dreuw, B.D., Dunietz, A.D. Dutoi, T.R. Furlani, S.R. Gwaltney, A. Heyden, S. Hirata, C-P. Hsu, G. Kedziora, R.Z. Khalliulin, P. Klunzinger, A.M. Lee, M.S. Lee, W.Z. Liang, I. Lotan, N. Nair, B. Peters, E.I. Proynov, P.A. Pieniazek, Y.M. Rhee, J. Ritchie, E. Rosta, C.D. Sherrill, A.C. Simmonett, J.E. Subotnik, H.L. Woodcock III, W. Zhang, A.T. Bell, A.K. Chakraborty, D.M. Chipman, F.J. Keil, A.Warshel, W.J. Hehre, H.F. Schaefer, J. Kong, A.I. Krylov, P.M.W. Gill and M. Head-Gordon, *Phys. Chem. Chem. Phys.,* **8**, 3172 (2006).
6. Databases comprising sets of diverse conformers pharmaceutical candidates as well as Maybridge and Life Chemicals collections may be licensed from Wavefunction.
7. Not available in the Essential Edition.
8. Chemical shifts only. HH coupling constants may be evaluated empirically.
9. A.V. Maronich, R.M. Olson, C.P. Kelly, C.J. Cramer and D.G. Truhlar, *J. Chem. Theory Comput.*, **3**, 2011 (2007).

Section II

Getting Started

The chapters in this section provide a brief introduction to the graphical user interface for ***Spartan'08***, and following this, a series of tutorials. The objective of the latter is to provide hands on experience, and in doing so, illustrate the way in which molecular mechanics, quantum chemical calculations and similarity analyses may be set up, performed and interpreted. This is the place to start for new users. Users of previous versions of ***Spartan*** are also advised to review this section.

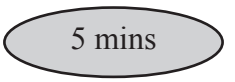
No attempt has been made to illustrate the full range of ***Spartan's*** capabilities. For this, the reader should consult **Section III** of this guide. Focus is on the use of ***Spartan*** to calculate equilibrium and transition-state geometries, to search conformation space, to align molecules and evaluate the similarities of molecules, to evaluate reaction thermochemistry and activation energetics and to obtain NMR, infrared and UV/visible spectra with particular emphasis on NMR. Likewise, no attempt has been made to assess the performance of the various molecular mechanics and quantum chemical models. The majority of tutorials make use of only molecular mechanics models, semi-empirical and Hartree-Fock molecular orbital models[1]. In some cases these may not lead to acceptable results. A thorough assessment of the most commonly used models in ***Spartan*** is available.[2]

Graphical models are illustrated in several of the tutorials that follow, and connections between specific models and chemical observables are pointed out, for example, between electrostatic potential maps and electrophilic character. A more complete account is provided in the aforementioned: ***A Guide to Molecular Mechanics and Quantum Chemical Calculations***.

Operating ***Spartan*** **(Chapter 2)** outlines the basic workings of ***Spartan's*** graphical user interface. It should be reviewed before starting the tutorials and referred to as needed. The tutorials that make up the bulk of this section are divided across seven chapters: ***Basic***

Operations, ***Organic Molecules***, ***Groups of Organic Molecules***, ***Organic Reactions***, ***Medicinal Chemistry, Polypeptides to Proteins*** and ***Inorganic/Organometallic Molecules***. The initial three cover a number of important general features and should be completed first, and in order. The remaining chapters cover more specialized topics.

Estimates for required computer time (in minutes, assuming a 2GHz mobile Pentium) are indicated at the top of the first page of each tutorial. Users with dual and quad-core processors and who access ***Spartan's*** parallel Hartree-Fock and density functional codes can expect significant reductions in these times.



Where two times are provided, the second corresponds to the total time including the optional parts of the tutorial.

1. This is a deliberate decision, as it minimizes the time required to complete the tutorials and allows nearly all tutorials to be completed using only the Essential Edition of ***Spartan*** (that lacks density functional, MP2 and other correlated calculations). Tutorials or sections of tutorials that cannot be completed with the essential edition are indicated as such.
2. W.J. Hehre, ***A Guide to Molecular Mechanics and Quantum Chemical Calculations***, Wavefunction, Irvine, CA, 2003. This is available as a PDF under the **Help** menu.

Chapter 2
Operating *Spartan*

This chapter describes the general operating features of ***Spartan'08****. It should be read prior to starting the tutorials.*

Opening and Exiting *Spartan*

To open on Windows, *click* on the **Start** button, then *click* on **All Programs**, and finally *click* on ***Spartan'08*** (or *double click* on the ***Spartan*** icon on your desktop). To open on Linux, bring up a terminal and *type* **spartan**. To open on Macintosh, *double click* on the ***Spartan'08*** icon in the Applications Folder. To exit, select **Exit** from the **File** menu (Select **Quit Spartan 08** from the **Spartan 08** menu on Mac), or *click* the **Close** button (⊠) at the top right (top left for Mac) of the ***Spartan*** interface.

Menus

Program functions may be accessed using pull-down menus under the headings in the menu bar, for example, the **Setup** menu.

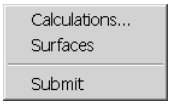
File	Allows you to create a new molecule or read in a molecule that you have previously saved, to print text and graphics, embed external files (for example Word and Excel files) into ***Spartan*** documents and to make QuickTime movies.
Edit	Allows you to transfer information to and from the clipboard, to undo the last operation, to find text strings and molecule fragments, to center molecules on screen, and to clear the active molecule by deleting it.
Model	Allows you to control the style of your model, to display

	hydrogen bonds and chemical function descriptors and to couple or decouple molecules in a multi-molecule file.
Geometry	Allows you to measure and constrain bond lengths, angles and dihedrals, define points and planes, setup frozen atoms, select conformational degrees of freedom, select centers for similarity analysis and align molecules.
Build	Allows you to build and edit molecules and to create lists of substituted molecules. Provides access to ChemDraw™. (**NAL**, **NAM**)
Setup	Allows you to specify the task to be performed and the theoretical model to be employed, to specify graphical surfaces and property maps and to submit jobs for calculation.
Display	Allows you to display text output, molecular and atomic properties, surfaces and property maps and infrared, NMR spectra and UV/visible, as well as to access experimental IR, NMR spectra and UV/visible over the internet. It allows you to present data in a spreadsheet and make plots from and perform regression analysis on these data, to access the results of a similarity analysis and to compute reaction energies based either on user data or from entries in the Spartan Molecular Database.
Search	Allows you to search the Spartan Molecular Database (SMD) of calculated structures, properties and spectra, the Spartan Infrared Database (SIRD) of calculated infrared spectra, the Spartan Reaction Database (SRD) of calculated transition-state structures and to retrieve structures, the Cambridge Structural Database (CSD) of experimental X-ray crystal structures, and extract ligands from the Protein Data Bank (PDB). Allows you to guess a transition-state geometry based on a library of reactions and to identify tautomers.
Options	Allows you to set display standards, specify databases, assign compute servers, monitor executing jobs and specify queues.
Help	Provides access to information on ***Spartan's*** general operation, the ***Spartan'08 Tutorial and User's Guide***, ***A Guide to Molecular Mechanics and Quantum Chemical Calculations***, and a number of computational FAQ's.

A complete listing of menu functions is provided in **Appendix B**.

Keystroke Equivalents (NAM)

Keystroke equivalents for menus are designated by the underlined letter: **File**, **Edit**, **Model**, **Geometry**, **Build**, **Setup**, **Display**, **Search**, **Options** and **Help**. To access a menu, *press* the **Alt** key and the appropriate letter key (case insensitive), for example, **Alt S** to access the **Setup** menu.



Each menu entry is designated by an underlined letter. To access the entry, *press* the appropriate letter key, for example, *press* the **C** key to access the **Calculations** dialog.

Icons

Icons provide convenient access to selected functions under the **File**, **Build, Geometry** and **Search** menus, as well as other specialized functions.

New	Minimize	Constrain Dihedral
Open	Measure Distance	Define Point
Close	Measure Angle	Define Plane
Save As	Measure Dihedral	Define CFD
View	Freeze Center	Align
Add Fragment	Set Torsions	Databases
Delete	Set Similarity Centers	Ligands
Make Bond	Constrain Distance	Transition States
Break Bond	Constrain Angle	Tautomers

Post to Spreadsheet	Guess Transition State
Lock/Unlock Constraints	Move Up/Down Dialog
Play / Pause	Step

Mouse and Keyboard Operations

The following functions are associated with the mouse.

	button	
keyboard	**left**	**right**
–	selection, X/Y rotate, atom/ fragment exchange[a], fragment insertion[a]	X/Y translate
Shift	range selection, Z rotate	scaling[b]
Ctrl	global X/Y rotate[c]	global X/Y translate
Ctrl + **Shift**	multiple selection, global Z rotate[c]	scaling[b]
Ctrl (add fragment mode)	fragment X/Y rotate chirality invert[a,e]	fragment X/Y translate
Ctrl + **Shift** (add fragment mode)	fragment Z rotate absolute configuration invert[a,e]	scaling[b]
Ctrl + **Shift**[d] (Mac & Linux)	bond rotation	bond stretching
Alt (Windows)	group selection, bond rotation[d]	bond stretching

a) Requires *double clicking*.
b) Scaling is always applied to all open molecules and all fragments. The center mouse wheel will also scale molecules.
c) Global rotations can be either molecule or screen centered. This is controlled by **Global Rotate** in the **Miscellaneous Preferences** dialog (**Preferences...** under **Options** menu).
d) In **Add Fragment** mode with bond selected (red arrow visible).
e) For Macintosh, **Command** (⌘) key replaces **Ctrl** key for chirality inversion.

Mouse/keyboard operations may be broadly separated into two categories: selection (picking) and manipulation (translation/rotation).

Selection. The left button is used for selection of objects on screen and/or of menu items. Left and right buttons together are used to define a selection box for copying/cutting to the clipboard, as well as for multiple model selection. Together with the **Shift** key, the left button allows for selection over a range. Together with the **Ctrl** (**Control**) key, the left button allows for multiple selection. Both range and multiple selection apply not only to text items in lists, but to atoms and bonds in molecules as well. Together with the **Alt**

key, the left button allows for selection of an entire group (detached molecular fragment). In add fragment mode, *double clicking* the left button leads to atom or group exchange. *Double clicking* on an atom with the **Ctrl** key depressed leads to inversion in chirality of the atom and *double clicking* on an atom with both **Ctrl** and **Shift** keys depressed inverts the absolute configuration of the molecule. Once an initial fragment, group or ring has been drawn, *double clicking* on the background will insert it alongside (but not bonded to) whatever fragments currently exist on screen.

Manipulation. The left button is used for rotation and the right button is used for translation and scaling of objects on screen. With no keys depressed, the left mouse button gives rise to rotation about the X and Y (screen) axes, while the right mouse button gives rise to translation in the X and Y (screen) directions. Together with the **Shift** key, the left mouse button gives rise to rotation about the Z direction and the right mouse button gives rise to scaling. The center (scroll) wheel on the mouse may also be used for scaling.

The **Ctrl** key in conjunction with the left or right mouse buttons and (optionally) the **Shift** key, signifies a change in focus away from the default for the purpose of rotations and translations. Outside of add fragment mode, the default is focus on a single molecule (the selected molecule). Use of the **Ctrl** key changes focus to the entire set of molecules on screen, meaning that rotations and translations are carried out globally. In add fragment mode, the default is focus on the full set of fragments that make up the molecule being constructed, and rotations and translations refer to this set of fragments as a whole. Use of the **Ctrl** key changes focus to a single fragment (the selected fragment), and rotations and translations now refer only to this fragment.

In add fragment mode, use of the **Alt** key with the left mouse button allows for rotation about a selected bond and, with the right mouse button, for changing the length of the selected bond.

Additional keys control various ***Spartan*** functions.

3	**3** selects red-cyan stereo display. Pressing again returns to non-stereo display.
Page Up, Page Down Home, End	Moves up (**Page Up**), down (**Page Down**), to the top (**Home**) and to the bottom (**End**) of the set of open molecules. Also, moves up and down pages in the **Output** dialog.
Insert	In add fragment mode only, inserts a new fragment on screen. This is accomplished by selecting the fragment from the model kit, holding down the **Insert** key and *clicking* on screen. **Alt** key for Macintosh. Insertion may also be accomplished by *double clicking* on the background following selection of a fragment.
Delete	Deletes a fragment, free valence, CFD, reaction arrow, the contents of a selection box, spectrum, curve, or plot. This is accomplished by holding down the **Delete** key and *clicking* on the fragment, etc.
Enter (**return** on Mac)	Required following text or data entry into spreadsheet or dialogs.

Selecting Molecules, etc.

Two or more molecules may be simultaneously displayed in ***Spartan's*** window. However, only one molecule may be selected. The selected molecule has access to all capabilities (molecule building, job setup and submission and text and graphical display and manipulation), while non-selected molecules may only be displayed as static images. The exceptions involve scaling and the use of the **Ctrl** key.

Selection of a molecule occurs by *clicking* on its structure model or on any of its associated graphical surfaces. This will also result in deselection of any previously selected molecule. Molecular properties for the selected molecule are available in the **Properties** dialog (**Display** menu). Atom, bond, substituent and surface display properties, pharmacophores and group descriptors, as well as information about geometrical constraints may be accessed by

subsequent *clicking* on an atom, bond, etc., associated with the selected molecule, that are then highlighted (colored gold). *Clicking* on the selected atom, bond etc. resets the display to molecular properties. *Clicking* on another molecule results in display of molecular properties for that molecule. Information about plots (including spectra) and changes to plot style and range are also available from the **Plot Properties** dialog.

Where the molecule belongs to a document with more than a single member, selection from among the different members in the document may be made using either the ◂ and ▸ buttons or the scroll bar at the bottom left of the screen. Alternatively, if the spreadsheet for the document is open on screen (see **Chapter 16**), selection can be made by *clicking* on the molecule label at the left of the spreadsheet. *Clicking* on ▸ at the bottom left of the screen animates the display of molecules in the document, that is, steps through them sequentially. This is useful for displaying a progression of graphical surfaces along a reaction coordinate. *Clicking* on ❚❚ (that replaces ▸) stops the animation.

Two or more molecules from the same document may be displayed at once (although only one may be selected). Molecules are marked for display by *checking* the box immediately to the left of the molecule label in the spreadsheet (**Chapter 16**).

Dialogs

Dialogs are either modal (as indicated by **...** following their reference in a menu) or non-modal. Modal dialogs must be dismissed before program operation can continue, while non-modal dialogs may be kept open on screen. Only one modal dialog may be open on screen, while several different non-modal dialogs and/or several copies of the same non-modal dialog (each copy referring to a different molecule in a document or different documents) may be open on screen.

Stereo Displays

Spartan supports red-cyan stereo. Red/blue glasses (supplied with ***Spartan***) must be worn. To enter stereo mode, *press* the **3** key. *Press* again to return to non-stereo mode.

Changing Colors and Setting Preferences

Colors and **Preferences...** under the **Options** menu (**Chapter 18**) allows for changing default background and graphical object colors, and for setting (and resetting) program defaults, respectively. Further control of color is discussed in **Chapter 16**.

Monitoring and Killing Jobs

Monitor under the **Options** menu (**Chapter 18**) allows for monitoring of executing jobs as well as for killing jobs.

Chapter 3
Basic Operations

*This tutorial introduces a number of basic operations in **Spartan** required for molecule manipulation, property query and spectra and graphics display. It should be completed by new **Spartan** users, but can probably be skipped by those who have used earlier versions. Specifically it shows how to: i) open molecules, ii) view different models and manipulate molecules on screen, iii) measure bond distances, angles and dihedral angles, iv) display energies, dipole moments, atomic charges and infrared, NMR spectra and UV/visible, and v) display graphical surfaces and property maps. Spreadsheet operations are not illustrated, no molecules are built and no calculations are performed.*

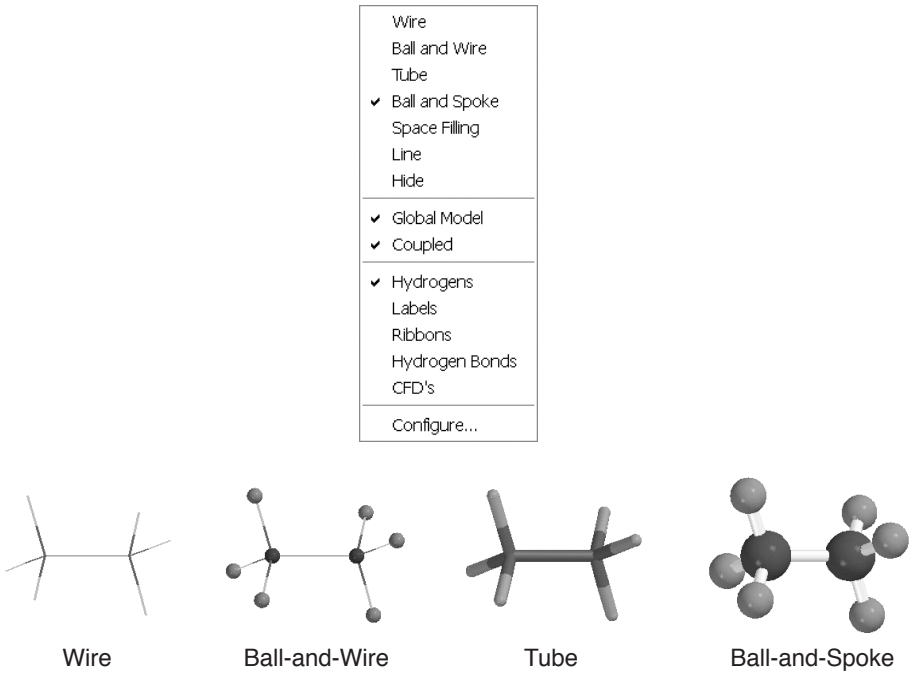
1. Start ***Spartan***. *Click* (left mouse button) on **File** from the menu bar that appears at the top of ***Spartan's*** main window. *Click* on **Open...** from the **File** menu that appears. Alternatively, *click* on the  icon at the top of the screen. A file browser appears.

   Move to the ***tutorials*** directory*, *click* on ***basic operations*** and *click* on **Open** (or *double click* on ***basic operations***). A single file containing ***ethane***, ***acetic acid dimer***, ***propene***, ***ammonia***, ***hydrogen peroxide***, ***acetic acid***, ***water***, ***cyclohexanone***, ***3-aminobenzophenone***, ***ethylene***, ***benzene***, ***aniline*** and ***cyclohexenone*** will be opened. A ball-and-spoke model for the first molecule (***ethane***) will be displayed, and its name appears at the bottom right of the screen.

2. Practice rotating (*move* the mouse while holding down the left button) and translating (*move* the mouse while holding down the right button). *Click* on **Model** from the menu bar.

---

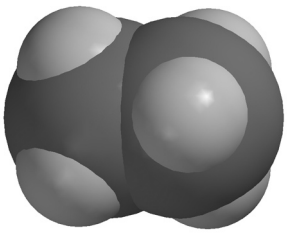
* For Windows, this is found in ***Program Files/Wavefunction/Spartan08***. For Linux, this is found in the directory where ***Spartan*** was installed. For Macintosh, this is located in the Spartan.08.full folder on the installation CDROM.



One after another, select **Wire**, **Ball and Wire**, **Tube** and finally **Ball and Spoke** from the **Model** menu. All four models for ***ethane*** show roughly the same information. The wire model looks the most like a conventional line formula. It uses color to distinguish different atoms, and one, two and three lines between atoms to indicate single, double and triple bonds, respectively.

The ball-and-wire model is identical to the wire model, except that atom positions are represented by small spheres, making it easy to identify atom locations. The tube model is identical to the wire model, except that bonds are represented by solid cylinders. The tube model is better than the wire model in conveying three-dimensional shape. The ball-and-spoke model is a variation on the tube model; atom positions are represented by colored spheres, making it easy to see atom locations.

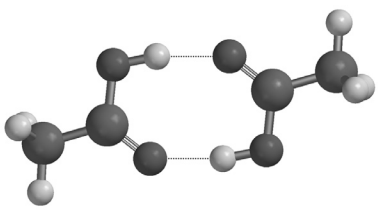
Select **Space Filling** from the **Model** menu.



Space-Filling

The space-filling model is different from the other models in that bonds are not shown. Rather, each atom is displayed as a colored sphere that represents its approximate size. Thus, the space-filling model for a molecule provides a measure of its size. While lines between atoms are not drawn, the existence (or absence) of bonds can be inferred from the extent to which spheres on neighboring atoms overlap. If two spheres substantially overlap, then the atoms are almost certainly bonded, and conversely, if two spheres barely overlap, then the atoms are not bonded. Intermediate overlaps suggest weak bonding, for example, hydrogen bonding.

3. *Click* once on the right arrow key ▶ at the bottom left of the screen. This will move to the next molecule in the document, ***acetic acid dimer***. Its name will appear at the bottom of the screen. If you make a mistake, use the backward ◀ or forward ▶ step keys to get to ***acetic acid dimer*** in the document. Switch to a space-filling model and look for overlap between the (OH) hydrogen on one acetic acid molecule and the (carbonyl) oxygen on the other. Return to a ball-and-spoke model and select **Hydrogen Bonds** from the **Model** menu.

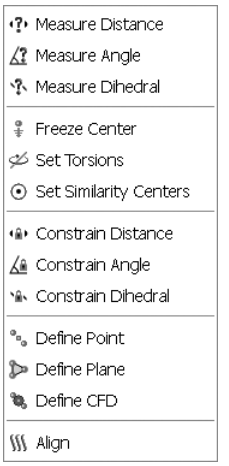


Ball-and-Spoke model for acetic acid dimer
with hydrogen bonds displayed

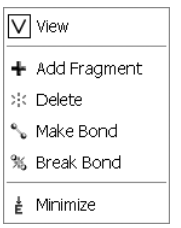
The two hydrogen bonds, that are responsible for holding the acetic acid molecules together, will be drawn.

Use the **3** key to toggle between stereo 3D and regular display. To view in 3D you will need to wear the red/blue glasses provided with ***Spartan***.

4. Distances, angles, and dihedral angles can easily be measured with ***Spartan*** using **Measure Distance**, **Measure Angle**, and **Measure Dihedral**, respectively, from the **Geometry** menu.



a) **Measure Distance**: This measures the distance between two atoms. *Click* once on ▶ to move to the next molecule, ***propene***, and then select **Measure Distance** from the **Geometry** menu (or *click* on the ? icon at the top of the screen). *Click* on a bond or on two atoms (the atoms do not need to be bonded). The distance (in Ångstroms) will be displayed at the bottom of the screen. Repeat the process for several atoms. When you are finished, select **View** from the **Build** menu (or *click* on the V icon at the top of the screen).

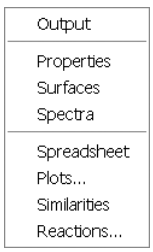


b) **Measure Angle**: This measures the angle around a central atom. *Click* once on ▶ to move to the next molecule, ***ammonia***, and then select **Measure Angle** from the

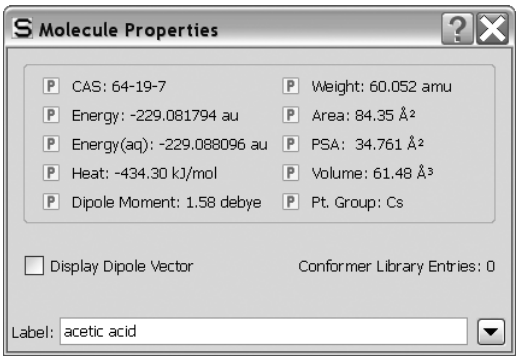
Geometry menu (or *click* on the icon at the top of the screen). *Click* first on H, then on N, then on another H. Alternatively, *click* on two NH bonds. The HNH angle (in degrees) will be displayed at the bottom of the screen. *Click* on V when you are finished.

c) **Measure Dihedral**: This measures the angle formed by two intersecting planes, one containing the first three atoms selected and the other containing the last three atoms selected. *Click* once on to move to the next molecule, ***hydrogen peroxide***, then select **Measure Dihedral** from the **Geometry** menu (or *click* on the icon at the top of the screen) and then *click* in turn on the four atoms (HOOH) that make up hydrogen peroxide. The HOOH dihedral angle will be displayed at the bottom of the screen. *Click* on V when you are finished.

5. Energies, dipole moments and atomic charges among other calculated properties, are available from **Properties** under the **Display** menu.



a) **Energy:** *Click* once on to move to the next molecule, ***acetic acid***, and then select **Properties** from the **Display** menu. The **Molecule Properties** dialog appears.

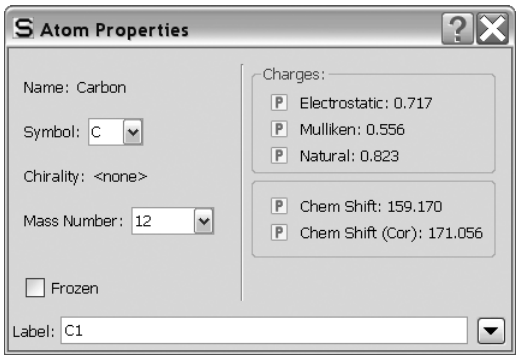


This provides the energy for acetic acid in atomic units (**Energy** in au). Also provided is an estimate of the energy in water (**Energy(aq)** in au).

b) **Dipole Moment:** The magnitude of the dipole moment (**Dipole Moment** in debyes) is also provided in the **Molecule Properties** dialog. A large dipole moment indicates large separation of charge. You can attach the dipole moment vector, ⟝ where the + side refers to the positive end of the dipole, to the model on the screen, by *checking* the box to the left of **Display Dipole Vector** near the bottom of the dialog.

> The vector will not be displayed if the dipole moment is zero. The dipole moment will not be reported if the molecule is charged because in this case it depends on the location and orientation of the molecule in space.

c) **Atomic Charges:** To display the charge on an atom, *click* on it with the **Molecule Properties** dialog on the screen. The **Atom Properties** dialog replaces the **Molecule Properties** dialog.

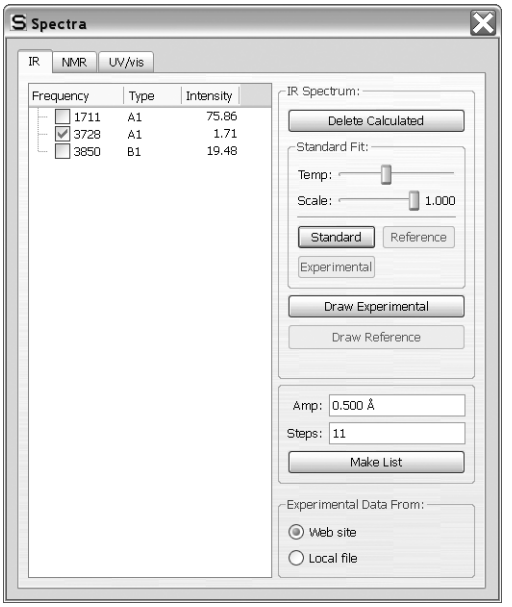


Three different atomic charges, **Electrostatic**, **Mulliken** and **Natural**, are given in units of electrons. A positive charge indicates a deficiency of electrons on an atom and a negative charge, an excess of electrons. Repeat for other atoms. Confirm that the positively-charged atom(s) lie at the positive end of the dipole moment vector. When you are finished, close the dialog by *clicking* on ☒ at the top of the dialog.

> The three sets of atomic charges for acetic acid are different because the charge on an atom in a molecule cannot be uniquely defined, let alone measured. The different calculated charges correspond to different ways of counting the number of electrons associated with an atom.

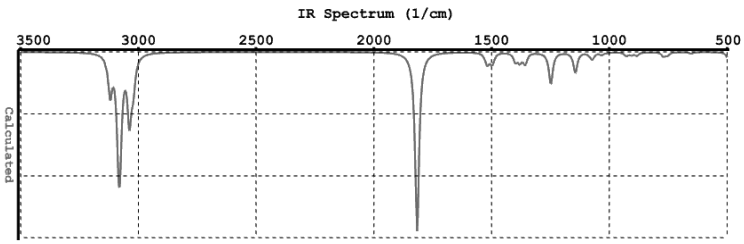
d) **Infrared Spectra:** Molecules vibrate (stretch, bend, twist) even if they are cooled to absolute zero. This is the basis of infrared spectroscopy, where absorption of energy occurs when the frequency of a particular molecular motion matches the frequency of the light. Infrared spectroscopy is important for identifying molecules as different functional groups vibrate at noticeably different and characteristic frequencies.

Click once on ▶ to move to the next molecule in the document, ***water***. To animate a vibration, select **Spectra** from the **Display** menu and *click* on the **IR** tab. This leads to the **IR Spectra** dialog.



This displays the three vibrational frequencies for the water molecule, corresponding to bending and symmetric and antisymmetric stretching motions. One after the other, *click* on each frequency and examine the motion. Turn "off" the animation when you are finished.

Click once on ▶ to move to the next molecule, ***cyclohexanone***. The **Spectra** dialog now lists its 45 vibrational frequencies. Examine each in turn (*click* on the entry in the dialog) until you locate the frequency corresponding to the CO (carbonyl) stretch. Next, *click* on **Draw Calculated** at the top of the dialog. The infrared spectrum of cyclohexanone appears.

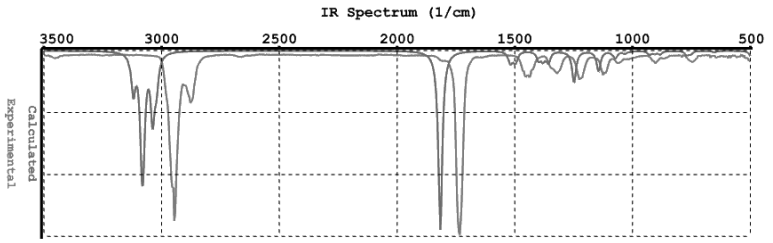


You can move the spectrum around the screen by first *clicking* on it to select it (it will turn yellow) and then *moving* the mouse while holding down the right button. You can size it by *moving* the mouse up and down while holding down both

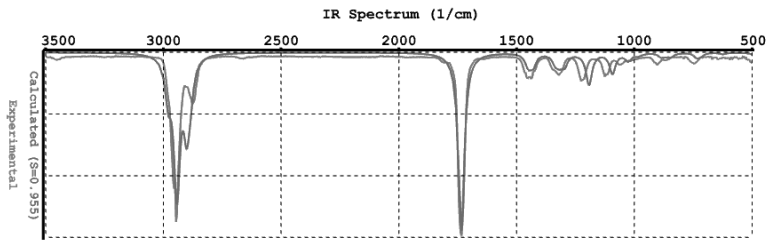
the **Shift** key and the right button.

Identify the line in the spectrum associated with the C=O stretch (a small red circle moves from line to line as you step through the frequencies in the **Spectra** dialog). Note that this line is isolated and that it is very intense, making it easy to find.

If your computer is connected to the internet, you can draw the experimental IR spectrum for cyclohexanone on top of the calculated spectrum. Select **Web Site** under **Experimental Data From:** at the bottom of the dialog and *click* on **Draw Experimental** in the middle of the dialog.



Note that the two spectra are broadly similar, but that the lines in the calculated spectrum are consistently of higher frequency. To see this more clearly, *click* on the calculated spectrum and move the slider bar to the right of **Scale** near the top of the dialog. The calculated spectrum will be uniformly scaled and it will be possible to bring it into close agreement with the experimental spectrum. Calculated and experimental spectra are automatically fit by selecting **Experimental** under **Calculated Fit** at the center of the **IR Spectra** dialog.



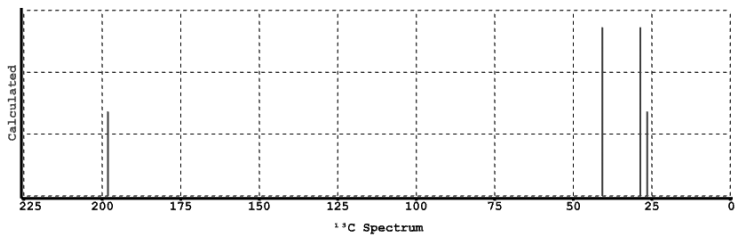
You can remove the plot by *clicking* on both **Delete Calculated** and **Delete Experimental** in the **Spectra** dialog.

(These buttons have replaced **Draw Calculated** and **Draw Experimental**, respectively.)

e) **NMR Spectra:** Along with mass spectrometry, NMR spectroscopy is the most powerful tool available with which to assign molecular structure. Many nuclei exhibit NMR spectra, but proton and ^{13}C are by far the most important.

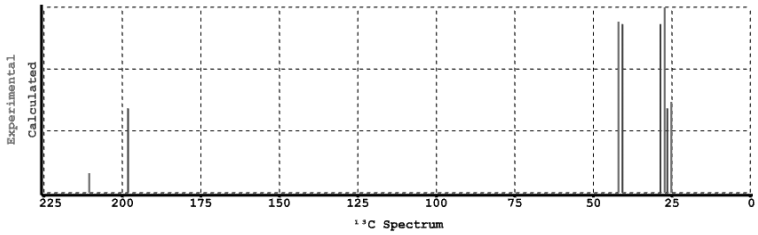
We will use cyclohexanone to illustrate ^{13}C NMR. This is already selected so there is no need to move in the list of molecules. With the **Spectra** dialog on screen, *click* on the **NMR** tab to bring up the **NMR Spectra** dialog.

Click on **Draw Calculated** under **^{13}C Spectrum** to show the calculated ^{13}C spectrum.



This comprises four lines, corresponding to the four distinct carbons. If you are connected to the internet, select **Web**

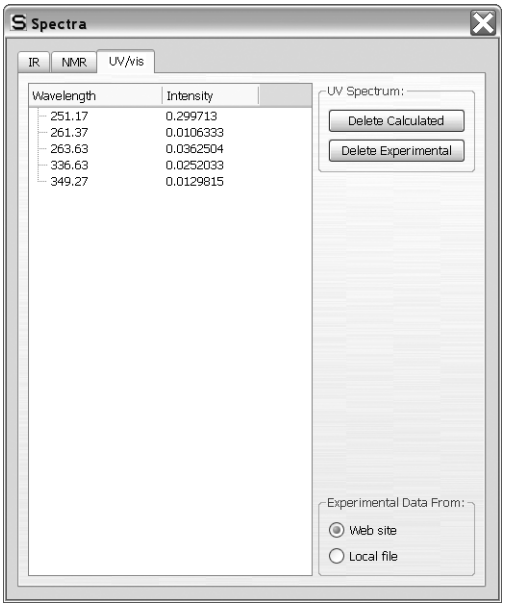
Site under **Experimental Data From:** at the bottom right of the dialog and *click* on **Draw Experimental** (under **^{13}C Spectrum**) to superimpose the experimental spectrum onto the calculated one.



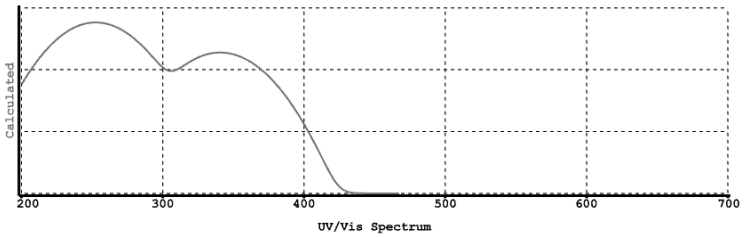
Remove the spectra by *clicking* on **Delete Calculated** and **Delete Experimental** in the **NMR Spectra** dialog. (These buttons have replaced **Draw Calculated** and **Draw Experimental**, respectively.)

f) **UV/visible Spectra:** Absorption of light in the visible or ultraviolet range of the electromagnetic spectrum leads to electronic excitation from the ground-state to excited-states and (in the case of absorption in the visible), is responsible for a molecule's color. UV/visible spectroscopy not only offers a valuable fingerprint but is also an important screen to identify molecules that may be damaged by exposure to light.

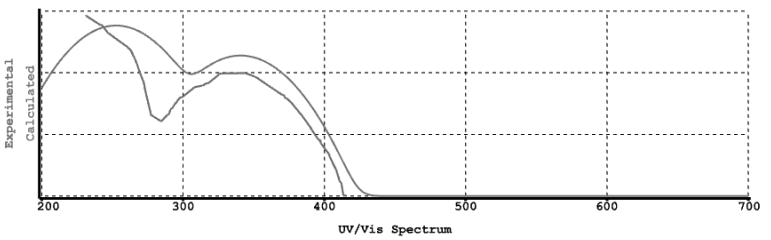
Click once on ▶ to move to the next molecule, ***3-aminobenzophenone***. Bring up the **Spectra** dialog and *click* on the **UV/vis** tab. This leads to the **UV/vis Spectra** dialog.



Click on **Draw Calculated** to show the calculated spectrum.



If you are on the internet, select **Web Site** under **Experimental Data From:** at the bottom of the dialog and *click* on **Draw Experimental** to superimpose the experimental spectrum onto the calculated one.

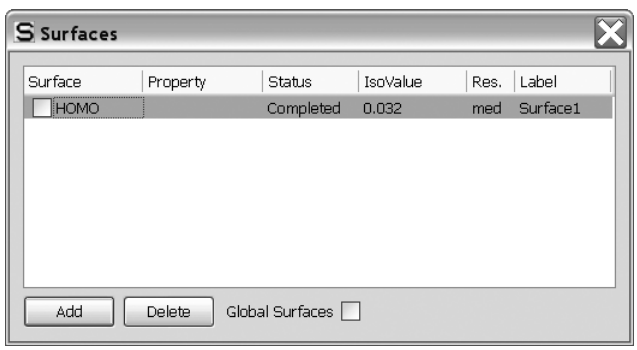


The two spectra are visually quite similar. Remove the spectra by *clicking* on **Delete Calculated** and **Delete Experimental**

in the **UV/vis Spectra** dialog. (These buttons have replaced **Draw Calculated** and **Draw Experimental**, respectively.)

6. ***Spartan*** permits display, manipulation and query of a number of important graphical quantities resulting from quantum chemical calculations. Most important are the ***electron density*** (that reveals how much space a molecule actually takes up), the ***bond density*** (that reveals chemical bonds), and key ***molecular orbitals*** (that provide insight into both bonding and chemical reactivity). In addition, the ***electrostatic potential map***, an overlaying of the electrostatic potential (the attraction or repulsion of a positive charge for a molecule) onto the electron density, is valuable for describing overall molecular charge distribution as well as anticipating sites of electrophilic addition. Another indicator of electrophilic addition is provided by the ***local ionization potential map***, an overlaying of the energy of electron removal (ionization) onto the electron density. Finally, an indicator of nucleophilic addition is provided by the ***LUMO map***, an overlaying of the lowest-unoccupied molecular orbital (the LUMO) onto the electron density.

   Click once on  to move to the next molecule, ***ethylene***, and then select **Surfaces** from the **Display** menu. The **Surfaces** dialog appears.



   Display ethylene's highest-occupied molecular orbital (the HOMO) as an opaque solid. *Click* inside the box to the left of the line ***homo*** inside the dialog. What you see is a π orbital, equally concentrated above and below the plane of the molecule. The

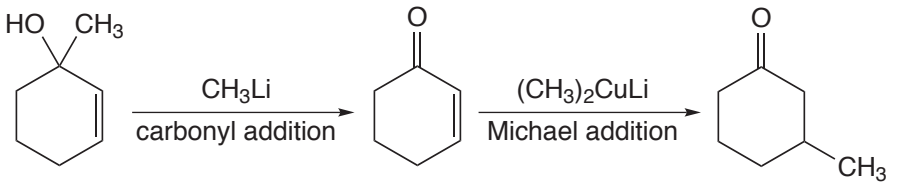
colors (red and blue) give the sign of the orbital. Changes in sign correlate with bonding or antibonding character. You can if you wish, turn "off" the graphic by again selecting ***HOMO***.

Click once on ▶ to move to the next molecule, ***benzene***, and select ***density potential*** inside the **Surfaces** dialog. An electrostatic potential map for benzene will appear. *Click* on the map. The **Style** menu will appear at the bottom right of the screen. Select **Transparent** from this menu. Making the map transparent allows you to see the molecular skeleton underneath. Go back to a **Solid** display (**Style** menu) in order to clearly see color differences. The surface is colored red in the π system (by convention, indicating negative potential and the fact that this region is attracted to a positive charge), and blue in the σ system (by convention, indicating positive potential and the fact that this region is repelled by a positive charge). Bring up the **Properties** dialog (**Display** menu) and *click* on the surface. *Click* inside the box to the left of **Bands** in the **Surface Properties** dialog to replace the continuous display with a series of color bands.

Click once on ▶ to move to the next molecule, ***aniline***, and select ***density ionization*** inside the **Surfaces** dialog. The graphic that appears, a local ionization potential map. By convention, red regions on the density surface indicate areas from which electron removal (ionization) is relatively easy, meaning that they are subject to electrophilic attack. These are easily distinguished from regions where ionization is relatively difficult (by convention, colored blue). Note that the *ortho* and *para* ring carbons are more red than the *meta* carbons, consistent with the known directing ability of the amino substituent.

Click once on ▶ to move to the next molecule, ***cyclohexenone***, and select ***LUMO*** in the **Surfaces** dialog. The resulting graphic portrays the lowest-energy empty molecular orbital (the LUMO) of cyclohexenone. This orbital is delocalized onto several atoms and it is difficult to tell where exactly a pair of electrons (a nucleophile) will attack the molecule.

A clearer portrayal is provided by a LUMO map, that displays the (absolute) value of the LUMO on the electron density surface. By convention, the color blue is used to represent maximum value of the LUMO and the color red, minimum value. First, remove the LUMO from your structure (select ***LUMO*** in the **Surfaces** dialog) and then turn on the LUMO map (select ***density*** **|*LUMO*|** in the dialog). Note that there are two blue regions, one directly over the carbonyl carbon and the other over the β carbon. This is entirely consistent with known chemistry. Enones may either undergo carbonyl addition or conjugate (Michael) addition.

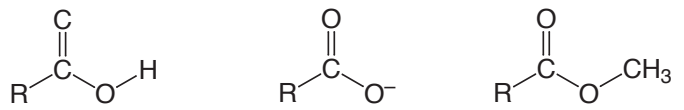


7. When you are finished, close the document by selecting **Close** from the **File** menu or alternatively by *clicking* on the icon at the top of the screen.

Chapter 4
Organic Molecules

This chapter shows how to: i) construct organic molecules from atomic fragments, functional groups and rings, ii) carry out molecular mechanics and quantum chemical calculations, iii) calculate and draw infrared and NMR spectra and iv) calculate and interpret graphical models. It also illustrates use of the Spartan Molecular Database of calculated structures, energies and molecular properties as well as* ***Spartan's*** *interface to the Cambridge Structural Database.*

The simplest building blocks incorporated into ***Spartan's*** organic model kit are atomic fragments. These constitute specification of atom type, for example, carbon, and local environment, for example, tetrahedral. However, much of organic chemistry is organized around functional groups, collections of atoms the structure and properties of which are roughly the same in every molecule. The organic model kit also incorporates a small library of functional groups that can easily be extended or modified. For example, the carboxylic acid group may be modified to build a carboxylate anion (by deleting a free valence from oxygen), or an ester (by adding tetrahedral carbon to the free valence at oxygen).

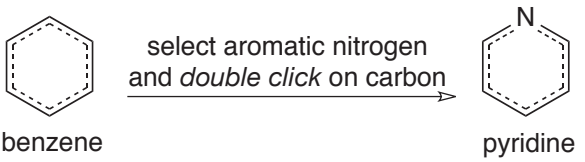


carboxylic acid carboxylate anion ester

Rings are also common components in organic molecules, and the organic model kit incorporates a library of commonly-encountered hydrocarbon rings that can easily be modified by atom replacement. For example, pyridine may be built starting from benzene by selecting

* ***Spartan'08*** includes a subset of approximately 5,000 molecules from the full Spartan Molecular Database (SMD) which comprises more than 150,000 molecules along with their structures, energies, spectra and properties calculated with up to 10 quantum chemical models.

aromatic nitrogen from the list of atomic fragments, and then *double clicking* on one of the carbons. This results in the substitution of aromatic nitrogen for aromatic carbon.

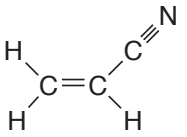


Functional groups may also be modified in this manner.

In addition to illustrating the construction of organic molecules, the tutorials in this chapter also illustrate the way in which molecular mechanics and quantum chemical calculations are set up, their results examined and interpreted, and infrared and both proton and ^{13}C NMR spectra computed and displayed. A variety of graphical models are introduced and illustrated. Finally, examples of the use of the Spartan Molecular Database of calculated molecular structures, energies and properties and of the Cambridge Structural Database of experimental X-ray crystal structures are provided.

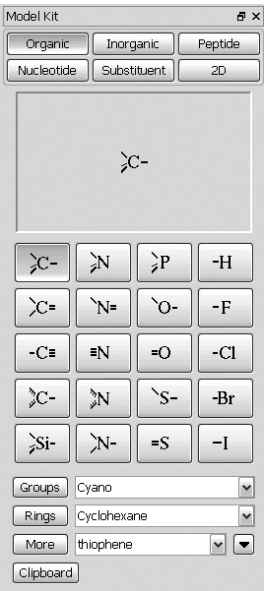
1 min

Acrylonitrile



Acrylonitrile provides a good opportunity to illustrate the basics of molecule building in ***Spartan***, as well as the steps involved in carrying out and analyzing a quantum chemical calculation.

1. *Click* with the left mouse button on **File** from the menu bar. Then *click* on **New** from the menu that appears (or *click* on the icon at the top of the screen). The organic model kit appears.



At the center of the kit is a library of atomic fragments. *Click* on trigonal planar sp^2 hybridized carbon from the fragment library. A model of the fragment appears at the top of the model kit. Bring the cursor anywhere on screen and *click*. Rotate the carbon fragment (*drag* the mouse while holding down the left button) so that you can clearly see both the double free valence (=) and the two single free valences (–).

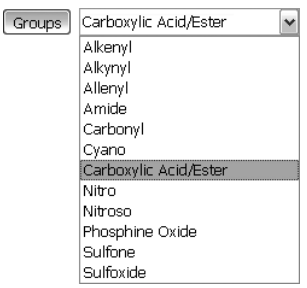
> ***Spartan's*** model kits connect atomic fragments (as well as groups, rings and ligands) through free valences. Any free valences that

remain upon exiting a model kit are automatically converted to hydrogen atoms; it is not necessary to explicitly add hydrogens to open valences.

2. sp^2 carbon is still selected. *Click* on the double free valence. The two fragments are connected by a double bond, leaving you with ethylene. The name "ethylene" will appear at the bottom right of the screen. If you make a mistake and *click* instead on the single free valence, select **Undo** from the **Edit** menu. You can also start over by selecting **Clear** from the **Edit** menu.

Spartan's organic model kit allow only the same type of free valences to be connected, for example, single to single, double to double, etc.

3. *Click* on **Groups** in the model kit, and select **Cyano** from the functional groups available from the menu.



Click on any of the four single free valences on ethylene (they are equivalent). This bonds the cyano group to ethylene, leaving you with acrylonitrile.* Its name will now appear at the bottom right of the screen.

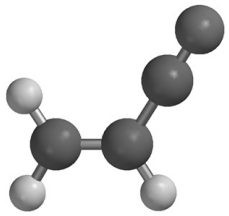
4. Select **Minimize** from the **Build** menu (or *click* on the icon at the top of the screen). The final molecular mechanics energy

---

* You could also have built acrylonitrile without using the **Groups** menu. Starting from scratch (**Clear** from the **Edit** menu), first build ethylene as above, then select sp hybridized carbon from the model kit and then *click* on one of the free valences on ethylene. Next, select sp hybridized nitrogen from the model kit and *click* on the triple free valence on the sp carbon. Alternatively, you could have built the molecule entirely from groups. Starting from scratch, *click* on **Groups**, select **Alkenyl** from the menu and *click* anywhere on screen. Then select **Cyano** from the menu of functional groups and *click* on one of the free valences on ethylene. In general, molecules can be constructed in more than one way.

(36.2 kJ/mol) and symmetry point group (C_s) are provided at the bottom right of the screen.

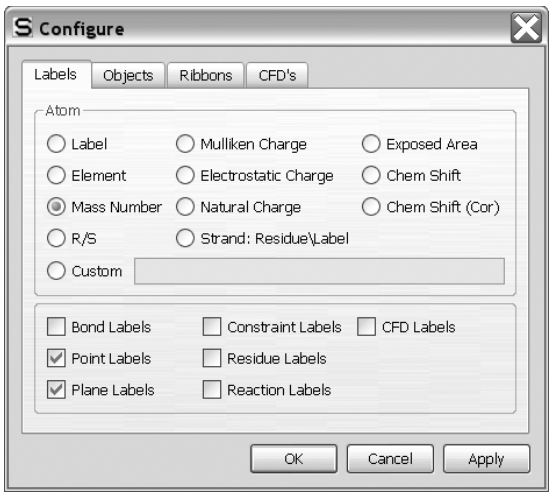
5. Select **View** from the **Build** menu (or *click* on the V icon at the top of the screen). The model kit disappears, leaving only a ball-and-spoke model of acrylonitrile on screen.



ball-and-spoke model

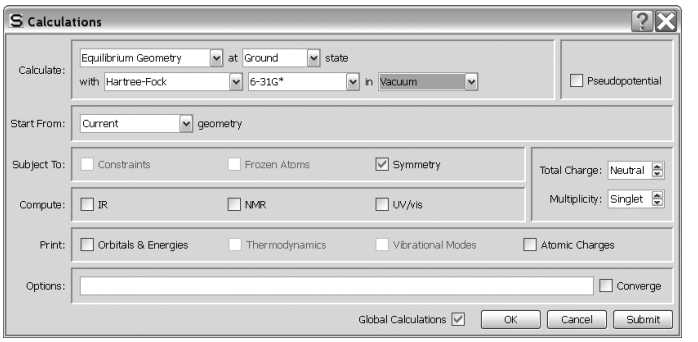
This model can be rotated, translated and zoomed by using the mouse in conjunction with keyboard functions. To rotate the model, *drag* the mouse while holding down the left button; to rotate in the plane of the screen also hold down the **Shift** key. To translate the model, *drag* the mouse with the right button depressed. To zoom the model (translation perpendicular to the screen), use the center mouse wheel (scroll wheel) if available, or hold down the **Shift** key in addition to the right button while *dragging* the mouse up (zoom in) or down (zoom out).

6. Select **Configure...** from the **Model** menu, and *click* to select **Mass Number** under **Atom** in the **Configure** dialog that appears.



Click on **OK** to remove the dialog. Mass numbers will appear next to the individual atoms. Remove the atom labels by *clicking* to deselect **Labels** from the **Model** menu.*

7. Select **Calculations...** from the **Setup** menu, and perform the following operations in the **Calculations** dialog which appears.



a. Select **Equilibrium Geometry** from the top menu to the right of **Calculate**. This specifies optimization of equilibrium geometry. **Ground** and not **First Excited** should appear in the menu to the left of **State**.

b. Select **Hartree-Fock** and then **3-21G** from the two bottom menus to the right of **Calculate**. This specifies a Hartree-Fock calculation using the 3-21G split-valence basis set. Keep the default **in Vacuum** setting.

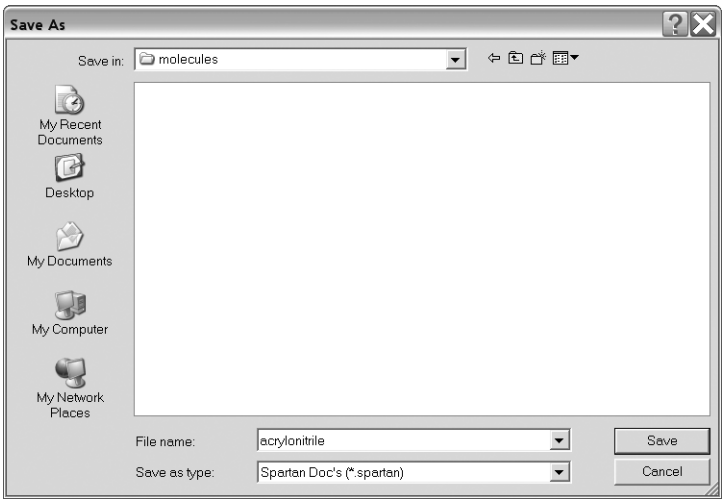
When you finish, *click* on **OK** to remove the dialog.

8. Select **Submit** from the **Setup** menu.** A file browser appears.

---

* **Labels** from the **Model** menu is automatically selected (turned "on") following a change in the **Configure** dialog by choosing **OK** or **Apply**.

** You could also have *clicked* on **Submit** inside the **Calculations** dialog.

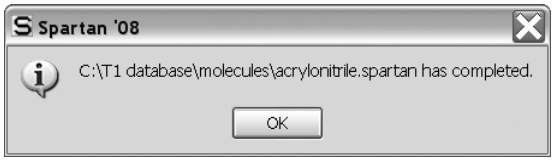


The name ***acrylonitrile*** will be presented to you in the box to the right of **File name**. Either use it or type in whatever name you like and then *click* on **Save**.* You will be notified that the calculation has been submitted.

Click on **OK** to remove the message from the screen.

> After a molecule has been submitted, and until the calculation has completed, you are not permitted to modify any dialogs or other information associated with it.

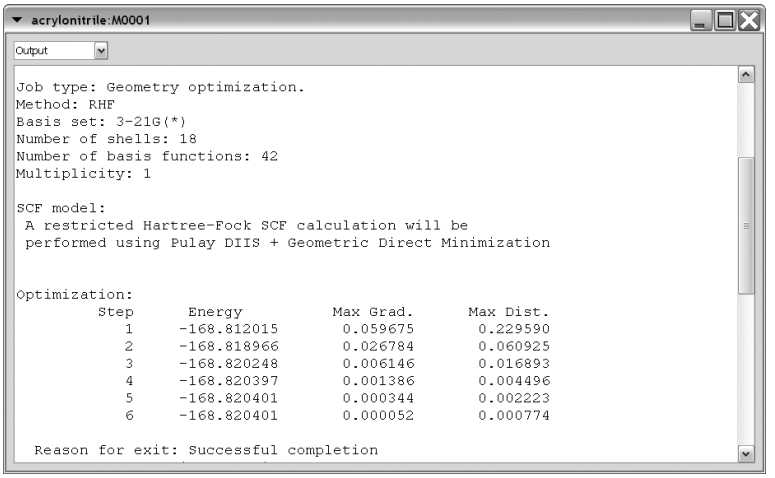
9. You will be notified when the calculation has completed.



---

* Proper names will automatically be provided for you to accept, modify or replace whenever the molecule exists in the Spartan Molecular Database and where the document being submitted contains only one molecule. Otherwise the names “spartan1”, “spartan2” will be provided.

Click on **OK** to remove the message from the screen. Select **Output** from the **Display** menu. A window containing text output for the job appears.

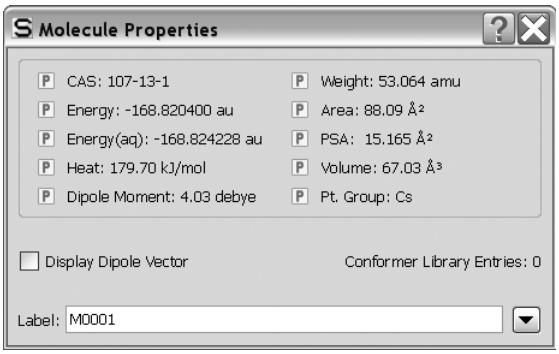


You can scan the output from the Hartree-Fock calculation by using the scroll bar at the right of the window or by *clicking* (left button) on or inside the output window and using the scroll wheel on your mouse. The information at the top of the dialog includes the task, basis set, number of electrons, charge and multiplicity, as well as further details of the calculation. Below this is the symmetry point group of the molecule that was maintained during the optimization.

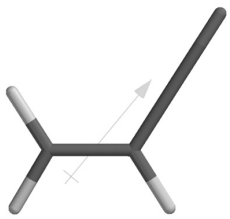
Eventually, a series of lines appear, under the heading "Optimization". These tell the history of the optimization process. Each line (or "Step") provides results for a particular geometry. Ideally, the energy will monotonically approach a minimum value for an optimized geometry. If the geometry was not optimized satisfactorily an error message, such as: "Optimization has exceeded N steps – Stop", will be displayed following the last optimization cycle. If this were the case, you would have been notified that the job had failed, rather than seeing the "completed" message dialog.

Near the end of the output is the final energy* (-168.82040 atomic units for acrylonitrile with the 3-21G basis set), and the computation time. *Click* on ☒ at the top of the output dialog to close it.

You may examine the total energy and dipole moment among other calculated properties without having to go through the output. Select **Properties** from the **Display** menu to bring up the **Molecule Properties** dialog.



To see the dipole moment vector (indicating the sign and direction of the dipole moment), *check* the box to the right of **Display Dipole Vector**. (wire, ball-and-wire or tube models are best for this display.)

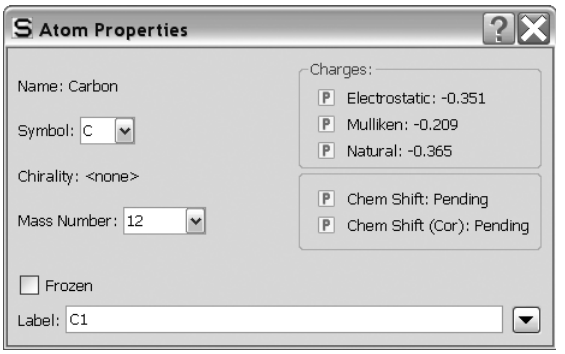


Uncheck the box to remove the dipole moment vector.

Click on an atom. The (**Molecule Properties**) dialog will be replaced by the **Atom Properties** dialog.

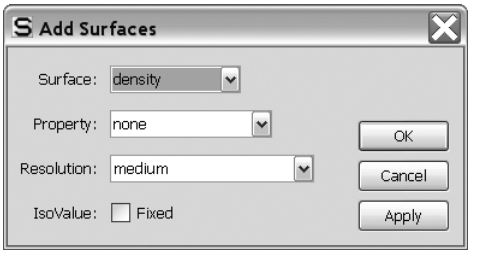
---

* See **Calculations...** (**Setup** menu; **Chapter 15**) for a discussion of how energy relates to the more familiar quantities, heat of formation and strain energy.



Among other things, this provides three different sets of atomic charges: **Electrostatic**, **Mulliken** and **Natural**. To obtain the charge on another atom, simply *click* on it. Inspect all the atomic charges on acrylonitrile (by *clicking* on the appropriate atoms). When you are finished, *click* on ☒ at the top of the **Atom Properties** dialog to close it.

10. Select **Surfaces** from either the **Setup** or **Display** menu. *Click* on **Add...** (at the bottom of the **Surfaces** dialog that results) to bring up the **Add Surface** dialog.



Select **density** from the **Surface** menu and **potential** from the **Property** menu.* This requests an electrostatic potential map (an electron density surface onto which the value of the electrostatic potential will be mapped). *Click* on **OK**. A line ***density potential*** appears at the top of the dialog. If you make a mistake, *click* on

---

* In this and most other tutorials, **Resolution** will be set to **medium**. Both lower and higher resolution images may be requested. The former provides rough images that require significantly less computer time to generate. The latter provides production quality images that require significantly more computer time and disk space. Computer time and storage can also be reduced by *checking* **Fixed** to the right of **Isovalue** in the **Add Surface** dialog. This results in constant isovalue surfaces that cannot be changed and cannot be used to generate "Slice" type displays.

this line (select ***density potential***) and then *click* on **Delete** at the bottom of the dialog.

11. The graphics calculation will run automatically (without needing to submit the job) following your request.* When it has completed (a few seconds), select ***density potential*** by *clicking* in the selection box in the **Surfaces** dialog. The surface itself corresponds to the electron density and provides a measure of the overall size and shape of acrylonitrile. The colors indicate values of the electrostatic potential on this surface; by convention, colors toward red correspond to negative potential (stabilizing interaction between the molecule and a positive charge), while colors toward blue correspond to positive potential. The nitrogen (the most electronegative atom) is red and the hydrogens (the most electropositive atoms) are blue.

12. Select **Close** from the **File** menu (or *click* on ) to remove acrylonitrile from the screen.** Also, close any dialogs that may be open.

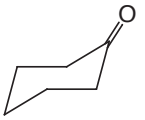
---

* If it does not, you need to select **Auto-Gen Graphics** in the **Settings Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 18**).

** While ***Spartan*** permits as many molecules as desired on screen at a given time, it will be less confusing for first-time users to keep only a single molecule open at a time.

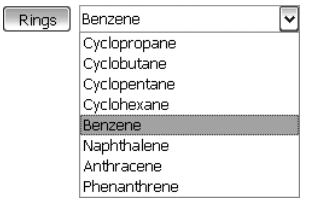
2 mins

Cyclohexanone



Cyclohexanone provides the opportunity to illustrate additional building features as well as graphical models for investigation of the stereoselectivity of an important class of organic reactions.

1. *Click* on  to bring up the organic model kit. *Click* on **Rings** and select **Cyclohexane** from the menu of rings.



Click anywhere on screen and cyclohexane will appear. Select sp^2 carbon  from the model kit. *Double click* on any carbon (***not free valence***). The sp^3 hybridized center will be replaced by an sp^2 hybridized carbon.

> ***Spartan's*** organic model kit allows replacement of fragments, subject both to the usual valence rules and to the availability of free valences. For example, replacement of an sp^3 carbon by an sp^2 carbon requires that at least two free valences be available. This is possible for cyclohexane (above), but would not have been allowed for the substituted ring carbon in methylcyclohexane.
>
> 
>

Select sp^2 oxygen  from the model kit. *Click* on the double free valence on the sp^2 carbon. You have made cyclohexanone (its name will appear at the bottom right of the screen). *Click* on  to produce a final structure with C_s symmetry, and then on  to remove the model kit.

2. Select **Calculations...** from the **Setup** menu. Specify **Equilibrium Geometry** from the top menu to the right of **Calculate**, and **Hartree Fock**, **3-21G** and **Vacuum** from the three bottom menus to the right of **Calculate**. *Click* on **OK**.

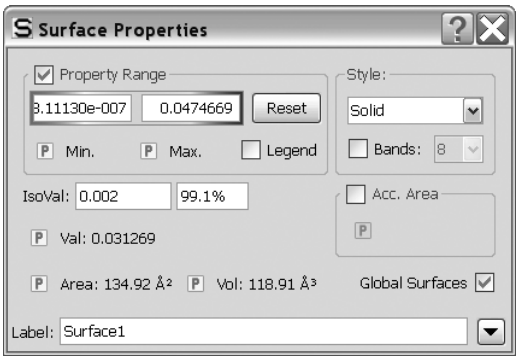
3. Cyclohexanone undergoes nucleophilic attack at the carbonyl carbon, and it is reasonable to expect that the molecule's lowest-unoccupied molecular orbital (the LUMO) will be localized here. To visualize the LUMO, enter the **Surfaces** dialog (**Setup** or **Display** menu). *Click* on **Add...**, and select **LUMO** from the **Surface** menu (**none** from the **Property** menu). *Click* on **Apply**. Also request an electron density surface onto which the (absolute) value of the LUMO has been mapped in color (a so-called LUMO map). Select **density** from the **Surface** menu and **|LUMO|** (absolute value of the LUMO) from the **Property** menu. C*lick* on **OK**. Submit the job (**Submit** from the **Setup** menu) with the name ***cyclohexanone*** (the selection that is offered).

> Graphical displays may either be requested at the time of geometry calculation (as in the present example), or at anytime after (as in the previous example). In the latter case, it is not necessary to submit the job. ***Spartan*** keeps track of what has been done and will not perform unnecessary calculations.

4. After the calculations have completed, reenter the **Surfaces** dialog and *click* in the selection box to select ***LUMO***. The LUMO of cyclohexanone appears. As expected, this is a π* orbital primarily localized on the carbonyl group.

5. Inside the **Surfaces** dialog, turn "off" the display of the LUMO (*click* to deselect ***LUMO***). Next, *click* to select ***density |LUMO|*** to display the electron density surface onto which the (absolute) value of the LUMO has been mapped. By convention, colors toward red indicate small (absolute) values of the LUMO (near zero), while colors toward blue indicate large (absolute) values of the LUMO. Note the blue spot directly over the carbonyl carbon. This corresponds to the maximum value of the LUMO and is where nucleophilic attack will occur.

6. You will see that the blue spot over the *axial* face of the carbonyl carbon is bigger (and "more" blue) than that over the *equatorial* face. This indicates preferential attack by nucleophiles onto the *axial* face. Quantify the difference by measuring the (absolute) value of the LUMO on these two faces. Bring up the (**Molecule**) **Properties** dialog (**Display** menu) and *click* anywhere on the LUMO map. The **Surface Properties** dialog appears.



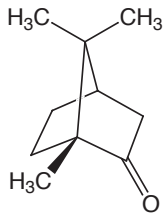
Click the selection box to the left of **Bands** to change from a continuous range of color values (to designate the value of LUMO) to a set of discrete color bands. *Click* the selection box next to **Legend** to display the property range on screen. To translate the legend, *click* on the legend to select, then hold down the right mouse button and move the mouse. The legend is useful when making qualitative comparisons of property values. Turn the map such that you can clearly see the *axial* face of the carbonyl carbon, and *click* on the area of maximum blue. The (absolute) value of the LUMO at the surface point you have selected is provided in the dialog to the right of **Val**. Note the value, and then turn the map over such that you can now see the *equatorial* face of the carbonyl carbon, and *click* on the region of maximum blue on this face*. Do these values support your qualitative conclusions from viewing the image?

7. Turn off the LUMO map (*click* to deselect). Keep cyclohexanone on screen, it will be used in the next tutorial.

---

* *Click* on the background to return to the **Molecule Properties** dialog.

Camphor



Camphor illustrates how one molecule may be used as the starting point for another. It also illustrates a search on the Cambridge Structural Database (CSD).* No calculations are involved in this example.

1. If you do not already have cyclohexanone on screen (see previous tutorial), select **Open...** from the **File** menu (or *click* on) and *double click* on ***cyclohexanone*** in the dialog which results. Select **Save As...** from the **File** menu (or *click* on the icon at the top of the screen), supply the name ***camphor*** and *click* on **Save**.

2. Select **Add Fragment** from the **Build** menu (or *click* on the icon at the top of the screen) to bring up the organic model kit. *Click* on sp^3 carbon from the model kit and then *click* on the *axial* free valence two positions removed from the carbonyl carbon. You have made *axial* 3-methylcyclohexanone. Its name will appear at the bottom right of the screen.

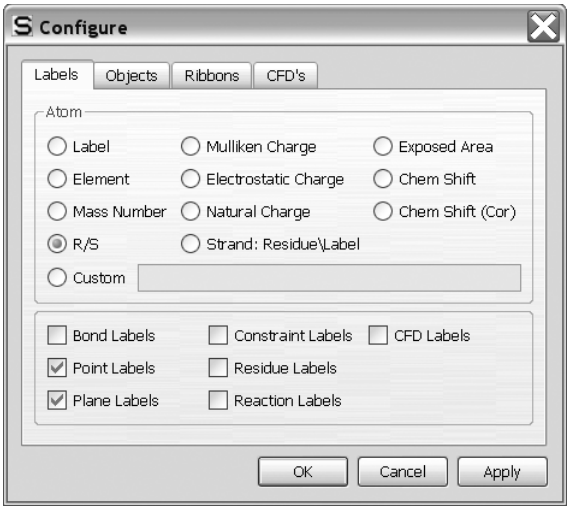
3. Select **Make Bond** from the **Build** menu (or *click* on the icon at the top of the screen). *Click* on one of the free valences of the sp^3 carbon which you have just added and then on the *equatorial* free valence on the opposite side of the six-membered ring (adjacent to the carbonyl group). A bond will be drawn. *Click* on to produce a refined (intermediate) structure. It is 2-norbornone, the name of which will appear at the bottom right of the screen. Finish making camphor by adding three methyl groups (*click* on from the model kit and, one after the other, *click* on the three appropriate free valences; refer to image above

---

* CSD will need to be licensed and installed in order to complete this tutorial. For information refer to the CCDC website (www.ccdc.cam.ac.uk). For installation instructions, see **Appendix F**. If you don't have access to CSD, a trial license may be obtained from the Cambridge Crystallographic Data Centre (www.ccdc.cam.ac.uk).

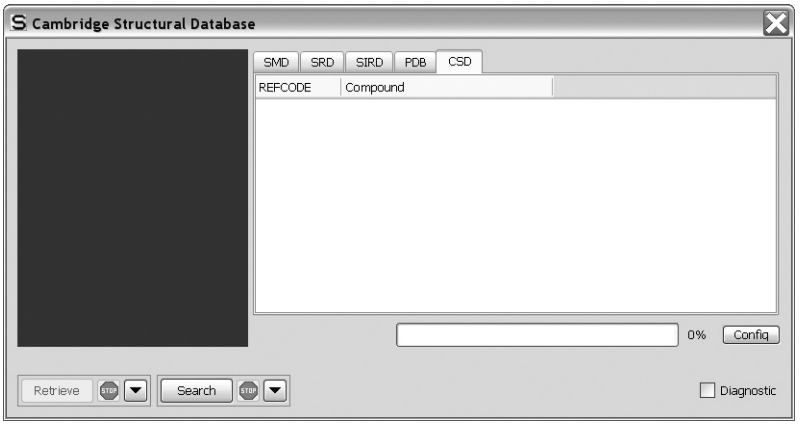
for CH_3 positions). Finally, *click* on  to give a minimized camphor geometry with C_1 symmetry. Its name will appear at the bottom right of the screen. *Click* on  to remove the model kit.

4. Select **Configure...** from the **Model** menu to bring up the **Configure** dialog.



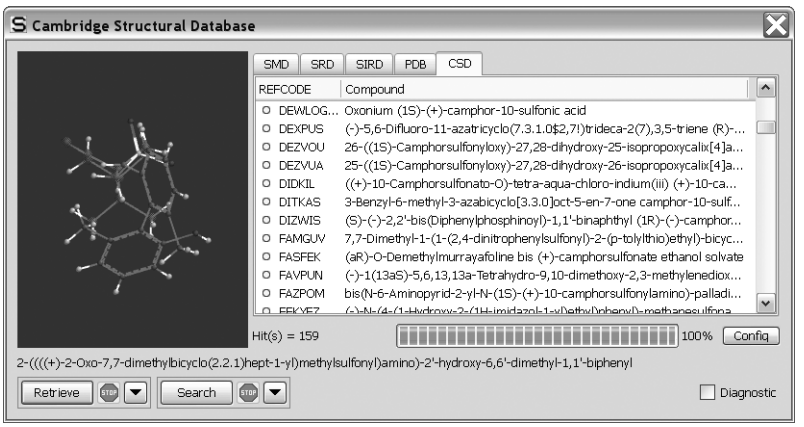
Check **R/S** under **Atom** in the dialog that results, and *click* on **OK**. R/S labels now appear (only) on the two bridgehead carbons. Both should be R (or both S). You can change from one to the other by *double clicking* on any atom while holding down both the **Ctrl** and **Shift** keys. You can turn "off" the labels by selecting **Labels** from the **Model** menu.

5. Move the molecule to the far right of the screen. Then, select **Databases** from the **Search** menu (or *click* on the icon at the top of the screen). *Click* on the **CSD** tab at the top of the **Databases** dialog that results. This brings up the **CSD** dialog.



Click on one of the free valences on the methyl group at the bridgehead carbon adjacent to the carbonyl group. The free valence is replaced by an orange cone. This restricts the search to molecules that derive from substitution of camphor at this carbon.

6. *Click* on the **Search** button at the bottom of the dialog. Hits will appear in the text box at the right of the **CSD** dialog.



7. Wait for the search to complete. *Click* on any line marked by a filled yellow circle.* *Click* on any one of these. Its name will be highlighted and a structure model will appear in the window at the top of the dialog. This model can be manipulated in the usual way.**

8. Close camphor and remove the **CSD** dialog (*click* on ☒).

---

* Unfilled circles correspond to entries for which structural data are unavailable.

** One (or more) structures may be transfered to ***Spartan's*** file system using the **Retreive** button. Full discussion is provided in **Chapter 17**.

Proton NMR Spectrum of 2-Norbornene
(cannot be completed with the Essential Edition)

Proton NMR spectroscopy was the first tool available to chemists to allow definitive assignment of the molecular structures of complex organic molecules. While it has been supplanted to some extent by ^{13}C NMR spectroscopy and more recently by *routine* X-ray crystallography, it remains indispensible.

NMR is based on the fact that nuclei possess spins that can either align parallel or antiparallel to an applied magnetic field, giving rise to different nuclear spin states. The relative energy of these states (ΔE) depends on the nucleus and on the strength of the applied magnetic field, by way of a simple relationship.

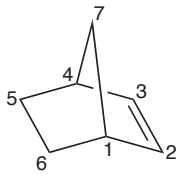
$$\Delta E = \gamma h/2\pi B_0$$

γ is the gyromagnetic ratio (a constant for a given type of nucleus), $h/2\pi$ is Planck's constant divided by 2π and B_0 is the strength of the magnetic field ***at the nucleus***. While the two nuclear spin states are normally in equilibrium, this equilibrium can be upset by applying a second magnetic field. The absorption of energy as a function of field strength (a ***resonance***) between the states can then be detected.

The key to the utility of the magnetic resonance experiment is that the energy at which a nucleus "resonates" depends on its location in the molecule, and is different for each (chemically) distinct nucleus. The reason for this is that the applied magnetic field is weakened by electrons around the nucleus. Nuclei that are well shielded by the electron cloud will feel a lesser magnetic field than those that are poorly shielded, and will show a smaller energy splitting. The difference, given relative to a standard, is termed a ***chemical shift***. By convention, both proton and ^{13}C chemical shifts (treated later in this chapter) are reported relative to tetramethylsilane (TMS) as a standard.

While each *unique* proton in a molecule gives rise to a single line (resonance) in the spectrum, the spins on nearby nuclei add and subtract to the external magnetic field. This leads to a "splitting" of lines, the splitting pattern depending on the number of neighboring

protons and their geometry. Discounting splitting, the intensity of the lines is approximately proportional to the number of equivalent protons that contribute. For example, the proton NMR spectrum of 2-norbornene shows six lines, two with unit intensity corresponding to the two different protons on the methylene bridge (C_7), and four with twice the intensity corresponding to protons at C_1 (C_4), C_2 (C_3) and the two different protons on C_5 (C_6).



In this tutorial, you will calculate the proton NMR spectrum of 2-norbornene and compare it with that contained in the Spectral Database of Organic Compounds (SDBS), freely-accessible on the web from the National Institute of Advanced Industrial Science and Technology (AIST) in Japan.*

You will use a COSY plot to see the relationship between spectral splittings (not calculated by ***Spartan*** but evaluated empirically) and three-dimensional geometry.

1. Build 2-norbornene. Start from cyclohexane (**Rings** menu), add an sp^3 carbon ( ) to an *axial* ring position and bond to ( ) to a free valence on the opposite side of the ring. Make a double bond between the appropriate ring carbons ( ). Minimize ( ) and *click* on ( ).

2. The name ***2-norbornene*** will appear at the bottom right of the screen. *Click* on ( ) to the left of the name, select ***B3LYP/6-31G**** from the models listed in the dialog that appears and *click* on **Replace**. Your model will be replaced by the entry from the Spartan Molecular Database.

3. Enter the **Calculations** dialog (**Calculations** from the **Setup** menu), and specify an **Energy** calculation using the **B3LYP**

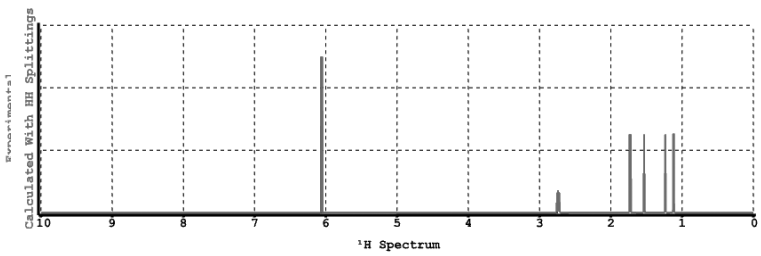
* Web address: http://riodb1.ibase.aist.go.jp/sbds.cgi-bin/cre_index.cgi. Unfortunately, electronic access to SDBS is prohibited by AIST, severely restricting the utility of this significant resource.

Density Functional model with the **6-31G*** basis set. *Check* **NMR** to the right of **Compute**. *Click* on **Submit**. Accept the name ***2-norbornene***.

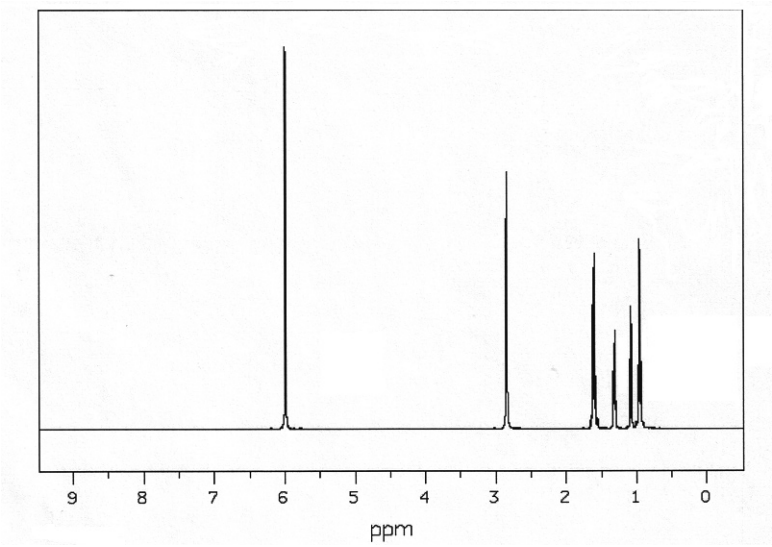
4. Calculation will require a few minutes. When it has completed, proton shifts can be found in **Atom Properties** dialog. Select **Properties** from the **Display** menu and one after the other *click* on the hydrogens. Propose assignments for the six resonances observed in the experimental spectrum: 5.99, 2.84, 1.61, 1.31, 1.08 and 0.95 ppm.

5. Select **Spectra** from the **Display** menu and click on the **NMR** tab. The **NMR Spectra** dialog appears.

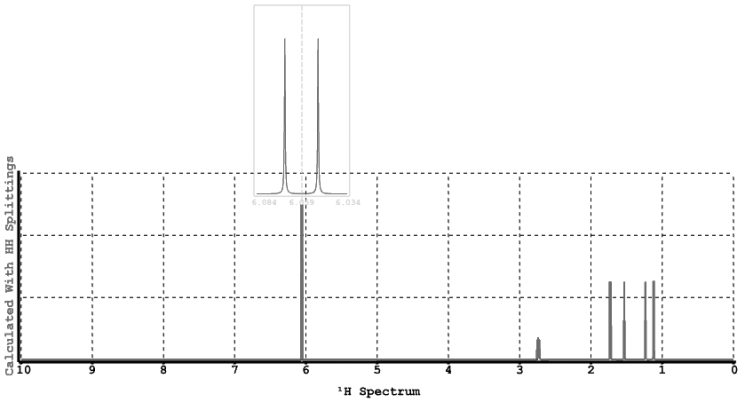
Click on **Draw Calculated With HH Splittings** (under ^{1}H Spectrum) at the top left of the dialog. The following proton spectrum appears.



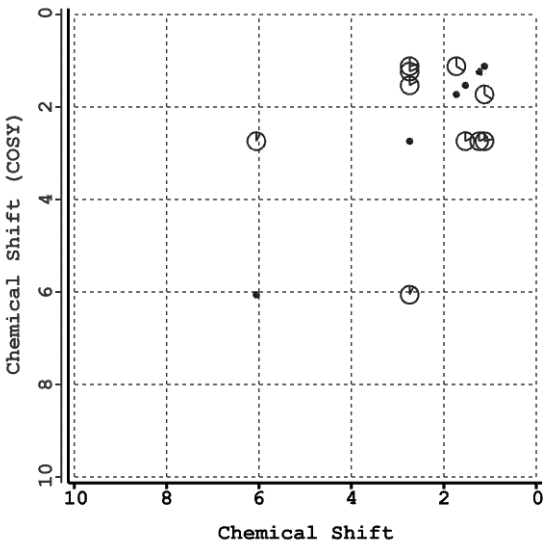
Compare the calculated proton spectrum with that in SDBS (see below or examine it on-line).



6. You may have a difficult time seeing that the lines are split (in both calculated and observed spectra). This is because the splittings are very small relative to differences in the magnitudes of the chemical shifts themselves. ***Spartan*** provides a magnifier tool to assist. With the **Properties** dialog on screen, click on the horizontal axis of the spectra plot. *Click* on the **Magnifier** button at the bottom right of the **Plot Properties** dialog that results. *Click* on the border of the green magnifier box that results to select it. Move it (as you would any graphical object) to a position directly above the resonance around 6 ppm. The single line (due to protons on the double bond) will be split in two due to interaction (coupling) with the magnetic spins of the protons on the bridgehead carbons.



7. Delete the proton spectrum (*click* on **Delete Calculated With HH Splittings** that has replaced **Draw Calculated With HH Splittings in the NMR Spectra** dialog). *Click* on **Draw COSY** at the bottom left of the dialog.



The points on the diagonal of the resulting two-dimensional plot (a COSY spectrum) indicate the six unique proton resonances (both axes are the chemical shift). The off-diagonal circles indicate interactions between neighboring protons. The magnitude of these interactions depends on the dihedral angle made by the interacting protons and the two carbons in between, and is very roughly indicated in the schematic.

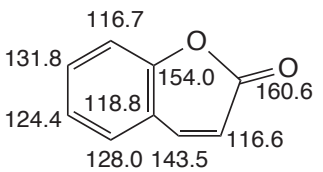
8. Close ***2-norbornene*** and any remaining dialogs when you are finished.

^{13}C NMR Spectrum of Coumarin
(cannot be completed with the Essential Edition)

There are several reasons why NMR spectroscopy, in particular ^{13}C NMR, is one of the most important analytical techniques for characterizing organic molecules. The experiment is straightforward and can be carried out rapidly. It requires relatively small samples and is non-destructive. The resulting (proton decoupled) spectrum is simple, comprising a single line (resonance) for each and every unique carbon. However, assigning ^{13}C spectra can be problematic and prone to error, in particular, where several carbons in a molecule may be in similar environments. HH and CH coupling constants that depend predictably on 3D geometry are commonly employed to assist in the one-to-one matchup of carbons to spectral lines.

Computational methods are able to provide an adequate account of ^{13}C chemical shifts, although in many cases, they are not sufficiently accurate to enable definitive assignments to be made. Part of the error is systematic and can easily be compensated for. However, there is a significant component that may not easily be taken into account.

Spartan provides the ability to empirically correct ^{13}C chemical shifts obtained from the B3LYP/6-31G* density functional model for the effects of local chemical environment. In practice, this involves a very simple linear regression where the variables are the numbers of the different kinds of directly-bonded atoms, for example, the number of sp^3 carbons. In this tutorial, you will first compare measured ^{13}C chemical shifts for coumarin with those obtained directly from B3LYP/6-31G* calculations and then correct these shifts from the empirical relationship available in ***Spartan***. Coumarin is a good example both because it is simple enough for the experimental assignments to be unambiguous, and because several of the carbons are closely related and therefore difficult to assign.



1. Build coumarin. *Click* on ▼ to the left of its name at the bottom of the screen, *check* **B3LYP/6-31G*** in the dialog that appears and *click* on **Replace**. The structure of coumarin from a B3LYP/6-31G* calculation will replace the one that you have built. You will still need to calculate its NMR spectrum as this is not available in SMD (for this molecule).

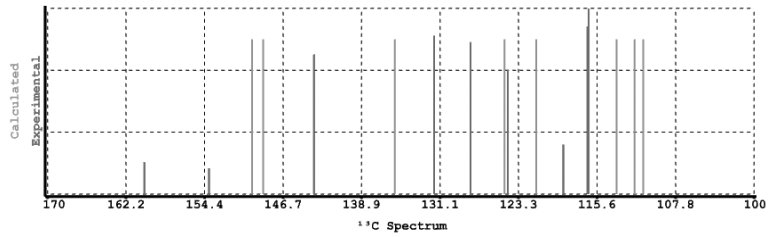
2. Bring up the **Calculations** dialog (**Setup** menu) and select **Energy** from the top menu to the right of **Calculate** and **Density Functional**, **B3LYP** and **6-31G*** from the three bottom menus. *Check* **NMR** to the right of **Compute**. *Click* on **Submit** and accept the name ***coumarin***.

3. The calculation will require several minutes. When it has completed, select **Spectra** from the **Display** menu and *click* on the **NMR** tab. The **NMR Spectra** dialog results.

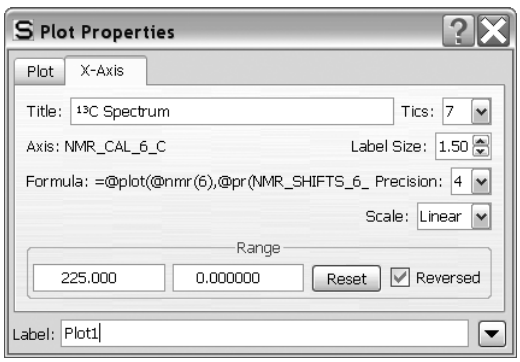
The experimental ^{13}C spectrum is available online.* Make certain that **Web site** under **Experimental Data From** at the bottom right of the dialog is checked, and *click* on **Draw Experimental** under **^{13}C Spectrum** at the right near the top of the dialog. In a few seconds, the experimental spectrum will appear. Make certain that

* Skip this and the next two steps if you are not connected to the internet.

the box to the left of **Use Corrected Shifts** is ***not checked*** (*click* inside of it if it is checked). Then, *click* on **Draw Calculated** at the top right of the dialog. A second spectrum in a different color will be superimposed onto the experimental spectrum.



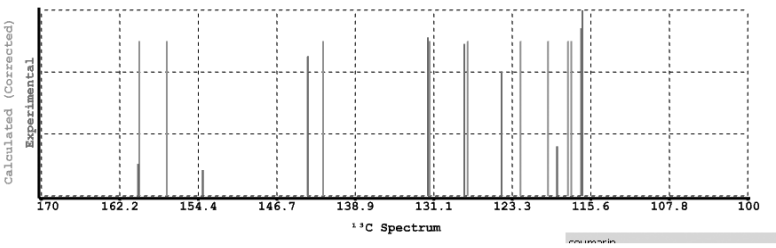
4. Change the scale of the plot to make comparison of calculated and experimental ^{13}C spectra easier (the scale in the above plot has already been changed). Select **Properties** from the **Display** menu and *click* on the horizontal plot axes. Both axes will turn gold to indicate that they are highlighted and then *click* on the **X-Scale** tab of the **Plot Properties** dialog that results (not shown) to pull up a dialog that allows you to change the scale.



Change the range from 225 to 0 ppm to 170 to 100 ppm. *Type* over each entry and *press* the **Enter** key (**return** key on Mac).

You will notice that it is very difficult to visually associate the lines in the two spectra.

5. *Click* on **Delete Calculated** inside the **NMR Spectra** dialog. *Check* the box to the left of **Use Corrected Shifts** and *click* on **Draw Calculated**. The calculated chemical shifts have now been empirically corrected for local environment.



You will notice that it is now much easier to visually associate lines in the two spectra.

Click on **Delete Calculated** and **Delete Experimental** to delete both calcualted and experimental spectra when you are done. Remove the **NMR Spectra** dialog.

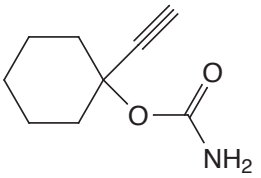
6. Select **Configure** from the **Model** menu to bring up the **Configure Labels** dialog. (If the **Labels** tab is not selected, *click* on it.) Select **Chem Shift** and *click* on **OK**. Calculated (uncorrected) chemical shifts are now attached to your model. Simplify the display by removing hydrogens (and chemical shifts attached to them). Select **Hydrogens** from the **Model** menu. (To revert to the original model with hydrogens, simply select **Hydrogens** again.)

7. Associate each of the calculated shifts with the corresponding experimental and compute a signed error. Repeat for the corrected shifts (select **Chem Shift (Cor)**) in place of **Chem Shift** inside the **Configure Labels** dialog.

8. Close ***commarin*** when you are done.

1 min

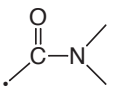
Ethinamate



Acrylonitrile, cyclohexanone and camphor are examples of molecules that adopt only a single conformation. Much more common are molecules that can adopt more than one conformation, usually due to rotation about single bonds or puckering of rings. The sedative ethinamate offers both rotation about the ring oxygen bond and the possibility for the carbonate functionality to occupy either the *equatorial* or *axial* position on the ring.

This tutorial and the one immediately following, illustrate the simplest approach to dealing with conformationally-flexible molecules, that is, finding the lowest-energy conformer (the global minimum). Tutorials in later chapters address the more difficult problem of identifying all reasonable low-energy conformers.

1. Bring up the organic model kit ( ). *Click* on **Rings**, select **Cyclohexane** from the menu and *click* anywhere on screen. *Click* on sp^3 oxygen ( ) in the model kit and add to either an *equatorial* or *axial* free valence on cyclohexane. *Click* on **Groups** and select **Amide** from the menu. Make certain that free valence on the amide group to be used for bonding is associated with carbon and not with nitrogen. This is indicated by a • on the icon at the top of the model kit.

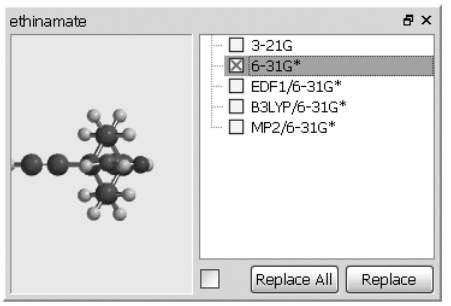


You can move the bonding position by repeatedly *clicking* on the icon. When you are satisfied, add the amide to oxygen to make the carbamate functionality. Finally, select **Alkynyl** from the **Groups** menu and add to the same carbon on the cyclohexane ring as occupied by the carbamate group (using the remaining *axial* or *equatorial* free valence). *Click* on and then on to remove the model kit.

2. Bring up the **Calculations** dialog (**Setup** menu). Select **Equilibrium Conformer** from the top menu to the right of **Calculate** and **Molecular Mechanics** and **MMFF** from the two bottom menus. You have requested that ***Spartan*** examine all possible conformers*, but then select only the most stable conformer (according to the MMFF model). *Click* on **Submit** at the bottom of the dialog. The name ***ethinamate*** will be suggested. Accept it or povide an alternative of your choosing.

3. When completed, the preferred conformer will replace the initial conformer (just like a calculated equilibrium geometry replaces a starting geometry). Which group, alkynyl or carbonate, occupies the *equatorial* position?

4. The name "ethinamate" appears at the bottom right of the screen. This indicates that the molecule exists in the Spartan Molecular Database (SMD) of structures and properties. *Click* on ⏏ immediately to the left of the name. This brings up the **SMD Preview** dialog.



Listed at the right are theoretical models for which ethanimate is available, while a ball-and-spoke model of the selected SMD entry is displayed at the left. Select **6-31G*** and *click* on **Replace**. The entry in SMD will replace your model. Is the conformation of ethanimate in the database entries the same as you have found?

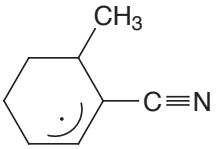
5. Close ***ethinamate*** as well as any open dialogs.

---

* This system is small, and the possible conformers will be systematically examined. This is not practical for molecules with a large number of degrees of conformational freedom. Here, ***Spartan*** employs Monte-Carlo methods.

1 min

3-Cyano-4-methylcyclohexenyl Radical



This tutorial illustrates how different models may be combined to achieve a desired result with minimum computation cost. MMFF molecular mechanics will be employed to establish equilibrium conformation, the semi-empirical AM1 model to establish equilibrium geometry and the Hartree-Fock 3-21G model to provide a basis for graphical calculations.

1. Build 3-cyano-4-methylcyclohexene. Start with **Cyclohexane** from the **Rings** menu, *click* on , and then *click* on *axial* free valences on any two adjacent carbons. Add **Cyano** from the **Groups** menu to one of the *equatorial* free valences on a carbon adjacent to the double bond. Finally, add a methyl group ( ) to the appropriate ring position using either the *equatorial* or *axial* free valence.

2. Select **Delete** from the **Build** menu (or *click* on the icon at the top of the screen), and *click* on the (*axial*) free valence on the carbon to which the cyano group is attached. *Click* on to produce a refined geometry with C_1 symmetry. *Click* on to remove the model kit.

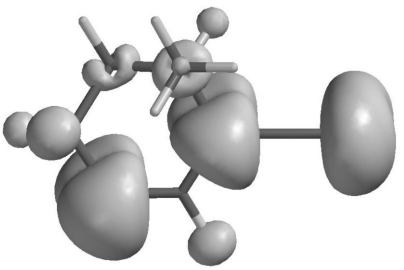
3. Select **Calculations...** (**Setup** menu). Specify **Energy** from the top menu to the right of **Calculate**, **Hartree-Fock** and **3-21G** from the two bottom menus. Next, specify **MMFF Conformer** from the **Start from** menu and **AM1** from the menu which then appears to the right. You have requested that a Hartree-Fock energy is to be obtained using an AM1 geometry, which in turn is to be based on the best conformation according to MMFF.

   This molecule has one unpaired electron. It is necessary to change **Multiplicity** from **Singlet** to **Doublet**. *Click* on **OK**.

4. Select **Surfaces** (**Setup** or **Display** menu). *Click* on **Add...**. Select **spin** from **Surface** menu and **none** from the **Property**

menu, and *click* on **Apply**. Select **density** from the **Surface** menu and **spin** from the **Property** menu. *Click* on **Apply**. You have requested two different representations of spin distribution. The first presents spin density as a surface of constant value, while the second color maps the value of the spin density onto an electron density surface. Also, request the singly-occupied molecular orbital (SOMO). Select **SOMO** from the **Surface** menu (**none** from the **Property** menu) and *click* on **OK**.

5. Submit the job. Name it ***3-cyano-4-methylcyclohexenyl radical***[*] When completed, examine the conformation. Is the methyl group *equatorial* or *axial*? Next, enter the **Surfaces** dialog. Select ***spin*** to display the spin density surface. Note that the spin density is delocalized over two of the ring carbons and onto the cyano group.



6. Again enter the **Surfaces** dialog. First, remove the spin density surface. Then, display the spin density map. Select ***density spin***. Note that the areas of maximum spin (colored blue) exactly match those where the surface is large in the previous image.

   Select **Properties** (**Display** menu) and *click* anywhere on the spin density map. This brings up the **Surface Properties** dialog. **Range** provides the range of spin density values. Note, in particular, that negative values of the spin density are possible.

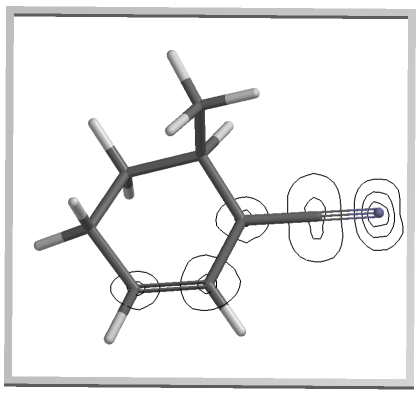
7. Enter the **Surfaces** dialog, turn "off" display of the spin density map and then select ***SOMO*** to turn "on" display of the SOMO. Note that the singly-occupied molecular orbital is nearly identical to the previously-displayed image of the spin.

---

* SMD contains very few radicals and/or charged species so a chemical name will only rarely be provided.

8. Turn "off" display of the SOMO. Next, *click* on **Add...** at the bottom of the (**Surfaces**) dialog, and select **Slice** from the **Surfaces** menu and **spin** from the **Properties** menu. *Click* on **OK**. A new line ***Slice spin*** appears in the window at the top of the dialog.* Select it.

9. A plane (a slice of spin density) surrounded by a frame appears in the middle of the model on screen. *Click* inside the frame to select. The frame will turn gold. Position the cursor outside the frame, then *press* both the **Shift** key and right button and move the mouse up and down (or use the scroll wheel). This will zoom the plane. You can also translate and rotate the plane independently of the molecule using the usual mouse operations. Alternatively, you can move the molecule and plane together by first *clicking* on the molecule (the frame will now turn white) and then using the mouse. For all operations, be certain to keep the cursor positioned outside of the frame. Size and orient the slice as you wish.**



10. Remove ***3-cyano-4-methylcyclohexenyl radical*** and any remaining dialogs from the screen.

---

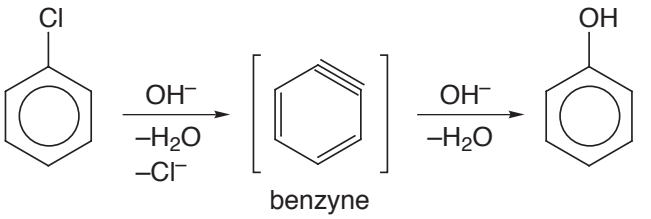
* No further calculations are involved as this is just another way of representing the spin density.

** You can change the display style from **Contours** to **Solid** or **Transparent** using the **Style** menu at the bottom right of the screen. This will appear only when the slice is selected.

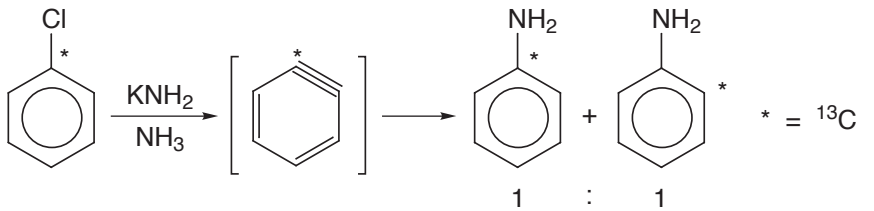
5 mins

Infrared Spectrum of Benzyne

Benzyne has long been implicated as an intermediate in nucleophilic aromatic substitution, for example,



While the geometry of benzyne has yet to be conclusively established, the results of a ^{13}C labeling experiment leave little doubt that two (adjacent) positions on the ring are equivalent.



Benzyne has reportedly been trapped in a low-temperature matrix and its infrared spectrum recorded.* Furthermore, a line in the spectrum at 2085 cm^{-1} has been assigned to the stretching mode of the incorporated "triple bond".

Here you will calculate the structure of benzyne using the Hartree-Fock 3-21G model and, following this, obtain an infrared spectrum for the molecule. Comparison with the experimental spectrum (specifically the line at 2085 cm^{-1} attributed to the C≡C stretch) should allow you to comment one way or another about the report.

1. Bring up the organic model kit, select **Benzene** from the **Rings** menu and *click* anywhere on screen. *Click* on  and *click* on two adjacent free valences on benzene. (Alternatively, hold down the **Delete** key while you *click* on the two free valences.) *Click* on  and then on  .

* O. Chapman *et al*, *J. Am. Chem. Soc.*, **95**, 6134 (1973).

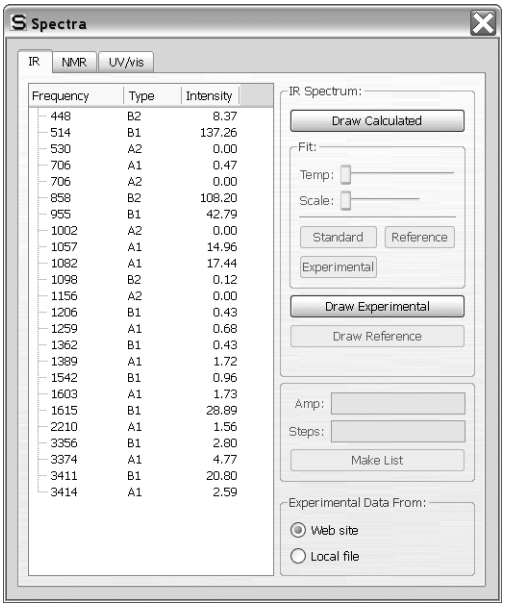
2. Enter the **Calculations** dialog (from the **Setup** menu). Select **Equilibrium Geometry** from the top menu to the right of **Calculate** and **Hartree-Fock** and **3-21G** from the two bottom menus. *Check* **IR** to the right of **Compute** in the center of the dialog. You have requested that an infrared spectrum be computed following optimization of geometry. *Click* on **Submit** and accept the name ***benzyne*** supplied to you or provide an alternative name.

3. While you are waiting for the calculation to complete, build 2-butyne, optimize its geometry with the same Hartree-Fock model and compute its infrared spectrum. Accept the name ***2-butyne***. 2-butyne will serve as a standard with which to adjust the infrared spectrum of the benzyne.

4. When both calculations have completed, examine the geometry of benzyne. Does it incorporate a triple bond? Compare the length with that in 2-butyne.

> You should have two molecules on screen, benzyne and 2-butyne. Go between them by *clicking* on them in turn. All manipulations (except scaling) and all dialog operations pertain only to the selected molecule.

5. Select (*click* on) benzyne and then select **Spectra** from the **Display** menu. *Click* on the **IR** tab in the dialog that results to bring up the **IR Spectra** dialog.



One after the other, *check* the selection boxes to the left of the frequencies in the dialog. In response, the vibrational motion associated with this frequency will be animated. Find the frequency which best fits the description of a C≡C stretch, and record its value. Do not close the **IR Spectra** dialog.

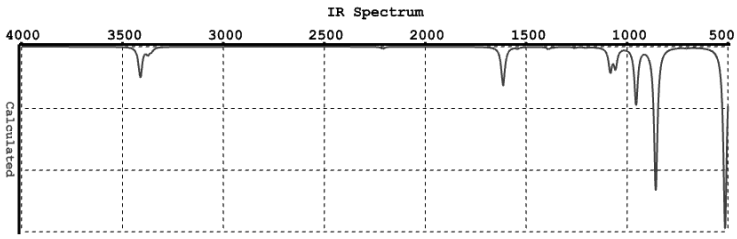
6. Select 2-butyne and find and record the frequency corresponding to the C≡C stretch. It should be on the order of 12% larger than the experimental frequency (2240 cm^{-1}) at this level of calculation.* Compute a scaling factor based on the ratio of measured to calculated frequencies, and apply this to the calculated frequency in benzyne. Is your (scaled) frequency in reasonable accord with the reported experimental value of 2085cm^{-1}?

7. Again, select benzyne and *click* on **Draw IR Spectrum** inside the **IR Spectra** dialog. The resulting spectrum** is like any other graphical object in that it can be translated and zoomed using the usual mouse commands. (It cannot be rotated.)

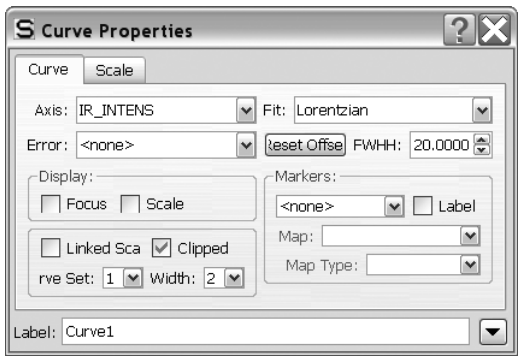
---

* Rationalization of this behavior together with performance statistics for a number of theoretical models available in ***Spartan*** is found in ***A Guide to Molecular Mechanics and Quantum Chemical Calculations*** available as a PDF under the **Help** menu.

** The calculated spectrum (frequencies and intensities) is fit to a Lorentzian function to give it a shape that is more recognizable as an infrared spectrum. Peak widths are controlled through the**TEMP** (temperature) slider bar in the **IR Spectra** dialog.



You need to *click* on the spectrum to select it. Locate the line corresponding to the C≡C stretch in the spectrum. Having done so, can you further comment one way or the other about the validity of the experimental result? To assist, bring up the **Curve Properties** dialog (**Properties** under **Display** menu) with the spectra selected and *click* inside the box to the left of **Focus**. A red circle will mark the line you have selected (*clicked* on) in the spectrum.

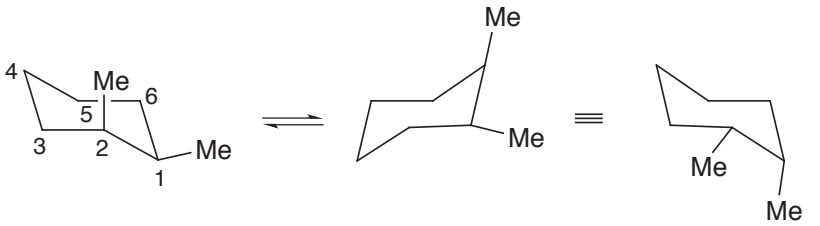


8. Close both molecules on screen together with any open dialogs.

10 mins/25mins

^{13}C Spectrum of *cis*-1,2-Dimethylcyclohexane
(cannot be completed with the Essential Edition)

At normal temperatures, the NMR spectrum of a conformationally-flexible molecule represents a (Boltzmann-weighted) average of the NMR spectra of all accessible conformers. Only when the temperature is lowered will the spectrum reveal its components, and its complexity. A simple example is *cis*-1,2-dimethylcyclohexane, a molecule with two equivalent conformers.



The room temperature ^{13}C NMR spectrum contains only four lines at 34.9, 31.9, 24.2 and 16.4 ppm relative to TMS, corresponding to an equal weighting of C_1 and C_2, C_3 and C_6, and C_4 and C_5 and the two methyl carbons, respectively. Only when the sample is cooled sufficiently to slow down motion between conformers does the spectrum reveal eight distinct lines.

In this tutorial, you will calculate the (static) ^{13}C NMR spectrum of *cis*-1,2-dimethylcyclohexane, and then average chemical shifts for the three pairs of equivalent ring carbons to yield the observed (room temperature) spectrum. Optionally, you will obtain a ^{13}C spectrum for *cis*-decalin to see to what extent it serves a model for static *cis*-1,2-dimethylcyclohexane.

1. Build and minimize *cis*-1,2-dimethylcyclohexane. Instead of calculating its equilibrium geometry, get it from the Spartan Molecular Database (SMD). *Click* on ▲ to the right of the name (***cis-1,2-dimethylcyclohexane***) at the bottom of the screen to bring up the **SMD Preview** dialog. *Check* **B3LYP/6-31G*** and *click* on the **Replace** button at the bottom of the dialog. The entry in SMD will replace the molecule you built. Inside the **Calculations** dialog, specify calculation of **Energy** using the **Density Functional B3LYP 6-31G*** model. *Check*

NMR to the right of **Compute**, and submit the job. Accept the name ***cis-1,2-dimethylcyclohexane***.

2. After the calculations have completed (several minutes), record ***corrected*** chemical shifts* for all ring carbons in *cis*-1,2-dimethylcyclohexane, either from the **Atom Properties** dialog (select **Properties** from the **Display** menu and *click* on a carbon) or by labeling the model (select **Configure...** from the **Model** menu and *check* **Chem Shift (Cor)** under **Atom** in the dialog that results). In the latter case, you may want to remove the hydrogens (*click* to deselect **Hydrogens** from the **Model** menu) to simplify the display. Finally, average the calculated chemical shifts to obtain a room temperature ^{13}C spectrum, and compare with measured values.

3 to 5 optional

3. *cis*-Decalin is a plausible model for static *cis*-1,2-dimethylcyclohexane, in that it is unable to undergo (rapid) ring inversion.

Build *cis*-decalin. Start with cyclohexane, add sp^3 carbons at adjacent *cis* free valences and connect these carbons by two additional sp^3 groups. Again, retrieve the molecule from SMD. *Click* on ▲ to the right of the name (***cis-decalin***) at the bottom of the screen to bring up the **SMD Preview** dialog. *Check* **B3LYP/6-31G*** and *click* on the **Replace** button at the bottom of the dialog.

---

* ***Spartan*** provides "corrected" ^{13}C chemical shifts from B3LYP/6-31G* calculations (only) to account for local environment (in addition to directly calculated shifts). In practice, this involves a linear regression using the numbers and kinds of neighboring atoms.

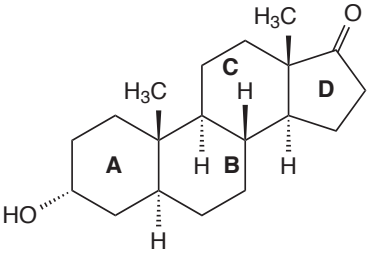
4. Enter the **Calculations** dialog. Specify an **Energy** calculation using the **Density Functional B3LYP 6-31G*** model. *Check* **NMR** to the right of **Compute** and *click* on **Submit** and accept the name ***cis-decalin***.

5. The calculation will require several minutes. When completed, select **Spectra** from the **Display** menu and *click* on the **NMR** tab. *Click* in the box to the left of **Use Corrected Shifts** and then *click* on **Draw Calculated** under **^{13}C Spectrum**. Because of the (C_2) symmetry of *cis*-decalin, your spectrum will contain only five lines, all of equal intensity. Relate these to the six lines corresponding to the ring carbons in static *cis*-1,2-dimethylcyclohexane. Pay particular attention to the differences in chemical shifts between carbons which average in the equilibrating molecule. Is *cis*-decalin a good model for *cis*-1,2-dimethylcyclohexane?

6. Close all molecules and any open dialogs.

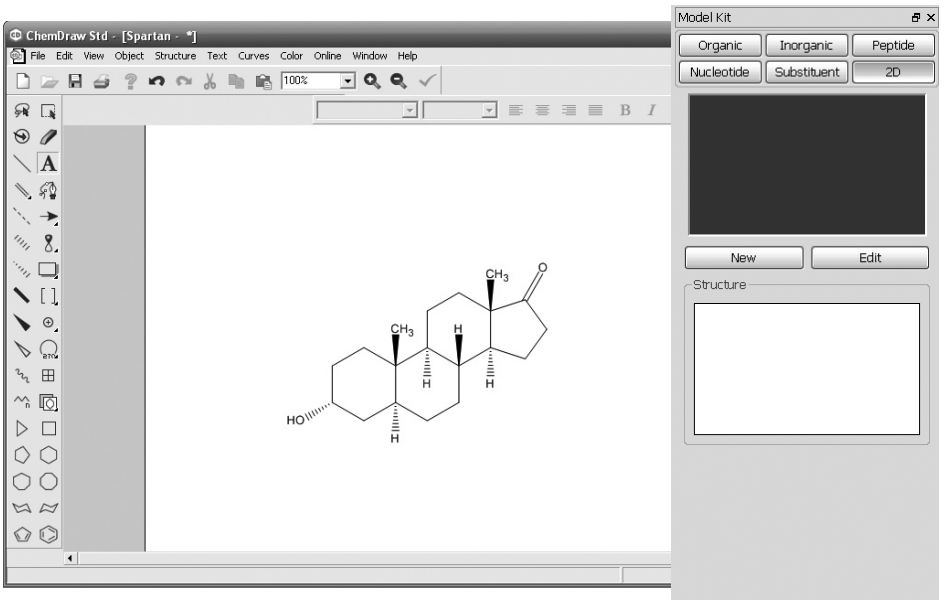
5 mins

Androsterone



The steroid androsterone illustrates the use of ChemDraw directly from ***Spartan***. Also, the molecule provides a good opportunity to compare space-filling models and electron density surfaces as metircs of molecular size and shape.

1. *Click* on and then *click* on the **2D** tab. *Click* on the **New** button to open ChemDraw.* Draw androsterone (pay close attention to stereochemistry) and close ChemDraw. You will now see the ChemDraw image of androsterone at the bottom of the 2D model kit and a 3D model at the top. *Click* in the workspace. The name ***androsterone*** will appear in the lower right hand corner. Click on ( ).



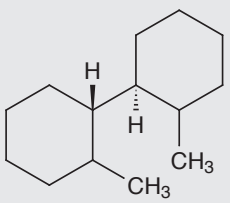
* The 2D model kit in ***Spartan*** will automatically open ChemDraw if it is available on your machine (**NAM**, **NAL**). If you do not have access to ChemDraw, contact CambridgeSoft (www.cambridgesoft.com). While ***Spartan*** for Macintosh and Linux cannot automatically access ChemDraw, both programs can open 2D molecule files in .CDX or .SKC formats.

If you are not running ***Spartan*** for Windows, or if you do not have ChemDraw installed on your machine, follow the directions below to build androsterone.

Click on . Select **Cyclohexane** from the **Rings** menu, and *click* anywhere on screen to create the A ring. **Cyclohexane** is still selected. The icon at the top of the model kit should indicate **eq** (*equatorial*) rather than **ax** (*axial*); attachment needs to be made using an *equatorial* free valence. If not, *click* on the icon to switch to **eq**. *Click* on an *equatorial* free valence of the **A** ring to join it to the **C** ring. You have made bicyclohexane. Its name will appear at the bottom right of the screen.

Rotate around the bond connecting the **A** and **C** rings to position the *axial* free valences on the connecting carbons opposite (180°) to each other. (This bond should be encircled by a red marker. If it is not, *click* on the bond.) To do so, hold down both the **Alt** key and the left button and *drag* the mouse up and down. *Equatorial* free valences on the cyclohexane carbons adjacent to those that connect the two rings are now positioned to form the **B** ring (in a chair geometry).

Select sp^3 carbon ( ) and *click* on the two opposing *equatorial* free valences on the rings you have just joined. The resulting structure should appear as follows.



Click on and, one after another, *click* on the (*equatorial*) free valences that nearly overlap to create the **B** ring. *Click* on . Check to see that all rings are chairs and that the stereochemistry is as in the figure in the previous page. If not, you can start over (**Clear** from the **Edit** menu), or undo the last operation (**Undo** from the **Edit** menu).

Next, form the **D** ring starting from the carbon that will become a carbonyl group. Select **Carbonyl** from the **Groups** menu and *click* on the appropriate free valence on the **C** ring. Select sp^3 carbon ( ) from the model kit and *click* first on the single free valence on the carbonyl carbon and then on a free valence of the sp^3 carbon just added. *Click* on and join the free valence from the second sp^3 carbon added in the previous step and the appropriate *equatorial* free valence on the **C** ring.

Click on . You have made one of the forms of perhydrophenanthrene and its name will appear at the bottom right of the screen.

Add two methyl groups. Select sp^3 carbon ( ) from the model kit. *Click* on the *axial* free valence at the appropriate juncture of the **C** and **D** rings, and then on the *axial* free valence at the proper juncture of the **A** and **B** rings. Add the hydroxy group. Select sp^3 oxygen ( ) from the model kit, and *click* on the appropriate *axial* position on the **A** ring. *Click* on and then on . The name adrosterone should appear at the bottom right of the screen.

Androsterone incorporates several chiral centers. To assign them as R or S, select **Configure...** (**Model** menu) and *check* **R/S** under **Atom** in the dialog that appears. *Click* on **OK**. R/S labels will appear next to each of the chiral centers. (To remove the labels, select **Labels** from the **Model** menu.)

2. Select **Calculations...** from the **Setup** menu and specify calculation of an energy using the **Hartree-Fock 3-21G** model. Select **AM1** from the menu to the right of **StartFrom**. *Click* on **Submit** and accept the name ***androsterone***.

3. When completed (several minutes), select **Properties** (**Display** menu). Write down the surface area of a space-filling (CPK) model given in the **Molecule Properties** dialog.

4. Select **Surfaces** (**Setup** or **Display** menu). *Click* on **Add**. Select **density** from the **Surface** menu, **none** from the **Property** menu and **high** from the **Resolution** menu, and *click* on **OK**.

5. When completed, *click* on ***density*** in the **Surfaces** dialog. *Click* on any portion of the density surface. The **Surface Properties** dialog which replaces the **Molecule Properties** dialog provides both the surface area and the volume of the electron density surface. The second box to the right of **IsoVal** contains the percentage of the total number of electrons that are enclosed by the density surface. Adjust this value (type over the existing value and *press* **Enter**) such that the surface area of the density surface is nearly the same as that for the CPK model. What is the percentage enclosure?

6. Remove ***androsterone*** from the screen.

Searching *Spartan's* Infrared Spectral Database
(cannot be completed with the Essential Edition)

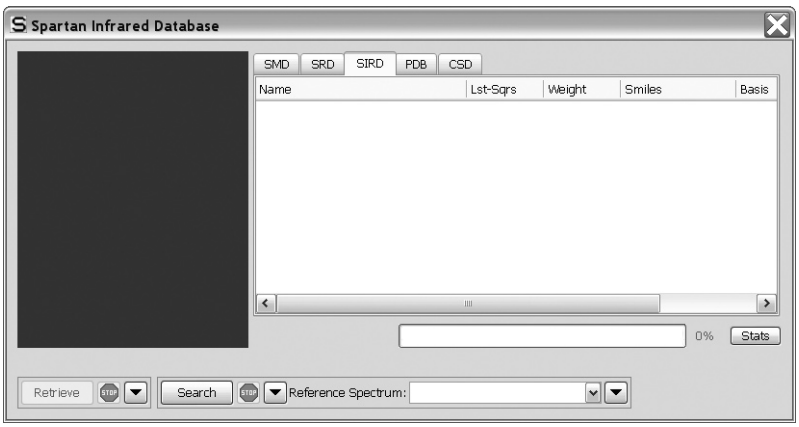
Pattern matching ("fingerprinting") of a measured infrared spectrum to experimental spectra contained in a library (database) has been common practice for many years. This has not previously been extended to comparisons of measured spectra to databases of calculated spectra. The likely reason is that the results of calculations (a set of vibrational frequencies and intensities to go along with these frequencies) do not "look like" experimental infrared spectra, at least, they do not look like spectra obtained at normal temperatures. However, a fit of the calculated data to a Lorentzian function in which peak width at half height is treated as a parameter (one value for all peaks) makes the two visually quite similar. One could argue that this parameter loosely corresponds to temperature, which is consistent with the fact that infrared spectra measured at low temperature comprise a series of sharp lines (rather than bands).

Frequency is proportional to (the square root of) the second derivative of the energy at a minimum in the potential surface for a particular vibrational coordinate, that is, the curvature of the surface. Practical frequency calculations replace the real potential at the energy minimum by a quadratic function. This leads to a surface that is too steep (and an energy second derivative that is too large), meaning that calculated frequencies will be larger than measured frequencies. For the most part, the error is systematic, with calculated frequencies being between 3% and 15% larger than measured frequencies, depending on the theoretical model. This can be compensated for by incorporating a single linear scaling parameter into the fitting function.

In summary, ***Spartan*** fits calculated and measured infrared using a Lorentzian function that incorporates two parameters, a non-linear parameter accounting for peak width and a linear scaling parameter.

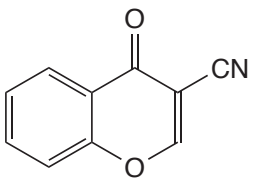
This tutorial illustrates the way in which searches of the Spartan Infrared Database (SIRD) are setup and performed and the results examined. It lets you choose from a small selection of measured infrared spectra (or supply your own spectra as a .dx file).

1. Select **Databases** from the **Search** menu (or *click* on). It is not necessary to have any molecules open on screen. *Click* on the **SIRD** tab in whatever database dialog results to bring up the **IR Spectra Database** dialog.

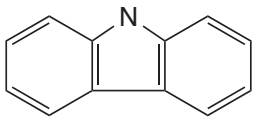


Click on ▾ to the far right of **Reference Spectrum** at the bottom of the dialog. Navigate to the sub directory ***expt files*** under the ***tutorials*** directory.*

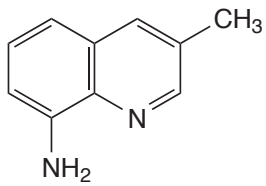
Select one of the following molecules and *click* on **Open**.



chromone-3-carbonnitrile



dibenzopyrrole

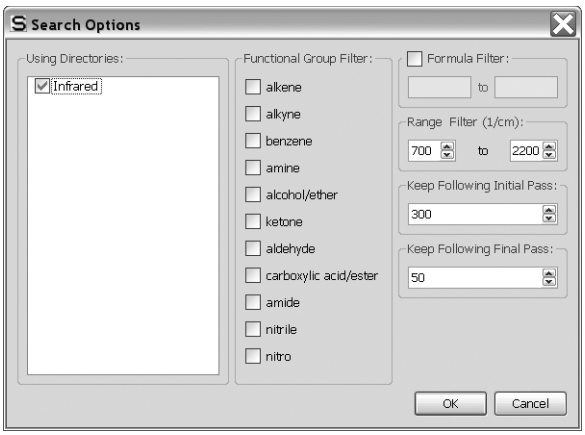


8-amino-2-methylquinoline

Its name will appear in the box to the right of **Reference Spectrum**.

2. *Click* on ▾ to the right of **Search** (to the left of **Reference Spectrum**) to bring up **Search Options** dialog.

* For Windows, this is found in ***Program Files/Wavefunction/Spartan08***. For Linux, this is found in the directory where ***Spartan*** was installed. For Macintosh, this is located in the Spartan.08.full folder on the installation CDROM.



Set the **Range Filter** from **700** to **2200** cm^{-1}, **Keep Following Initial Pass** to **300** and **Keep Following Final Pass** to **50**. The first limits the range of the fit, the second limits the number of entries passed between the first (rough) search and the final search (the larger this number, the less likely the search will miss a "hit", but the more time the search will take) and the third limits the "hits" returned. You can if you wish set the appropriate **Functional Group Filter**, for example, identify **nitrile** for chromone-3-carbonnitrile. *Click* on **OK**.

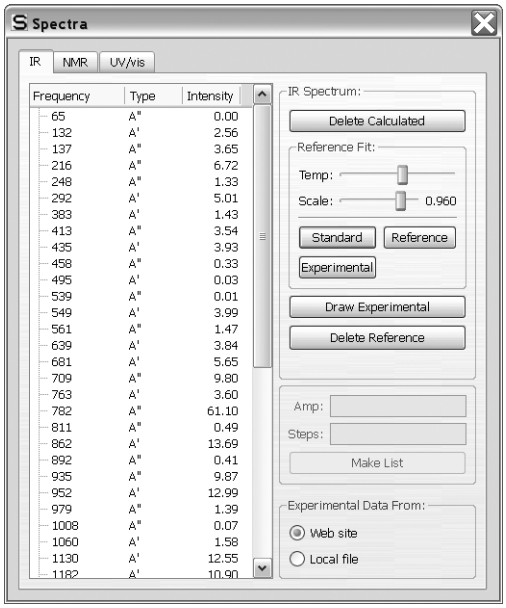
3. *Click* on the **Search** button at the bottom of the **IR Spectra Database** dialog. The search will require a minute or two. When it has completed, scan the list of hits in the window at the top of the dialog for the name of your query. It will be at or near the top for all the molecules provided.* *Click* on this entry. *Click* on ▼ to the right of the **Retrieve** button (to the left of the **Search** button), and make certain that **Current Document** is selected. Close the dialog and *click* on **Retrieve**. Close the **IR Spectra Database** dialog (*click* on ☒ at the top).

4. Select **Spectra** from the **Display** menu and *click* on the **IR** tab to bring up the **IR Spectra** dialog. The data actually correspond to calculations performed during the EDF2/6-31G* density

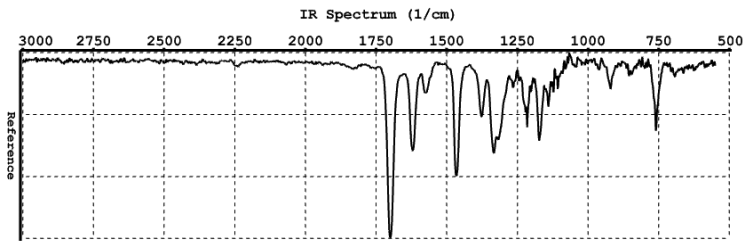
---

* This will not always be the case. The obvious exception occurs where the query is not in the database. Problems will also occur with the (query) infrared spectrum lacks sufficient "structure" to distinguish it from similar molecules.

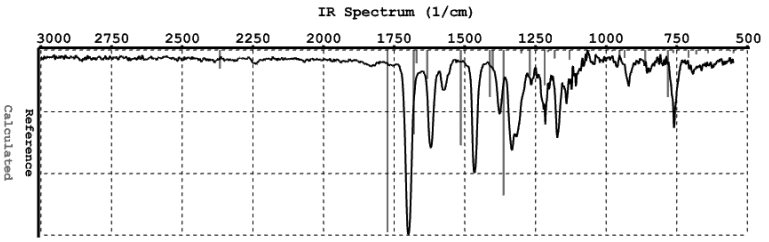
functional model with best conformation assigned from the T1 model.



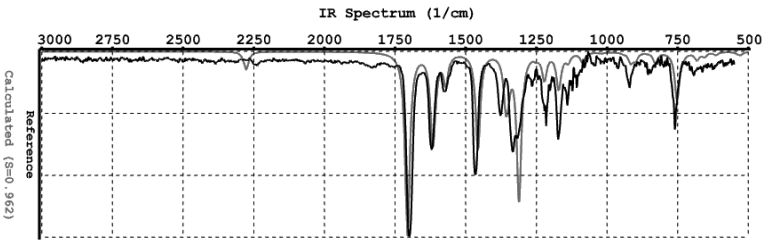
Click on **Draw Reference** in the center right of the dialog to draw the reference (query) spectrum. It will range from 500 to 3000 cm^{-1} despite the fact that spectra matching has been carried out over a more limited range.



Click on **Draw Calculated** at the top right of the dialog. This superimposes the "raw data" from the calculations onto the experimental data. Visually, the two "spectra" are quite different.



This is because the peak width is set to zero ("low temperature") and no scaling connection has been applied to account for systematic error. *Click* on **Reference** (under **Fit** near the top riht of the dialog). This applies the two parameter Lorentzian fit to the calculated spectrum. Note that the two spectra are now visually quite similar.



5. Close the query molecule and all dialogs when you are done.

Chapter 5
Groups of Organic Molecules

This chapter introduces and illustrates a number of basic operations involved in processing groups of molecules, as well as the associated spreadsheet for organizing and fitting data and facilities for making plots.

Computational investigations like experimental investigations are rarely restricted to single molecules, but typically involve a series of related molecules. Here, it may be of interest to compare geometries, energies or other calculated properties, or to compare trends in calculated and measured properties. ***Spartan*** provides facilities for these purposes. In particular, it allows molecules to be grouped, either manually, or automatically as a result of a conformational search, from following a particular vibrational motion, or from a scan of one or more geometrical variables. Once put into a group, molecules may be aligned based either on their structures or chemical functionalities. Calculations may be performed either on individual molecules or, just as simply, on the complete group of molecules. The results of calculations on a group can be examined and analyzed individually or alltogether to seek out trends.

Associated with a group (including a group of one molecule) is a spreadsheet. This allows convenient access to virtually any calculated quantity that can be given a numerical value. Additionally, data may be entered manually into the spreadsheet or transferred from another application such as Excel. Data in the spreadsheet may be manipulated, linear regression analyses performed and both 2D and 3D plots constructed. Alternatively, the data in a ***Spartan*** spreadsheet may be transferred to Excel for further analysis. The Excel worksheet can then be embedded in the ***Spartan*** document (see **Embedded Data** under the **File** menu; **Chapter 10**).

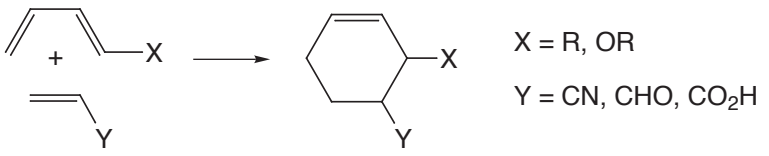
The tutorials in this chapter introduce and illustrate a number of the basic group operations available in ***Spartan***. These include

building a group from scratch, and processing groups resulting from a conformational search and from varying a torsion angle. Also provided is an example of fitting an experimental observable to one or more calculated properties by way of linear regression. The use of the **Reactions** dialog (new to ***Spartan'08***) is illustrated both as it operates on molecules in a group and on molecules derived from substitution of grouped molecules.

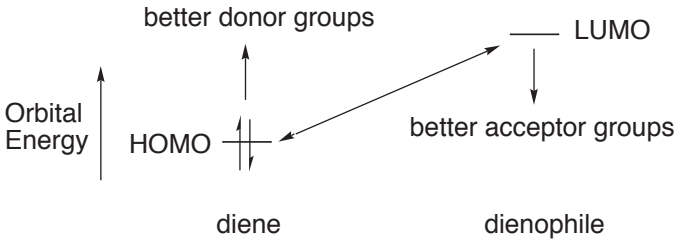
2 mins

Dienophiles in Diels-Alder Cycloadditions

The most common Diels-Alder reactions involve electron-rich dienes and electron-deficient dienophiles.

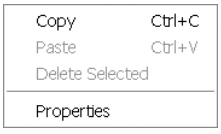


The rate of these reactions generally increases with increasing π-donor ability of the diene substituent, and with increasing π–acceptor ability of the dienophile substituent. This can be rationalized by noting that donor groups raise the energy of the highest-occupied molecular orbital (HOMO) on the diene, while acceptor groups lower the energy of the lowest-unoccupied molecular orbital (LUMO) on the dienophile. Thus, the HOMO-LUMO gap is reduced, leading to enhanced stabilization.



To test such an hypothesis, you will examine the extent to which experimental relative rates of Diels-Alder cycloadditions involving cyclopentadiene and a variety of cyanoethylenes correlate with dienophile LUMO energies.

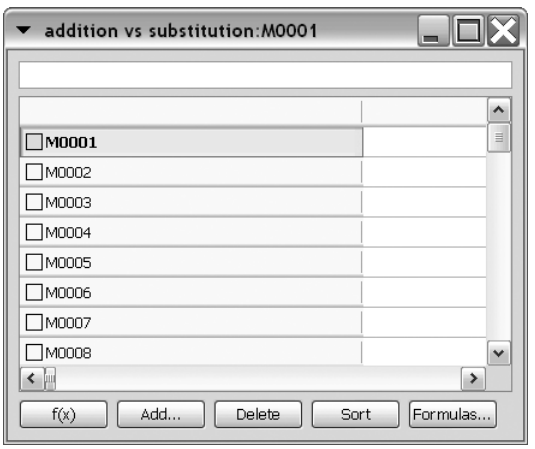
1. Build acrylonitrile, H_2C=C(H)CN, (select **New** from the **File** menu or *click* on), minimize ( ) and place it on the clipboard. *Right click* on the background and choose **Copy** from the resulting menu.



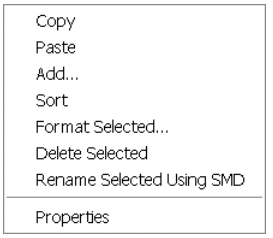
2. Select **New Molecule** (not **New**) from the **File** menu. This requests that a molecule that has yet to be built is to be appended onto the end of a group of molecules in which the currently selected molecule (acrylonitrile) is a member. The screen will be cleared. *Click* on **Clipboard** at the bottom of the model kit and *click* on screen. Acrylonitrile will appear. *Click* on **Groups** in the model kit, select **Cyano** and add to the appropriate free valence on acrylonitrile to make 1,1-dicyanoethylene. *Click* on .

3. Repeat this procedure (**New Molecule**, followed by **Clipboard**, followed by **Groups**, followed by) four more times to build *cis* and *trans*-1,2-dicyanoethylene, tricyanoethylene and tetracyanoethylene. When you are all done (six molecules in total), *click* on to remove the model kit.

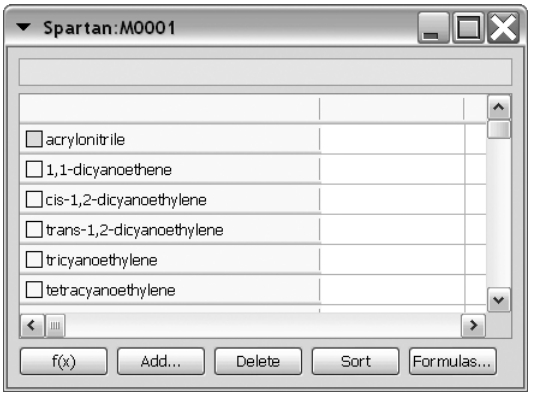
4. The molecules have been put into a single group, allowing calculated properties to be accessed via a spreadsheet. Select **Spreadsheet** from the **Display** menu.



To select individual molecules, *click* on their labels (**M0001**, **...**) in the left hand column, or use the and keys at the bottom left of the screen. Position the cursor in the header field above the left-hand column, *right click* and select **Rename Selected Using SMD** in the menu that results.



This results in replacement of the default labels by proper names in the Spartan Molecular Database.



5. Enter the **Calculations** dialog (**Calculations** under the **Setup** menu), select **Energy** using the **Hartree-Fock 3-21G** model, and **AM1** from the **Start from** menu. A 3-21G energy calculation will be preceded by calculation of geometry using the AM1 model. Make certain that **Global Calculations** is *checked* to ensure that the calculations are to be applied to all the molecules.

6. Submit the job. Name it ***Diels-Alder dienophiles***. After it completes, enter experimental relative rates into the spreadsheet.

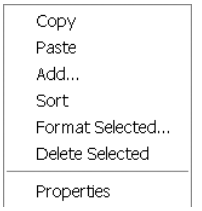
Experimental relative rates for Diels-Alder cycloadditions of cyclopentadiene*

dienophile	$\log_{10}$ (relative rate)	dienophile	$\log_{10}$ (relative rate)
acrylonitrile	0	1,1-dicyanoethylene	4.64
trans-1,2-dicyanoethylene	1.89	tricyanoethylene	5.66
cis-1,2-dicyanoethylene	1.94	tetracyanoethylene	7.61

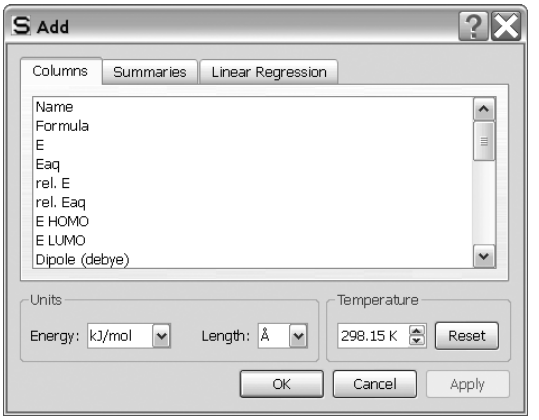
* J. Sauer, H. Weist and A. Mielert, *Chem. Ber.*, **97**, 3183 (1964).

Double click on the header cell for a blank column in the spreadsheet, type **Log(rate)=** and *press* the **Enter** (**return**) key. Enter the rate data in the cells below (*press* **Enter** (**return**) key after each entry). Sort by relative rate. *Click* on the column header **Log(rate)**, and then *click* on **Sort** at the bottom of the spreadsheet.

7. *Click* on a header cell for a blank column and *click* on **Add...** at the bottom of the spreadsheet. Alternatively, *right click* inside the spreadsheet, and select **Add...** from the menu that results.



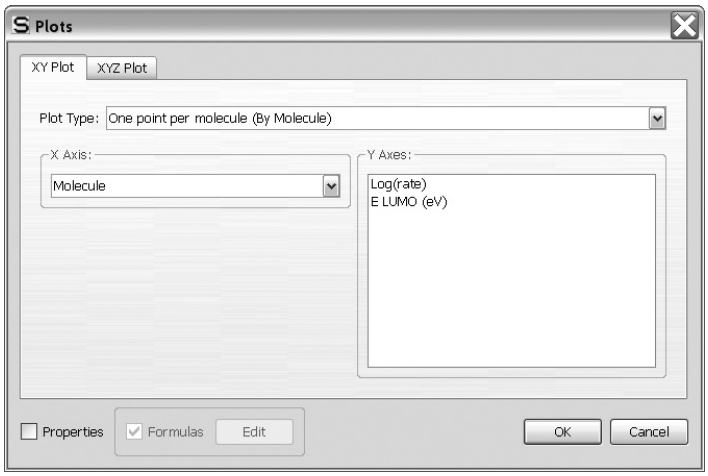
Select **E LUMO** from the list in the dialog that results and **eV** from the **Energy** menu. *Click* on **OK**.



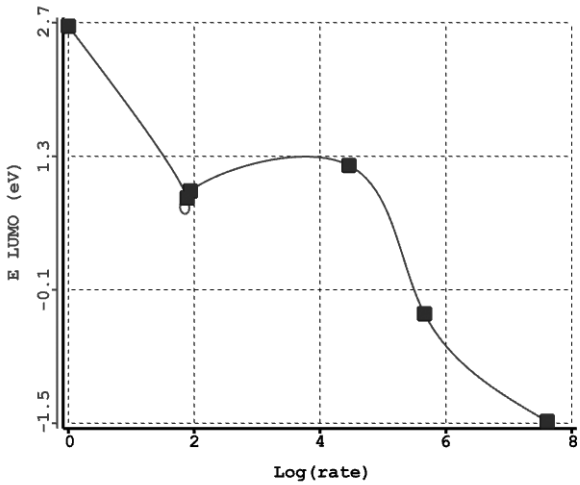
Two data columns are added to the spreadsheet.

The spreadsheet has now served its purpose by collecting both calculated (LUMO energy) and experimental (relative rate) data for all six dienophiles. You can, if you wish, now remove it from the screen (*click* on ☒ at the top).

8. Select **Plots...** from the **Display** menu. This leads to the **Plots** dialog.

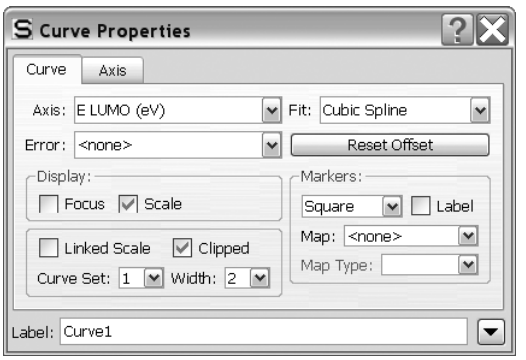


Make certain that the **XY Plot** tab at the top of the dialog is selected. Select **Log(rate)** from the list of items in the **X Axis** menu and **E LUMO(eV)** from the **Y Axes** list. *Click* on **OK** to remove the dialog and show the plot.

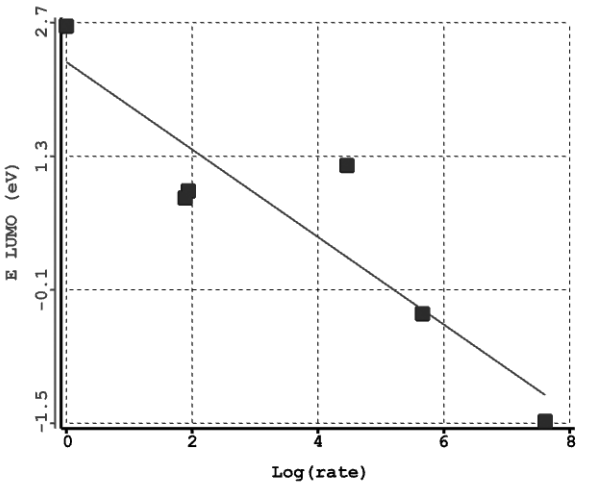


The default curve is a so-called cubic spline that smoothly connects the data points. To establish a correlation, use a simple

linear least-squares fit. Select **Properties** from the **Display** menu and *click* anywhere on the curve that you have just drawn. The **Curve Properties** dialog appears.



Select **Linear LSQ** from the **Fit** menu. A new plot appears.

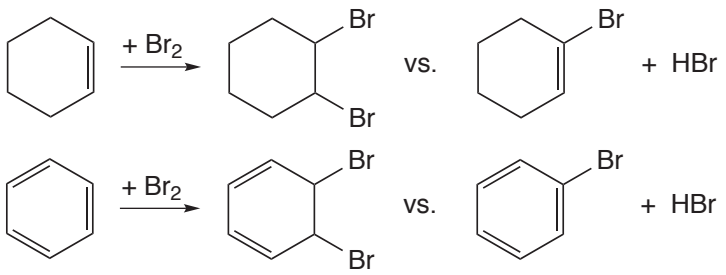


Note the correlation between relative rates and LUMO energy.

9. Remove all molecules and any remaining dialogs from the screen.

Addition vs. Substitution

Alkenes normally undergo addition reactions whereas aromatic compounds normally undergo substitution reactions. For example, bromine reacts with cyclohexane to give *trans*-1,2-dibromocyclohexane (the addition product) not 1-bromocyclohexene, whereas it reacts with benzene to give bromobenzene (the substitution product) not *trans*-5,6-dibromo-1,3-cyclohexadiene.

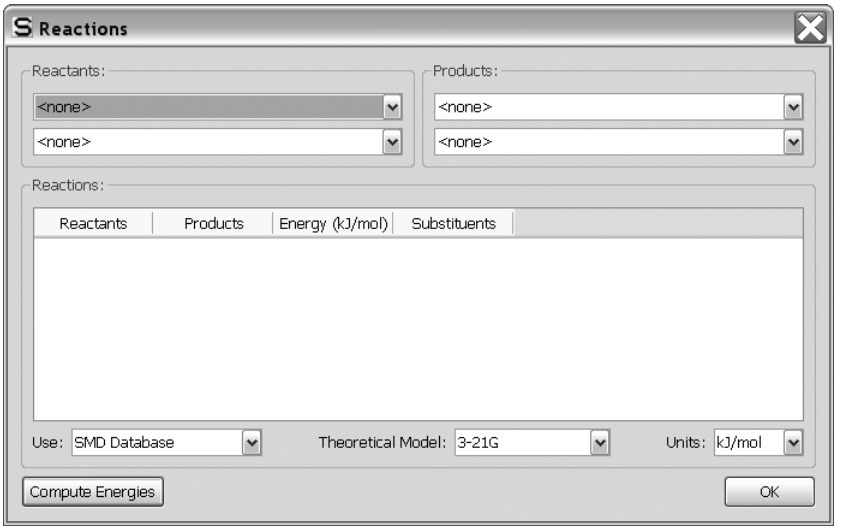


In this tutorial you will use results from the Spartan Molecular Database as accessed from ***Spartan's*** reaction energy calculator in an effort to establish the change in preferred product. Specifically, heats of formation from the T1 thermochemical recipe will be employed. No calculations are involved.

1. One after the other, build cyclohexene, *trans*-1,2-dibromocyclohexane, 1-bromocyclohexene, benzene, *trans*-5,6-dibromo-1,3-cyclohexadiene, bromobenzene, bromine (Br_2) and hydrogen bromide (eight molecules in total). Put all in the same document. Use **New** from the **File** menu ( ) to enter the builder for the first molecule, and **New Molecule** (**File** menu) for each successive molecule.

2. Select **Spreadsheet** from the **Display** menu. *Right click* inside the header cell for the leftmost column and select **Rename Selected Using SMD** in the menu that appears. *Click* on **OK** in the informative dialog that results. The molecules can now be referenced using chemical names. Close the spreadsheet (*click* on at the top).

3. Select **Reactions** from the **Display** menu.



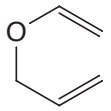
Choose **SMD Database** from the menu to the right of **Use** and **T1** from the menu to the right of **Theoretical Model** at the bottom of the dialog. Compute the energy for Br_2 addition to cyclohexene: select ***cyclohexene*** and ***bromine*** as **Reactants** and ***trans-1,2-dibromocyclohexane*** (and ***<none>***) as **Products** and *click* on the **Compute Energies** button at the bottom left of the dialog. Record the reaction energy. Repeat for the corresponding substitution energy (same reactants but the products are 1-bromocyclohexene and hydrogen bromide) and for both addition and substitution reactions of benzene (reactants are ***benzene*** and ***bromine*** and products are ***trans-5,6-dibromomo-1-3-cyclohexadiene*** and ***<none>*** for addition and ***bromobenzene*** and ***hydrogen bromide*** for substitution).

Are all reactions thermodynamically favorable (*exothermic*)? Identify any reactions that are not and provide a rationale as to why they are not. Why is there a change in preferred reaction in moving from the alkene to the arene?

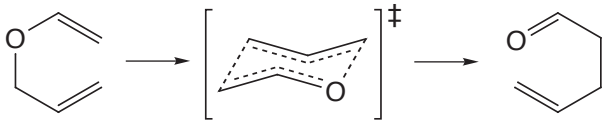
4. Close the (unnamed) document and any open dialogs.

5 mins

Allyl Vinyl Ether



Allyl vinyl ether undergoes Claisen rearrangement, the mechanism presumes a chair arrangement of the reactant.



Is this the lowest-energy conformer (global minimum) or is additional energy required to properly orient the molecule for reaction? To find out, you need to locate all the conformers of allyl vinyl ether, identify the best chair structure and evaluate its energy relative to that of the actual global minimum. You will first carry out a conformational search using the MMFF molecular mechanics model, and then obtain relative conformer energies using single-point 6-31G* Hartree-Fock calculations based on 3-21G Hartree-Fock equilibrium geometries.

1. Bring up the organic model kit and build allyl vinyl ether from a sequence of sp^2 and sp^3 carbons and an sp^3 oxygen. *Click* on E to produce a refined geometry. *Click* on V.

2. Bring up the **Calculations** dialog and select **Conformer Distribution** from the top menu to the right of **Calculate** and **Molecular Mechanics** and **MMFF** from the two bottom menus. *Click* on **OK**.

3. Submit the job. Accept the name ***allyl vinyl ether***. When completed, it will give rise to a series of low-energy conformers* placed in a new file ***allyl vinyl ether.Conf.M0001***.** A prompt will ask you if you want to open this file. *Click* on **OK**.*** Select **Spreadsheet** from the **Display** menu.

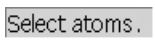
---

* By default, only conformers within 40 kJ/mol of the global minimum will be kept. This can be changed (see **Conformational Search** in **Appendix C**).

** **M0001** is the default label of the molecule you built. You can change it by altering the **Labels** field in the **Molecule Properties** dialog (**Properties** under the **Display** menu; **Chapter 15**).

*** To avoid confusion, it is a good idea to close the original file ***allyl vinyl ether***.

Size the spreadsheet such that all rows (corresponding to different conformers) are visible at one time. *Click* on **Add...** at the bottom of the spreadsheet. Select **E** from the list at the top of the dialog which appears, **kJ/mol** from the **Energy** menu and *click* on **OK**. Energies for each of the different conformers will be added to the spreadsheet. Examine the lowest-energy conformer (the top entry). Is it in a chair conformation suitable for Claisen rearrangement? If not, identify the lowest-energy conformer that is suitable. You can keep two or more conformers on screen at the same time by *checking* the boxes immediately to the left of the molecule labels (the leftmost column) in the spreadsheet. To get a clearer idea of structural similarities and differences, align the conformers. Select **Align** from the **Geometry** menu (or *click* on the  icon at the top of the screen) and select **Structure** from the **Align by** menu at the bottom right of the screen. A message will appear at the bottom left of the screen.



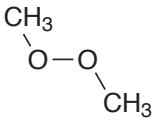
Click on the two carbons of the vinyl group and on the oxygen to designate them as alignment centers. Each will be marked by a red circle. If you make a mistake, *click* on the circle and it will disappear. When you are done, *click* on the **Align by** button at the bottom right of the screen. *Click* on  .

4. To obtain a better estimate of conformational energy differences, perform 6-31G* energy calculations using 3-21G geometries. Make a copy of ***allyl vinyl ether.M0001*** (), and name it ***allyl vinyl ether Hartree-Fock***. Using the copy, delete all conformers except the global minimum and the lowest-energy conformer poised for Claisen rearrangement. Select each conformer to be discarded, and then select **Delete Molecule** from the **File** menu. Alternatively, *click* on the cell in the spreadsheet containing the label for the molecule to be deleted, and then *click* on **Delete** at the bottom of the spreadsheet. When you are done, the spreadsheet should contain only two entries.

5. Enter the **Calculations** dialog, and specify calculation of an energy using the **Hartree-Fock 6-31G*** model. Also, specify **3-21G** from the **Start from** menu. Make certain that **Global Calculations** at the bottom of the dialog is checked to specify that the dialog settings apply to both conformers. *Click* on **Submit**.
6. When completed, examine the conformer energies. How much energy is needed to go from the global minimum to a conformer poised to undergo Claisen rearrangement?
7. Remove all molecules and any remaining dialogs from the screen.

15 mins/25 mins

Internal Rotation in Dimethylperoxide



Quantum chemical calculations, in particular, Hartree-Fock molecular orbital calculations, density functional calculations and MP2 calculations, may be called on to furnish data to parameterize empirical energy functions for use in molecular mechanics and/or molecular dynamics calculations. Most important are data relating to torsional motions, for it is here that experimental data are most scarce. Note that the empirical energy function needs to reflect the inherent periodicity. One suitable choice is a truncated Fourier series.

$$E^{torsion}(\omega) = k^{torsion1}(1 - \cos(\omega - \omega^{eq})) + k^{torsion2}(1 - \cos 2(\omega - \omega^{eq})) + k^{torsion3}(1 - \cos 3(\omega - \omega^{eq}))$$

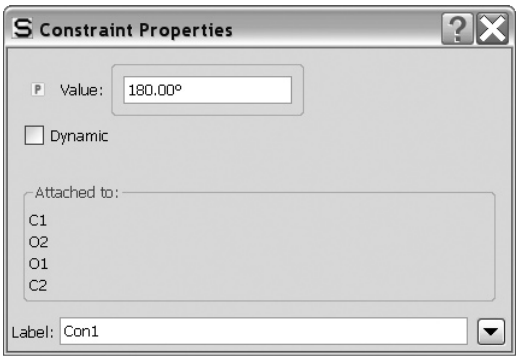
Here, ω^{eq} is the ideal dihedral angle and $k^{torsion1}$, $k^{torsion2}$ and $k^{torsion3}$ are parameters. The first (one-fold) term accounts for the difference in energy between *syn* and *anti* conformers, the second (two-fold) term for the difference in energy between planar and perpendicular conformers, and the third (three-fold) term for the difference in energy between eclipsed and staggered conformers.

In this tutorial, you will generate a potential energy function for rotation about the oxygen-oxygen bond in dimethylperoxide using 6-31G* Hartree-Fock calculations (and optionally from B3LYP/6-31G* density functional calculations) and then fit this potential to a truncated Fourier series.

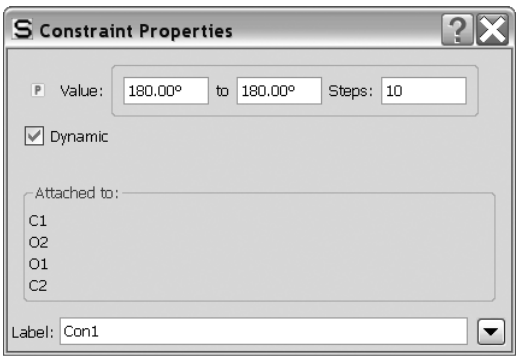
1. Build dimethylperoxide. If the molecule is not already in an *anti* conformation, *click* on  and set the COOC dihedral angle to **180** (180°) by *typing* **180** in the box at the lower right of the screen and *pressing* the **Enter** key (**return** key on Mac). ***Do not minimize***.

2. Select **Constrain Dihedral** from the **Geometry** menu (or *click* on the  icon at the top of the screen). Select the COOC torsion, and then *click* on at the bottom right of the screen.

The icon will change to  indicating that a dihedral constraint is to be applied. Select **Properties** (**Display** menu) and *click* on the constraint marker on the model on screen. This leads to the **Constraint Properties** dialog.



3. *Check* **Dynamic** inside the dialog. This leads to an extended form of the **Constraint Properties** dialog that allows the single (dihedral angle) constraint value to be replaced by a sequence of constraint values.



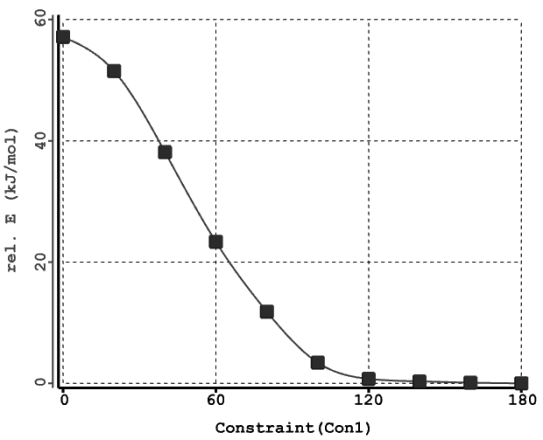
Leave **180** (180°) in the box to the right of **Value** alone, but change the contents of the box to the right of **to** to **0** (0°). You need to *press* the **Enter** key after you type in the value. **Steps** should be **10**. (If it is not, *type* **10** and *press* the **Enter** key.) What you have specified is that the dihedral angle will be constrained first to 180°, then to 160°*, etc. and finally to 0°. *Click* on to dismiss the dialog.

* The difference between constraint values is given by: (final-initial)/(steps-1).

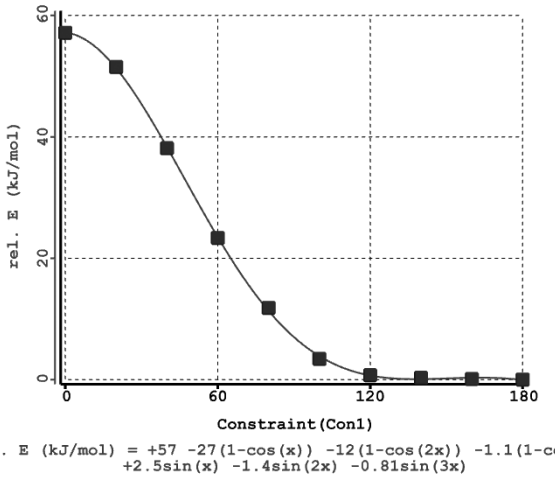
4. Bring up the **Calculations** dialog and select **Energy Profile** from the top menu to the right of **Calculate**, and **Hartree-Fock** and **6-31G*** from the two bottom menus. *Click* on **Submit** and accept the name ***dimethylperoxide***.

5. When the calculations have completed, they will go into a file named ***dimethylperoxide.Prof.M0001***. A prompt will ask you if you want to open this file. *Click* on **OK**. Align the conformers. *Click* on , select **Structure** from the **Align by** menu at the bottom right of the screen and, one after the other, *click* on both oxygens and on one of the carbons. Then *click* on the **Align by** button at the bottom right of the screen. Bring up the spreadsheet (**Display** menu), and enter both the energies relative to the 180° conformer, and the COOC dihedral angles. First *click* on the label (**M0001**) for the top entry in the spreadsheet (the 180° conformer), then *click* on the header cell for the leftmost blank column, and finally, *click* on **Add...** at the bottom of the spreadsheet. Select **rel. E** from the quantities in the dialog that results, **kJ/mol** from the **Energy** menu and *click* on **OK**. To enter the dihedral angle constraints, *click* on , *click* on the constraint marker attached to dimethylperoxide and *click* on P at the bottom of the screen (to the right of the value of the dihedral angle). *Click* on V.

6. Select **Plots...** (**Display** menu). Select **Constraint(Con1)** from the **X Axis** menu and **rel. E(kJ/mol)** from the **Y Axes** list. *Click* on **OK**.



A smooth curve (called a cubic spline) connects the data points. To fit the points to a Fourier series, bring up the **Properties** dialog (**Display** menu) and *click* anywhere on the curve. Select **Fourier LSQ** from the **Fit** menu and *click* on ✕ to remove the dialog. A new plot appears.



The actual fit expression appears at the bottom of the plot.

7 and 8 optional *(cannot be completed with the Essential Edition)*

7. Use energies from B3LYP/6-31G* density functional calculations, together with the Hartree-Fock geometries to provide a better fitting function. First, make a copy of ***dimethylperoxide.Prof.M0001***. Name it ***dimethylperoxide density functional***. Next, enter the **Calculations** dialog and specify calculation of **Energy** using the **B3LYP/6-31G* Density Functional** model. Make certain that **Global Calculations** (at the bottom of the dialog) is *checked* to signify that energy calculations are to be performed on all conformers. *Click* on **Submit**.

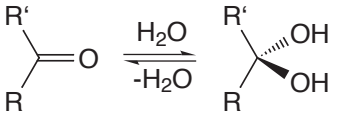
8. It involves energy calculations for all ten confromers and will require 10-20 minutes to complete. When it is done, draw a new energy plot and compare it to the energy plot produced earlier.

9. Remove all molecules and dialogs from the screen.

15 mins

Hydration of Carbonyl Compounds

The hydration of carbonyl compounds has been extensively studied primarily because it serves as a model for a number of important reactions, nucleophilic addition to carbonyl compounds among them.



$$K_{eq} = \frac{[\text{hydrate}]}{[H_2O]\,[\text{carbonyl}]} = \frac{[\text{hydrate}]}{55.5\,[\text{carbonyl}]}$$

Experimental K_{eq} for hydration of carbonyl compounds*			
	$\log(k_{eq}/55.5)$		$\log(k_{eq}/55.5)$
$PhCOCH_3$	-6.8	CF_3COCH_3	-0.2
CH_3COCH_3	-4.6	$PhCOCF_3$	0.1
PhCHO	-3.8	H_2CO	1.6
t-BuCHO	-2.4	CF_3CHO	2.7
CH_3CHO	-1.7	CF_3COCF_3	4.3

* J.P. Guthrie, Can. J. Chem., **53**, 898 (1975); **56**, 962 (1978).

In this tutorial, you will use ***Spartan's*** linear regression analysis tool to correlate calculated properties of carbonyl compounds with measured equilibrium constants for their hydration.

1. Build the compounds* listed above. Use **Carbonyl** (**Groups** menu) as a starting point for each to ensure consistent atom numbering. Put in one list; select **New Molecule** following construction of each molecule, that is, build and minimize acetophenone, select, **New Molecule**, build and minimize acetone, select **New Molecule**, etc. *Click* on V.

2. Enter the **Calculations** dialog and specify calculation of **Equilibrium Geometry** using the **Hartree-Fock 3-21G** model.**

---

* Keep the phenyl ring and the carbonyl group coplanar for benzaldehyde, acetophenone and trifluoroacetophenone.

* * You can save computer time by specifying calculation of an **Energy** using the **Hartree-Fock 3-21G** model, and then selecting **AM1** from the **Start from** menu. This requests a semi-empirical AM1 geometry in place of the 3-21G geometry. As all of these molecules are in SMD, you can save even more time by bringing up the **SMD Preview** dialog (*click* on ▲ at the bottom right of the screen), *clicking* on the yellow square to the left of **3-21G** and then *clicking* on **Replace All** at the bottom of the dialog. In this case, skip the calculation step as all the results you require are already available.

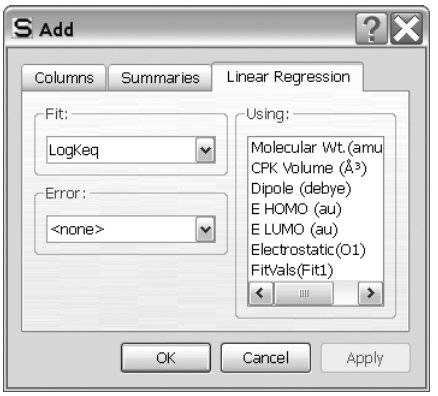
Make certain that **Global Calculations** is *checked*. *Click* on **Submit**. Name the job ***carbonyl compounds***.

3. After the calculations complete, bring up the spreadsheet. *Double click* inside the header cell of an empty column, *type* **Log(Keq)** and *press* the **Enter** key (**return** key on Mac). Then enter the experimental equilibrium constants from the table on the previous page. You need to *press* the **Enter** (**return**) key following each data entry. Sort the list according to the value of Log(Keq). *Click* inside the header cell **LogKeq** and then *click* on **Sort** at the bottom of the spreadsheet. Also, replace the molecule identifiers in the spreadsheet (***M0001***, ...) by correct names (***acetophenone***, ...). Position the cursor inside the header cell of the leftmost column of the spreadsheet, *right click* and choose **Rename Selected Using SMD** from the menu that appears.

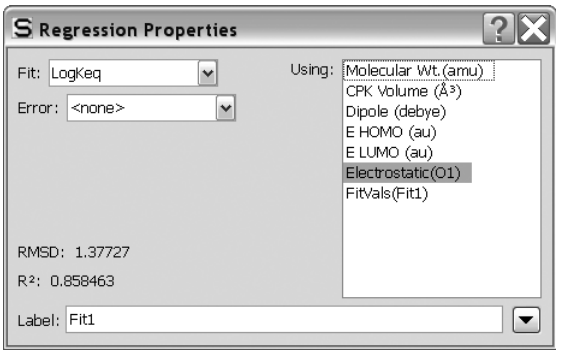
4. *Click* inside the header cell for an empty column, and then *click* on **Add...** at the bottom of the spreadsheet. Select (*click* on) the following quantities in the dialog: **E HOMO**, **E LUMO**, **Dipole**, **CPK Volume** and **Molecular Weight**. *Click* on **OK**.

5. Select **Properties** (**Display** menu) and *click* on the oxygen atom for whatever compound is displayed. *Click* on P to the left of **Electrostatic** (under **Charges**) in the **Atom Properties** dialog (that has replaced the **Molecular Properties** dialog), to place oxygen charges into the spreadsheet.

6. *Click* on **Add...** at the bottom of the spreadsheet, and then *click* on the **Linear Regression** tab at the top of the dialog that results. This brings up the **Add Linear Regression** dialog.



Select **Log(Keq)** from the **Fit** menu and **Electrostatic (Ox)*** from the **Using** list. *Click* on **OK**. Two new rows will be added at the bottom of the spreadsheet. With a **Properties** dialog on screen, *click* anywhere on the row in the spreadsheet labelled **Fit1** (the first fit performed on this data set). The **Regression Properties** dialog appears.



This provides information about the fit of Log(Keq) to the charge on oxygen, in particular, the value of R^2. This will tend to unity as the quality of the fit improves.

7. Inside the **Regression Properties** dialog, try fitting to different (single) properties in the **Using** list. (The entries toggle "on" and "off" with repeated *clicking*.) Maximize the value of R^2. Next, try all combinations of two properties. The combination of **E LUMO** and **Electrostatic (Ox)** should give the highest value of R^2. With this combination selected, bring up the **Plots**

* Numbering depends on the order that atoms were introduced during building, but the reference is to the carbonyl oxygen.

dialog (**Display** menu) and select **Log(Keq)** from the **X Axis** menu, and **FitVals (Fit1)** from the **Y Axes** list and *click* on **OK**. Next, *click* on the curve (not on the axes) to bring up the **Curve Properties** dialog (which replaces the **Regression Properties** dialog). Select **Linear LSQ** (to replace **Cubic Spline**) from the **Fit** menu. The resulting plot should show good correlation.

8. Close ***carbonyl compounds*** and remove any remaining dialogs from the screen.

10 mins

Acidities of Carboxylic Acids

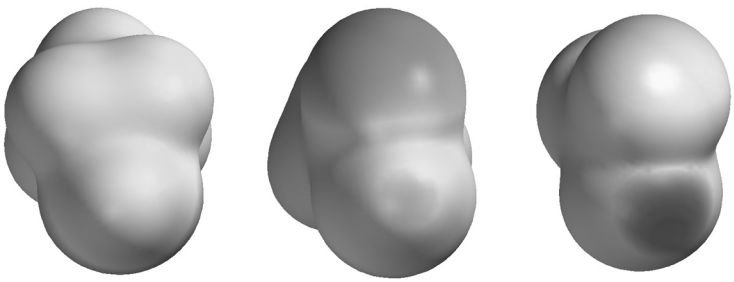
Acid strength is among the most important molecular properties. It is readily available from calculation, either in terms of absolute deprotonation energy,

$$AH \longrightarrow A^- + H^+$$

or, more commonly, as the deprotonation energy relative to that of some standard acid (A°H).

$$AH + A^{\circ -} \longrightarrow A^- + A^\circ H$$

The value of the electrostatic potential in the vicinity of the acidic hydrogen in the neutral acid provides a possible additional measure of (relative) acid strength. Electrostatic potential maps are known to uncover gross trends in acidity, for example, the acidic hydrogen in a strong acid, such as nitric acid, is more positive than that in a weak acid, such as acetic acid, which in turn is more positive than that in a very weak acid, such as ethanol.



ethanol acetic acid nitric acid

In this tutorial, you will explore the use of electrostatic potential maps to quantify changes in acid strength due to subtle variations in structure. You will use the Spartan Molecular Database of calculated structures, energies and properties may be used to shorten the time required for an investigation.

1. One after the other, build trichloroacetic, dichloroacetic, chloroacetic, formic, benzoic, acetic and pivalic acids. Put all molecules into the same document (use **New Molecule** from the **File** menu following the first molecule). *Click* on V when you are done.

2. All of the molecules that you have built are available in the Spartan Molecular Database. *Click* on ⏏ to the left of the name of the presently selected molecule at the bottom of the screen. The **SMD Preview** dialog will appear. This displays the structure of the selected molecule on the left and a list of theoretical models for which calculated results are available on the right. *Check* the yellow square to the left of **3-21G**, *click* on **Replace All** at the bottom of the dialog and respond **OK** to the message that follows. Structures obtained from Hartree-Fock 3-21G calculations will replace those you have built. However, you need to calculate a wavefunction for each of the molecules in order to produce electrostatic potential maps.

3. Enter the **Calculations** dialog (**Calculations** from the **Setup** menu). Request a **Hartree-Fock 3-21G Energy** calculation, and make certain that **Global Calculations** is *checked*. *Click* on **OK**. Enter the **Surfaces** dialog (**Surfaces** from the **Setup** menu). Make certain that **Global Surfaces** at the bottom of the dialog is checked. *Click* on **Add...** at the bottom of the dialog, select **density** from the **Surface** menu and **potential** from the **Property** menu (**medium** from the **Resolution** menu) in the **Add Surface** dialog that appears and then *click* on **OK**. Leave the **Surfaces** dialog on screen. Submit the job (**Submit** from the **Setup** menu). Name it ***carboxylic acids***.

4. After the calculations have completed, bring up the spreadsheet (**Spreadsheet** from the **Display** menu). Expand it so that you can see all seven molecules, and that two data columns are available. Position the cursor in the leftmost header cell of the spreadsheet, *right click* and select **Rename Selected Using SMD** in the menu that appears. *Double click* inside the header cell of the next available data column, *type* **pKa** and *press* the **Enter** key (**return** key on Mac). Enter the experimental pK_a's (given on the next page) into the appropriate cells under this column. You need to *press* the **Enter** (**return**) key following each entry.

acid	pK_a	acid	pK_a
trichloroacetic (Cl_3CCO_2H)	0.7	benzoic ($C_6H_5CO_2H$)	4.19
dichloroacetic (Cl_2CHCO_2H)	1.48	acetic (CH_3CO_2H)	4.75
chloroacetic ($ClCH_2CO_2H$)	2.85	pivalic ($(CH_3)_3CCO_2H$)	5.03
formic (HCO_2H)	3.75		

Experimental data from: E.P. Sargeant and B. Dempsey, ***Ionization Constants of Organic Acids in Aqueous Solution***, IUPAC no. 23, Permagon Press, 1979.

5. Arrange the seven molecules on screen such that you can clearly see the acidic hydrogen on each. To display all molecules at once, *check* the box on the right hand side of the **Label** column (in the spreadsheet) for each entry. To manipulate them independently of one another, *click* to deselect **Coupled** from the **Model** menu.

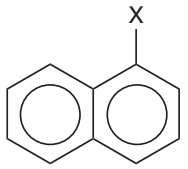
6. Bring up the **Surfaces** dialog (**Surfaces** from either the **Setup** or **Display** menu) and select ***density potential*** to turn on the electrostatic potential map for each molecule. Bring up the **Properties** dialog (**Properties** from the **Display** menu). If the **Surface Properties** dialog is not already displayed, *click* on the electrostatic potential map to display it. Change to a discrete or banded display. *Click* inside the **Surface Properties** dialog. Finally, *click* on P to the right of **Maximum**. The maximum value of the electrostatic potential (corresponding to the acidic proton) will be pasted to the spreadsheet.

7. Plot experimental pK_a vs. maximum in the electrostatic potential. Bring up the **Plots** dialog (**Plots** from the **Display** menu), select **pKa** under the **X Axis** menu and **Property Max (Surface)** from the **Y Axes** list, and *click* on **OK**. The data points are connected by a smooth curve (a so-called cubic spline). To get a least squares fit, select **Properties** from the **Display** menu, *click* on the curve, and select **Linear LSQ** from the **Fit** menu in the **Curve Properties** dialog. Does there appear to be a correlation between pK_a and the maximum value of the electrostatic potential?

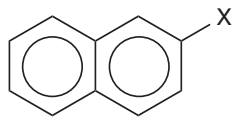
8. Close ***carboxylic acids*** and any remaining dialogs from the screen.

Positional Selectivity of Substituted Naphthalenes

Thermochemical stabilities of positional isomers may depend on several factors, including a tendency to minimize unfavorable non-bonded intramolecular interactions (***sterics***) and to minimize overall dipole moment (***electrostatics***). Naphthalene offers a good example with only two different positions.



1-substituted naphthalene

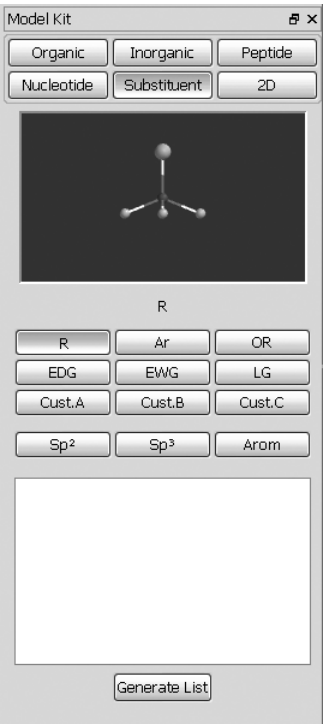


2-substituted naphthalene

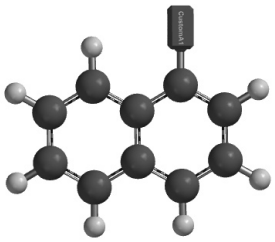
In this tutorial you will use results from the Spartan Molecular Database as accessed from ***Spartan's*** reaction energy calculator to assign in an effort to establish the change in preferred product. Specifically, heats of formation from the T1 thermochemical recipe will be employed. Instead of building a series of substituted naphthalenes, you will use ***Spartan's*** substituent model kit to construct a pair of molecules, one marked for substitution on the 1-position and the other on the 2-position. No calculations are involved.

1. One after the other build methane, ammonia, water hydrogen fluoride, hydrogen cyanide and formic acid. Put all on the same list (use **New** to enter the builder for the first molecule and **New Molecule** for each successive molecule). Bring up the spreadsheet (**Spreadsheet** under the **Display** menu), *double click* inside the leftmost cell for each molecule and replace the identifier (***M0001***, …) by proper functional group names (***methyl***, ***amino***, ***hydroxy***, ***fluoro***, ***cyano*** and ***carboxylic acid***). Put the contents on the clipboard. *Right click* inside the header cell for the leftmost column of the spreadsheet and select **Copy** for the menu that results. Close the document without saving.

2. Build naphthalene. *Click* on the **Substituent** tab at the top of the organic model kit to bring up the substituent model kit.



Click on **Cust. A** near the middle of the model kit, then *right click* inside the box near the bottom and select **Paste** from the menu that results. Names of the six functional groups that you built in the previous step will appear in the box. As you *click* through the list, ball-and-wire models of their structures will appear in the small screen at the top of the model kit. Note that one of the hydrogens (white balls) for each is highlighted in gold. This hydrogen will be removed to make a free valence for attachment. The carboxylic acid group has two different hydrogens, one on carbon (leading to a carboxylic acid) and one on oxygen (leading to a formate). If the hydrogen on the carbon is not already highlighted, *click* on it and it will be highlighted. *Click* on the free valence on naphthalene that will lead to 1-substituted naphthalenes. The molecule on screen will now appear with a marker to indicate substitution at the 1-position.



3. Select **New Molecule** (**File** menu). Bring up the organic model kit and again build naphthalene. Reenter the substituent model kit, make certain that **Cust. A** is selected (if not, *click* on it) and *click* on the free valence on naphthalene that will lead to 2-substituted naphthalenes. *Click* on V to remove the model kit.

4. Bring up the spreadsheet (**Spreadsheet** from the **Display** menu). *Double click* inside the leftmost cell for the first molecule and change the default identifier (**M0001**) with ***1-substituted naphthalenes***. *Double click* inside the corresponding cell for the second molecule and replace the default name by ***2-substituted naphthalenes***.

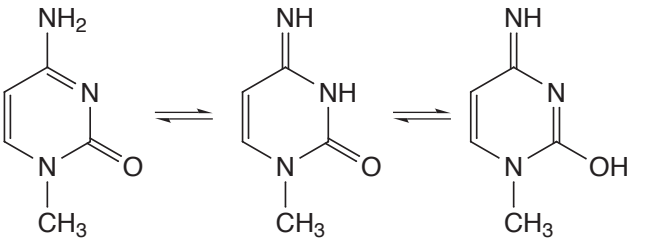
5. Bring up the **Reactions** dialog (**Reactions** from the **Display** menu). Choose **SMD Database** from the menu to the right of **Use** and **T1** from the menu to the right of **Theoretical Model**. Select ***1-substituted naphthalenes*** as **Reactants** and ***2-substituted naphthalenes*** as **Products**, and *click* on **Compute Energies** at the bottom of the dialog. Reactions will be written (and reaction energies computed) for all six substituents.

6. Close the (unnamed) document and any open dialogs.

5 mins/30 mins

Tautomers of Nucleotide Bases

Protons bound to heteroatoms in heterocyclic compounds are likely to be very mobile in solution and, where two or more heteroatoms are present in a structure, different isomers (tautomers) may be in equilibrium. As a case in point, consider the nucleotide base cytosine (where a methyl group has replaced the sugar-phosphate backbone at the 1-position).



The existence of a low-energy tautomer could have far-reaching consequences, given that the valence structure of cytosine is key to hydrogen bonding in DNA. In this tutorial, you will examine the possible tautomers of 1-methylcytosine for evidence of low-energy structures.*

1. Build 1-methylcytosine. Start with **Amide** from the **Groups** menu. Select planar trigonal nitrogen for the external amino group. Minimize (*click* on) and *click* on .

2. Note that the word **Tautomer** appears at the bottom right of the screen, indicating that tautomers exist. Select **Tautomers** from the **Search** menu (or *click* on the icon at the top of the screen). Step through the tautomers using the and keys that appear at the bottom right of the screen. To put the tautomers in a list, *click* on at the bottom right of the screen and *click* on **OK** in the dialog that results.

3. Enter the **Calculations** dialog with the list of tautomers (close the molecule you built to get the tautomers). Specify calculation of

* Note, however, that were the energy of an alternative tautomer only 10 kJ/mol higher than that for the normal structure, this would translate into a relative abundance of only about 1% at room temperature. Thus, any alternative tautomers would need to be very close in energy to the lowest-energy tautomer to have noticeable effect.

Equilibrium Geometry using the **3-21G Hartree-Fock** model. Submit the job with the name ***cytosine tautomers***.

4. After the calculations have completed, bring up the spreadsheet and *click* on the line corresponding to the structure for cytosine that you originally built. *Click* on **Add** at the bottom of the spreadsheet, select **rel.E** and **Boltzmann Distribution** from the available quantities and **kJ/mol** from the **Energy** menu and *click* on **OK**. Are either of the alternatives close in energy to the normal form of cytosine?

5 and 6 optional ***(cannot be completed with the Essential Edition)***

To get a better estimate of relative tautomer energies, perform energy calculations using the T1 model.

5. Make a copy of ***cytosine tautomers***. Name it ***cytosine tautomers T1***. Enter the **Calculations** dialog (**Setup** menu) with this copy. Specify **Energy** from the top menu to the right of **Calculate** and **Thermochemical Recipes** and **T1** from the two bottom menus. Submit the job.

6. When the calculations have completed (several minutes), bring up the spreadsheet and again examine relative tautomer energies and Boltzmann populations. Are either of the alternative structures competitive?

7. Remove all molecules and remaining dialogs from the screen.

30 mins

Stereochemical Assignments from ^{13}C Spectra
(cannot be completed with the Essential Edition)

NMR spectroscopy, in particular ^{13}C spectroscopy, is without doubt the method of choice to establish the three-dimensional structure of organic molecules. Only X-ray diffraction provides more definitive results, although the requirement of a crystalline sample severely limits its application. It is now practical to routinely calculate the NMR spectra of organic molecules. In time, the availability of a "virtual NMR spectrometer" will offer organic chemists an entirely new paradigm for structure determination, that is direct comparison of a measured spectrum with calculated spectra for one or more chemically reasonable candidates.

Here, you will obtain ^{13}C chemical shifts for *endo* and *exo* stereoisomers of 2-methylnorbornane, and compare these to experimental shifts. The B3LYP/6-31G* density functional model will be employed, but you will get all structures from the Spartan Molecular Database. You will establish the extent to which the calculations are able to reproduce differences in chemical shifts as a function of stereochemistry.

1. Build norbornane. Start with **Cyclohexane** from the **Rings** menu. Add sp^3 carbon  to one of the *axial* free valences (leaving you with *axial*-3-methylcyclohexane). *Click* on  and then *click* on one of the free valences of the methyl group and the *equatorial* free valence on cyclohexane that is directly across. *Click* on  . The name norbornane should appear at the bottom right of the screen. Put norbornane onto the clipboard. *Right click* and select **Copy** from the menu that appears.

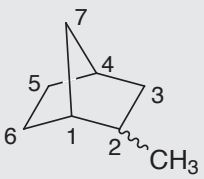
2. Select sp^3 carbon  from the model kit and *click* on the *exo* free valence of the carbon in the 2-position. The name ***exo-2-methylnorbornane*** should appear at the bottom right of the screen.

3. Select **New Molecule** from the **File** menu. *Click* on **Clipboard** at the bottom of the model kit and *click* on screen. Norbornane appears. Select sp^3 carbon  and *click* on the *endo* free valence of the carbon in the 2-position. The name ***endo-2-methylnorbornane*** should appear at the bottom right of the screen. *Click* on .

4. *Click* on ⏏ at the bottom of the screen (to the left of the molecule name), to bring up the **SMD Preview** dialog. *Click* on the box to the left of **B3LYP/6-31G*** and *check* on **Replace All**. *Click* on **OK** in the dialog that results. You have replaced both your structures with entries from the Spartan Molecular Database.

5. Inside the **Calculations** dialog, specify calculation of **Energy** using the **B3LYP 6-31G* Density Functional** model. *Check* **NMR** (to the right of **Compute**). Submit the job with the name ***2-methylnorbornane***.

6. When completed (several minutes), compare the ^{13}C chemical shifts with experimental values with particular attention to differences between *endo* and *exo* stereoisomers.



	endo	***exo***	Δ
C_1	11.4	16.5	5.1
C_1	11.4	16.5	5.1
C_1	43.5	42.2	-1.3
C_2	36.8	34.6	-2.2
C_3	40.2	40.7	0.5
C_4	37.3	38.2	0.9
C_5	30.3	30.6	0.3
C_6	29.0	22.4	-6.6
C_7	35.0	38.9	3.9
CH_3	22.3	17.4	-4.9

Can the calculations be used to distinguish between the two isomers?

8. Close ***2-methylnorbornane***.

Chapter 6
Organic Reactions

This chapter outlines and illustrates strategies for locating and verifying transition states for organic reactions as well as exploring changes in product distributions as a function of substituents and reactant stereochemistry.

The treatment of chemical reactions adds an entirely new dimension to the application of quantum chemical models. Whereas unique valence structures may generally be written for most molecules and, based on these structures, reasonable guesses at bond lengths and angles may be made, it is often difficult to designate appropriate valence structures for transition states, let alone specify detailed geometries. While there is a complete absence of experimental data for the structures of transition states, calculated transition-state geometries are now commonplace. ***Spartan*** provides both an extensive library of calculated transition-state geometries and a facility for automatically matching an entry in this library with the reaction at hand.* This library is also available as the Spartan Reaction Database (SRD), and may be searched by substructure to yield all available transition states of reactions related to the one of interest.

Spartan also provides a procedure for driving user-defined coordinates. Aside from conformational analysis (see discussion in previous chapter), the major application of this is to force reactions, thereby permitting identification of transition states.

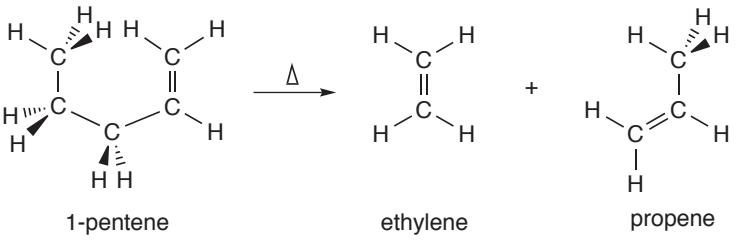
The first four tutorials in this chapter illustrate ***Spartan's*** automatic procedure for guessing transition-state geometries based on its library of organic reactions. The last two of these show how a transition state for one reaction may be used to provide a good guess at the transition state of a closely-related reaction. These tutorials also illustrate the

* Where a reaction is unknown to ***Spartan's*** library, a fallback technique which averages reactant and product geometries (similar to the so-called linear synchronous transit method) is invoked.

use of vibrational analysis (infrared spectroscopy) to verify that a particular structure corresponds to a transition state and to show the motion connecting it to reactants and products. The fifth tutorial illustrates how a reaction may be "driven" through a transition state. This tutorial, along with the fourth tutorial, draw the connection between relative activation energies and kinetic product distributions. The last tutorial in this chapter illustrates how transition states may be extracted from the Spartan Reaction Database.

An example of a transition state calculation for an organometallic reaction is provided in **Chapter 9**.

Ene Reaction of 1-Pentene

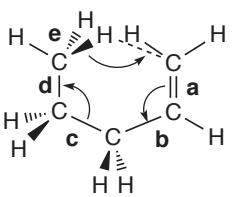


The proposed mechanism of the ene reaction involves simultaneous transfer of a hydrogen atom and CC bond cleavage. Here, you will obtain the transition state for the ene reaction of 1-pentene from 3-21G Hartree Fock calculations, and examine the reaction coordinate for evidence of concerted motion. Optionally, you will perform B3LYP/6-31G* density functional energy calculations to estimate the activation energy for the reaction.

1. Bring up the organic model kit and build 1-pentene in a conformation in which one of the terminal hydrogens on the ethyl group is poised to transfer to the terminal methylene group. Minimize and *click* on V.

2. First, save 1-pentene as ***1-pentene density functional*** for optional use later ( )*. Also make a copy for immediate use; name it ***ene reaction 1-pentene***. In both cases you will need to replace the suggested name (***1-pentene***).

3. Select **Transition States** from the **Search** menu (or *click* on the ↷ icon at the top of the screen). *Click* on bond **a** in the figure on the following page and then *click* on bond **b**. A curved arrow from bond **a** to bond **b** will be drawn.



* This is not necessary for users of the essential edition of ***Spartan***, where density functional models are not available.

Next, *click* on bond **c** and then on bond **d**. A second curved arrow from bonds **c** to **d** will be drawn. Finally, *click* on bond **e** and, while holding down the **Shift** key, *click* on the (methyl) hydrogen to be transferred and on the terminal (methylene) carbon to receive this hydrogen. A third curved arrow from bond **e** to the center of a dotted line that has been drawn between the hydrogen and oxygen will appear.

If you make a mistake, you can remove an arrow by selecting **Delete** from the **Build** menu (*click* on ) and then *clicking* on the arrow. (You will need to select  to continue.) Alternatively, hold down the **Delete** key as you *click* on an arrow. With all three arrows in place, *click* on  at the bottom right of the screen. Your structure will be replaced by a guess at the ene transition state. If the resulting structure is unreasonable, then you have probably made an error in the placement of the arrows. In this case, select **Undo** from the **Edit** menu to return to the model with the arrows and modify accordingly.

4. Enter the **Calculations** dialog (**Setup** menu), and specify calculation of transition-state geometry using the 3-21G Hartree-Fock model. Select **Transition State Geometry** from the top menu to the right of **Calculate**, and choose **Hartree-Fock** and **3-21G** from the two bottom menus. Finally, *check* **IR** to the right of **Compute**. This will allow you to confirm that you have found a transition state, and that it smoothly connects reactant and product. *Click* on **Submit**.

5. When the job completes, animate the motion of atoms along the reaction coordinate. Select **Spectra** from the **Display** menu and *click* on the **IR** tab. *Click* on the top entry in the list in the **IR** dialog that results. It corresponds to an imaginary frequency, and will be designated with an **i** in front of the number.

A vibrational frequency is proportional to the square root of a quantity that reflects the curvature of the potential surface along a particular (normal) coordinate corrected for the masses of atoms involved in motion along that coordinate. At a transition state (the "top of a hill"), the curvature is negative (it "points down"). Since mass is positive, the quantity inside the square root is negative and the frequency is an imaginary number.

Is the vibrational motion consistent with an ene reaction of interest and not with some other process?

6. Controls at the bottom of the **IR** dialog allow for changing both the amplitude of vibration (**Amp**) and the number of steps that make up the motion (**Steps**). The latter serves as a speed control. Change the amplitude to 0.3. *Type* **0.3** in the box to the right of **Amp** and *press* the **Enter** key (**return** key on Mac). Next, *click* on **Make List** at the bottom of the dialog. This will give rise to a group of structures that follow the reaction coordinate down from the transition state both toward reactant and product. Remove the original transition state from the screen. *Click* on ***ene reaction 1-pentene*** (the vibrating molecule) and close it, along with the **IR** dialog.

7. Enter the **Calculations** dialog and specify calculation of **Energy** using the **3-21G Hartree-Fock** model (the same level of calculation used to obtain the transition state and calculate the frequencies). Make certain that **Global Calculations** is checked. Next, enter the **Surfaces** dialog and specify evaluation of two surfaces: a bond density surface and a bond density surface onto which the electrostatic potential has been mapped. *Click* on **Add . . .**, select **density (bond)** for **Surface** and **none** for **Property** and *click* on **Apply**. Select **density (bond)** for **surface** and **potential** for **Property** and *click* on **OK**. Make certain that **Global Surfaces** is checked before you request the surfaces.

8. Submit for calculation. Name it ***ene reaction 1-pentene sequence***. Once the job has completed, enter the **Surfaces** dialog and examine the surfaces that you have calculated. For

each, step through the sequence of structures (and) keys at the bottom of the screen) or animate the reaction ( ). Note, in particular, the changes in bonding revealed by the bond density surface. Also pay attention to the value of the potential on the migrating atom. This reflects its charge. Is it best described as a proton (blue), hydrogen atom (green) or hydride anion (red)?

9. Close ***ene reaction 1-pentene sequence*** as well as any remaining dialogs.

10 to 14 optional *(cannot be completed with the Essential Edition)*

Methods that account for electron correlation are generally needed to furnish accurate estimates of absolute activation energies. Perform B3LYP/6-31G* density functional energy calculations on both 1-pentene and on the ene reaction transition state (using 3-21G geometries).

10. Open ***ene reaction 1-pentene*** ( ) and make a copy ( ). Name it ***ene reaction 1-pentene density functional***.

11. Enter the **Calculations** dialog, and specify calculation of **Energy** using the **B3LYP/6-31G* Density Functional** model. Remove the checkmark from **IR**. Submit the job.

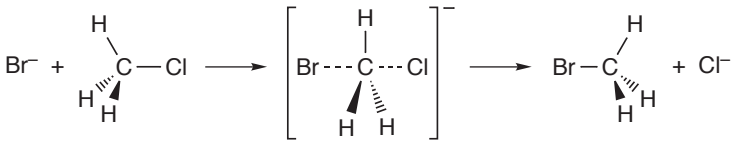
12. Open ***1-pentene density functional*** ( ). Specify calculation of **Energy** using the **B3LYP/6-31G* Density Functional** model. Also select **3-21G** from the **Start from** menu. This designates that a 3-21G equilibrium structure is to be used (that is, the same as employed for the transition state geometry). Submit the job.

13. Obtain the activation energy (difference in total energies between 1-pentene and the ene reaction transition state).

14. Remove any molecules and dialogs from the screen.

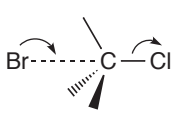
2 mins/5mins

S_N2 Reaction of Bromide and Methyl Chloride



1. The S_N2 reaction passes through a transition state in which carbon is in a trigonal bipyramid geometry and the entering and leaving groups are colinear. To build it, first construct methyl chloride. Then select bromine from the palette of icons in the model kit, *click* on the **Insert** button at the bottom right of the screen or hold down the **Insert** key (**alt** key on Mac) and *click* anywhere on screen. Alternatively, *double click* in a blank area of the screen following selection of bromine. Two detached fragments, methyl chloride and hydrogen bromide, appear on screen. *Click* on  and then *click* on the free valence on bromine. *Click* on  . Alternatively, hold down the **Delete** key and *click* on the free valence on bromine. You are left with methyl chloride and bromine atom (bromide). Manipulate the two such that bromide is poised to attack methyl chloride from the backside (as in the transition state above). (Recall that translations and rotations normally refer to both fragments, but can be made to refer to a single fragment by first *clicking* on the fragment and then holding down on the **Ctrl** key while carrying out the manipulations.) ***Do not minimize*** (If you do so by accident, select **Undo** from the **Edit** menu). *Click* on .

2. *Click* on  . *Click* on bromide and, while holding down the **Shift** key, *click* again on bromide and then on carbon. A dotted line will be drawn from bromine to carbon, together with an arrow from bromine to the center of this line. Next, *click* on the CCl bond and then *click* on the chlorine. A second arrow from the carbon-chlorine bond to the chlorine will be drawn.



Click on  at the bottom right of the screen. Your structure will be replaced by a guess at the transition state.

3. *Click* on  and then on the CBr bond. Replace the current CBr distance in the box at the bottom right of the screen by **3.8** (3.8Å) and *press* the **Enter** key (**return** key on Mac). You have now made a complex representing the reactant.

4. Select **Constrain Distance** from the **Geometry** menu (or *click* on the  icon at the top of the screen). *Click* on the CBr bond, and then *click* on  at the bottom right of the screen. The icon will change to  indicating a constraint is to be applied to this distance. Next, bring up the **Properties** dialog and *click* on the constraint marker. The **Constraint Properties** dialog appears. *Click* on **Dynamic**. Leave the value **3.8** (3.8Å) in the box to the right of **Value** alone, but change the number in the box to the right of **to** to **1.9** (1.9Å) and *press* the **Enter** (**return**) key. Change the number in the box to the right of **Steps** from **10** (the default) to **20**. 20 Calculations with CBr bond lengths constrained from 3.8Å (the starting point) to 1.9Å (the ending point) will be performed. The transition state should have a CBr distance in between these values. Dismiss the **Constraint Properties** dialog.

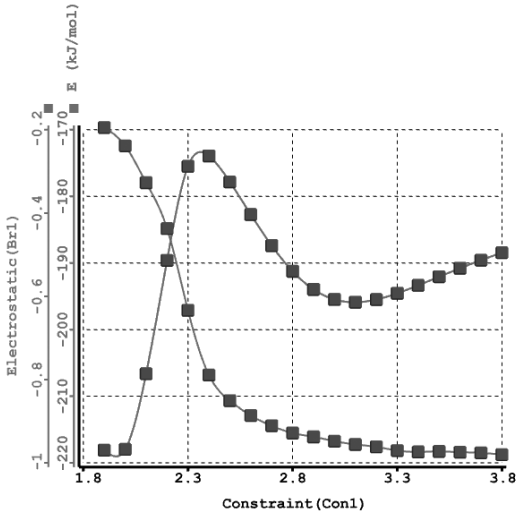
5. Enter the **Calculations** dialog, and select **Energy Profile**, **Semi-Empirical** and **AM1** from the appropriate menus to the right of **Calculate**. You need to change **Total Charge** to **Anion**.

6. Submit the job. Name it ***bromide+methyl chloride***. When completed, it will give rise to a sequence of calculations placed in ***bromide+methyl chloride.Prof.M0001***. You will be prompted as to whether you want to open this file. *Click* on **OK**. Align the molecules. *Click* on  , select **Structure** from the **Align by** menu at the bottom right of the screen and, one after the other, *click* on the three hydrogens. Finally *press* the **Align by** button at the bottom right of the screen. *Click* on  .

7. Bring up the **Spreadsheet** (**Spreadsheet** from the **Display** menu) and *click* on **Add...**. Select **E** from among the quantities listed at the top of the dialog, **kJ/mol** from the **Energy** menu, and *click*

on **OK**. Next, enter the (constrained) CBr distances and bromine charges in the spreadsheet. *Click* on , *click* on the constraint marker and *click* on P at the bottom right of the screen. *Click* on V. Bring up the **Properties** dialog. *Click* on bromine and *click* on P to the left of **Electrostatic** under **Charges** in the **Properties** dialog. Finally, bring up the **Plots** dialog, and select **Constraint (Con1)** (the distance at which the CBr bond has been constrained) from the **X Axis** menu, and both **E (kJ/mol)** and **Electrostatic (Br1)** from the **Y Axes** list. *Click* on **OK**.

One plot gives the energy as the reaction proceeds and the other gives charge on bromine. Are the two related? Explain.



S_N2 reactions involving charged species normally need to be carried out in highly-polar media, for example, water. ***Spartan*** provides two solvation models in which the solvent is represented by a reaction field (**Appendix A**). The SM5.4 model only alters the energy and not the charges. The more recent SM8 model does affect charges but is available only for Hartree-Fock and density functional models.

8. Add aqueous phase data (based on the SM5.4 model) to the spreadsheet. *Click* on an empty column header, *click* on **Add...**, select **Eaq** from the list of available quantities (**kJ/mol** from the **Energy** menu), and *click* on **OK**. Bring up the **Plots** dialog and

select **Constraint (Con1)** from the **X Axis** and **Eaq(kJ/mol)** from the **Y Axes** list. *Click* on **OK**.

9 to 12 optional

9. With ***bromide+methyl chloride.Prof.M0001*** selected, enter the **Calculations** dialog. Specify **Energy** (**Semi-Empirical** and **AM1** should be selected) and *click* on **OK**.

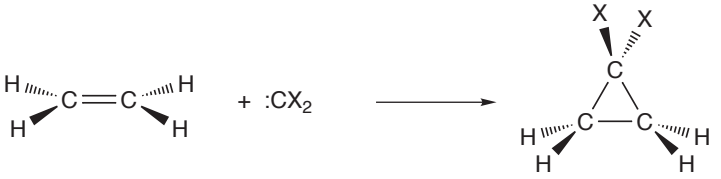
10. Enter the **Surfaces** dialog. *Click* on **Add...**. Select **density (bond)** from the **Surface** menu and **none** from the **Property** menu and *click* on **Apply**. Select **density (bond)** from the **Surface** menu, but this time **potential** from the **Property** menu. *Click* on **OK**.

11. Submit the job. When completed, *select* **density** inside the **Surfaces** dialog. *Click* on ▶| at the bottom left of the screen to animate the display. Note, that bonds are smoothly broken and formed during the course of reaction. *Click* on ⏸ at the bottom of the screen when you are done.

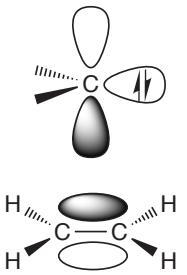
12. Reenter the **Surfaces** dialog. Turn "off" display of the bond density (select ***density***), and turn "on" display of the electrostatic potential mapped onto the bond density (*select* ***density potential***). *Click* on ▶|. Relate the migration of negative charge during reaction as indicated by colors in the electrostatic potential map to the plot you constructed earlier in step 7. Recall, that colors near red indicate maximum negative potential.

13. Remove all molecules and any remaining dialogs from the screen.

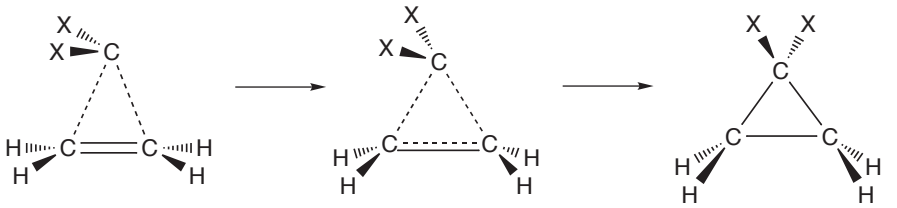
Carbene Additions to Alkenes



Singlet carbenes add to alkenes to yield cyclopropanes. Since a singlet carbene possesses both a high-energy occupied molecular orbital in the plane of the molecule, and a low-energy, out-of-plane unoccupied molecular orbital, this reaction presents an interesting dilemma. Clearly it would be more advantageous for the low-lying vacant orbital on the carbene, and not the high-lying filled orbital, to interact with the olefin π system during its approach.



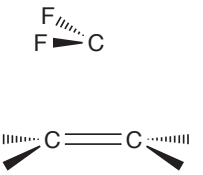
However, this leads to a product with an incorrect geometry. The carbene must twist by 90° during the course of reaction.



In this tutorial, you will use the Hartree-Fock 3-21G model to find the transition state and to analyze the motion of the fragments.

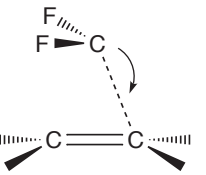
1. Bring up the organic model kit and build ethylene.
2. Select  from the model kit. *Click* on the **Insert** button at the bottom right of the screen or hold down the **Insert** key (**alt** key on Mac) and then *click* anywhere on screen. Alternatively, *double click* on a blank area of the screen following selection of sp^3

carbon. Next, select -F from the model kit and *click* on two of the free valences on the sp^3 carbon. Next, *click* on and, one after the other, *click* on the remaining two free valences on the sp^3 carbon. Finally, *click* on +. You are left with two fragments, ethylene and difluorocarbene. Orient the two as to be poised for reaction.*

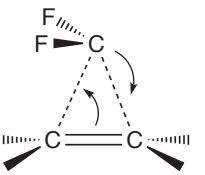


Translations and rotations normally refer to the complete set of fragments, but if you *click* on a fragment (not on a free valence) to select it, and then hold down the **Ctrl** key they will refer to an individual fragment.

3. *Click* on . *Click* on the carbon on the CF_2 fragment and then, while holding down the **Shift** key, *click* on the CF_2 carbon and on one of the carbons on the ethylene fragment. A dotted line is drawn between the two carbons that are to be bonded along with an arrow from the CF_2 carbon to the center of this line.



Click on the CC double bond and then, while holding down the **Shift** key, *click* on the other ethylene carbon and on the CF_2 carbon. A second dotted line and arrow will be drawn.



* Proper orientation of the two fragments is not crucial in this case, but is primarily to allow you to associate the arrows with the intended reaction. Proper orientation is, however, essential where different stereochemical outcomes are possible.

Click on  at the bottom right of the screen. Your structure will be replaced by a guess at the transition state.

4. Enter the **Calculations** dialog. Specify calculation of a **Transition State Geometry** using the **3-21G Hartree-Fock** model. *Check* **IR** to the right of **Compute**, and *click* on **Submit**. Name the job ***difluorocarbene+ethylene***.

5. When the job is complete, examine the geometry of the transition state. In light of the previous discussion, would you describe your structure as corresponding to an early or late transition state? Animate the vibration corresponding to the reaction coordinate. Bring up the **IR** dialog (**Spectra** under the **Display** menu and *click* on the **IR** tab). *Click* on the imaginary frequency at the top of the list of frequencies. Does the animation show that the carbene reorients as it approaches the double bond? Turn "off" the animation by again *clicking* on the imaginary frequency.

6. Select **Properties** (**Display** menu) and, in turn, *click* on each of the four hydrogens in the transition state. Change the value in the **Mass Number** menu in the **Atom Properties** dialog (**Properties** under the **Display** menu) from **1** to **2 Deuterium**. Resubmit the job. (No additional quantum chemical calculations are involved, but the vibrational analysis needs to be repeated.) When complete, examine the new set of vibrational frequencies. Note that they are uniformly smaller than those for the undeuterated system, and that the largest changes are for vibrational motions where hydrogens are involved.

7 to 10 optional

7. Make a copy of ***difluorocarbene+ethylene*** () name it ***difluorocarbene+cyclohexene***.

8. Bring up the organic model kit with the copy (*click* on ). Select **Freeze Center** from the **Geometry** menu (or *click* on the  icon at the top of the screen). One after the other, *click* on all atoms in the transition state except two of the hydrogens on one side

of the ethylene fragment.* Each of the selected atoms will be given a magenta colored marker, indicating that it is to be frozen (not moved during optimization).** Build a chain of four sp^3 carbons off one of the free valences that you have not frozen. Select and *click* on the free valence on the terminal sp^3 carbon and on the remaining free valence that you have not frozen to make a cyclohexene ring. *Click* on . Note that the only four methylene groups move and not the underlying transition state. You have produced an excellent guess at the transition state for difluorocarbene addition to cyclohexene based on your previous transition state for the simpler process.***

9. Enter the **Calculations** dialog. It should already specify calculation of **Transition State Geometry** using the **3-21G Hartree-Fock** model. Remove the checkmark on **IR** (to the right of **Compute**). *Click* on **Submit**.

10. To compare the geometries of the two transition states, you need to put them into the same document. Select **Append Molecule(s)...** from the **File** menu and *double click* on ***difluorocarbene+ethylene*** in the browser that results. Select **Spreadsheet** from the **Display** menu and *check* each of the boxes to the right of their identifiers. This allows the two transition states to be displayed simultaneously on screen. *Click* on **Add** at the bottom of the spreadsheet and select **Alignment Scores** from the menu that results. *Click* on , select **structure** from the **Align by** menu at the bottom right of the screen, *click* on the three carbons that make up the cyclopropane ring and *click* on **Align by** at the bottom right of the screen. *Click* on . Are the two transition states very similar as expected (an alignment score of 1.0 is perfect alignment)?

11. Remove all molecules and dialogs from the screen.

---

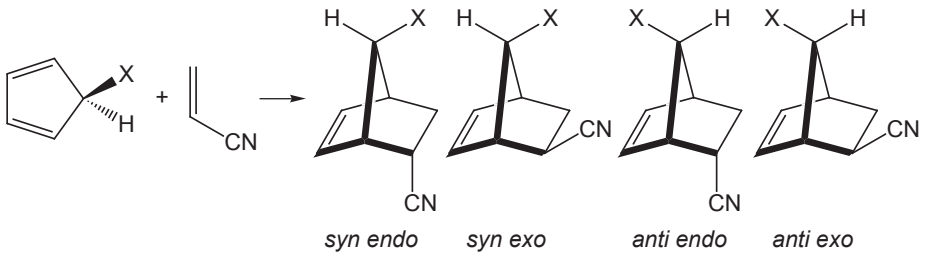
* You could do this more quickly by first *clicking* on , then drawing a selection box around all atoms to be frozen, and then *clicking* inside this box.

** You can remove the frozen atom markers from the model by bringing up the **Molecule Properties** dialog (**Display** menu), *clicking* on at the bottom right of the dialog and then removing the checkmark from **Frozens** in the **Molecule Utilities** dialog that results.

*** It is not necessary to replace the deuteriums by hydrogens. You will not request an IR spectrum and calculated geometry does not depend on atomic mass.

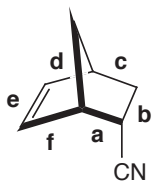
5 mins/20 mins

Stereospecific Diels-Alder Reactions



Diels-Alder cycloaddition of 5-substituted cyclopentadienes with acrylonitrile can lead to four stereoproducts, in which the substituent, X, at the 5 position is *syn* or *anti* to the dienophile, and the nitrile is *endo* or *exo*. *Anti* products are preferred when X is alkyl (consistent with sterics), while *syn* products are favored when X is halogen or alkoxy. In general, *endo* adducts are kinetically favored over *exo* adducts (see following tutorial). In this tutorial, you will use AM1 calculations to obtain both *syn* and *anti* transition states for *endo* addition of both 5-methylcyclopentadiene and 5-fluorocyclopentadiene and acrylonitrile, and then use Hartree-Fock 3-21G calculations to estimate relative activation energies. All four transition states are likely to be very similar to the transition state for the parent cycloaddition (cyclopentadiene and acrylonitrile). You will take advantage of this, and obtain a transition state for the parent reaction, and then to use it as a starting point for locating transition states for the substituted systems.

1. Construct the following substituted norbornene (the product of *endo* addition of cyclopentadiene and acrylonitrile).



2. *Click* on  . *Click* on bond **a** (see figure above) and then on bond **b**. A curved arrow will be drawn from **a** to **b**. Next, *click* on bond **c** and then on bond **d**; a second curved arrow from **c** to **d** be drawn. Finally, *click* on bond **e** and then on bond **f**, leading to a third curved arrow. The model on screen should now appear as follows.

Click on  at the bottom right to produce a guess at the transition state.

3. Inside the **Calculations** dialog, specify calculation of **Transition State Geometry** using the **Semi-Empirical AM1** model. *Check* **IR** under **Compute** to specify calculation of vibrational frequencies. This will allow you to verify that you have indeed located a transition state that is consistent with a Diels-Alder reaction.

4. Submit the job. Name it ***cyclopentadiene+acrylonitrile***. When completed, examine the resulting structure and calculated frequencies. Verify that there is one and only one imaginary frequency. Animate the motion associated with the imaginary frequency (**Spectra** under the **Display** menu, *click* on the **IR** tab and *check* the box to the left of the imaginary frequency in the dialog that results) to verify that it corresponds to the expected motion along the reaction coordinate.

5. Place the transition-state structure onto the clipboard. Right *click* on the background and select **Copy** from the resulting menu. Close ***cyclopentadiene+ acrylonitrile*** ().

6. Bring up the organic model kit (). *Click* on **Clipboard** (at the bottom of the model kit) and then *click* anywhere on screen. The transition-state structure will appear. Add a fluorine to the methylene group of the cyclopentadiene fragment of the transition state either *syn* or *anti*. ***Do not minimize***. Your starting structure should provide an excellent guess at the transition state for the substituted system.

7. Select **New Molecule** (not **New**) from the **File** menu. The screen will clear. *Click* on **Clipboard** and then *click* anywhere on screen. Add fluorine to the other methylene group position. Repeat the

process two more times (starting with **New Molecule**), to add methyl both *syn* and *anti* on the methylene group. When you are done (four substituted transition states in total), *click* on  .

Bring up the spreadsheet and replace the molecule identifiers in the leftmost column (***M0001***, etc.) but more descriptive names (***F syn***, etc.).

8. Enter the **Calculations** dialog and specify calculation of **Transition State Geometry** using the **Semi-Empirical AM1** model. *Check* **IR** to the right of **Compute**. Make certain that **Global Calculations** is checked. Submit the job (name it ***Diels-Alder stereochemistry***. When it has completed, verify that all four structures correspond to transition states (using **Spectra** from the **Display** menu).

9. Make a copy of ***Diels-Alder stereochemistry*** (). Name it ***Diels-Alder stereochemistry Hartree-Fock***. Enter the **Calculations** dialog with this copy and specify a **3-21G Hartree-Fock Energy** calculation. Make certain that you remove the checkmark on **IR** and that **Global Calculations** is checked before you exit the dialog.

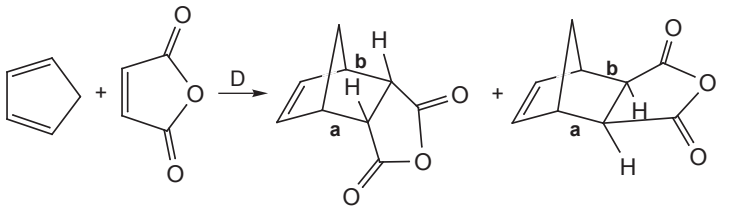
10. Submit the job. When it completes, bring up the **Reactions** dialog (**Reactions** from the **Display** menu). Select *syn* and *anti* F compounds (and then *syn* and *anti* Me compounds) for **Reactants** and **Products**, respectively, select **Current Document** from the **Use** menu, **kJ/mol** from the **Units** menu and *click* on the **Compute Energies** button at the bottom of the dialog.

11. Remove all molecules and any remaining dialogs from the screen.

5 mins/15 mins

Thermodynamic vs. Kinetic Control

Chemical reactions may yield different products depending on the conditions under which they are carried out. High temperatures and long reaction times favor the most stable (thermodynamic) products, whereas low temperatures and short reaction times favor the most easily formed (kinetic) products. For example, the kinetic product in Diels-Alder cycloaddition of cyclopentadiene and maleic anhydride is the *endo* adduct, whereas it seems likely that the less-crowded *exo* adduct is the thermodynamic product.



endo *exo*

Here you will first obtain pathways for reactions leading to both *endo* and *exo* adducts using the PM3 semi-empirical model and then follow these by 3-21G Hartree-Fock energy calculations to get a better estimate of the difference in activation energies.

1. Build the *endo* Diels-Alder adduct of cyclopentadiene and maleic anhydride and minimize ( ).

2. Select **Constrain Distance** from the **Geometry** menu (or *click* on), and *click* on bond **a** in the figure above. *Click* on the icon at the bottom right of the screen (it will then turn to). *Click* on .

3. Select **Properties** from the **Display** menu, and *click* on the constraint marker for bond **a**. The **Constraint Properties** dialog appears, with the value of the constraint for bond distance **a** (1.54Å) given in a box to the right of **Value**. Change this to 1.5Å by *typing* **1.5** inside the box and then *pressing* the **Enter** key (**return** key on Mac). *Check* **Dynamic** inside the dialog. This leads to an extended form of the **Constraint Properties** dialog, that allows the single constraint value to be replaced by

a sequence of constraint values. *Type* **2.7** (2.7Å) into the box to the right of **to** and *press* the **Enter** (**return**) key. Set **Steps** to **13**. *Type* **13** and *press* the **Enter** (**return**) key. You have specified that bond **a** will be constrained first to 1.5Å, then to 1.6Å, then to 1.7Å, etc. and finally to 2.7Å.

You need to repeat the process starting with step 2 for bond **b**. When you are done, both bonds **a** and **b** will be constrained from 1.5Å to 2.7Å in 13 equal steps. Close the **Constraint Properties** dialog.

4. Bring up the **Calculations** dialog and select **Energy Profile** from the top menu to the right of **Calculate**, and **Semi-Empirical** and **PM3** from the two bottom menus. *Click* on **Submit** at the bottom of the dialog. Name it ***cyclopentadiene+ maleic anhydride endo***.

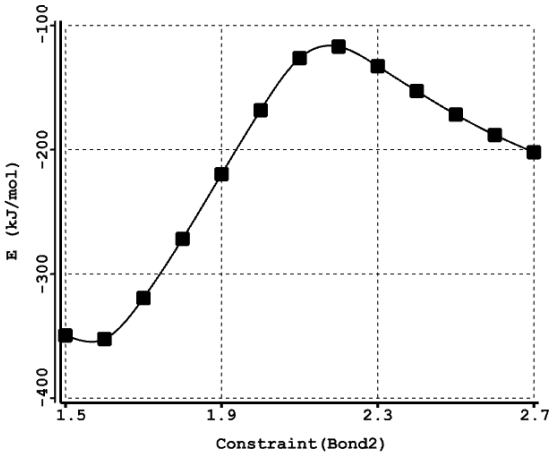
5. When completed the job will give rise to a new document ***cyclopentadiene+maleic anhydride endo.Prof.M0001***. This contains 10 calculations corresponding to the 10 steps that make up the energy profile. You will be prompted as to whether you want to open this file. *Click* on **OK**.*

6. Bring up the **Spreadsheet**, and *click* on **Add...** at the bottom. Select **E** from among the entries, **kJ/mol** from the **Energy** menu, and *click* on **OK**. Next, *click* on  and *click* on one of the two CC bonds varied in the energy profile. *Click* on  at the bottom right of the screen. *Click* on  . Finally, bring up the **Plots** dialog (**Plots** from the **Display** menu), and select **Constraint (BondX)**** from among the items in the **X Axis** menu and **E (kJ/mol)** from the **Y Axes** list. *Click* on **OK**.

---

* To avoid confusion, it is a good idea to close the original file ***cyclopentadiene+maleic anhydride endo***.

** Bonds are numbered in the order they were formed upon initial construction of the molecule.



Identify both the reactant and transition state from the plot and estimate the activation energy for the cycloaddition reaction.

7. Repeat steps 1 to 6 for the *exo* adduct. Compare the activation energy for *exo* addition to that for *endo* addition (above). What is the kinetic product?

8. Remove all molecules and dialogs from the screen.

9 to 11 optional

9. Open ***cyclopentadiene+maleic anhydride endo.Prof.M0001*** () and make a copy (). Name it ***cyclopentadiene +maleic anhydride endo Hartree-Fock***. Enter the **Calculations** dialog and specify a **Hartree-Fock 3-21G Energy** calculation. Submit. When completed, perform the same spreadsheet and plot operations you did for the PM3 calculations.

10. Repeat the above procedure for the *exo* adduct and compare the two activation energies. What is the kinetic product?

11. Remove all molecules and dialogs from the screen.

20 mins

Activation Energies of Diels Alder Reactions

In an earlier tutorial ***Dienophiles in Diels-Alder Reactions*** (**Chapter 5**), you examined the extent to which LUMO energies for a series of related dienophiles correlated with relative rates of Diels-Alder cycloadditions involving cyclopentadiene. In this tutorial, you will compare calculated activation energies for this same set of reactions with the experimental rates. You will use the previous set of calculations on the dienophiles and obtain transition states from the Spartan Reaction Database (SRD). The only quantum chemical calculation that is required is for cyclopentadiene.

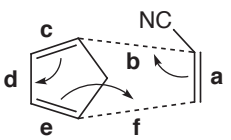
1. Build cyclopentadiene and minimize ( ). Before you proceed with the next step, copy cyclopentadiene onto the clipboard. *Right click* and select **Copy** from the menu that appears.

2. Enter the **Calculations** dialog (**Setup** menu) and specify an **Equilibrium Geometry** calculation using the **Hartree-Fock 3-21G** model. Submit the job with the name ***cyclopentadiene***. Move it off to the side of the screen for use later on.

3. Enter the builder ( ). *Click* on **Clipboard**. The structure of cyclopentadiene will appear at the top of the model kit. Click anywhere on screen and cyclopentadiene will appear.

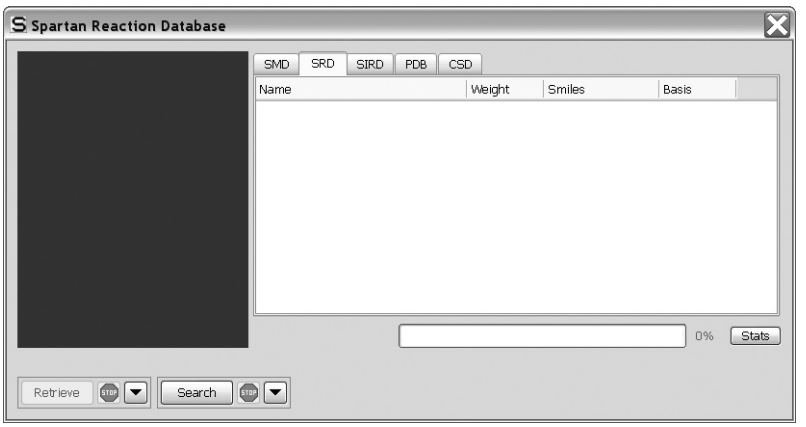
4. Select **Alkenyl** from the **Groups** menu, *press* the **Insert** button (at the bottom right of the model kit) or hold down the **Insert** key (**alt** key on Mac) and *click* on screen, or *double click* in a blank area on screen. Both cyclopentadiene and ethylene will appear on screen, but they will not be connected. Select **Cyano** from the **Groups** menu and *click* on one of the free valences on ethylene. Both cyclopentadiene and acrylonitrile will appear on screen.

5. Orient the two molecules such that they are poised for a Diels-Alder reaction leading to an *endo* product.



Select **Transition States** from the **Search** menu (*click* on). *Click* on the bond marked **a** in the above figure and, while holding down the **Shift** key, *click* on the two carbons that when connected will lead to the bond marked **b**. A dotted line will be drawn between these carbons and an arrow drawn from the center of bond **a** to the center of bond **b**. Next, *click* on bond **c** and then on bond **d**. A second arrow will be drawn. Finally, *click* on bond **e** and, while holding down the **Shif**t key, *click* on the two carbons that when connected will lead to the bond marked **f**. A second dotted line and a third arrow will be drawn. ***Do not click on*** ***as the query is based on the reactants and arrows and not the resulting transition state.***

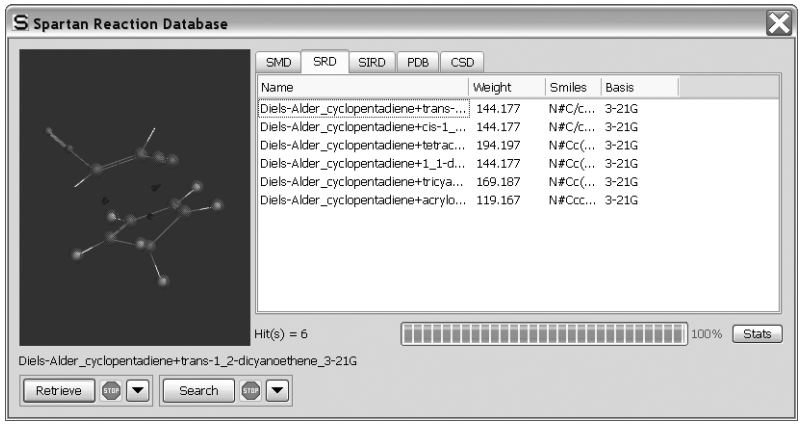
6. Select **Databases** from the **Search** menu (*click* on) and *click* on the **SRD** tab at the top of the dialog that results. This leads to the **Spartan Reaction Database (SRD)** dialog.



Click on all three free valences on acrylonitrile fragment. They will be replaced by orange cones, meaning that anything (including hydrogen) may be attached to these positions. Before you begin the search, *click* on to the right the **Search** button at the bottom of the dialog. This brings up the **Search Options** dialog.

Check **3-21G** under **Method Filters** and *click* on **OK**. *Click* on the **Search** button to search the Spartan Reaction Database for Diels-Alder transition states between cyclopentadiene and substituted acrylonitriles.

7. When the search completes, a listing of hits appears at the right of the dialog.



This should contain all six Diels-Alder transition states for *endo* additions of cyclopentadiene and acrylonitrile, *trans*-1,2-dicyanoethylene, *cis*-1,2-dicyanoethylene, 1,1-dicyanoethylene, tricyanoethylene and tetracyanoethylene. All need to be retrieved into a single document. *Click* on ▼ to the right of the **Retrieve** button at the bottom of the dialog to bring up the **Retrieve Options** dialog. Make sure that **New Document** is selected in this dialog and *click* on **OK**. Hold down the **Ctrl** key, select (*click* on) all six Diels-Alder transition states and *click* on **Retrieve** at

the bottom of the dialog. A new (unnamed) document will be created. Open a spreadsheet for this document.

8. Open ***Diels-Alder dienophiles*** that you created in an earlier tutorial* and bring up a spreadsheet for this document. Select all entries in this document. *Click* inside the lefthand cell of the first entry, hold down the **Shift** key and *click* on the last entry. *Right click* and select **Copy** from the menu that results. Move the cursor over the leftmost cell of the first empty column of the new (unnamed) document, *right click* and select **Paste** from the menu that results. The contents of ***Diels-Alder dienophiles*** (the dienophiles and cyclopentadiene) will be copied into the new document. Close ***Diels-Alder dienophiles***.

9. Bring up the **Reactions** dialog (**Reactions** from the **Display** menu). One after the other, calculate activation energies for the six Diels-Alder reactions. Use cyclopentadiene and one of the dienophile for **Reactants** and the appropriate transiton state and ***none*** for **Products**. Select **Current Document** under **Use** at the bottom of the dialog and *click* on **Compute Energies**. Is there a correlation between calculated activation energies and relative rates (experimental realative rates are available in the tutorial **Diels-Alder Dienophiles** in **Chapter 5**)?

10. Close all molecules and any open dialogs.

---

* If you have not completed this tutorial, a copy of the document ***Diels-Alder dienophiles*** may be found in the ***organic reactions*** sub-directory (***tutorials*** directory). For Windows, this is found in ***Program Files/Wavefunction/Spartan08***. For Linux, this is found in the directory where ***Spartan*** was installed. For Macintosh, this is located in the Spartan.08.full folder on the installation CDROM.

Chapter 7
Medicinal Chemistry

*This chapter illustrates applications of **Spartan** to problems of relevance to medicinal chemists. The first involves an investigation of the relationship between polar surface area (defined both by atom counts and by electrostatic potential maps) and the rate of transport between blood and brain. The second uses a published pharmacophore to decide whether the antihistamine terfenadine should be considered a potential potassium channel blocker. The third seeks to find molecules with similar analgesic activity as morphine, based first on similar structure and then on similar chemical functionality.*

Many molecular properties may now be calculated using quantum chemical methods, some more accurately than they can be measured. Still, some quantities of key interest to medicinal chemists are not yet subject to calculation. A good example is the rate at which a molecule passes through a membrane. Here, overall molecular polarity plays an obvious role. Too polar and the lipid will not take up the molecule; not sufficiently polar and it will be unable to dissolve in the "aqueous" environment on either side of the membrane. One descriptor commonly used to account for polarity is the so-called polar surface area, defined as the area of space-filling model due to nitrogen and oxygen together with any bonded hydrogens. The first tutorial in this chapter examines the correlation between polar surface area and blood-brain transport and then proposes an alternative based on the electrostatic potential.

Many (if not most) drugs act by binding to a protein or nucleotide, either promoting a particular process (agonists) or inhibiting one (antagonists). New drugs can be found by identifying molecules that are structurally and/or functionally similar to compounds that are known or presumed to bind. The second tutorial examines whether the antihistamine terfenadine is likely to show serious side effects

because of a binding affinity in potassium ion channels. Here, the similarity analysis is not carried out between molecules but between a molecule and an environment. The third and final tutorial makes use of similarity analysis to suggest compounds that look like morphine and hence might exhibit similar action.

Anticipating Blood-Brain Transport

A drug must be able to be transported throughout the body. This includes intestinal absorption and passage across the blood/brain barrier. In this tutorial, you will first examine how well the polar surface area, the area of space-filling model due to nitrogen and oxygen atoms together with any attached hydrogens, correlates with log (C_{brain}/C_{blood})*.

1. Open ***blood brain transport*** from the ***medicinal chemistry*** sub-directory (***tutorials*** directory).** This comprises a list of drugs/drug candidates for which experimental data relating to the ratio of concentrations in the brain and in the blood (C_{brain}/C_{blood}) are available. These span a range of nearly 5 log units.

2. Bring up a spreadsheet (**Spreadsheet** under the **Display** menu) and size it to show all compounds, several extra rows and three columns, in addition to the column of experimental data **log(brain/blood)** that is already displayed.

3. *Click* on the **Add** button at the bottom of the spreadsheet. Select **PSA** (polar surface area) from the quantities listed in the dialog that appears and *click* on **OK**. Polar surface area values will fill one column in the spreadsheet. Again, *click* on the **Add** button, and select (*click* on) the **Linear Regression** tab at the top of the dialog that appears (most likely the **Add Columns** dialog). This leads to the **Add Linear Regression** dialog.

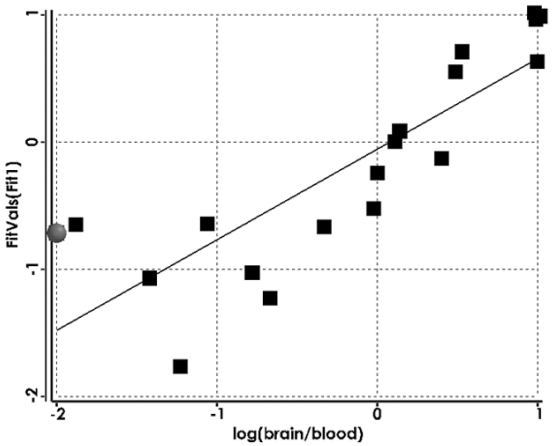
   Select **log(brain/blood)** from the **Fit** menu and *click* on **PSA** from the quantities in the box below **Using**. *Click* on **OK**. Bring up the **Regression Properties** dialog by first selecting **Properties** from the **Display** menu and then *clicking* on the cell labelled **Fit1** under **Label** in the spreadsheet. RMSD and R^2 values are reported. The smaller RMSD and the closer to unity R^2, the better the fit. Leave the **Regression Properties** dialog on screen.

---

* J. Kelder *et al*, Pharmaceutical Res., **16**, 1514 (1999). The 19 named compounds considered in this paper have been included.

** For Windows, this is found in ***Program Files/Wavefunction/Spartan08***. For Linux, this is found in the directory where ***Spartan*** was installed. For Macintosh, this is located in the Spartan.08.full folder on the installation CDROM.

4. Make a plot of log(brain/blood) vs. the regression fit. Select **Plots** from the **Display** menu to bring up the **Plots** dialog. Select **log (brain/blood)** from the **X Axis** menu, **Fit Vals (Fit 1)** from the **Y Axis** list and *click* on **OK**. Initial presentation uses a smooth curve (cubic spline) to connect the data points. *Click* on the curve (not on the axes) to bring up the **Curve Properties** dialog. Select **Linear LSQ** from the **Fit** menu to give a least squares line.

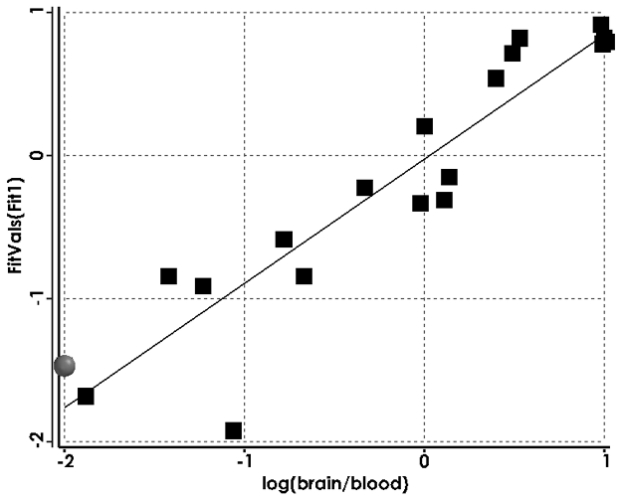


Next, you will consider an alternative definition of polar surface area, based on electrostatic potential maps.

5. Select **Surfaces** from the **Display** menu and select **density potential** in the dialog that results. An electrostatic potential map for the selected drug appears (maps for the remaining drugs may be seen by stepping through the list). The overall size and shape is that of the electron density and corresponds roughly to a conventional space-filling model. The colors indicate the value of the electrostatic potential. Colors toward red designate areas of negative potential (where a positive charge is most likely to be attracted), while colors toward blue designate areas of positive potential (where a positive charge is least likely to be attracted). To see the molecular skeleton underneath the electrostatic potential map, change to a transparent or mesh display. *Click* on the map to select it and select **Transparent** or **Mesh** from the **Style** menu that appears at the bottom right of the screen.

6. Select **Properties** from the **Display** menu. *Click* on the map. *Click* on the P button to the left of **P-Area** inside the **Surfaces Properties** dialog that results. Polar areas, defined as the area for which the absolute value of the electrostatic potential is > 100 kJ/mol*, will be added to the spreadsheet.

7. *Click* on the **Add** button at the bottom of the spreadsheet and *click* on the **Linear Regression** tab at the top of the dialog that results. Select **log(brain/blood)** from the **Fit** menu and *click* on **Polar Area** from the quantities in the box below **Using**. *Click* on **OK**. Bring up the **Regression Properties** dialog by *clicking* on the cell labelled **Fit2** under **Label** in the spreadsheet. Note that the R^2 value is better (closer to unity) for this fit than for the previous fit (to PSA defined in the usual manner). Make a plot of log (brain/blood) vs. the new regression fit. Select **Plots** from the **Display** menu, choose **log (brain/blood)** from the **X-Axis** menu and **FitVals (Fit 2)** (instead of **Fit Vals (Fit 1)** from the **Y Axis** list and *click* on **OK**. Switch to a least square fit by *clicking* on the curve that results and selecting **Linear LSQ** from the **Fit** menu in the **Curve Properties** dialog. Which property provides better correlation with transport across the blood/brain barrier, PSA or polar area?



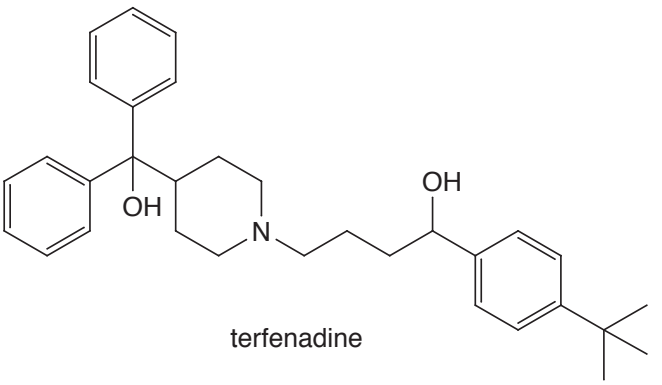
8. Close ***blood brain transport*** when you are finished.

* This value may be changed in the **Settings Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 18**).

5 mins

Terfenadine. A Potassium Channel Blocker?

Cardiovascular toxicity due to blocking of potassium ion channels is commonly screened early in the drug development process. One way this is done is to compare drug candidates to a pharmacophore deduced from 3D QSAR studies. This tutorial uses a published pharmacophore* to see whether the antihistamine, terfenadine should be considered a potential channel blocker.



terfenadine

1. Open ***terfenadine conformer library*** from the ***medicinal chemistry*** sub-directory (***tutorials*** directory).** This contains a selection of several hundred conformers of terfenadine (the number is reported in the **Molecule Properties** dialog) obtained from a conformer library calculation. It has been supplied in order to save computer time. Close ***terfenadine conformer library*** ( ).

2. Open ***potassium channel blocker pharmacophore*** from the ***medicinal chemistry*** sub-directory (***tutorials*** directory). This is a five-point pharmacophore comprising four hydrophobes (blue spheres) and one positive ionizable center (red sphere).

3. Select **Set Similarity Centers** from the **Geometry** menu (or *click* on the icon at the top of the screen) and then select **CFD** from the menu at the bottom right of the screen. In turn, *click* on each of the (five) pharmacophore elements. In response,

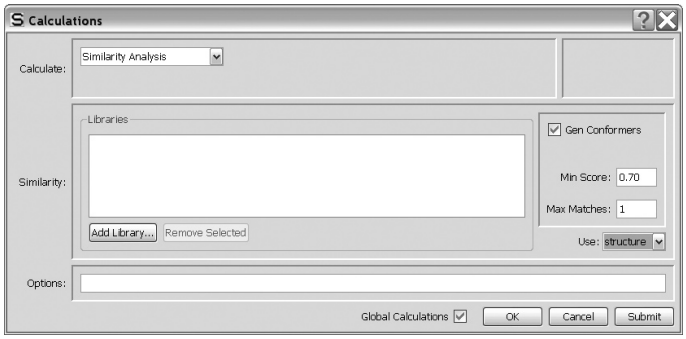
---

* S. Ekins *et al*, J. Pharmacology and Experimental Therapeutics, **301**, 427 (2002).

** For Windows, this is found in ***Program Files/Wavefunction/Spartan08***. For Linux, this is found in the directory where ***Spartan*** was installed. For Macintosh, this is located in the Spartan.08.full folder on the installation CDROM.

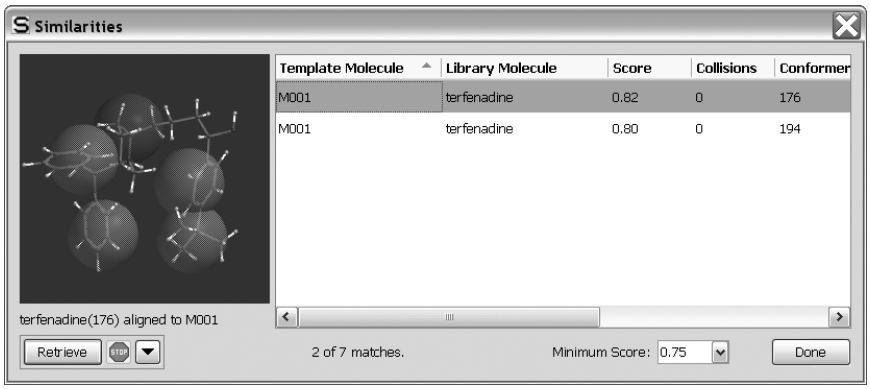
each will be surrounded by a violet circle indicating that it is to be used in the similarity analysis.

4. Select **Calculations...** from the **Setup** menu. Select **Similarity Analysis** from the top left menu to the right of **Calculate** inside the **Calculations** dialog. This leads to a new form of the **Calculations** dialog.



Click on **Add Library** to bring up a file browser. Locate ***terfenadine conformer library*** in the ***medicinal chemistry*** sub-directory (***tutorials*** directory) and *click* on **OK**. This indicates that the library of terfenadine conformers will be searched for a match to the pharmacophore. *Click* on **Submit**.

5. When the analysis has completed, select **Similarities...** from the **Display** menu to bring up the **Similarities** dialog.



One or more hits will appear in the box at the right of the dialog, together with a similarity score (limiting on 1.0 which means a perfect match). If there is more than one hit, *click* on that

with the largest score. A composite graphic of terfenadine and the pharmacophore will appear in a window at the left of the dialog. You can manipulate the two (as a single object) in the usual way (you need to position the cursor inside the window). Note that the positive ionizable center is positioned above the nitrogen, while the four hydrophobes are above the phenyl rings and the *tert*-butyl group.

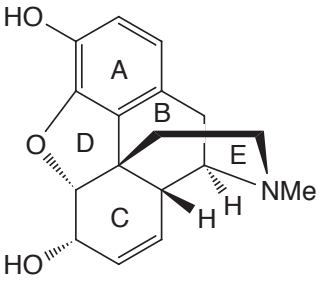
> The scoring algorithm not only accounts for position but also for direction of any nitrogen or oxygen centers that overlap with (non-hydrophobe) pharmacophore elements.

6. You can, if you wish, retrieve your hit (without retrieving the pharmacophore), by *clicking* on the **Retrieve** button at the bottom left of the **Similarities** dialog. This can be used for further analysis.

9. Close all open documents on screen.

5 mins

Morphine. Structure vs. Pharmacophore



morphine

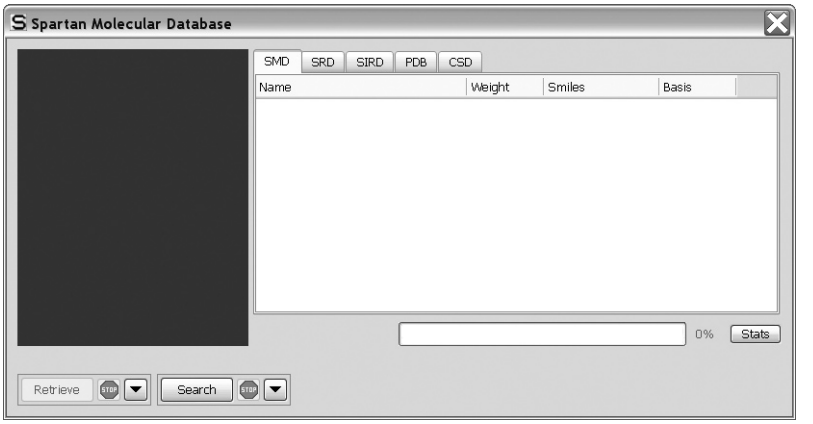
Three elements of morphine appear to be required in order for the molecule to act as an analgesic. These are: i) the nitrogen center (assumed to be protonated in the protein-bound complex), ii) the aromatic ring and iii) the hydroxyl group attached to the aromatic ring. Loss of any of these results in significant loss of activity. This knowledge may either be used directly, to identify other likely analgesics based on structure, or indirectly, to construct a simple 3-point pharmacophore that in turn may be employed to find compounds with potentially comparable analgesic activity.

This tutorial comprises three parts. In the first part, you will identify molecules in SMD that incorporate the three required elements. In practice, you will look for molecules that incorporate a nitrogen center substituted by a methyl group and two other (sp^3) carbons and a substituted phenol. You will then select a single hit and generate a conformer library for similarity analysis. (In a comprehensive investigation, you would likely use several hits or all hits; the restriction here is simply in order to save computer time.)

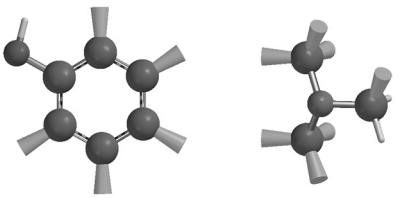
1. Bring up the organic model kit ( ). Select **Benzene** from the **Rings** menu and *click* anywhere on screen. Select sp^3 oxygen ( ) and *click* on one of the free valences of benzene. You have made phenol. Select sp^3 nitrogen ( ), *click* on the **Insert** button at the bottom of the screen or hold down the **Insert** key (**alt** key on Mac) and *click* anywhere on screen. Alternatively, *double click* on a blank region on screen following selection of sp^3 nitrogen. Two molecules (phenol and ammonia) appear on screen. *Click* on sp^3 carbon ( ) and *click* on all three free

valences on the nitrogen. Phenol and trimethylamine now appear on screen. *Click* on remove the model kit.

2. Select **Databases** from the **Search** menu ( ). *Click* on the **SMD** tab in the dialog that results.

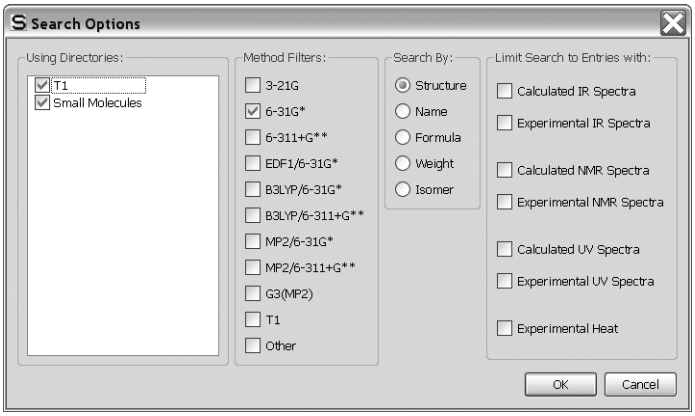


Click on all five free valences on the carbons in phenol and on all three free valences on two of the three methyl groups and on one free valence for the remaining methyl group in trimethylamine. Orange cones will appear at all (twelve) selected positions.



Anything can be grown from these positions (including hydrogen) but any hit must contain phenol and a nitrogen bonded to three sp^3 carbon centers.

Click on to the right of the **Search** button in the **SMD** dialog to bring up the **Search Options** dialog.



Check **6-31G*** under **Method Filters** and **Structure** under **Search by** in the **Search Options** dialog. *Click* on **OK**. Finally, *click* on the **Search** button at the bottom of the **SMD** dialog.

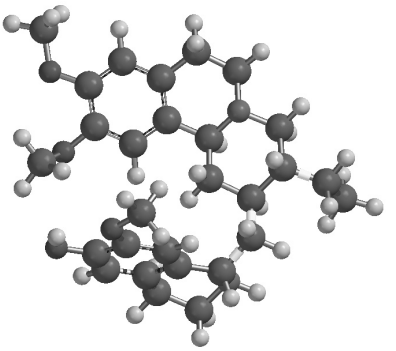
3. The search will result in several hits. Find and *click* on ***cephaeline*** from among those hits listed inside in the scroll box at the right of the **SMD** dialog. Its structure will be displayed in the window at the left of the dialog*. You can manipulate the model in the usual way by positioning the cursor inside this window. *Click* on ▾ to the right of the **Retrieve** button at the bottom of the **SMD** dialog to bring up the **Retrieve Options** dialog. *Check* **New Document** under **Retrieve Options** and *click* on **OK**. Finally, *click* on the **Retrieve** button. Dismiss the **SMD** dialog by *clicking* on ☒ at the top right-hand corner.

   Close the molecule used for the search (phenol and trimethylamine) without saving (🗁). Save the molecule you retrieved under the name ***cephaeline*** (🖫).

4. To shorten computer time, remove some of the degrees of conformational freedom from ***cephaeline***. Select **Set Torsions** from the **Geometry** menu (or *click* on ⟲). Your model will be augmented with yellow cylinders to designate flexible bonds. Remove the markers from the bonds connecting the three methoxy groups to the aromatic rings by *double clicking* on each in turn. These bonds will no longer be included in the

* *Click* on **Name** at the top of the list to sort in alphabetical order.

conformational search. Your model should now appear as follows with only three single bonds selected.



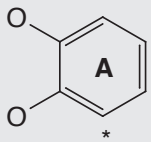
If more are selected, *double click* on the appropriate torsion markers to remove them.

5. Bring up the **Calculations** dialog (**Calculations...** from the **Setup** menu). Select **Conformer Library** from the top left menu to the right of **Calculate**. No further information is required. *Click* on **Submit** at the bottom of the dialog. Name it ***cephaeline library***. You can continue with the tutorial while you are waiting for the job to complete. When ***cephaeline library*** completes, close this file ( ).

In the second part of the tutorial, you will identify the key structural components in morphine and perform a similarity analysis based on structure on the conformer library for cephaeline.

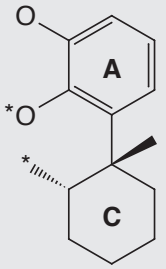
6. You can either build morphine or (to save time) open it: ***morphine*** from the ***medicinal chemistry*** sub-directory (***tutorials*** directory).

To build morphine, start with **Benzene** from the **Rings** menu. Add two (*ortho*) sp^3 oxygens ( ) to make catechol and orient as below (free valences not drawn).

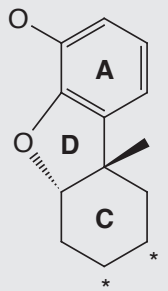


Select **Cyclohexane** from the **Rings** menu and *click* on the free valence on the *'ed carbon in catechol. Orient as below, with the free valence on

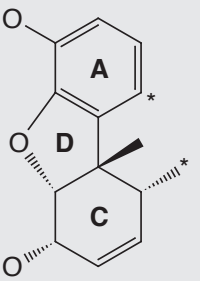
the cyclohexane carbon used to attach to benzene pointing straight up.



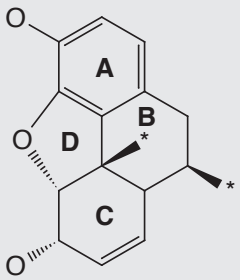
Select **Make Bond** from the **Build** menu (or *click* on). *Click* on the *'ed free valences indicated in the structure above, one on oxygen and one on carbon pointing down (relative to the free valence on the carbon used to attach the two rings). A bond will be drawn to make the **D** ring.



Make Bond is still selected. *Click* on *axial* free valences on the *'ed carbons in the cycohexane ring to introduce a double bond. Also, add a hydroxyl group to this ring ( ) adjacent to the double bond. (Pay attention to stereochemistry.) Select **Minimize** from the **Build** menu (or *click* on) to refine your structure.

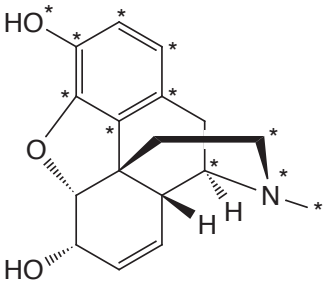


Add sp^3 carbons ( ) to free valences at the *'ed positions in the figure above (pay attention to stereochemistry on the cyclohexene ring). *Click* on and then *click* on free valences on each of the two sp^3 carbons that you have just added. You have made the **B** ring. Check that the stereochemisty is as it should be.



Connect the two *'ed free valences in the figure above with a chain of two sp^3 carbons ( ) and an sp^3 nitrogen ( ). Add an sp^3 carbon to the nitrogen. You have made morphine (its name should appear at the bottom of the screen). *Click* on to produce a refined structure and on to remove the model kit.

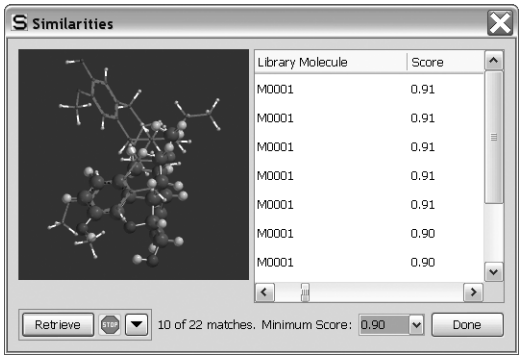
7. Select **Set Similarity Centers** from the **Geometry** menu (or *click* on) and select **structure** from the menu at the bottom right of the screen. *Click* on the eleven *'ed atoms in the figure below. A red circle will be drawn around each. If you make a mistake and select the wrong atom, *click* on the circle to deselect it.



8. Select **Calculations...** from the **Setup** menu. Select **Similarity Analysis** from the top left menu to the right of **Calculate** inside the **Calculations** dialog. *Click* on **Add Library** to bring up a file browser, locate and select ***cephaeline library***, and *click* on **OK**. Make certain that **Structure** is selected from the **Use** menu at the lower right of the dialog. You have requested that the library of cephaline conformers be searched for a match to the key structural elements of morphine. *Click* on **Submit**.* Supply the name ***morphine***.

* Make sure ***cephaeline library*** is closed before carrying out the similarity analysis.

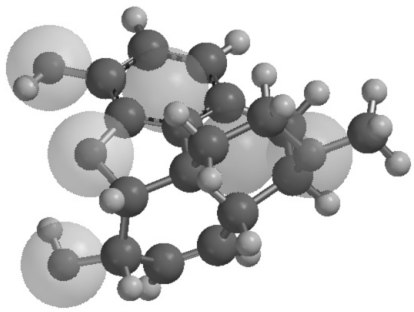
9. The similarity analysis will require a few seconds. When it completes, select **Similarities...** from the **Display** menu.



10. One or more hits appear in the box at the right of the dialog together with a score. Sort according to score (best on top) by *clicking* on **Score** at the top of the box. Select (*click* on) one of the best matches. A composite graphic consisting of morphine and the matching cephaeline conformer from the library will appear in a window at the left. You can manipulate the two structures (as a single object) in the usual way.

Repeat the similarity analysis using CFD's rather than structure.

11. Select **Set Similarity Centers** from the **Geometry** menu (), but this time select **CFD** from the menu at the bottom right of the screen. CFD's will augment the morphine structure.



12. Designate three of the CFD's as similarity centers. Select (*click* on) the CFD over the phenolic oxygen. A violet circle will surround this center indicating that it is to be used in the similarity analysis. (If you make a mistake and select the wrong

CFD, *click* on the circle and it will disappear.) Select (*click* on) the CFD at nitrogen and the CFD at the middle of the phenol ring.

13. Enter the **Calculations** dialog. This should already designate **Similarity Analysis** and the appropriate library (***cephaeline***). Change the entry in the **Use** menu to **CFD**. *Click* on **Submit**.

14. When the analysis is complete, bring up the **Similarities** dialog (**Similarities...** under the **Display** menu). One or more hits should appear in the box at the right of the dialog. Select a hit to get a composite graphic (CFD and structure from the library). You can easily see the extent to which the two are matched.

15. Close all documents open on screen.

Chapter 8
Polypeptides to Proteins

This chapter introduces the peptide model kit and shows how a structure from the Protein Data Bank (PDB) may be brought into ***Spartan*** *for further analysis.*

At present, calculations on biopolymers (proteins and RNA/DNA strands) are limited to molecular mechanics models. However, quantum chemical calculations are now routinely applicable to molecules bound into biopolymers and to the investigation of the electronic and steric factors that are responsible.

This chapter comprises three tutorials. The first illustrates ***Spartan's*** peptide model kit to construct triglycine and decaglycine in idealized helical arrangements. Comparisons are made between zwitterion and non-zwitterion forms in and out of water and electrostatic potential maps are employed to investigate charge separation.

The second tutorial illustrate ***Spartan's*** on-line access to PDB. A small molecule is extracted from a protein structure and the question is posed as to the energy of its (bound) conformer compared to that of the free conformer.

The third tutorial also involves extraction from a PDB file, but this time the quantity of interest is not the molecule itself, but rather its environment. This is turned into a so-called pharmacophore (a simplified representation of the environment), and a search for other molecules that may also be consistent with this environment is carried out.

Polyglycine

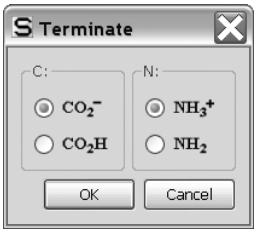
Triglycine and decaglycine may be used to illustrate construction of simple amino acid sequences using ***Spartan's*** peptide model kit, as well as to investigate the role of aqueous media in altering the equilibrium between neutral and zwitterionic forms of polypeptides and to show the utility of electrostatic potential maps in conveying information about overall charge distribution.

1. Bring up the peptide model kit by *clicking* on , and then *clicking* on the **Peptide** tab at the top of the organic model kit.

2. Make certain that the box to the left of **Sequence** is selected, and then *click* three times on gly (glycine) from the selection of amino acid codes. The sequence **gly-gly-gly** will appear in the box at the top of the model kit.

3. α **Helix** near the bottom of the model kit should be selected. If it is not, *click* on it. The backbone ψ and ϕ angles which define an α helix will appear at the right. *Click* anywhere on screen. Triglycine will appear, although it will not be properly terminated.

4. First consider the non-zwitterionic structure. *Click* on **Terminate** in the model kit. The **Terminate** dialog will appear allowing you to specify C and N terminating groups.



Choose **CO_2H** for the C terminating group and **NH_2** for the N terminating group. *Click* on **OK**.

> Amino acids replace atomic fragments, functional groups, rings and ligands as the basic building blocks in the peptide model kit, and these other building blocks are missing. Therefore, most modifications of peptides, aside from modifications in sequence of amino acids and in overall conformation, need to be carried out using either the organic or inorganic model kits.

5. Select **Freeze Center** from the **Geometry** menu (or *click* on). *Click* on **Freeze Heavy** at the bottom right of the screen. This indicates that all heavy (non-hydrogen) atoms are not to be moved during molecular mechanics minimization.* *Click* on to produce a refined structure (subject to the restriction that heavy atoms are kept in place).

6. Select **New Molecule** from the **File** menu. *Click* anywhere on screen. *Click* on **Terminate** inside the model kit and this time select **CO_2^-** and **NH_3^+** for C and N terminating groups, respectively. *Click* on **OK**. *Click* on and then *click* on **Freeze Heavy** at the bottom of the screen. Finally, *click* on .

7. *Click* on . Both forms of triglycine now occupy the same group. Enter the **Calculations** dialog and specify an **Energy** calculation using the **Hartree-Fock 3-21G** model. Do not change **Total Charge** (from **Neutral**) for the zwitterionic form of triglycine. Even though its Lewis structure incorporates formal + and - charges, the molecule is neutral. Also make certain that **Global Calculations** is checked. *Click* on **Submit** at the bottom of the dialog. Name the job ***triglycine***.

---

* This is necessary to maintain the idealized helix structure.

8. After the calculations complete, bring up the spreadsheet, *click* on the header cell of an empty column, and *click* on **Add...** at the bottom of the spreadsheet. Select both **E** and **Eaq** from the list of available properties at the top of the dialog which appears, and **kJ/mol** from the **Energy** menu, and then *click* on **OK**.

   Which form of triglycine, zwitterion or non-zwitterion, is favored in the gas phase? Why? Does solvent better stabilize the neutral or zwitterionic form? Why?

9. Remove ***triglycine*** and remaining dialogs from the screen.

10 to 14 optional

An electrostatic potential map for the zwitterionic form of a larger polypeptide, decaglycine, clearly shows charge separation, and explains why such a structure is not favored in the gas phase.

10. Select **New** from the **File** menu ( ) and bring up the peptide model kit. *Click* seven times on gly (glycine) from the selection of amino acid codes.* A sequence of ten glycines has now been requested. Make certain that α **Helix** is selected. *Click* anywhere on screen. *Click* on **Terminate** in the model kit and select $\mathbf{CO_2^-}$ and $\mathbf{NH_3^+}$ for C and N terminating groups, respectively, and then *click* on **OK**. *Click* on and then *click* on **Freeze Heavy** at the bottom of the screen. *Click* on , and then *click* on .

11. Enter the **Calculations** dialog and specify **Energy** using the **AM1 Semi-Empirical** model. *Check* **Converge** (near the bottom right of the dialog). Semi-empirical calculations on molecules of this complexity often have trouble converging and this option invokes a number of special convergence techniques.

12. Bring up the **Surfaces** dialog and request calculation of an electrostatic potential map. *Click* on **Add...**, select **density** from the **Surface** menu and **potential** from the **Property** menu, and *click* on **OK**.

---

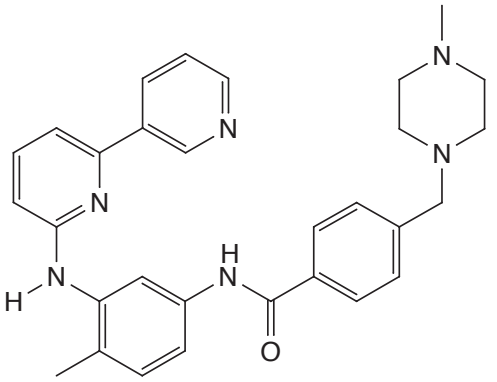
* This assumes that the box at the top of the peptide model kit already contains a sequence of three glycines. If you are starting from scratch, first make certain that **Sequence** is checked, and then *click* ten times on gly.

13. Submit the job. Name it ***decaglycine zwitterion***. When completed, again enter the **Surfaces** dialog. *Click* on the line ***density potential***. Examine the electrostatic potential map. Choose **Properties** from the **Display** menu and then *click* on the electrostatic potential map. *Click* on the selection box next to **Bands** to display the electrostatic potential in terms of a series of discrete color bands. If you wish, also *click* on the selection box next to **Legend** to provide a graphical legend for the value of electrostatic potential. Recall that colors near red correspond to areas on the surface that are negatively charged (attracted to a point-positive charge), while colors toward blue designate regions that are positively charged. Intermediate colors (greens) represent neutral regions. What do you conclude about the environment in which decaglycine wishes to exist?

14. Remove decaglycine and any remaining dialogs from the screen.

10 mins

Gleevec. Protein Bound vs. Free Conformer
(cannot be completed with the Essential Edition)



gleevec

Gleevec (Glivec, Imatinib) is a protein kinase inhibitor used in anticancer therapy. It specifically targets a protein kinase coded for in the rogue gene "bcr-abl". Several crystal structures for gleevec docked in protein kinases have appeared in the literature and are available in the PDB. In this tutorial, you will examine one of them* to establish whether or not the conformation of gleevec in the protein is identical (or similar) to that of the free (gas-phase) molecule, and if not, what energy penalty is to be paid to adopt the protein-bound conformation. The next tutorial will return to this same structure and extract a pharmacophore (gleevec's footprint).

1. Retreive a protein structure from the PDB. Select **Databases** from the **Search** menu and *click* on the **PDB** tab to bring up the **PDB** dialog.

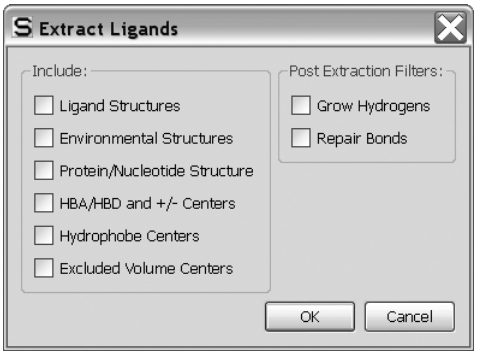
> If you are not online, you may skip this step as the structure is available as ***1opj*** in the ***polypeptides to proteins*** sub-directory (***tutorials*** directory). Open this document. For Windows, it is found in ***Program Files/Wavefunction/Spartan08***. For Linux, it is found in the directory where ***Spartan*** was installed. For Macintosh, it is located in the Spartan.08.full folder on the installation CDROM.

* B. Nager *et al*, Cell, **112**, 859 (2003): PDB identification 1opj.

Type **1opj** into the box to the right of **By PDB ID** at the bottom of the dialog and *click* on the **Search** button. In a few seconds, the PDB entry ***1opj*** will appear on the large box at the right of the dialog. Select (*click* on) it and then *click* on the **Retrieve** button. The protein will be represented by a ribbon model and incorporated molecules by sets of translucent spheres. Gleevec is the larger of the two incorporated molecules and is designated by its ***PDB HET*** code ***STI***. Close the dialog by *clicking* on ☒ at the top right-hand corner.

> The monochrome ribbon display may be modified, by selecting **Configure...** from the **Model** menu and *clicking* on the **Ribbons** tab. Coloring **By Secondary Structure** will display helices in red and β sheets in blue, while coloring **By Residue** will give each amino acid its own unique color.

2. Select **Ligands** from the **Search** menu (or *click* on the icon at the top of the screen). Select (*click* on) one of the gleevec molecules inside the protein structure and *click* on the **Extract Ligands** button at the bottom right of the screen. This leads to the **Extract Ligands** dialog.



Check **Ligand Structures** under **Include** and **Grow Hydrogens** under **Extraction Filters** inside the dialog and *click* on **OK**.

3. A ball-and-spoke model for gleevec will appear on screen. (You no longer need the protein; select and close it.) Put a copy of gleevec on the clipboard. With the molecule selected, select **Copy** from the **Edit** menu.

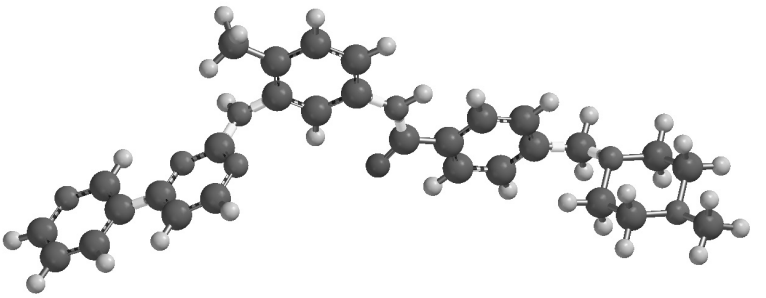
4. Select **New Molecule** (not **New**) from the **File** menu. Select **Clipboard** from the organic model kit and *click* anywhere on screen. *Click* on V. You now have two copies of gleevec in a single document. You can step between them using the step buttons at the bottom left of the screen.

5. Select the first copy and then select **Calculations...** from the **Setup** menu. **Equilibrium Geometry** from the top left menu to the right of **Calculate** and **Molecular Mechanics** and **MMFF** from the two bottom menus. The conformation will remain unchanged, but bond lengths and angles will be optimized.

> This is necessary in order to establish the energy difference between protein-bound and free conformations of gleevec. The resolution of protein X-ray crystallography is not adequate to establish bond lengths and angles to chemical accuracy.

Remove the checkmark from **Global Calculations** at the bottom of the dialog to signify that your choice (geometry optimization) only applies to this copy of gleevec. *Click* on **OK** (not on **Submit**).

6. Select the second copy of gleevec which will be used as a starting point for a conformational search. First, remove some degrees of conformational freedom in order to shorten the computer time for the conformational search. *Click* on  . Your model will be augmented with yellow cylinders and circles to designate flexible bonds and rings, respectively. Remove the marker from the bond connecting the phenyl ring to the carbonyl end of the amide group, by *double clicking* on it. Also, remove the (two) circles on the piperazine ring by *double clicking* on each in turn. The conformation of the piperazine ring will not change during the search. The model should now appear as below.



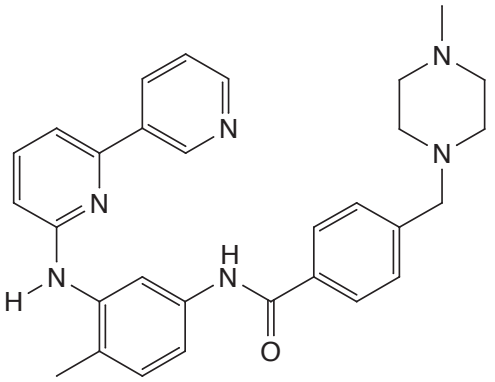
7. Re-enter the **Calculations** dialog. Select **Equilibrium Conformer** from the top left menu to the right of **Calculate** and (as before) **Molecular Mechanics** and **MMFF** from the two bottom menus. *Click* on **Submit**, and provide the name ***gleevec protein bound vs. free conformer***.

8. When the calculations have completed, bring up the spreadsheet (**Spreadsheet** under the **Display** menu), *click* inside leftmost the cell for the second molecule (the one on which a conformational search was performed) then *click* inside the header cell for a blank data column. *Click* on **Add** at the bottom of the dialog and select **Rel. E** from the scroll box in the dialog that appears and **kJ/mol** from the **Energy** menu. The relative energy of the protein bound conformer is provided in the spreadsheet. It is close to zero meaning that the protein-bound conformer will be present in significant amount in a sample of free gleevec?

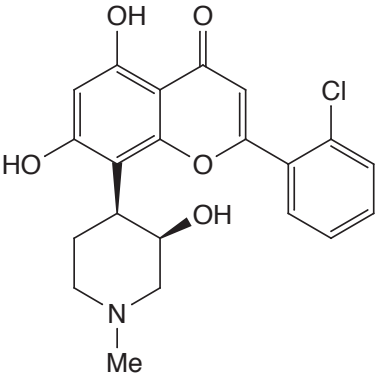
9. Close ***gleevec protein bound vs. free conformer***.

5 mins

Gleevec. Making a Pharmacophore from PDB
(cannot be completed with the Essential Edition)



gleevec



flavopiridol

In the previous tutorial, you extracted gleevec from a PDB structure of a protein kinase in order to see if the bound conformer was the same as the free conformer. Here, you will start from the same protein structure*, but instead of abstracting the molecular structure of gleevec, you will abstract its "footprint", that is, the locations from which it interacts with its protein host in terms of hydrogen-bond or charge-charge interactions (non-steric contacts). This information, together with knowledge of either the steric requirements of the guest or the space occupied by the protein (excluded volumes), constitutes a structural pharmacophore, that is, a template for scrutiny of other possible guests. In this tutorial, you will use a subset of the non-steric contacts together with excluded volume elements to see to what extent this pharmacophore fits another protein kinase inhibitor, flavopiridol.

1. Either obtain the PDB file ***1opj*** from the PDB database (see previous tutorial) or open ***1opj*** from the ***polypeptides to proteins*** sub-directory (***tutorials*** directory)**. Select **Ligands** from the **Search** menu (or *click* on), *click* on the model for gleevec inside the protein structure (displayed as a set of transparent spheres and look much like a space-filling model) to select

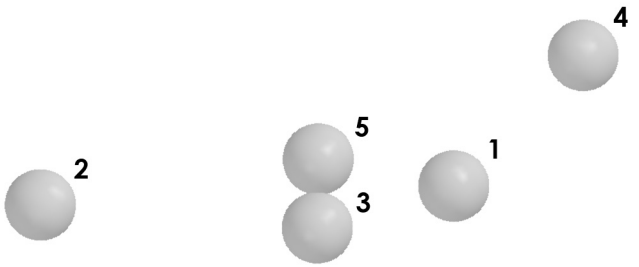
* B. Nager *et. al*., Cell, **112**, 859 (2003): PDB identification 1opj.

** For Windows, this is found in ***Program Files/Wavefunction/Spartan08***. For Linux, this is found in the directory where ***Spartan*** was installed. For Macintosh, this is located in the Spartan.08.full folder on the installation CDROM.

it, and *click* on the **Exact Ligands** button. Inside the **Extract Ligands** dialog, *check* both ***HBA/HBD and +/- Centers*** and **Excluded Volume Centers** and *click* on **OK**.

2. You are finished with ***1opj*** and can close it ( ). Also, simplify the display of the extracted information by removing the excluded volume elements from view. Select **Configure...** from the **Model** menu and *click* on the **CFDs** tab. *Uncheck* **Excluded Volumes** and *click* on **OK**.

3. The model that now appears comprises five blue spheres representing non-steric contacts between gleevec and its protein host. Turn on labels. Select **Configure...** from the **Model** menu, *click* on the **Labels** tab, *check* **CFD Labels** and *click* on **OK**.



These have been assigned based on the observation that each is within a predefined distance of (one or more) complementary groups from amino acids in the protein host. Because hydrogen positions are not established (in the experimental X-ray structure) it may be impossible to say whether a particular site on the guest is acting as a (hydrogen-bond) donor or acceptor. In fact, it may not even be possible to say whether a particular site is protonated or deprotonated (although pH may dictate this). This being the case, all non-steric contacts have initially been given four CFD definitions: hydrogen-bond acceptor (**HBA**), hydrogen-bond donor (**HBD**), positive ionizable (+) and negative ionizable (–). Examination of the structure of gleevec suggests more focused assignments: **1** (**HBA**, +), **2** (**HBA**), **3** (**HBD**), **4** (**HBA**) and **5** (**HBA**). (See above figure for numbering.)

4. You will use only three elements of this pharmacophore (together with all the excluded volume elements) to assess whether or not flavopiridol is likely to fit into the same host environment as gleevec. Select **Properties** from the **Display** menu. Then select **Set Similarity Centers** from the **Geometry** menu (or *click* on ). *Click* on the pharmacophore element designated as **1** (see figure on previous page). A violet circle will surround the element designating that it is to be used in the similarity analysis. If you make a mistake and select the wrong element, *click* on the circle and the designation will be removed. Inside the **CFD Properties** dialog, remove checkmarks from **HBD** and **– Ionizable** (you are left with **HBA** and **+ Ionizable**). Repeat the process for the pharmacophore element **3**, designating it as **HBD** only and **5**, designating it as **HBA** only. When you are done, only three of the five pharmacophore elements should be marked. *Click* on  .

5. Select **Calculations...** from the **Setup** menu and then **Similarity Analysis** from the top left menu to the right of **Calculate**. *Click* on **Add Library...** and then locate and select ***flavopiridol*** in the ***polypeptides to proteins*** sub-directory (***tutorials*** directory) and *click* on **OK**. *Click* on **Submit**. Name the document ***flavopiridol fit to gleevec pharmacophore***.

6. When the similarity analysis has completed, bring up the **Similarities** dialog (**Similarities...** from the **Display** menu). Sort the hits according to similarity score (*click* on **Score** at the top of the dialog). One after the other, *click* on the top scoring entries in the box to the right of the dialog and examine the fits in the window at the left.

7. Close any open documents on screen.

Chapter 9

Inorganic and Organometallic Molecules

*This chapter shows how to construct inorganic and organometallic molecules using **Spartan's** inorganic model kit. It also identifies quantum chemical models that appear to be suitable for calculations on molecules incorporating transition metals.*

Many molecules are made up of a relatively few elements and obey conventional valence rules. They may be easily built using the organic model kit. However, many others cannot be assembled with this model kit either because they incorporate other elements, or do not conform to normal valence rules or they involve ligands. Important among these are inorganic and organometallic compounds involving transition metals. They may be constructed using the inorganic model kit.

Transition-metal inorganic and organometallic compounds may also require different quantum chemical methods from those that have proven to be satisfactory for organic molecules. In particular, Hartree-Fock and MP2 models are known to produce poor results where transition metals are involved. The PM3 semi-empirical model has been parameterized for most transition metals, and generally provides a good account of equilibrium geometries. Density functional models are also successful in this regard and are also believed to provide a satisfactory account of the energies of reactions involving transition-metal systems. Note, however, that there are very little experimental thermochemical data with which to compare.

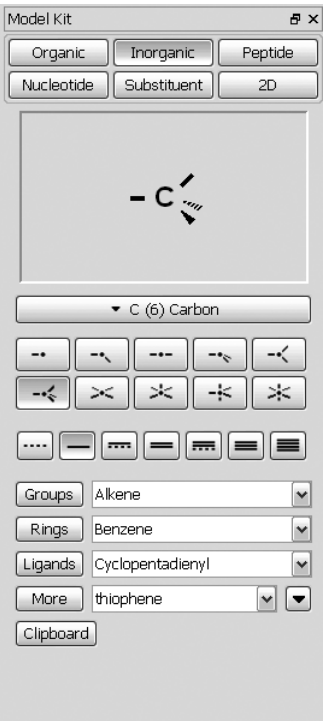
The tutorials in this chapter illustrate construction of inorganic and organometallic molecules, as well as the selection of suitable quantum chemical models.

1 min

Sulfur Tetrafluoride

Sulfur tetrafluoride cannot be constructed using ***Spartan's*** organic model kit. This is because sulfur is not in its normal bent dicoordinate geometry, but rather in a trigonal bipyramid geometry with one of the *equatorial* positions vacant. However, the molecule can easily be made using the inorganic model kit.

1. Bring up the inorganic model kit by *clicking* on and then *clicking* on the **Inorganic** tab at the top of the (organic) model kit.

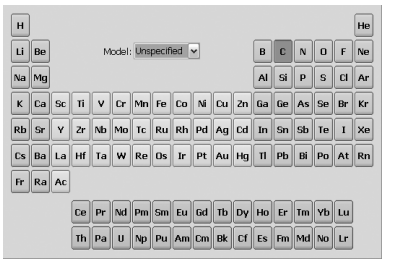


The inorganic model kit comprises an atom bar (*clicking* on which bring up the *Periodic Table**) followed by a selection of atomic hybrids, then bond types, and finally **Rings**, **Groups**,

* Not all methods are available for all elements listed. Elements for which a specific method (selected in the **Calculations** dialog) are available will be highlighted following selection of a theoretical model from the **Model** menu that appears in the center of the *Periodic Table*.

Ligands, **More** and **Clipboard** menus (all except for **Ligands** are the same as found in the organic model kit).

2. *Click* on the atom bar to bring up the *Periodic Table*.



Select (*click* on) **S** in the *Periodic Table* and the five coordinate trigonal bipyramid structure from the list of atomic hybrids. *Click* on screen. A trigonal bipyramid sulfur will appear at the top of the model kit.

3. Again, *click* on the atom bar, select **F** in the *Periodic Table* and the one-coordinate entry from the list of atomic hybrids. One after the other, *click* on both *axial* free valences of sulfur, and two of the three *equatorial* free valences.

4. It is necessary to delete the remaining free valence (on an *equatorial* position); otherwise it will become a hydrogen. *Click* on and then *click* on the remaining *equatorial* free valence.

5. *Click* on . *Click* on to remove the model kit.

6. Select **Calculations...** from the **Setup** menu. Specify calculation of **Equilibrium Geometry*** using the **Hartree-Fock 3-21G** model.**

7. Submit the job. Accept the name ***sulfur tetrafluoride***.When completed, select **Properties (Display** menu) and *click* on an

---

* It should be noted that were an incorrect geometry specified at the outset, optimization would lead to the correct structure, as long as the starting geometry possessed no symmetry (C_1 point group). Thus, square planar SF_4 in D_{4h} symmetry would remain square planar, while an almost square planar structure (distorted only slightly from D_{4h} symmetry to C_1 symmetry) would collapse to the proper structure.

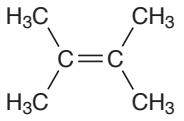
** The 3-21G basis set is extended by a set of d-type valence functions for sulfur (and other second-row elements).

atom, for example, sulfur. Three different atomic charges will appear in the (**Atom Properties**) dialog (corresponding to different methods for establishing atomic charge). Of these, the procedure based on fitting the electrostatic potential is generally considered to be the best. Are the electrostatic charges consistent with covalent or ionic bonding?

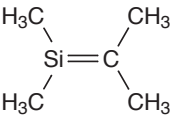
8. Remove ***sulfur tetrafluoride*** and any remaining dialogs from the screen.

5 mins

Reactivity of Silicon-Carbon Double Bonds



2,3-dimethyl-2-butene



tetramethylsilaethylene

With the exception of so-called phosphorous ylides, compounds incorporating a double bond between carbon and a second-row element are quite rare. Most curious perhaps is the absence of stable compounds incorporating a carbon-silicon double bond. This can be rationalized by using local ionization potential and LUMO maps to compare the reactivities of olefins and silaolefins.

1. Build 2,3-dimethyl-2-butene, minimize ( ) and copy onto the clipboard. (*Right click* on the molecule and select **Copy** from the resulting menu.) Select **New Molecule** (not **New**) from the **File** menu, *click* on **Clipboard** inside the organic model kit and *click* anywhere on screen. *Click* on the **Inorganic** tab (at the top of the organic model kit) to bring up the inorganic model kit. *Click* on the atom bar and select **Si** from the *Periodic Table*. *Double click* on one of the sp^2 carbons to substitute a silicon atom for a carbon atom and to make tetramethylsilaethylene. *Click* on and then on . You are left with 2,3-dimethyl-2-butene and tetramethylsilaethylene in the same document.

2. Use the structures in SMD to save computer time. *Click* on to the left of the molecule name (either 2,3-dimethyl-2-butene or tetramethylsilaethylene) at the bottom of the screen, select **3-21G** from the entries in the **SMD Preview** dialog and *click* on **Replace All**.

3. Select **Calculations** from the **Setup** menu. Specify calculation of **Energy** using the **Hartree-Fock 3-21G** model. *Click* on **OK**.

4. Select **Surfaces** from the **Setup** menu. First request a local ionization potential map. *Click* on **Add**, select **density** from the **Surface** menu and **ionization** from the **Property** menu and

click on **Apply**. Then request a LUMO map. Change **ionization** (**Property** menu) to **|LUMO|** and *click* on **OK**.

5. Submit the job. Name it ***tetramethylsilaethylene***. (Even though both molecules are in SMD, ***Spartan*** does not attempt to assign names to documents that contain more than one molecule.) When completed, bring up the spreadsheet (**Spreadsheet** under the **Display** menu) and *check* the box at the far left for each entry in the spreadsheet. This allows simultaneous display of both molecules. Select **Coupled** (**Model** menu) to uncouple the motions of the molecules. Position the two molecules side by side on screen.

6. Inside the **Surfaces** dialog, select ***density ionization...***. Compare local ionization potential maps for the olefin and silaolefin, recognizing that the more red the color, the lower the ionization potential and the more susceptible toward electrophilic attack. Which molecule do you conclude is likely to be more reactive? Turn "off" local ionization potential maps by again selecting ***density ionization...***. Select ***density |LUMO|...*** to turn on LUMO maps. Here, the more blue the color, the greater the concentration of the LUMO and the more susceptible toward nucleophilic attack. Which molecule do you conclude is likely to be more reactive? Close ***tetramethylsilaethylene*** when you are finished.

7 - 10 Optional *(Requires Access to the Cambridge Structural Database)*

Identify molecules in the Cambridge Structural Database (CSD) that incorporate a carbon-silicon double bond.

7. Build parent silaethylene, $H_2Si{=}CH_2$, by starting from ethylene and (using the inorganic model kit) substituting a silicon for one of the carbons.

8. Bring up the **Cambridge Structural Database** (**CSD**) dialog by first selecting **Databases** from the **Search** menu and then *clicking* on the **CSD** tab. *Click* on all four free valences in silaethylene.

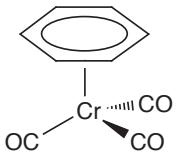
They will be marked by orange cones. *Click* on the **Search** button at the bottom of the dialog. In a few seconds, hits will begin to appear in the box at the right of the dialog. Examine each in turn (in the window at the left of the dialog) by *clicking* on its CSD reference code (**REFCODE**). Identify a structure that appears to be representative for closer scrutiny. Before you retrieve it, *click* on ▼ to the right of the **Retrieve** button at the bottom of the dialog. Make certain that all five boxes under **Post Retrieval Filters** in the **Retrieve Options** dialog are *checked*. *Click* on **OK**. *Click* on the selected **REFCODE** and then *click* on the **Retrieve** button. Dismiss the **CSD** dialog.

9. *Click* on  . Except for hydrogens, all atoms in the structure you have retrieved will be fixed and energy minimization will only change hydrogen positions. This is to compensate for the fact that hydrogens are generally not well located in X-ray crystal structure determinations.

   Change the display to a space-filling model. Pay particular attention to the incorporated silicon-carbon double bond. Is it accessible in all of the compounds or is it encumbered by the surrounding groups? What, if anything, do your observations suggest about why this particular compound can be (has been) made?

10. Close silaethylene (without saving) and any open dialogs.

Benzene Chromium Tricarbonyl

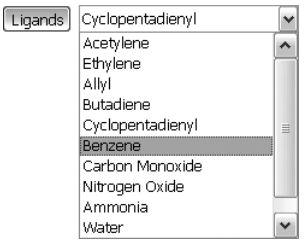


benzene chromium tricarbonyl

Comparison of electrostatic potential maps for this system and that of free benzene will allow you to classify $Cr(CO)_3$ as an electron donor or an electron acceptor substituent.

1. *Click* on and bring up the inorganic model kit. *Click* on the atom bar and select **Cr** from the *Periodic Table*. Select the four-coordinate tetrahedral structure from the list of atomic hybrids. *Click* anywhere on screen.

2. *Click* on **Ligands** in the model kit, select **Benzene** from the menu of available ligands.



   Click on one of the free valences on the four-coordinate chromium center.

3. Select **Carbon Monoxide** from the **Ligands** menu, and *click* on the remaining (three) free valences on chromium. *Click* on to produce a refined structure.

4. Select **New Molecule** from the **File** menu. The screen will blank. Build benzene and *click* on . *Click* on . The document now contains two molecules, benzene chromium tricarbonyl and benzene.

5. Select **Calculations...** (**Setup** menu). Specify calculation of **Equilibrium Geometry** with the **Semi-Empirical PM3** model. Make certain that **Global Calculations** (at the bottom of the

dialog) is checked. You want the calculations to apply to both benzene chromium tricarbonyl and benzene. *Click* on **OK**.

6. Select **Surfaces** (**Setup** or **Display** menu). *Click* on **Add...**. Specify **density** from the **Surface** menu, and **potential** from the **Property** menu, and *click* on **OK**. Make certain that **Global Surfaces** is checked.

7. Submit the job. Name it ***benzene chromium tricarbonyl***. When completed bring up the spreadsheet (**Spreadsheet** from the **Display** menu), and *check* the box to the right of the label for both entries. This allows the two molecules to be displayed simultaneously on screen. If **Coupled** (**Model** menu) is *checked*, remove the checkmark by selecting it. The two molecules may now be moved independently. Orient each molecule so that you can clearly see the benzene face (exposed face in the case of the organometallic).

8. Select ***density potential*** from the **Surfaces** dialog. Compare electrostatic potential maps for both free and complexed benzene, with attention to the exposed benzene face.* Does the $Cr(CO)_3$ group donate or withdraw electrons from the ring? Would you expect the aromatic ring in benzene chromium tricarbonyl to be more or less susceptible to electrophilic attack than free benzene? More or less susceptible to nucleophilic attack?

9 optional ***(cannot be completed with the Essential Edition)***

9. Repeat the two calculations using density functional theory. Make a copy of ***benzene chromium tricarbonyl*** (); name it ***benzene chromium tricarbonyl density functional***. Inside the **Calculations** dialog, specify an **Energy** calculation using the **BP/6-31G* Density Functional** model. Submit. When completed, examine the electrostatic potential maps. Are they qualitatively similar to those from the PM3 calculations?

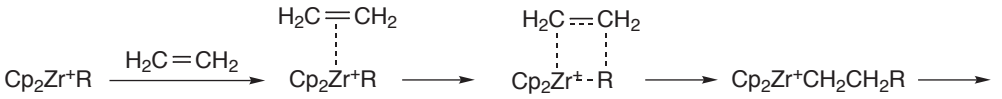
10. Remove all molecules and any remaining dialogs from the screen.

---

* Electrostatic potential maps (as well as other maps) for molecules in a group will be put onto the same (color) scale. This allows comparisons to be made among different members.

Ziegler-Natta Polymerization of Ethylene

Ziegler-Natta polymerization involves a metallocene. This first complexes an olefin, which then inserts into the metal-alkyl bond.

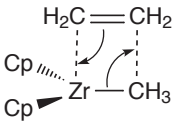


In this tutorial, you will use PM3 calculations to obtain a transition state for insertion of ethylene into $Cp_2ZrCH_3^+$ and, (optionally), estimate the activation energy using density functional calculations.

1. Bring up the inorganic model kit. *Click* on the atom bar and select **Zr** from the *Periodic Table*. Select from the list of hybrids and *click* on screen. Select **Cyclopentadienyl** from the **Ligands** menu and *click* on two of the free valences on zirconium.

2. Move to the organic model kit. Select sp^3 carbon and *click* on the remaining free valence on zirconium. Select **Alkenyl** from the **Groups** menu, *click* on the **Insert** button at the bottom of the model kit or hold down the **Insert** key (**alt** key on Mac) and *click* anywhere on screen or *double click* in a blank area on screen.

3. Orient the two fragments (Cp_2ZrCH_3 and ethylene) as shown below:



   (To move the two fragments independently, hold down the **Ctrl** key.)

4. Select **Transition States** from the **Search** menu ( ). *Click* on the ZrC (methyl) bond, hold down the **Shift** key and, one after another, *click* on the methyl carbon and on one of the ethylene carbons. Next, *click* on the ethylene double bond, hold down the **Shift** key and, one after another, *click* on the other ethylene carbon and on zirconium. *Click* on at the bottom right of the screen. In a few seconds, a guess at the transition state appears.

5. Bring up the **Calculations** dialog. Specify a **Transition State Geometry** calculation using the **PM3 Semi-Empirical** model. Change **Total Charge** to **Cation** and *check* **IR** to the right of **Compute**. *Click* on **Submit** and name the job ***Cp2ZrMe cation + ethylene***.

6. When the job has completed, bring up the **IR** dialog (**Spectra** from the **Properties** menu and *click* on the **IR** tab) and *click* on the imaginary frequency at the top left of the dialog. Would you describe the process as concerted or occurring in discrete steps?

7 to 9 optional *(cannot be completed with the Essential Edition)*

7. Perform BP/6-31G* density functional energy calculations using PM3 geometries to obtain a better estimate for the energy barrier for ethylene insertion. Make a copy of ***Cp2ZrMe cation+ethylene*** (); name it ***Cp2ZrMe cation+ethylene density functional***. Enter the **Calculations** dialog with this copy, and specify calculation of **Energy** using the **BP/6-31G* Density Functional** model. *Check* **Pseudopotential** at the far lower right of **Calculate** to specify use of a pseudopotential for Zr associated with the 6-31G* basis set. Remove the checkmark on **IR** (to the right of **Compute**). **Total Charge** should still be set to **Cation**. *Click* on **Submit**.

8. Build both ethylene and $Cp_2ZrCH_3^+$ (name them ***ethylene density functional*** and ***Cp2ZrMe cation***, respectively). For $Cp_2ZrCH_3^+$, start with three-coordinate trigonal Zr, and then add two cyclopentadienyl ligands and a four-coordinate tetrahedral carbon. For each, enter the **Calculations** dialog, and specify calculation of **Energy** using the **BP/6-31G* Density Functional** model. Select **PM3** from the **Start from** menu to designate use of a PM3 geometry. For ***Cp2ZrMe cation density functional*** (only) *check* **Pseudopotential**, and set **Total Charge** to **Cation**.

9. Submit both jobs. When they have completed, calculate an activation energy for the insertion reaction.

10. Remove all molecules and any remaining dialogs from the screen.

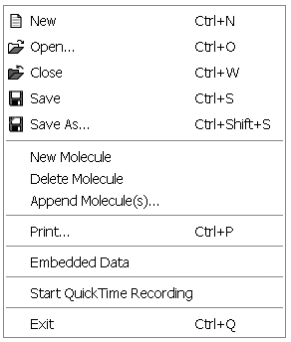
Section III
Features and Functions

This section describes the functions available under the menus incorporated into ***Spartan'08***, and is intended to serve as a general reference to the program. The coverage follows the order of the menus presented in ***Spartan's*** user interface: **File**, **Edit**, **Model**, **Geometry**, **Build**, **Setup**, **Display**, **Search**, **Options** and **Help**. In turn, it describes the functions and usage of each of the menu entries. Of particular note, are **Calculations** under the **Setup** menu which enumerates the available molecular mechanics and quantum chemical methods as well as the similarity analysis tools, and **Surfaces** under both **Setup** and **Display** menus which enumerates the available graphical models. While general commentary is provided about the performance and requirements of different computational methods and the utility of the different graphical models, this is not our primary purpose. The reader is referred to ***A Guide to Molecular Mechanics and Quantum Chemical Calculations***, available as a PDF under the **Help** menu.

Chapter 10

The File Menu

*Operations under the **File** menu access model kits to build, edit and substitute molecules and the file system to read and write both native and non-native files, print text and on-screen graphics, to embed text and other files into **Spartan** documents and to create QuickTime movies.*



New ( )

Brings up a model kit and clears the screen. Model kits are discussed in **Chapter 14**.

Open... ( )

Opens a file that contains all information associated with a particular molecule or pharmacophore (or list of molecules and/or pharmacophores). In addition to native (.spartan) files (documents), supported are files containing 2D drawings: CambridgeSoft (ChemDraw) CDX (.cdx), and MDL SDF (.sdf)*, SKC (.skc), and

* SDF files may contain information in addition to geometry, for example, the name of the molecule or its CAS number. If available, each quantity appears in a line following <**xyz**> where **xyz** is a character string denoting the quantity, for example, **CAS** refers to the CAS number.

<CAS>
22090-26-2

This additional information may be accessed in the spreadsheet by typing the character string into a header cell, for example, typing **CAS** will result in the CAS number.

TGF (.tgf), files containing 3D structures: MacroModel (.mac), SYBYL Mol (.mol), SYBYL Mol2 (.mol2), and PDB (.pdb), and files containing 1D "strings": SMILES (.smi). Finally, infrared, NMR and UV/visible spectra may be input (discussion of file formats is provided in **Appendix I**).

> We use the word "file" to refer to any information contained in the "file system" (the disk on your computer) and the word "document" to refer to files that are native to ***Spartan***. Thus, an MDL SDF ***file*** may be read into ***Spartan*** and then written as a ***document***. Documents may contain information on a single molecule or single pharmacophore or on a collection of molecules and/or pharmacophores. This collection is loosely referred to as a "list" or as a "spreadsheet" (in the context that it is presented in a spreadsheet).

Cartesian coordinates from any of these files (except spectra files which do not contain coordinates) may be replaced by coordinates generated based on atomic connectivity using **Replace Coords.** in the **Molecule Utilities** dialog (**Properties** under the **Display** menu; **Chapter 16**). Non-native files are normally hidden from view, but may be seen by selecting **All Files** from the **Files of type** menu at the bottom of the dialog.

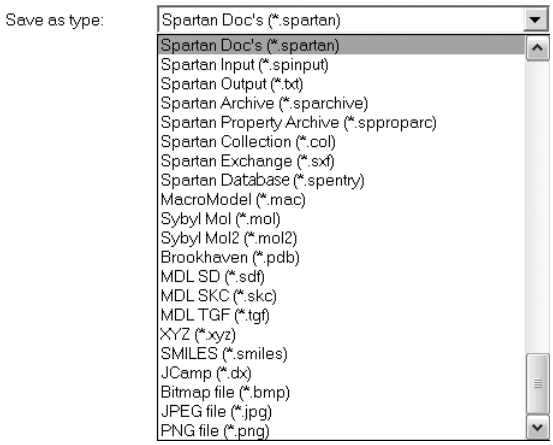
Close ()

Closes the document containing the selected molecule, as well as the spreadsheet, graphics, plots, spectra and document specific dialogs. If the document has not previously been saved, a name is requested. If a previously-saved document has been altered, verification that the changes are to be saved is requested.

Save
Save As... ()

Saves the document containing the selected molecule ***exactly as it appears on screen***. Opening the document will bring it on screen exactly as it was last saved. If the document has not previously been named, **Save** behaves as **Save As...**. Documents may be either be saved in native format or in one of the formats listed under **Open**.

addition, bitmap (.bmp), JPEG (.jpg) and PNG (.png) graphics file formats are supported as is the Spartan Database (.spentry) format for creating custom databases (see **Appendix G**). Support is also provided for writing QuickTime (movie) files (see discussion later in this chapter). Selection is made under the **Save as type** menu in the **Save As** dialog.



New Molecule

Brings up a model kit and clears the screen. This differs from **New**, in that the resulting molecule is appended to the end of the list associated with the molecule that is presently selected.*

Delete Molecule

Deletes the selected molecule(s) from a document. Deleting of the last molecule leads to an empty document.

---

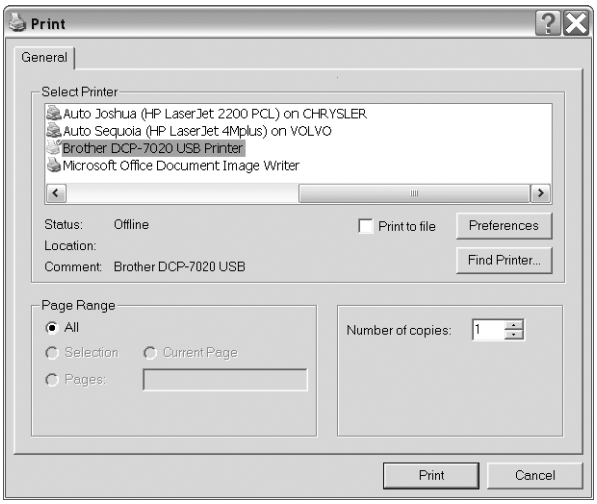
* Examples are provided in the tutorials ***Dienophiles in Diels-Alder Cycloadditions and Hydration of Carbonyl Compounds*** in **Chapter 5**, ***Stereospecific Diels-Alder Reactions*** in **Chapter 6**, ***Polyglycine*** in **Chapter 7** and ***Benzene Chromium Tricarbonyl*** in **Chapter 9**.

Append Molecule(s)...

Appends one or more documents onto the end of the document that contains the selected molecule.* **Append Molecule(s)...** leads to a file browser from which one or more documents need to be selected.**

Print/Print Output/Print Spreadsheet

Selection leads to a dialog in order to designate a printer, specify print layout and number of copies. It also allows "printing" to a file.



By default, prints whatever is presently displayed on screen. To print text output, bring up and select an output window (**Output** under the **Display** menu) and select **Print Output...** (that has replaced **Print...**) from the **File** menu. To print the contents of the spreadsheet, bring

---

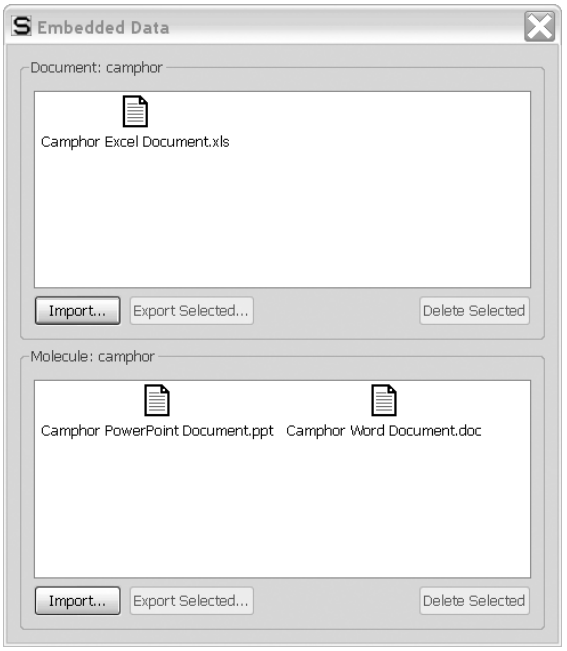
* An example is provided in the tutorial ***Carbene Additions to Alkenes*** in **Chapter 6**.

** Alternatively, molecules may be appended onto an existing document either by copy/paste operations using the clipboard or by *dragging* from an external window. Both require that the spreadsheet associated with the destination document be open on screen.
To copy a molecule open on screen onto the clipboard, first select (*click* on) it, and then select **Copy** from the **Edit** menu. Alternatively, *click* on its label in its spreadsheet (in the leftmost column), and then select **Copy** from the **Edit** menu. The latter permits several molecules to be selected (and copied) at once using the **Shift** and **Ctrl** keys in the usual manner. Once on the clipboard, the molecule or molecules may be moved to the destination list by *clicking* on an empty row header in the spreadsheet (for the destination document), and then selecting **Paste** from the **Edit** menu.
To copy a document from an external window, *drag* it onto the open spreadsheet (associated with the destination document) inside of ***Spartan***. Several documents can be *dragged* at once using the **Shift** and **Ctrl** keys in the usual manner.

up a spreadsheet (**Spreadsheet** under the **Display** menu) and select **Print Spreadsheet...** (that has replaced **Print...**) from the **File** menu. The results of a reaction energy calculation (**Reactions...** under the **Display** menu; **Chapter 16**), may be printed using **Print** from the contextual menu.

Embedded Data

This allows non-native files, for example, Word, Excel or PowerPoint files, to be embedded in a ***Spartan*** document. These may be associated either with an individual molecule or with all the molecules (or both). Selection results in the **Embedded Data** dialog, the upper half of which relates to document level files and the lower half of which relates to molecule level files.



A file may be imported (at the document or molecule level) by *clicking* on the appropriate **Import** button. This leads to a file browser from which a file may be selected. Once embedded in a ***Spartan*** document, a non-native file may be exported by first selecting (*clicking* on) it and then *clicking* on **Export Selected**. It may be deleted by first selecting it and then *clicking* on **Delete Selected**.

Note that ***applications cannot be launched*** from the **Embedded Data** dialog. They need to be copied outside prior to launching.

Start/Stop QuickTime Recording

This allows QuickTime movies to be created. To start making a movie, select **Start QuickTime Recording**. Any motions of all molecules in ***Spartan's*** main screen will be captured. Dialogs (including the builders) will not be captured. Note the use of "tumbling" (see **Settings Preferences** under **Preferences** in the **Options** menu; **Chapter 18**) in making QuickTime movies. To stop making a movie, select **Stop QuickTime Recording** (which has replaced **Start QuickTime Recording** in the **File** menu) and supply the requested file name.

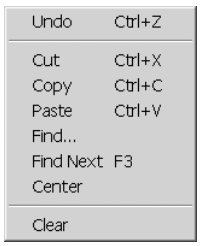
Exit

Exits ***Spartan***, that is, clears the screen and closes all open documents. A prompt for a name is provided for each document that has not previously been saved.

Chapter 11

The Edit Menu

*Operations under the **Edit** menu provide for undoing commands, copying items to and from the clipboard, finding text and graphics, centering molecules on screen and clearing the selected molecule.*



Undo

Undoes the last operation from the **Build** and **Edit** menus and from the **Molecule Utilities** dialog. Undoes transition-state formation and retrieval from the Spartan Molecular Database.

Cut, Copy, Paste

Cut moves the selected item to the clipboard and removes it from the document. **Copy** copies the item to the clipboard. The item is unaffected. **Paste** transfers the contents of the clipboard to the selected location. The contents of the clipboard are unaffected.

Among the important uses of the clipboard within ***Spartan*** are:

(i) Transferring on-screen graphics into other applications such as Microsoft Word® and PowerPoint®.

(ii) Temporary storage of a molecular structure for use in molecule building.

(iii) Transferring data between ***Spartan*** spreadsheets and between a ***Spartan*** spreadsheet and other applications such as Microsoft Excel®.

(iv) Making multi-molecule documents and/or transferring molecules between documents.

Cut operations for (i) and (ii) require drawing a selection box. First position the cursor slightly above and slightly to the left of the item to be transferred. Then, while holding down both buttons, *drag* the mouse to a location slightly below and slightly to the right of the item to be transferred. Finally, release both buttons. **Copy** operations for (i) and (ii) also refer to the contents of a selection box ***if one has been drawn***, but to the selected molecule if a box has not been drawn. **Copy** operations from a ***Spartan*** spreadsheet refer to all information associated with a molecule if selection is made on the header cell of the leftmost column, but only to the selected (text) information if selection is made on any other column. Further discussions relating to use of the clipboard spreadsheet in molecule building is provided in **Chapter 14** and for operations involving multi-molecule documents in **Chapter 15**.

Find..., Find Next

Find locates a text string defined in the **Find** dialog if an output window or a spreadsheet is selected, or a structure sequence defined on the clipboard if an on-screen model is selected. **Find Next** locates the next occurrence of a text string or a structure sequence.

Center

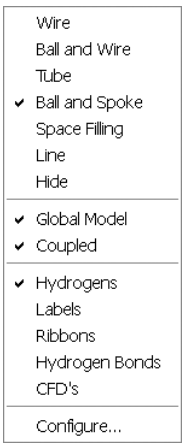
Centers on screen all molecules in the document for which the selected molecule is a member (only the selected molecule is displayed).

Clear

Clears (deletes) the structure and other information for the selected molecule, and brings up a model kit. No information is actually removed from the file system until the document is saved.

Chapter 12
The Model Menu

*Structure models available under the **Model** menu include wire, ball-and-wire, tube, ball-and-spoke, space-filling (CPK) and line models, with or without hydrogens, with or without hydrogen bonds indicated, with or without chemical function descriptors (CFD's) shown and with or without atom labels, as well as ribbon displays for polypeptides and polynucleotides, with or without labels and with or without hydrogen bonds indicated. The menu also provides for configuring atom labels to display element name, R/S chirality, mass number, charge or chemical shift, and for specifying color coding and display style for ribbon labels, as well as turning "on" and "off" a variety of other labels. Finally, it allows model style to be applied globally (to all molecules in a document) and models to be manipulated in concert.*



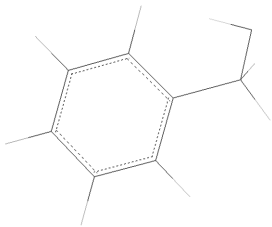
Only one model style **Wire**, **Ball and Wire**, **Tube**, **Ball and Spoke**, **Space Filling** or **Hide**) may be selected. The selected model is designated by a check mark ✔ in front of its entry in the menu. **Global Model**, **Coupled**, **Hydrogens**, **Labels**, **Ribbons**, **Hydrogen Bonds** and **CFD's** operate as toggle switches. A ✔ in front of the entry in the menu indicates that it is turned "on".

All structure models, and graphics may be displayed either in orthogonal or perspective projections. The latter may be valuable in helping to visualize large molecules. Selection is done in the **Settings Preferences** dialog (**Preferences...** under the **Options** menu; **Chapter 18**). Both structure models and graphics may be presented in 3D stereo. This is also controlled from the **Settings Preferences** dialog as well as from the **3** key. Stereographic displays require perspective projections.

Wire

This represents the molecule as a wire model where the vertices represent the atoms.

Wire Model



The bonds are drawn in two colors, one for each of the atoms making up the bond. Default atom colors are given in **Table 12-1**.

Atom colors apply globally (to all atoms of given type), and may be changed using the **Set Colors** dialog (**Colors** under the **Options** menu; **Chapter 18**). Colors of individually selected atoms may be set using the **Atom Style** dialog (**Properties** under the **Display** menu; **Chapter 16**). All models use the same color scheme for atoms, and provide for the same mechanism of changing colors globally or individually.

Ball and Wire

This represents atoms by small balls and bonds by wires.

Ball-and-Wire Model

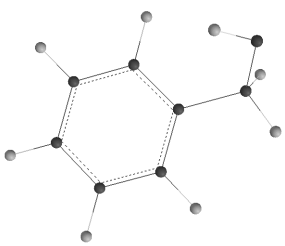


Table 12-1: Default Atom Colors

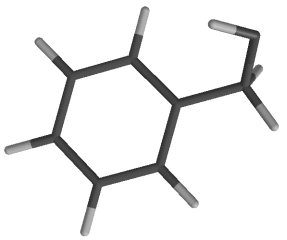
main group		main group (con't)	
Hydrogen	white	Indium	orange
Lithium	orange	Tin	gray
Beryllium	green	Antimony	orange
Boron	orange	Tellurium	orange
Carbon	gray	Iodine	violet
Nitrogen	blue-gray	**transition metals**	
Oxygen	red	Scandium-Zinc	green
Fluorine	pale yellow	Yttrium-Cadnium	green
Sodium	yellow	Lanthanum-Mercury	green
Magnesium	purple	**lanthanides**	
Aluminum	magenta	Cerium-Lutetium	violet
Silicon	gray	**actinides**	
Phosphorus	orange	Thorium-Lawrencium	dark blue
Sulfur	yellow	**noble gases**	
Chlorine	green	Helium	dark orange
Potassium	red	Neon	dark orange
Calcium	red	Argon	dark orange
Gallium	orange	Krypton	dark orange
Germanium	gray	Xenon	orange
Arsenic	dark orange		
Selenium	red orange		
Bromine	dark red		
Rubidium	red		
Strontium	red		

The balls are color coded according to atom type, and the wires representing bonds are drawn in two colors (as in wire models).

Tube

This is similar to the wire model, except that tubes instead of wires are used to represent bonds.

Tube Model

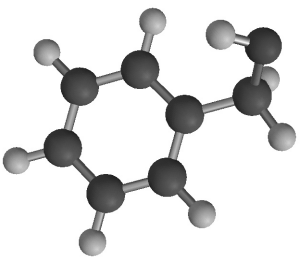


Tubes may either be solid cylinders or be split to represent multiple bonds depending on whether **Split Tubes** in the **Settings Preferences** dialog (**Preferences...** under the **Options** menu; **Chapter 18**) is "off" or "on". As with wire models, bonds are drawn in two colors.

Ball and Spoke

This represents atoms by balls (the size and color of which depend on atom type), and bonds by spokes.

Ball-and-Spoke Model

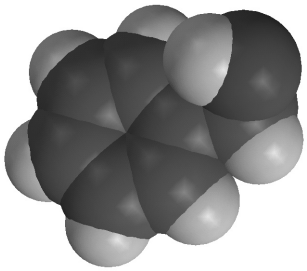


Spokes may either be cylinders or be split to represent multiple bonds depending on whether **Split Tubes** in the **Settings Preferences** dialog (**Preferences...** under the **Options** menu; **Chapter 18**) is "off" or "on". Bond (spoke) color is gray by default but it may be changed using the **Set Colors** dialog (**Colors** under the **Options** menu; **Chapter 18**). Colors of individually selected bonds may be set using the **Bond Style** dialog (**Properties** under the **Display** menu; **Chapter 16**).

Space Filling

This represents the molecule as a composite of spheres, the radii of which have been chosen to approximate van der Waals contact distances. Also known as CPK models, space-filling models are intended to portray overall molecular size and shape.

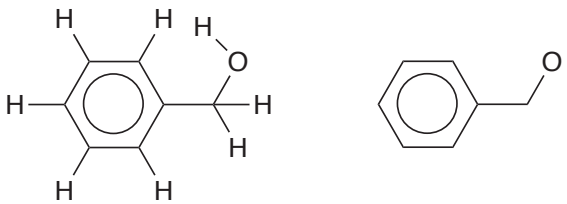
Space-Filling Model



Volume, surface area and polar surface area (PSA)* displayed in the **Molecule Properties** dialog (**Properties** under the **Display** menu; **Chapter 16**) correspond to a space-filling model.

Line

The **Line** model has been provided primarily to allow ***Spartan*** to display native structures from 2D drawing programs such as ChemDraw and IsisDraw. It designates carbons as the termini or the intersection of lines and non-carbons by their elemental symbols without any use of color.



These models are perhaps best shown minus hydrogens (see **Hydrogens** later in this chapter), and include stereochemical cues (— and ⋯) if these have been provided in the 2D drawing or if they are specified from the **Bond Properties** dialog (**Properties** under the **Display** menu; **Chapter 16**).

---

* By default, polar surface area is defined as the area due to nitrogen and oxygen and any hydrogens attached to nitrogen and oxygen. The user may employ other combinations of atoms to define a polar surface area. See **Spreadsheets** under the **Display** menu (**Chapter 16**).

Hide

This removes the structure model from the screen. This may be desirable where its display may lead to unnecessary crowding, for example, in proteins where ribbon displays are more appropriate. A structure model may be restored by selecting it from the **Model** menu.

Different parts of a molecule may be rendered in terms of different model styles and colors. This allows focus on specific "interesting" regions of a molecule. Regions may be individual atoms and/or bonds or any collection of atoms and/or bonds contiguous or not. They are selected either by *clicking* on an individual atom or bond or, with the aid of the **Ctrl** (select multiple) or **Command** key (Macintosh), **Shift** (select over a range), and **Alt/Alt Option** (select all bonded) keys, by *clicking* on a set of atoms and/or bonds, or by defining a selection box. Discussion has already been provided in **Chapter 2**. Style and color of selected model components is by way of the "**Selected**" **Style** dialog under **Properties** in the **Display** menu (**Chapter 16**).

Global Model

If *checked* (turned "on"), this signifies that all molecules in a document will share attributes. These include presentation of hydrogens, atom and other labels, hydrogen bonds, CFD's and ribbon displays. Global model style is controlled from the **Molecule Preferences** dialog (**Preferences...** under the **Options** menu; **Chapter 18**). **Global Model** acts in a toggle manner, switching between global and local display. **Global Model** is normally "on".

Coupled

If *checked* (turned "on"), this signifies that all molecules in a document selected for simultaneous display will be moved together. **Coupled** is turned "on" following molecule alignment (see **Align** under the **Geometry** menu; **Chapter 13**). **Coupled** acts in a toggle manner, that is, repeated selection couples and decouples the molecules.

Hydrogens

If *checked*, this signifies that hydrogens are to be included in the model. Note that structures retreived from the Cambridge Structural Database or the Protein Data Bank may lack hydrogens. These need to be "grown" before they can be displayed, before CFD's can be assigned (see discussion later in this chapter) and before calculations may be performed. This may be done either upon retrieval (see **Databases** and **Ligands** under the **Search** menu; **Chapter 17**) or using **Grow Hydrogens** in the **Molecule Utilities** dialog (**Properties** under the **Display** menu; **Chapter 16**). **Hydrogens** acts in a toggle manner, that is, repeated selection turns "on" and "off" the display of hydrogens.

Labels

If *checked*, this signifies that labels associated with atoms, ribbons and bonds as well as with other attributes specified in **Configure...** (see discussion later in this chapter) are to be displayed in the model. **Labels** acts in a toggle manner, that is, repeated selection turns "on" and "off" display of labels. **Labels** is automatically turned "on" following selection of **Apply** or **OK** in the **Configure** dialog.

Ribbons

If *checked*, this signifies that ribbons are to be displayed along with the selected model. (If only ribbons are desired, for example, in proteins, select **Hide** for the model.) **Ribbons** acts in a toggle manner, that is, repeated selection turns "on" and "off" display of ribbons.

Hydrogen Bonds

If *checked*, this signifies that hydrogen bonds are to be drawn as part of the model. Hydrogen bonds are defined as non-bonded contacts between a nitrogen or oxygen and a hydrogen attached to nitrogen or oxygen separated by a distance ranging from 1.6 to 2.1Å and making an X–H--Y (X,Y = N, O) angle of >120°. **Hydrogen Bonds** acts in a toggle manner, that is, repeated selection turns "on" and "off" display of hydrogen bonds.

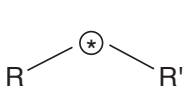
CFD's

If *checked*, this signifies that **C**hemical **F**unction **D**escriptors (**CFD's**) are to be displayed along with the structure model. CFD's are descriptors given to a molecule in order to characterize or anticipate its chemical behavior or to identify commonality among molecules with different structures. They parallel terms in a chemist's vocabulary such as ***lone pair*** (to suggest the role of a hydrogen-bond acceptor) and ***sterically crowded*** (to suggest that getting close may be difficult).

Spartan uses seven different kinds of CFD, the first six of which may be though of as attributes of a molecule.

	CFD	chemical language
1.	hydrophobe	sterically-crowded region
2.	aromatic	aromatic π system
3.	hydrogen-bond donor	acidic hydrogen
4.	hydrogen-bond acceptor	lone pair
5.	positive ionizable site	basic site
6.	negative ionizable site	acidic site

Default CFD assignments follow from atomic connectivity and rules covering common organic functional groups. They may be modified on a case-by-case basis using the **CFD Properties** dialog (**Properties** under the **Display** menu; **Chapter 16**). CFD's (represented by *'s in the figures below) may either be atom centered, or they may be centered in between atoms. For example, a hydrogen-bond acceptor CFD is placed on an ether oxygen, while an aromatic CFD is placed at the center of a benzene ring.

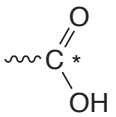


hydrogen-bond acceptor

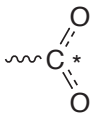
aromatic

CFD's may also be associated with a number of common functional groups. For example, a single CFD is provided for the carboxylic acid group, sited in-plane and equidistant from the carbon and two oxygens.

The CFD's for this and other functional groups may depend on pH. At neutral (or basic) pH this CFD should be designated both as a hydrogen-bond donor and a hydrogen-bond acceptor, whereas at acidic pH it should be designated as a negative ionizable site.



neutral pH: hydrogen-bond donor
hydrogen-bond acceptor



acidic pH: negative ionizable site

Finally, note that default assignments for hydrogen-bond donors and acceptors depend not only on atomic type (N, O, . . .) and (in the case of hydrogen-bond donors) on the availability of hydrogens, but also on known (or presumed) chemical behavior. For example, because an amide nitrogen would generally not be considered to be a hydrogen-bond acceptor, no (hydrogen-bond acceptor) CFD is provided. It is assigned to be a hydrogen-bond donor (if hydrogens are available) and the amide oxygen is assigned as a hydrogen-bond acceptor.

Note that, hydrogen-bond acceptor and donor CFD's may be related to the electrostatic potential obtained from quantum chemical calculations. For example, a negative potential associated with an oxygen center suggests that it might serve as a hydrogen-bond acceptor, while a positive potential associated with a hydrogen attached to such a center suggests its role as a hydrogen-bond donor. These are, of course, the same conclusions that a chemist would draw simply by looking at the molecule. However, the calculations may be able to say which acceptor or donor sites are likely to be strong and which are likely to be weak.

Calculated electrostatic potentials may offer significant advantage over CFD's and other qualitative descriptions when there are no heteroatoms and hydrogen bonding is not possible. The π systems of benzene and other aromatics provide the most common examples. The same conditions that make the ring susceptible to electrophilic aromatic substitution mean that it repels other electron-rich regions (and attracts electron-poor regions).

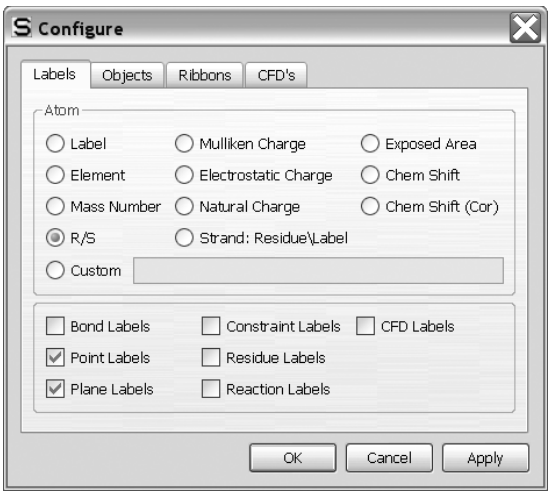
A seventh CFD type, excluded volume, derives from knowledge of a molecule incorporated into a host. Excluded volumes may be obtained

from PDB files using the **Ligands** dialog (**Ligands** under the **Search** menu; **Chapter 17**).

Configure...

This selects the types of labels attached to atoms, ribbons and CFD's.

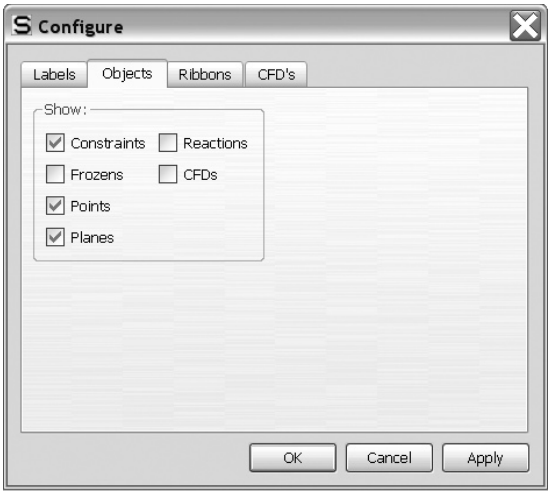
Configure Labels



Atom labels may be selected from among the following: **Labels**, a unique element/number combination that may be changed from the **Atom Properties** dialog (**Properties** under the **Display** menu; **Chapter 16**), **Element**, **Mass Number**, **R/S** (chirality), **Mulliken Charge**, **Electrostatic Charge**, **Natural Charge**, **Strand: Res\ Label** (polypeptides and polynucleotides), **Exposed Area** (of an atom in a space-filling model), **Chem Shift** or **Chem Shift (Cor)**. Selection of **Custom** leads to a text box into which a formula defining an atomic label may be entered. Details are provided in **Spreadsheet** under the **Display** menu (**Chapter 16**). In addition, **Bond Labels**, **Point Labels**, **Plane Labels**, **Constraint Labels**, **Residue Labels**, **Reaction Labels** and/or **CFD Labels** may be provided. Default settings (for a new molecule) are made in the **Molecule Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 18**).

Configure Objects

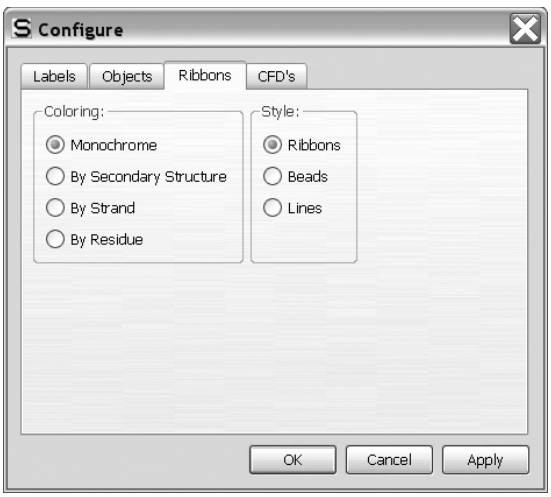
Clicking on the **Objects** tab leads to the **Configure Objects** dialog.



If checked, **Constraint** and **Frozen** markers, **Points** and **Planes**, **Reaction** arrows and **CFD's** attach to the model. If not checked, these are shown only in the respective modes, for example, **Frozen** markers are shown only if **Freeze Center** is selected.

Configure Ribbons

Clicking on the **Ribbons** tab leads to the **Configure Ribbons** dialog.

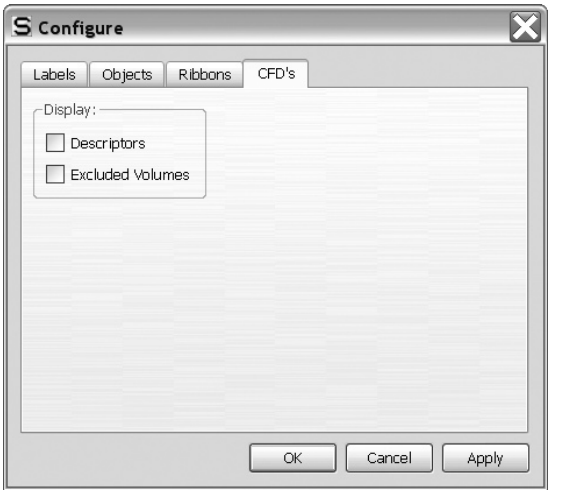


Ribbon coloring may be selected from among the following: **Monochrome**, **By Secondary Structure**, **By Strand** or **By**

Residue. Ribbon style may be selected from among the following: **Ribbons**, **Beads** or **Lines**.

Configure CFD's

Clicking on the **CFD's** tab leads to the **Configure CFDs** dialog.



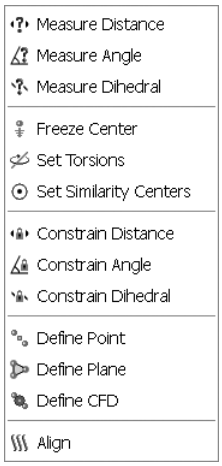
This allows for turning "on" CFD descriptors (identifying the type of CFD) identifiers and CFD's corresponding to excluded volumes.

The **Configure** dialog is removed from the screen with all selections maintained by *clicking* on **OK**. *Clicking* on **Cancel** or on ☒ removes the dialog but selections are lost. *Clicking* on **Apply** maintains the selections but leaves the **Configure** dialog on screen. Note, that **Labels** (from the **Model** menu) will be turned "on" following either *clicking* on **OK** or on **Apply**.

Chapter 13

The Geometry Menu

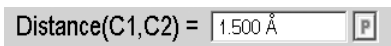
*Functions available under the **Geometry** menu allow querying and changing bond lengths, angles and dihedral angles, defining points and ligand points, planes and CFD locations, setting geometrical constraints, freezing atomic centers, altering default bond and ring conformer assignments, introducing NOE's in conformational searching, selecting atoms or CFD's for similarity analysis and aligning molecules in a list.*



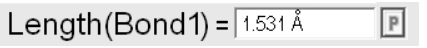
Measure Distance ( )
Measure Angle ( )
Measure Dihedral ( )

Measure Distance displays the distance (in Ångstroms) between two atoms, whether or not they are bonded. Selection results in a message at the bottom left of the screen.

Clicking on two atoms displays the distance at the bottom right of the screen.



Alternatively, *clicking* on a bond displays its length.



Measure Distance may also be used to alter the distance between atoms (as long as both are not incorporated into the same ring), by altering the contents of the box to the right of **Distance (A,B) =** or **Length (A)=**, and then *pressing* the **Enter** key (**return** key on Mac). The distance (length) may be entered into the spreadsheet by *clicking* on P to the right of its display (see **Spreadsheet** under the **Display** menu; **Chapter 16**).

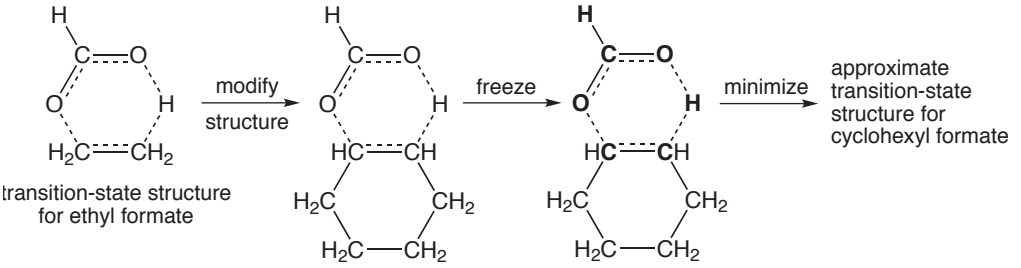
Angle and dihedral angle queries are handled in a similar manner. Angles require that three atoms or two bonds be identified in the proper order while dihedral angles require that four atoms or three bonds be identified in the proper order.

Freeze Center ( )

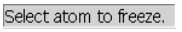
This allows one or more atoms to be held in place during minimization (in build mode) or (optionally) during equilibrium or transition-state geometry optimization, conformational searching, or energy profile generation. The latter needs to be explicitly indicated in the **Calculations** dialog (**Chapter 15**).

Atom freezing may be useful in a number of situations, among them guessing a transition-state geometry for a reaction that is closely related to one for which a transition state is available. For example, a good guess at the transition state for pyrolysis of cyclohexyl formate will be obtained by modifying the transition state for pyrolysis of ethyl formate, freezing all but the modified sections (designated in bold in the figure below) and then minimizing.*

* An example of this is presented in the tutorial ***Carbene Additions to Alkenes*** in **Chapter 6**.



Selection of **Freeze Center** leads to a message at the bottom left of the screen.



Clicking on an atom or free valence*, freezes it; *clicking* again thaws it. Buttons at the bottom right of the screen allow for freezing all atoms (**Freeze All**), freezing all heavy (non-hydrogen) atoms (**Freeze Heavy**) and for thawing all atoms (**Thaw All**).

Another important use of frozen atoms is to "correct" XH bond lengths in structures resulting from a search of the Cambridge Structural Database or the ligands database (see **Databases** under the **Search** menu; **Chapter 17**). Hydrogen positions are more often than not poorly located in small molecule X-ray structures, and XH bond lengths are commonly as much as 0.1 to 0.2Å shorter than they should be. Structures incorporating such bond lengths are clearly inappropriate for energy and property calculations and may also be problematic as starting geometries in quantum chemical calculations. Protein crystal structures always lack hydrogens, and hydrogen positions of bound ligands extracted from PDB files may be poorly defined. A reasonable solution to both problems is to freeze all heavy (non-hydrogen) atoms (**Freeze Heavy**) and then to carry out molecular mechanics minimization using either **Minimize** from the **Build** menu (see discussion in **Chapter 14**) or from the **Calculations** dialog under the **Setup** menu (see discussion in **Chapter 15**).

Frozen atoms are indicated by magenta colored markers ( ). Whether or not these are included with the model (outside of freeze center mode) for an individual molecule is controlled from the **Molecule Utilities** dialog available under the **Display** menu (**Chapter 16**). Global settings are controlled from the **Molecule Preferences** dialog under **Preferences...** in the **Options** menu (**Chapter 18**).

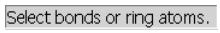
* The bond distance in this case is that appropriate for hydrogen being added to the free valence.

Set Torsions ()

Spartan automatically identifies bonds and rings for conformational searching and specifies default step sizes.

> Two different rule sets have been provided. **Normal** is intended for use in determining the lowest-energy conformation or in establishing a distribution of low-energy conformers, whereas **Trimmed** is intended for use in building conformer libraries. Selection of rule set is made in the **Settings Preferences** dialog (**Preferences...** under the **Options** menu; **Chapter 18**).

Set Torsions allows these defaults to be altered. Selection results in rotatable bonds each being marked by a gold cylinder, and flexible rings each being marked by a gold circle around one or more atoms, and a message at the bottom left of the screen.



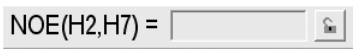
Clicking on a bond or an atom contained in a ring selects it for rotation. In the case of a ring, rotation means that the atom is to be puckered up and down (restricted rotation). The default rotation is provided in a box to the right of **Fold** at the bottom right of the screen. This is typically 2 or 3 for a single bond (step size of 180° and 120°, respectively) and 3 for a flexible ring. A value of either 0 or 1 indicates that the bond (ring) is not to be rotated (rotation by 360°). Other integer values may be entered into the box, followed by *pressing* the **Enter** key (**return** key on Mac). The original settings may be retrieved by *clicking* on **Defaults** at the bottom right of the screen. *Double clicking* on a marked bond or ring atom removes the marker and eliminates the conformational degree of freedom.

> Saturated six-membered rings for example, cyclohexane, are treated as special cases with only chair conformers examined. This is indicated by two circles on opposite atoms. Other conformers, for example, twist-boat conformers, can be generated by selecting additional ring atoms.

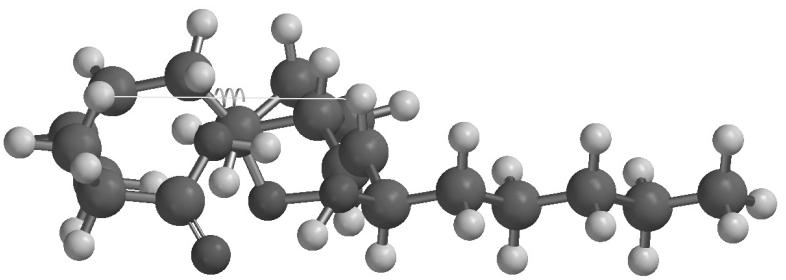
 and  buttons at the bottom right of the screen are available to

step through the possible single-bond conformers* (ring conformers are not provided). Geometries have not been optimized and some conformers may be severely crowded. Any conformer can be selected in lieu of the initial structure. That is, the conformer in vew upon exiting set torsions mode will replace the conformer in view at entry. The full set of single-bond conformers may be generated and placed in an unnamed document by *clicking* on ▼ (to the right of the step buttons), and then *clicking* **OK** in the dialog that results.** Note, that duplicate conformers have not been removed.

Set Torsions is also used to specify non-bonded distances that need to be kept below a threshold value. These follow from NOE (Nuclear Overhauser Effect) measurements and may be referred to as NOE conditions or simply NOEs. NOEs are specified by *clicking* on two (non-bonded) atoms while holding down the **Shift** key. In response, a message appears at the bottom of the screen.



Clicking on 🔓 changes it to 🔒 and enters the default value for the NOE threshold into the box. This value can be changed. A line is drawn between atoms that are to be kept within the threshold value.



Once set, NOEs are used in conformational searching (conformer distribution only) without further user intervention. They act as a post-processor filter to eliminate conformers that do not satisfy the constraints and do not affect the speed of the search.

---

* A limit of 1000 conformers is enforced.

** The total number of single-bond conformers is given to the right of **Conformers** at the bottom of the screen.

Set Similarity Centers (⊙)

This specifies the set of atomic centers, or **C**hemical **F**unction **D**escriptors (**CFD's**) or pharmacophore elements that are to be used in similarity analysis. Choice of whether the similarity analysis is to be based on structure or CFD's (pharmacophore elements) is made from a menu that appears at the bottom right of the screen following selection of **Set Similarity Centers**. This menu is also accessible from the **Similarity Analysis** task in the **Calculations** dialog (**Calculations...** under the **Setup** menu; **Chapter 15**).

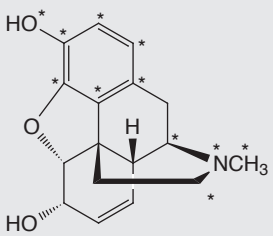
Similarity Based on Structure

Selection of **structure** from the **Similarity by** menu at the bottom right of the screen results in a message at the bottom left of the screen.

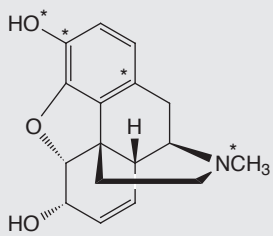
Select atoms.

Clicking on an atom designates it as an alignment center and it marks it with a violet circle. *Clicking* on the circle removes the designation (and the circle). Hydrogens may not be selected.

The computer time required for similarity analysis based on structure depends on the number of molecules in the query, the number of molecules in the library to which comparisons are being made and the choice of similarity centers. Single atoms should never be selected as similarity centers unless they are unique. Rather contiguous groups of atoms that are likely to be unique should be chosen. For example, analysis carried out on morphine based on selection of its phenol and trialkylamine components (left) will be much faster and much more informative than that based on selection of the phenol oxygen, two carbons on benzene and the nitrogen (right).



good



bad

This is because there will be many few occurences of phenol and trialkylamine substructures in the library entries than there will be aromatic carbons and nitrogen atoms and fewer permutations that need to be examined and fewer false positive results.

Similarity Based on CFD's

Similarity analysis may also be based on CFD's. These indicate whether a particular center is likely to act as a hydrogen-bond donor or hydrogen-bond acceptor, is likely to take on positive or negative charge or is a hydrophobe. This is accomplished by selecting **CFD** from the **Similarity by** menu at the bottom right of the screen. In response, the selected molecule is augmented with CFD's, and a message appears at the bottom left of the screen.

Clicking on a CFD designates it as a similarity center and marks it with a violet circle. *Clicking* on the circle removes the designation (and the circle).

The computer time required for similarity analysis based on CFD's depends on the number of molecules and/or pharmacophores in the query, the number of molecules and/or pharmacophores in the library and the nature of the selected CFD's in the query. For each query/library pair, where the query is represented by i CFD's of type A, j CFD's of type B, k CFD's of type C, etc., the number of different alignment combinations that need to be considered is given by:

$$i!j!k! \ldots$$

(! is the factorial symbol.) This rapidly becomes unmanageable. for example, if there are 6 CFD's of type A, 3 of type B and 2 of type C, the number of combinations is 6!3!2! or 8640. Considerable savings can be achieved by limiting the number of CFD's of a given type in the overall representation.

Similarity Based on a Pharmacophore

Finally, it is possible to base similarity on a pharmacophore. This is set up in exactly the same way as similarity based on CFD's insofar as the elements of a pharmacophore are CFD's (see previous section).

Selection of similarity centers either needs to be repeated for the remaining molecules in the list or the selections may be designated as global, that is, applying to all molecules in the list. The latter is applicable where the molecules in the list are conformers or are closely related at least insofar as the alignment centers are concerned. Global designation is accomplished by *checking* the **Global Similarity Settings** at the bottom of the screen.

Constrain Distance ()
Constrain Angle ()
Constrain Dihedral ()

These introduce of one or more geometrical constraints during structure minimization (in build mode), and (optionally) during equilibrium or transition-state geometry optimization, conformational searching or energy profile generation using methods under the **Calculations** dialog (**Chapter 15**). Constraints may be useful in a number of situations, among them:

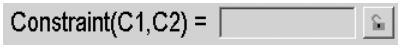
(i) constructing conformational energy profiles where one or more dihedral angles need to be fixed while other geometrical variables are optimized,

(ii) optimizing molecular structures where the values of certain key parameters are known, for example, optimizing the geometry of a molecule with an intramolecular hydrogen bond or a disulfide linkage, and

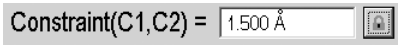
(iii) building molecules with unusual geometries, for example, molecules with very long bonds, as might be required in the construction of transition states and intermolecular complexes.

Selecting **Constrain Distance** results in a message at the bottom left of the screen.

Clicking on two atoms, or a bond results in a message at the bottom right of the screen.



Clicking on changes it to and shows the current distance.



This (constraint) distance can now be changed by altering the contents of the box and then *pressing* the **Enter** key (**return** key on Mac). Alternatively, the existing distance may be used as the constraint distance. If the selected distance had previously been constrained, the icon would have been initially displayed. In this case, *clicking* on turns "off" the constraint and returns the icon to the . Finally, the value of the constraint, that may be different from the value of the current distance*, may be entered into the spreadsheet by *clicking* on to its right.

This sequence of operations (bond identification followed by turning "on" and "off" the constraint) may be repeated as many times as necessary. Any bonds or non-bonded distances on which constraints are to be imposed are indicated by magenta colored markers. Any constraints introduced are automatically enforced only upon energy minimization in add fragment mode ( ), but are optional using methods under the **Calculations** dialog (**Chapter 15**).

Angle and dihedral angle constraints are handled in a similar manner. Note that points and planes may not be used to define constraints.

Constraints may also be modified as well as posted to the spreadsheet from the **Constraint Properties** dialog (available under **Properties** in the **Display** menu; **Chapter 16**). In addition, a sequence of constraints may also be modified as well as posted to the spreadsheet from the **Constraint Properties** dialog (available under **Properties** in the **Display** menu; **Chapter 16**). In addition, a sequence of constraints

---

* Note, however, that it may be problematic to carry out a constrained geometry optimization starting from a structure that is very different from that satisfying one or more constraints.

may be defined (from some initial value to some final value in a given number of steps). This allows generation of an energy profile along a predefined set of coordinates* (**Energy Profile**; **Calculations** in **Chapter 15**).

Whether or not constraint markers are included as part of the model (outside of constrain distance, constrain angle or constrain dihedral mode) for an individual molecule is controlled from the **Molecule Utilities** dialog available under the **Display** menu (see **Chapter 16**). Global settings are controlled from the **Molecule Preferences** dialog (**Preferences...** under the **Options** menu; **Chapter 18**).

Define Point ()

This defines a point as the geometric (unweighted) center of selected atoms (or points) previously defined. Selection results in display of a message at the bottom left of the screen.

Clicking on atoms (or points) in any order, and *clicking* a second time on any one of the atoms (or points) defines a point (depicted as a small sphere). As many points as desired can be defined and these are treated in the same way as an atom in defining distances, angles, etc. Points move with the molecule as its geometry is altered.

Selecting **Define Point** (or *clicking* on ) while holding down on the **Shift** key, followed by *clicking* on the appropriate atoms, leads to a ***ligand point***. This is a point of attachment directed perpendicular to the geometric center of the plane defined by three atoms (or best plane in the case of four or more atoms). A ligand point shares all the characteristics of a normal point, but may also be used to bond to atomic fragments, functional groups, etc. See **Make Bond** under the **Build** menu (**Chapter 14**) for a discussion. Ligand points move with the molecule as geometry is altered.

---

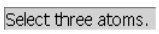
* Examples of this are provided in the tutorial ***Internal Rotation in Dimethylperoxide*** in **Chapter 5**, and the tutorial ***Thermodynamic vs. Kinetic Control*** in **Chapter 6**.

Delete from the **Build** menu () or the **Delete** key may be used to remove a point or ligand point.

Whether or not points and ligand points are shown as part of the model for an individual molecule is controlled from the **Molecule Utilities** dialog available under the **Display** menu (**Chapter 16**). Global settings are controlled from the **Molecule Preferences** dialog (**Preferences...** under the **Options** menu; **Chapter 18**).

Define Plane ()

This defines and displays a reference plane. Selection results in display of a message at the bottom left of the screen.



Clicking on three atoms or points defines a plane. As many planes as desired may be defined, and these may be used in defining distances, angles, etc. Planes move with the molecule as its geometry changes.

Delete from the **Build** menu () or the **Delete** key may be used to remove a plane.

Whether or not planes are included as part of the model for an individual molecule is controlled from the **Molecule Utilities** dialog available under the **Display** menu (**Chapter 16**). Global settings are controlled from the **Molecule Preferences** (**Preferences...** under the **Options** menu; **Chapter 18**).

Define CFD ()

This defines the position of a new CFD. The **CFD Properties** dialog (**Properties** in the **Display** menu; **Chapter 16**) allows CFD size and type to be selected. Selection results in display of a message at the bottom left of the screen.

Clicking on atoms (or points) in any order, and *clicking* a second time

on any one of the atoms (or points) defines the position of a CFD (depicted as a transparent sphere labelled "Null"). *Clicking* twice on only one atom locates the CFD on the atom. As many CFD's as desired can be defined. These move with the molecule as its geometry is altered.

Align ()

This aligns the selected molecule to all other molecules in the same document based either on structure or on CFD's. If the selected molecule is a pharmacophore, aligns the pharmacophore to all molecules in the same document (based on CFD's). Choice of whether alignment is based on structure or CFD's is made from the **Align by** menu that appears at the bottom right of the screen.

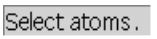
The terms "alignment" and "similarity analysis" have subtly different meanings inside of ***Spartan***. Alignment refers to the operation in which the molecules in a document are reoriented to best coincide with the selected molecule (pharmacophore). Where a molecule is selected, alignment may either be based on a set of atoms or on a set of CFD's. Where a pharmacophore is selected, alignment is based on a set of CFD's. A single reorientation is provided for each molecule in the document that can be aligned to the selected molecule (pharmacophore).

Similarity analysis refers to the operation in which all molecules in a user-specifed library comprising one or more documents are compared one-on-one to all molecules and/or pharmacophores in the selected document (in the graphical user interface). As with alignment, either structural elements or CFD's may be employed. The result is a similarity score from 0 to 1 (1 is perfect) and a manipulatable model of the two molecules (or molecule and pharmacophore). It is likely that most comparisons will not be successful, and also that some comparisons may lead to more than one score (corresponding to different orientations of the two components).

In short, the role of alignment is to orient molecules and/or pharmacophores in order to highlight their common features, while the role of similarity analysis is to rank pairs of molecules or molecule/pharmacophore pairs according to the extent of their common features.

Align Based on Structure

Selection of **Structure** from the **Align by** menu results in a message at the bottom left of the screen.



Clicking on an atom designates it as an alignment center, and marks it with a red circle. *Clicking* on the circle removes the designation (and the circle). Following selection of alignment centers, *clicking* on the **Align by** button at the bottom right of the screen aligns the molecules. If no atoms are selected prior to *clicking* on **Align by**, then alignment is based on all (non-hydrogen) atoms.

Following alignment, two or more molecules may be displayed at once using spreadsheet functions (see **Spreadsheet** under the **Display** menu; **Chapter 16**). Their motions (coordinates) will be coupled following alignment, but may be uncoupled allowing the aligned molecules to move independently (see **Coupled** under the **Model** menu; **Chapter 12**). Note that alignment center selections are kept and molecules can be realigned by again selecting **Align** from the **Geometry** menu (or *clicking* on ) followed by *clicking* on the **Align by** button.

Align Based on CFD's

It is also possible to base molecule alignment on CFD's. These indicate whether a particular center is likely to act as a hydrogen-bond donor or acceptor, is likely to take on positive or negative charge or is a hydrophobe. This is accomplished by choosing **CFD** from the **Align by** menu at the bottom right of the screen followed by *clicking* on the **Align by** button. In response, the selected molecule is augmented with CFD's, and a message appears at the bottom left of the screen.

Clicking on a CFD designates it as an alignment center and marks it with a red circle. *Clicking* on the circle removes the designation (and the circle).

The **CFD Properties** dialog (**Properties** in the **Display** menu; **Chapter 16**) allows definitions of the individual CFD's to be altered.

Following selection of CFD's, *clicking* on the **Align by** button at the bottom right of the screen aligns the molecules. If no CFD's are selected prior to *clicking* on **Align by**, then alignment is based on all CFD's.

Following alignment, one or more molecules may be displayed at once using spreadsheet functions (see **Spreadsheet** under the **Display** menu; **Chapter 16**). Their motions (coordinates) will be coupled following alignment, but may be uncoupled allowing the aligned molecules to move independently (see **Coupled** under the **Model** menu; **Chapter 12**). Note that alignment center selections are kept and molecules can be realigned by again selecting **Align** from the **Geometry** menu (or *clicking* on ) followed by *clicking* on the **Align by** button.

Align to a Pharmacophore

Finally, it is possible to align molecules to a pharmacophore contained in the same list. This is set up in exactly the same way as alignment based on CFD's insofar as a pharmacophore element is a CFD (see previous section).

For both structure and CFD alignment (and for alignment to a pharmacophore), an alignment score from 0 to 1 (perfect alignment), is available in the spreadsheet. This is accessed by *clicking* on the **Add** button at the bottom of the spreadsheet, and selecting **Alignment Score** from the list of available properties (see **Spreadsheet** under the **Display** menu; **Chapter 16**). A score of 0 is assigned to molecules that cannot be aligned to the selected molecule.

Chapter 14
The Build Menu

*The **Build** menu provides model kits and tools for building and editing organic, inorganic and organometallic molecules, as well as polypeptides and polynucleotides, for generating lists of substituted molecules, and a molecular mechanics procedure for structure refinement. This menu also provides seamless access to ChemDraw.*

Model Kits

Spartan provides five different model kits for assembling a variety of molecular systems: an organic model kit for most organic molecules, an inorganic model kit for organic molecules not easily represented in terms of classical valence structures, as well as inorganic and organometallic molecules, model kits for polypeptides and polynucleotides and a model kit for attaching substituents (lists of chemical groups) to previously-built molecules. The organic and inorganic model kits utilize atomic fragments, functional groups and rings (and ligands in the inorganic model kit), while the peptide model kit uses the set of natural amino acids as building blocks, and the nucleotide model kit the set of nucleotide bases. The substituent model kit is different in that it adds substituents to a previously constructed molecule and/or alters one or more incorporated atoms.

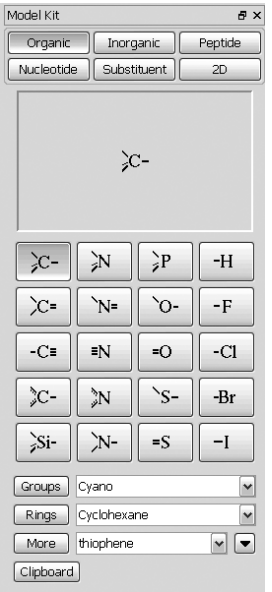
Two-dimensional molecular structures (drawings) produced using ChemDraw* can also be brought into ***Spartan*** and converted to three-dimensional structures.

Molecule construction in ***Spartan*** proceeds much in the same manner as a chemist would assemble a structure from a model kit, that is, pieces are taken from the kit one at a time and added sequentially to the molecule under construction.

* Seamless access to ChemDraw is provided in the Windows version only. All versions are able to read ChemDraw files. ChemDraw is not included with ***Spartan*** and must be obtained from CambridgeSoft at http://www.cambridgesoft.com.

Organic Model Kit

The organic model kit contains a suite of molecule building/editing tools specifically designed to construct organic molecules.



In the center of the model kit are a selection of atomic fragments, which from left to right and then top to bottom, correspond to:

$C(sp^3)$	$N(sp^3)$	$P(sp^3)$	H
$C(sp^2)$	$N(sp^2)$	$O(sp^3)$	F
C(sp)	N(sp)	$O(sp^2)$	Cl
C(aromatic)	N(aromatic)	$S(sp^3)$	Br
$Si(sp^3)$	N(planar)	$S(sp^2)$	I

A fragment is chosen by *clicking* on its icon, which is then displayed at the top of the model kit. Once selected, the fragment may be used to initiate building, to add alongside of an existing structure or appended onto an existing structure. To initiate building, *click* anywhere on screen. To add alongside of an existing structure, first *click* on the **Insert** button at the bottom right of the screen or hold down the **Insert** key (**alt** key on Mac), and then *click* anywhere on screen or *double click* in a blank area on screen. To bond to an existing structure, *click* on a free valence (***not an atom***). (Free valences are colored yellow on the selected molecule.) Bond type in the case of atomic fragments

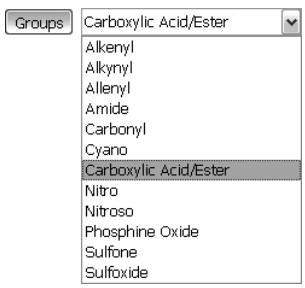
with multiple bond types, for example, sp^2 carbon, depends on the nature of the free valence selected.

> While only H, C, N, O, F, P, S, Cl, Br and I are available from the organic model kit, other elements may be substituted using atom replacement feature available in the inorganic model kit (see **General Molecule Building Functionality** later in this chapter). Similarly, bond types may be altered in the inorganic model kit. The latter may be necessary in order to alter bonding in structures retrieved from the Cambridge Structural Database or ligands obtained from PDB files. Atom and bond types may also be altered using the **Atom** and **Bond Properties** dialogs, respectively (**Properties** under the **Display** menu; **Chapter 16**).

Menus inside the model kit provide access to a number of pre-built fragments corresponding to functional groups (**Groups**) and rings (**Rings**), and to additional libraries of rings (as well as any user-defined structures) stored in ***Spartan's*** file system (**More**). The model kit also accesses the clipboard (**Clipboard**).

Groups

Clicking on **Groups** brings up a menu of groups, and displays an icon of one group at the top of the model kit.

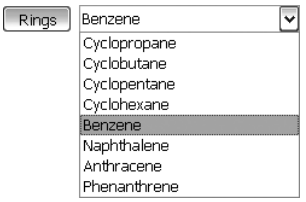


Once selected from the menu, a group may be used to initiate building, to add alongside of an existing structure on screen, or to add to an existing structure.

The amide and carboxylic acid/ester groups have more than one different free valence. The free valence that is to be used is marked with a gold • (in the icon at the top of the model kit). The marked position circulates among the possible positions with repeated *clicking* on the icon.

Rings

Clicking on **Rings** brings up a menu of hydrocarbon rings, and displays an icon of one ring at the top of the model kit.



Once selected from the menu, a ring may be used to initiate building, to add alongside of an existing structure on screen, or to add to an existing structure.

Cyclohexane, napthalene, anthracene and phenanthrene have more than one different free valence. The one that is to be used is marked with a gold • (in the icon). The marked position circulates among the available positions with repeated *clicking* on the icon. Selection of an *axial* or *equatorial* free valence in cyclohexane is indicated by the label **ax** or **eq** appearing alongside the icon. All rings in this menu are hydrocarbons, but heteroatoms may be substituted (see **General Molecule Building Functionality** later in this chapter).

More

This provides access to a broader selection of rings as well as to access user-defined entities (rings, groups, ligands, etc.). Upon initial entry, the menu to the right of **More** will be empty. It can be populated, by *clicking* on ▾ to the far right. This brings up a file browser that has been set to point toward a directory containing several ***Spartan*** files of common rings.

nitrogen heterocycles	saturated nitrogen rings
oxygen heterocycles	saturated oxygen rings
sulfur heterocycles	saturated sulfur rings
mixed heterocycles	saturated mixed rings

Clicking on a file followed by *clicking* on **Open** or *double clicking* on a file fills the menu to the right of **More**. Menu entries are

selected in the usual way. In response, a ball-and-wire model of the selected ring will appear at the top of the model kit. This may be manipulated (rotated, translated, zoomed) using the usual mouse/keyboard commands (you need to position the cursor inside the box). The ring may be used to initiate building, to add alongside of an existing structure, or to add to an existing structure. In the latter case, the attachment point (on the ring in the window) needs to be identified by *clicking* on the appropriate free valence.

Note that three other collections (ligands, chelates and high-coordination fragments) appear in the selected directory. Their use will be discussed in the next section. Note also that user-defined libraries may be accessed (these are normal ***Spartan*** documents).

Clipboard

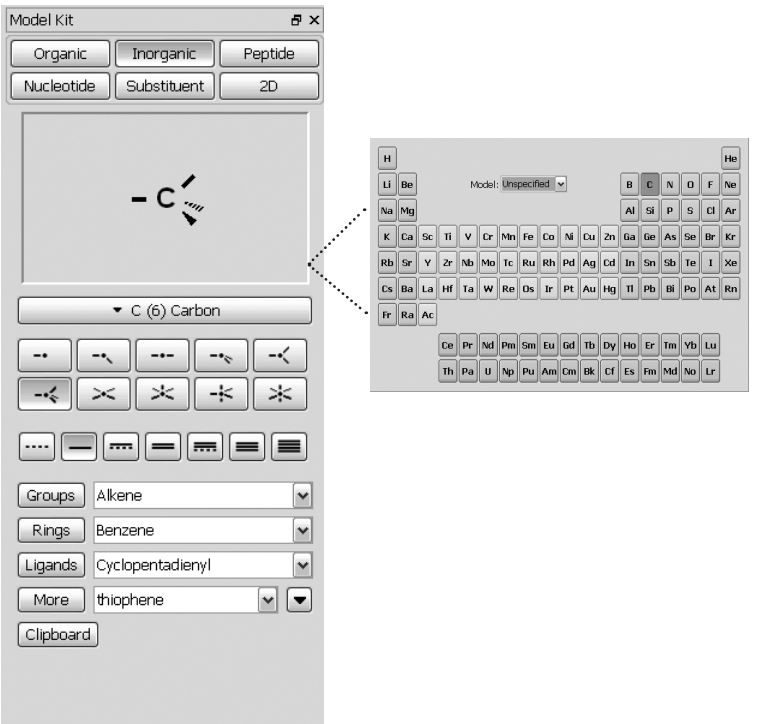
Clicking on **Clipboard** accesses the clipboard. A ball-and-wire model of whatever is on the clipboard is displayed at the top of the model kit. This may be manipulated using the usual mouse/keyboard commands (you need to position the cursor inside the box). Once selected, the molecule may be used to initiate building, to add alongside of an existing structure, or to add to an existing structure. In the latter case, the attachment point needs to be identified by *clicking* on the appropriate free valence in the clipboard.

An empty clipboard will be signaled by:

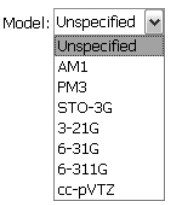
Inorganic Model Kit

Spartan's inorganic model kit allows construction of a much wider class of molecules (including inorganic and organometallic species) than possible with the organic model kit. Structures that violate conventional bonding rules may be constructed, as this model kit purposefully provides no checking. The inorganic model kit is reached by *clicking* on the **Inorganic** tab, located at the top of the organic (peptide, nucleotide or substituent) model kit. (Return to the

organic, peptide, nucleotide or substituent model kit or to ChemDraw is accomplished by *clicking* on the **Organic**, **Peptide**, **Nucleotide**, **Substituent** or **2D** tab, respectively, in the inorganic model kit).



Atoms may be selected by *clicking* on the center atom button. This leads to a full *Periodic Table*. Main-group elements are colored red, transition metals are colored green and lanthanides and actinides are colored blue. The **Model** menu inside the *Periodic Table* contains a selection of theoretical models (basis sets except for semi-empirical models).



Selecting an entry from this menu leads to recoloring of the *Periodic Table*. A light green color is used to indicate elements for which the selected model may be used.* Immediately below is a selection of

* While molecular mechanics models are available for all elements, they have been carefully parameterized for only a relatively few of these.

atomic hybrids. Following this is a selection of bond types. **Groups**, **Rings**, **More** and **Clipboard** are the same as in the organic model kit; **Ligands** is new to the inorganic model kit.

Selection of atom type is effected by *clicking* on the appropriate element in the *Periodic Table*. The entry will be highlighted. Selection of an atomic hybrid follows by *clicking* on the appropriate icon which will then be highlighted.* This combination (atom type + atomic hybrid) may be used to initiate building, to add alongside of an existing structure or to append onto an existing molecular fragment. To initiate building, *click* anywhere on screen. To add alongside of an existing structure, first *click* on **Insert** at the bottom right of the screen or *press* the **Insert** key (**alt** key on Mac), and then *click* anywhere on screen, or *double click* in a blank area on screen. To bond to an existing fragment, *click* on the appropriate free valence.

Two of the hybrids (trigonal bipyramidal and square-based pyramidal) may bond either *axially* or *equatorially*. Selection of the appropriate bonding point, marked by a •, is effected by repeatedly *clicking* on the icon; the bonding point alternates between the two sites.

Atoms are connected with whatever bond type (partial single, single, aromatic, double, triple or quadruple) is selected in the model kit. A bond type may be changed by first selecting a type and then *double clicking* on the bond. Bond types have no impact on quantum chemical calculations, but do affect molecular mechanics calculations (including minimization in the builder; see discussion later in this chapter).

No valence checking is performed in the inorganic model kit, and the user is free to construct any arrangement of atoms.

Three fragment collections are located inside the inorganic model kit:

Groups

This is the same as described in the organic model kit.

Rings

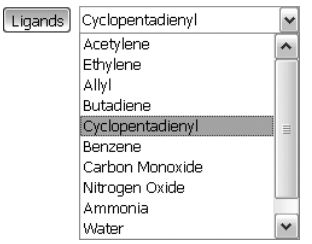
This is the same as described in the organic model kit.

---

* Additional hybrids for high-coordination centers are available as a library reachable from **More** (see discussion under organic model kit).

Ligands

This provides access to a number of pre-built ligands, useful in the construction of inorganic and organometallic molecules. Its operation is analogous to that for the **Groups** and **Rings** menus. *Clicking* on **Ligands** brings up a menu of available ligands, and results in an icon of one ligand from this menu being displayed at the top of the model kit.



A ligand may be used to initiate building or to add alongside or to an existing structure. Additional ligands are accessible from **More** (see previous discussion). Ligands may also be built with the aid of ligand points (**Define Point** in the **Geometry** menu; **Chapter 13**).

More

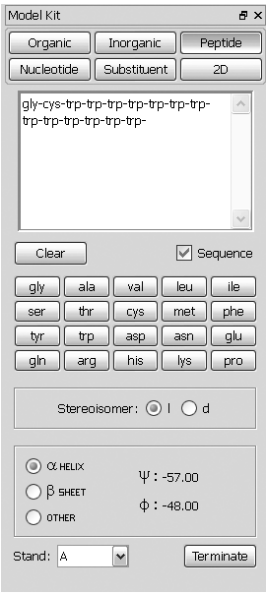
This is the same as described in the organic model kit.

Clipboard

This is the same as described in the organic model kit.

Peptide Model Kit

A model kit for construction of polypeptides is reached by *clicking* on the **Peptide** tab located at the top of the organic, inorganic, nucleotide and substituent model kits. (Return to the organic, inorganic, nucleotide or substituent model kit or to ChemDraw is accomplished by *clicking* on the **Organic**, **Inorganic**, **Nucleotide**, **Substituent** or **2D** tab, respectively, in the peptide model kit.)



At the middle of the peptide model kit are icons designating the amino acids (specified by their usual three-letter codes). An amino acid is selected by *clicking* on its three-letter code, following which either an icon of the amino acid is displayed in the box at the top of the model kit, or the three-letter code for the amino acid is appended to the sequence of codes in the box. Amino acids replace atoms, functional groups, rings and ligands as the building blocks in the peptide model kit. Because these other building blocks are missing, modifications of peptides, aside from modifications in sequence and in overall conformation, need to be carried out using the organic or inorganic model kits.

There are two different modes of operation: single amino acid mode and polypeptide mode. The former is used to initiate building with a single amino acid, to add a single amino acid alongside of an existing structure or to add a single amino acid to an existing structure, while the latter is used to construct amino acid sequences (polypeptides). **Sequence** "off" (unchecked) corresponds to single amino acid mode, and "on" (checked) corresponds to polypeptide mode.

Peptide construction (**Sequence** "on") is accomplished in three steps:

Specification of Amino Acid Sequence

This is accomplished by *clicking* in the desired order on the amino acid codes. Building occurs from the N end to the C end of the peptide. In response to each selection, the three-letter code is appended to the sequence of codes in the box at the top of the model kit. The stereochemical configuration of the amino acid is by default the l configuration; this may be changed to the d configuration prior to selection of the amino acid, by *checking* **d** to the right of **stereoisomer** in the model kit. (It may be changed back to l by *checking* **l**). d amino acids are indicated by **.d** following the code in the box.

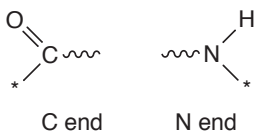
The sequence may be altered by changing the text in the box. Existing amino acid codes may be deleted or changed or new codes can be added. The entire sequence may be specified in this way if desired. Specification of a non-existent code will result in an error message. The sequence can be cleared by *clicking* on **Clear**.

Specification of Macroscopic Structure

Once sequencing is complete, macroscopic structure (ψ and ϕ angles), is specified by *clicking* on one of α **Helix**, β **Sheet** or **Other**. In the case of the first two, preset angle values are displayed on the right. In the case of specification of **Other**, boxes appear, into which the desired dihedral angles need to be entered.

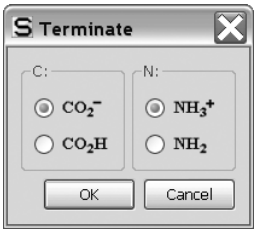
Termination

The peptide is not yet terminated, and the two ends are still set up for addition of further amino acids.



C end N end

where the * indicates a free valence. *Clicking* on **Terminate** at the bottom of the model kit leads to the **Terminate** dialog.

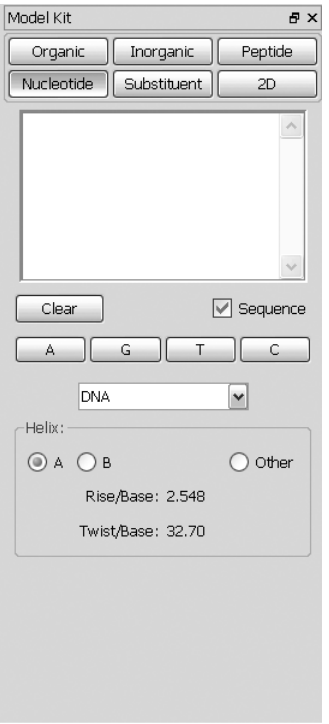


C and N terminating groups may be selected by repeated *clicking* on the C and N icons, respectively. Selection will rotate among the available terminating groups. *Clicking* on **OK** removes the dialog and terminates the polypeptide. *Clicking* on **Cancel** or ☒ removes the dialog but does not terminate the polypeptide.

The peptide (or single amino acid) may now be used either to initiate building, by *clicking* anywhere on screen or added alongside of an existing structure, by first *clicking* on **Insert** at the bottom right of the screen or *pressing* the **Insert** key (**alt** key on Mac), followed by *clicking* anywhere on screen, or *double clicking* in a blank area on screen. If unterminated, it may also be joined onto an existing structure by *clicking* on a free valence. In the latter case, attachment is made from the N end, unless the free valence corresponds to an unterminated peptide fragment, in which case the appropriate end required to make an amide bond is used.

Nucleotide Model Kit

Finally, ***Spartan*** provides a model kit for construction of polynucleotides. It is reached by *clicking* on the **Nucleotide** tab which is located at the top of the organic, inorganic, peptide and substituent model kits. (Return to the organic, inorganic, peptide or substituent model kit or to ChemDraw is accomplished by *clicking* on the **Organic**, **Inorganic**, **Peptide**, **Substituent** or **2D** tab, respectively, in the nucleotide model kit.)



At the middle of the model kit is a menu designating the type of polynucleotide.

DNA
DNA (single strand)
DNA-RNA
RNA
RNA (double strand)
RNA-DNA

Immediately below this menu are icons, designating the nucleotide bases. Selection of DNA, DNA (single strand) or DNA-RNA from the menu leads to one set of icons.

Selection of RNA, RNA (double strand) or RNA-DNA leads to a second set, the only difference is that uracil (U) has been substituted for thymine (T).

A nucleotide base is selected by *clicking* on its letter, following which either an icon of the base is displayed in the box at the top of the model kit, or the letter for the base is appended to the sequence of letters in the box. Nucleotide bases replace atomic fragments, functional groups, rings and ligands as the building blocks in the nucleotide model kit. Because these other building blocks are missing, modifications of nucleotides, aside from modifications in sequence and helical structure, need to be carried out using either the organic or inorganic model kits.

There are two different modes of operation: single base mode and polynucleotide mode. The former is used to place a single base or base pair on screen, to add a single base or base pair alongside of an existing structure, or to add a single base or base pair to an existing structure, while the latter is used to construct strands of DNA or RNA (or mixed strands). **Sequence** "off" (unchecked) corresponds to single base (base pair) mode and "on" (checked) corresponds to polynucleotide mode.

Polynucleotide construction (**Sequence** "on") is accomplished in two steps:

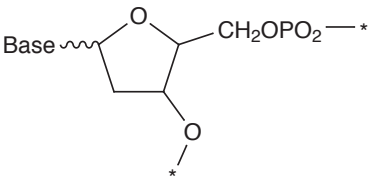
Specification of Base Sequence

This is accomplished by *clicking* in order on the base codes. In response to each selection, the letter code is appended to the sequence of codes in the box at the top of the model kit. The sequence may be altered by editing the contents of the box. Existing base codes may be deleted or changed or new codes added. The entire sequence can be specified in this way if desired. The sequence may be cleared by *clicking* on **Clear**.

Specification of Helical Structure

Once sequencing is complete, a helical structure may be specified by *clicking* on **A** or **B**. These correspond to A and B helices, respectively. Selecting **Other** allows user modification of the rise (in Å) per base (**Rise/Base**) and twist (in degrees) per base (**Twist/Base**).

Note that the polynucleotide is not yet terminated, and the two ends are still set up for addition of further bases or base pairs.



* indicates a free valence. Hydrogens occupy all free valences (except the *'ed positions at the two ends of the chain).

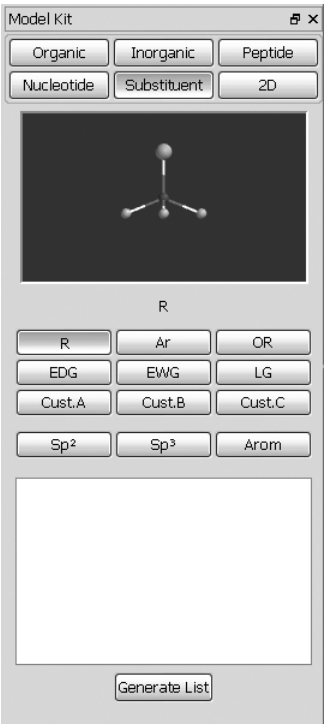
The polynucleotide (or single base pair) may now be used to either initiate building, by *clicking* anywhere on screen, added alongside of an existing structure, by first *clicking* on **Insert** at the bottom right of the screen or *pressing* the **Insert** key (**alt** key on Mac) followed by *clicking* on screen, or *double clicking* on a blank area on screen, or joined onto an existing structure by *clicking* on a free valence. In the latter case, attachment is made from the phosphate end.

Substituent Model Kit

The substituent model kit designates one or more locations for libraries of substituents to be added to a molecule. Additionally, it designates one or more carbon hybrids for replacement by their heteroatom "equivalents", for example, sp^3 carbon by sp^3 nitrogen. There are two uses for the **Substituent** model kit. The first is to facilitate assembling lists of molecules that differ only in substitution and/or heteroatom. The second application is to use "molecules" incorporating one or more substituent libraries as queries in searches of the Spartan Molecular Database (see **Databases** under the **Search** menu; **Chapter 17**) or as elements of reaction energy calculations (see **Reactions** under the **Display** menu; **Chapter 16**). Here, the search or reaction energy calculation is not confined to a single molecule but rather is open to a set of molecules that are related by substitution. Further discussion is provided in the appropriate sections of **Chapters 16** and **17**.

The substituent model kit is accessed by *clicking* on the **Substituent** tab located at the top of the organic, inorganic, peptide and nucleotide model kits. (Return to the organic, inorganic, peptide or nucleotide

model kit or to ChemDraw is accomplished by *clicking* on the **Organic**, **Inorganic**, **Peptide**, **Nucleotide** or **2D** tab, respectively, in the substituent model kit.)



At the center of the model kit are four rows of buttons. Buttons in the first two rows provide access to predefined substituent libraries: **R** (alkyl groups), **Ar** (aromatic groups), **OR** (alkoxy groups), **EDG** (electron donor groups), **EWG** (electron withdrawing groups) and **LG** (leaving groups). Buttons in the third row allow access to user-defined substituent libraries (**Cust. A**, **Cust. B**, **Cust. C**), while those in the last row permit substitution of heteroatom hybrids for sp^3, sp^2 and aromatic carbons incorporated in a molecule. Immediately below the buttons, a dialog lists the entries in the individual libraries (the three custom libraries should be blank upon initial entry). As entries are selected, they will appear in the window at the top of the dialog. Predefined libraries may not be altered and individual entries may not be renamed, but existing library entries in custom libraries may be renamed or deleted and new entries added. Modifications to existing substituent libraries need to be performed ***following attachment*** using

the **Substituent Properties** dialog (**Properties** under the **Display** menu; **Chapter 16**).

Creating lists of substituted molecules starting from an unsubstituted molecule occurs in three stages: library preparation, attachment and (optionally) list generation.

Library Preparation

Six substituent libraries have been provided.

R (alkyl group): methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *sec*-butyl, isobutyl, *t*-butyl, *t*-amyl

Ar (aromatic ring): phenyl, p-tolyl, p-nitrophenyl, 1-naphthyl, 2-naphthyl, 2-pyridyl, 3-pyridyl, 4-pyridyl

OR (alkoxy group): methoxy, ethoxy, *n*-propoxy, isopropoxy, *t*-butoxy

EDG (electron donor group): hydroxy, methoxy, ethoxy, phenoxy, amino, dimethylamino, mercapto, methylthio, phenylthio

EWG (electron withdrawing group): cyano, nitro, trifluoromethyl, carboxy, formyl, methoxycarbonyl, ethoxycarbonyl, methyl sulfonyl, methyl sulfoxyl

LG (leaving group): bromo, iodo, tosylate, triflate, mesylate

The contents of an existing library may be displayed in the scroll box at the bottom of the model kit by *clicking* the appropriate button. A ball-and-wire model of a substituent may be seen at the top of the model kit by *clicking* on its name in the scroll box. The model may be manipulated in the usual way. The attachment point is indicated by a gold ball surrounding a "hydrogen".

To create a custom substituent library, first *click* on the **Cust. A**, **Cust. B** or **Cust. C** buttons inside the model kit. The scroll box is empty. If a substituent (set of substituents) has previously been copied to the clipboard, it may be pasted into the library by *right clicking* inside the scroll box and selecting **Paste** from the menu that appears. Alternatively, a substituent (set of substituents) may be appended to the document (including an empty document) by

right clicking inside the scroll box, selecting **Append** from the menu that results and selecting a ***Spartan*** document comprising one or more molecules from the browser. Finally, any ***Spartan*** document may be *dragged* onto the scroll box from the file system. As many substituents (sets of substituents) as desired may be pasted, appended and/or *dragged* into a library. To remove a substituent from the library, *right click* on its name in the scroll box and select **Delete** from the menu that appears. **Ctrl** (**Apple** for Mac) and **Shift** keys may be used in the usual way.

Substituents need to be specified in terms of both their structure and their connection point. One way to do this is to specify the latter is to replace all free valences ***except the one that will serve as the connection point*** by hydrogens. For example, to designate a 2-pyridyl substituent, add hydrogens (

from the organic model kit) to four of the five free valences in pyridine, leaving only one free valence in the 2-position available. Alternatively, a connection point can be specified by *clicking* on the appropriate position as the substituent displayed at the top of the model kit.

The three buttons in the last row designate heteroatom equivalents of carbon-based hybrids.

sp^2: carbon (

), nitrogen (

)

sp^3: carbon (

), nitrogen (

), oxygen (

)

Arom.: carbon (

), nitrogen (

)

Attachment

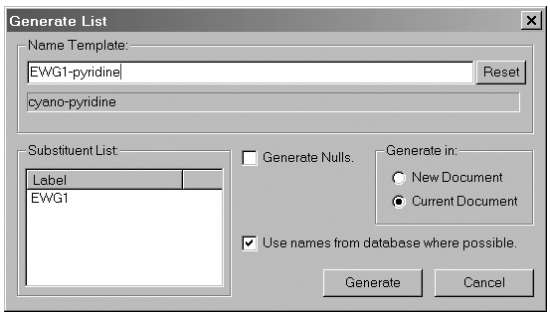
Following selection, a substituent library may be used to initiate building, to add alongside of an existing structure or to attach to an existing structure. The model is augmented by an icon containing the library name from the button, for example **EWG**, followed by a number. As many substituent libraries as desired may be attached to a molecule. Each library is given a unique label, that can be changed prior to list generation using the **Substituent Properties** dialog. A substituent library may be removed from the molecule

by selecting **Delete** from the **Build** menu and then *clicking* on the library icon.

A document containing one or more molecules with attached substituent libraries may be saved prior to list generation just as any other ***Spartan*** document. Any queries (bond distances, angles, etc.) and calculations refer to the unsubstituted molecule, that is, hydrogens replacing substituent markers. Database searches (on the Spartan Molecular Database) and reaction energy calculations based on entries in the Spartan Molecular Database pertain to all possible substituted molecules.

List Generation

Generation of a list of substituted molecules follows specification and (optionally) editing of one or more substituent libraries and attachment of each of these libraries at one or more positions in a molecule. Duplicates are eliminated. *Clicking* on **Generate List** at the bottom of the model kit, leads to the **Generate List** dialog.



A box at the top of the dialog allows the user to specify (or modify) a **Name Template**, from which the names of substituted molecules will be derived. Upon entry, this takes the form.

Ax-By-Cz-molecule name

A, B, C are the names of the substituent libraries (**R**, **EWG**, etc.), x, y, z are integers and molecule name is the name of the unsubstituted molecule if and as it appears in the Spartan Molecular Database (SMD). If the unsubstituted molecule is not in the database, then the box below **Name-Template** will be blank. If the substituted molecule is in SMD, its name in the database may be used (instead

of the name generated from the template) by *checking* the box to the left of **Use names from SMD where possible**.

Nulls are molecules with less than the full complement of substituents, for example, monosubstituted molecules in the case where two different positions have be substituted. These will not be included in the list of substituted molecules unless **Generate Nulls** is checked.

Note, that the conformations of the substituted molecules are arbitrary and they have not been subjected to energy minimization to relieve unfavorable crowding interactions. Therefore, the list entries should be examined, conformations adjusted as appropriate and minimized with molecular mechanics prior to using them for quantum chemical calculations.

The list may either be written to a **New Document** or appended to the end of the **Current Document**.

Accessing ChemDraw (NAL, NAM)*

The ChemDraw program may be seamlessly accessed inside of ***Spartan***** allowing chemical drawings to be produced in a familiar environment and then brought over as 3D structures***. The conversion is unambiguous as long as all stereochemical cues are in place in the 2D drawing. Note that the conformation implied by the 2D drawing may be ambiguous and need to be altered.

To access ChemDraw, enter the builder, *click* on the **2D** tab and *click* on **Edit** at the bottom of the panel that results. ChemDraw will appear. To return to ***Spartan***, close ChemDraw. The 2D drawing will appear at the center of the panel and manipulatable 3D structure will appear at the top of the panel. *Clicking* on screen will move the 3D structure into ***Spartan's*** main window.

---

* ChemDraw files (.cdx) may be read with all versions of ***Spartan***.

** ChemDraw is not provided with ***Spartan*** but must be obtained from CambridgeSoft at http://www.cambridgesoft.com.

*** Transfer is one directional only. 3D structures that have been altered may not be transferred back to ChemDraw.

General Molecule Building Functionality

Additional capabilities are available with **Add Fragment** selected:

Multiple Fragments

Multiple fragments may result either from bond breakage (see **Break Bond** later in this chapter) or from use of the **Insert** button or **Insert** key (**alt** key on Mac), or *double clicking* in a blank area on screen. A fragment is selected by *clicking* on it, following which the associated free valences are colored gold (free valences for any non-selected fragments are colored white). All builder functions apply to the selected fragment only. Rotation and translation also apply to the selected fragment, but may be made to apply to the entire set of fragments by holding down the **Ctrl** key while carrying out these operations.

Fragments may be attached using **Make Bond** (see discussion later in this chapter).

Bond Rotation/Bond Stretching

In addition to molecule rotation, translation and scaling, the mouse is used to rotate about and stretch bonds not incorporated into rings. This is accomplished via the following sequence of operations:

(i) *Clicking* on the bond, which is then marked by a red cylindrical arrow. (The bond connecting the last atom, group, ring or substituent added to the molecule is automatically selected.)

(ii) Simultaneously holding down the **Alt** key and the left mouse button while *dragging* the mouse up and down, for bond rotation, or the **Alt** key and the right mouse button for bond stretching.

Atom/Fragment Replacement

Another function of the mouse is atom replacement. *Double clicking* on an atom (not a free valence) while an atomic fragment in the organic model is highlighted, replaces the atom by selected

fragment. Free valences are adjusted to accommodate the replacement, for example, replacement of sp^3 carbon by sp^3 oxygen results in two free valences being removed. Atom replacements that violate valence rules or that would disrupt substituents are not permitted. *Double clicking* on an atom (not a free valence) while an element in the *Periodic Table* from the inorganic model kit is selected, replaces the atom by the selected element. The latter merely changes the atomic number. No changes in the number or arrangement of free valences is made, and no checking is done. Atom replacement is not available in the peptide, nucleotide and substituent model kits.

Chirality Inversion

In the **Add Fragment** mode, *double clicking* on a chiral atom with the **Ctrl** key (**Command** key on Mac) depressed inverts the chirality of the atom (R→S or S→R). This is not available in fused ring systems. *Double clicking* on any atom with both **Ctrl** (**Apple** key on Mac) and **Shift** keys depressed inverts the absolute configuration of the molecule.

Building/Editing Menu Functions

Molecule building/editing functions are found under the **Build** menu.

View (V)

This exits build mode, and removes the model kit from the screen.

> Initial entry into build mode is by way of **New** or **New Molecule** under the **File** menu (**Chapter 10**). **Add Fragment**, **Delete**, **Make Bond**, **Break Bond** and **Minimize** are for modifying existing structures.

Add Fragment (+)

In addition to the capabilities discussed under **General Molecule Building Functionality**, this allows access to the libraries of atomic fragments, groups, rings, ligands and substituents, as well as the file system and the clipboard. *Clicking* on any buttons or menus in the organic, inorganic, peptide, nucleotide or substituent model kits, leads to **Add Fragment**. (If a model kit is not on screen, selection brings up the last-accessed model kit.) A fragment may be used to initiate building by *clicking* anywhere on screen, to add alongside an existing structure on screen by *clicking* on the **Insert** button at the bottom right of the screen or *pressing* the **Insert** key (**alt** key on Mac) followed by *clicking* anywhere on screen, or by *double clicking* in a blank area on screen, or be added to an existing structure by *clicking* on the appropriate free valence. Fragment addition can be terminated by selection of any other function.

Delete ( )

This allows atom, free valence and substituent library removal from a structure. Selection leads to a message at the bottom left of the screen.

Subsequent *clicking* on an atom, free valence or substituent library results in its deletion. Deletion of an atom results in deletion of all of its associated free valences. Free valences for any atoms to which the deleted atom was previously connected are restored. Note that atom deletion may result in one or more detached fragments. *Clicking* inside a selection box results in deletion of everything inside the box. Selection of **Delete** does not bring up a model kit nor does it remove a model kit that is present on screen. **Delete** may be terminated by selection of any other function.

> **Delete** may also be used to delete points and planes, CFD's and pharmacophore elements as well as spectra and plots (or individual curves that make up plots).

Deletion may also be accomplished by holding down on the **Delete** key while *clicking* on the item to be deleted.

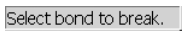
Make Bond ()

This allows bonds to be drawn between free valences and/or atoms on different centers. Selection leads to a message at the bottom left of the screen.

Clicking on two free valences (on different atoms) will cause these atoms to be linked by a single bond. Alternatively, *double clicking* on each of two atoms will bond them, and *clicking* on a free valence on one atom and *double clicking* on a different atom will bond the two atoms. Note that available free valences are consumed as a result of bond formation, irrespective of whether free valences or atoms are selected.* If the selected atoms are already bonded, this will result in the bond order being increased by one, that is, single → double, double → triple. Selection of **Make Bond** does not bring up a model kit nor does it remove a model kit that is already present on screen. **Make Bond** may be terminated by selection of any other function.

Break Bond ()

This allows breaking an existing bond resulting in free valences. Selection leads to a message at the bottom left of the screen.



Clicking on a bond breaks it and restores free valences. Note that bond breaking may result in detached fragments. Selection of **Break Bond** does not bring up a model kit nor does it remove a model kit that is present on screen. **Break Bond** may be terminated by selection of any other function.

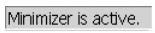
---

* Free valences can be "protected" without altering the molecule by adding hydrogens to them (

from the organic model kit). One use of this is in conjunction with custom substituent libraries (see discussion of substituent model kit earlier in this chapter).

Minimize ( )

This uses MMFF molecular mechanics to refine the geometry. Selection leads to a message at the bottom left of the screen.



The molecular mechanics energy* in kJ/mol, displayed at the bottom right of the screen, is continually updated during the minimization process. Minimization may be stopped at any time by *clicking* on the STOP icon at the bottom right of the screen. Any geometrical constraints imposed on the structure (see **Constrain Distance**, **Constrain Angle**, **Constrain Dihedral** in **Chapter 13**) are enforced during minimization. NOE's are not taken into account. Also, any frozen atoms in the structure (see **Freeze Center** under the **Geometry** menu; **Chapter 13**) remain frozen.

Following completion, **Minimize** reverts back to **Add Fragment** only if a model kit is on screen.

---

* The mechanics energy is a combination of the strain energy which is necessarily positive and the non-bonded or intramolecular interaction energy which can be either positive or negative.

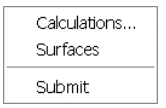
Chapter 15
The Setup Menu

This chapter describes functions available under the ***Setup*** *menu.* ***Calculations*** *is used to specify molecular mechanics calculations, semi-empirical calculations, Hartree-Fock molecular orbital calculations, and correlated calculations, including local density calculations, density functional calculations, Møller-Plesset calculations, coupled cluster calculations and quadratic configuration interaction calculations for ground-state species, and configuration interaction calculations, local density calculations and (time dependent) density functional calculations for excited-state species. Tasks include calculation of energy, equilibrium structure and conformation, transition-state structure and constructing energy profiles, although not all tasks are available for some methods. In addition, Hartree-Fock and density functional calculations may be extended to include an empirical solvation term, allowing "solvated" equilibrium and transition-state structures, infrared spectra and molecular properties to be obtained. A wide variety of all-electron Gaussian basis sets are supported for Hartree-Fock and correlated calculations as are pseudopotentials for calculations on molecules incorporating elements for which all-electron basis sets are not available. Also provided are a number of thermochemical recipes, including G3 and G3(MP2), as well as the T1 recipe that has been parameterized to closely reproduce G3(MP2) heats of formation with 2-3 orders of magnitude less computation time.* ***Calculations*** *also requests IR, NMR and UV/visible spectra, and calculation and printing of a variety of molecular properties. Finally,* ***Calculations*** *specifies conditions for identifying similar molecules, based either on molecular structure or chemical functionality, as well as for identifying molecules that are compatible with a particular molecular environment ("pharmacophore").*

Surfaces *is used to designate graphical surfaces, including electron and spin densities, electrostatic potentials, local ionization potentials and molecular orbitals, for later display as surfaces, property maps and contour plots. Accessible regions on electron density surfaces and property maps based on these surfaces may be demarked.*

Submit *is used to direct jobs for calculation, either locally or on a remote Linux server. Dual and quad-core processors may either be used to simultaneously process different jobs or (for certain tasks involving Hartree-Fock and density functional models) may be directed at a single job.*

The **Setup** menu provides access to dialogs for specifying molecular mechanics and quantum chemical calculations, for specifying conditions for similarity analysis as well as specifying surfaces and property maps and for submitting jobs for calculation.



Calculations...

Calculations modules perform: molecular mechanics calculations using the SYBYL and MMFF force fields; MNDO (MNDO/d for second-row elements), AM1, RM1 and PM3 semi-empirical calculations, including PM3 calculations on transition-metal inorganic and organometallic systems; Hartree-Fock molecular orbital calculations; local density calculations; BP, BLYP, EDF1 and B3LYP among other density functional calculations, including calculations based on a user-specified mix of existing exchange and correlation functionals; MP2 and RI-MP2 Møller-Plesset calculations; configuration interaction calculations (for excited-state species). All are applicable to energy calculation, equilibrium and transition-state geometry determination, conformational searching and making an energy profile, although some methods may not be practical for some tasks. Available for energy calculations only, are a series of high-order correlated models, including MP3 and MP4 Møller-Plesset models, coupled cluster models and quadratic configuration interaction models.

A wide selection of basis sets is available, ranging from the STO-3G minimal basis set (recommended for Hartree-Fock models only), to the 3-21G split-valence basis set (recommended for Hartree-Fock models only), to 6-31G*, 6-311G*, and cc-pVDZ, cc-pVTZ and cc-pVQZ polarization basis sets. In addition, pseudopotentials for elements which all-electron basis sets do not exist are available. All basis sets, except STO-3G and 3-21G may be supplemented with polarization functions on hydrogen, additional and/or higher-order polarization functions on heavy atoms and/or diffuse functions on hydrogens and/or heavy atoms.

Finally, a series of recipes for estimating heats of formation are available. These include the G3 and G3 (MP2) recipes as well as the T1 recipe that is based on G3(MP2) and provides nearly identical heats of formation, but is two to three orders of magnitude less costly in terms of computation.

Solvent may be introduced in one of two ways, both of which involve empirical formulations. The simpler and less costly in terms of computation is to add the aqueous solvation energy obtained by the SM5.4 model to the energy of any molecular mechanics or quantum chemical calculation. Equilibrium and transition-state geometries as well as spectra and any properties derived from the wavefunction are unaffected. A second and computationally more costly method is based on the more recent SM8 model. Here, the correction affects the wavefunction in addition to the energy. This means that it can be used to obtain equilibrium and transition-state geometries as well as infrared spectra and most properties that depend on the wavefunction. The solvation correction is limited to Hartree-Fock and density functional models, and applies to a variety of solvents (including water). A listing of available solvents is provided in **Appendix C**.

The Essential Edition of ***Spartan'08*** permits calculations on ground-state species only, and with molecular mechanics, semi-empirical and Hartree-Fock models only. STO-3G, 3-21G, 6-31G* and 6-311G* basis sets only, with extensions for polarization and/or diffuse functions are supported. Some properties and capabilities are not supported, and only IR

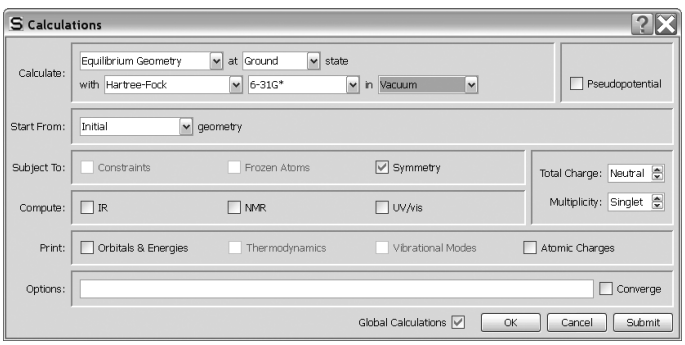
spectra are available. Differences between the full and essential editions of ***Spartan'08*** are detailed at the end of this section.

Discussion of molecular mechanics and quantum chemical methods in general, focusing on the specific methods available in ***Spartan,*** is provided in ***A Guide to Molecular Mechanics and Quantum Chemical Calculations*** which is provided as a PDF under the **Help** menu. The guide also discusses how different methods may be combined, as well as illustrates graphical modeling techniques.

Quantum chemical calculations also result in a variety of atomic and molecular properties as well as IR, NMR spectra and UV/visible. These are available for selected models.

Finally, the **Calculations** dialog accesses a procedure to quantify the extent to which two molecules or a molecule and a pharmacophore are similar. This does not involve either molecular mechanics or quantum chemical calculations but depends only on molecule (pharmacophore) geometry.

Selection of **Calculations...** results in the **Calculations** dialog*.



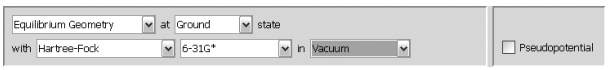
This contains a number of pull-down menus, buttons, check boxes and a space to enter additional information (keywords):

---

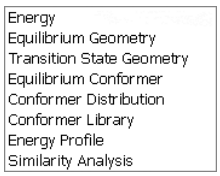
* The dialogs associated with similarity analysis and with conformer library generation are different from that associated with the other entries under the **Calculate** menu, and will be presented later in this section.

Calculate

This section contains a series of menus and check boxes that specify the task to be accomplished, and where appropriate, the electronic state (ground or first excited) the type of calculation, method, basis set and solvent to be employed, as well as details specific to each calculation type.



Specification of **Task** is by way of a pull-down menu:



Energy specifies calculation of energy (and in the case of quantum chemical methods, a wavefunction) at a single geometry.

Spartan reports energies from molecular mechanics calculations in kJ/mol, from semi-empirical calculations and from thermochemical recipes as heats of formation in kJ/mol, and from Hartree-Fock and correlated calculations as total energies in atomic units (hartrees).

The molecular mechanics energy comprises two parts: the strain energy and the non-bonded energy. The strain energy is the difference in energy between a molecule and its "strain free" analog. It is nearly always positive and less than a few hundred kJ/mols in magnitude. The non-bonded energy accounts for attraction or repulsion between atomic centers that are not connected due to van der Waals and Coulombic interactions. Because the strain energy of every molecule relates to a different standard, molecular mechanics energies cannot be used to obtain reaction energies (unless there are no changes in bonding between reactants and products).

The heat of formation is to the enthalpy at 298K of a balanced chemical reaction in which a molecule is converted to a set of standard products. For example, the heat of formation of ethylene is given by reaction,

$$C_2H_4 + \rightarrow 2C\ (\text{graphite}) + 2H_2\ (\text{gas})$$

where graphite and hydrogen molecule are the carbon and hydrogen standards, respectively. In practice, the actual measurement is typically

carried out for a combustion reaction, for example, for ethylene:

$$C_2H_4 + 3O_2 \rightarrow 2CO_2 + 2H_2O$$

Heats of formation may either be positive or negative quantities and generally span a range of only a few hundred kJ/mol.

Heats of formation are not suitable for presenting energy data from quantum chemical calculations, simply because the standards for several important elements (most notably, carbon) are not well-defined isolated species. In its place is the energy of a reaction that splits a molecule into isolated nuclei and electrons, for example, for ethylene:

$$C_2H_4 \rightarrow 2C^{+6} + 4H^{+} + 14e^{-}$$

Total energies, as the energies of such reactions are termed, are always negative and may be very large (tens of thousands of kJ/mol). They are most commonly given in atomic units (hartrees).

$$1 \text{ atomic unit} = 2625 \text{ kJ/mol}$$

Aside from a difference in units, it makes no difference whatsoever which standard is employed to investigate thermochemistry.

To summarize, the heat of formation differs from the total energy both with regard to the standard reaction and with regard to units. Either provides a suitable basis for thermochemical calculations.

Equilibrium Geometry specifies that the nearest energy minimum will be located, and **Transition State Geometry** that the nearest ***first-order saddle point*** (energy maximum in one dimension and energy minima in all other dimensions or commonly known as a ***transition state***) will be located. **Equilibrium Conformer**, **Conformer Distribution** and **Conformer Library** attempt to characterize the conformers available to a molecule based on different criteria (see discussion following). **Energy Profile** steps along a user-defined coordinate. **Similarity** quantifies the likeness among molecules based either on structure or chemical functionality or between molecules and a template (a pharmacophore).

Three different conformational analysis procedures are available in ***Spartan***. The first two (**Equilibrium Conformer** and **Conformer Distribution**) may be employed with a variety of molecular mechanics and quantum chemical models (although practical concerns may limit application to molecular mechanics), while the third (**Conformer Library**) is restricted to the MMFF molecular mechanics model.

Equilibrium Conformer replaces the initial conformer of a molecule by the lowest-energy conformer. This procedure is typically used to get a good guess at the best (lowest-energy) conformer for calculation of properties using quantum chemical methods.

Conformer Distribution provides a selection of low-energy conformers. This procedure is commonly used to obtain a set of energetically accessible conformers for calculation of average molecular properties.

Except for molecules with very few degrees of conformational freedom full systematic searches are rarely practical. A ***Monte-Carlo*** search follows a path that biases in favor of low-energy conformers (but that does not completely exclude high-energy conformers). While there is no guarantee that the lowest-energy conformer (the ***global minimum***) will actually be located, it can be shown that the set of conformers examined (and kept) approach a Boltzmann distribution. (This is strictly true only for the hypothetical case of an infinite number of search steps, but is closely approximated for searches with reasonable numbers of steps.)

A ***restricted systematic search*** walks through all possible conformers but randomly discards some without further minimization in order to enforce a fixed number of conformers. **Conformers Examined** sets this number and **Conformers Kept** ($\leq$ to **Conformers Examined**) sets the number that are actually kept.

Conformer Library replaces the initial conformer of a molecule by the lowest-energy conformer and attaches the coordinates of a set of conformers spanning all shapes accessible to the molecule without regard to energy. Used to build libraries for similarity analysis.

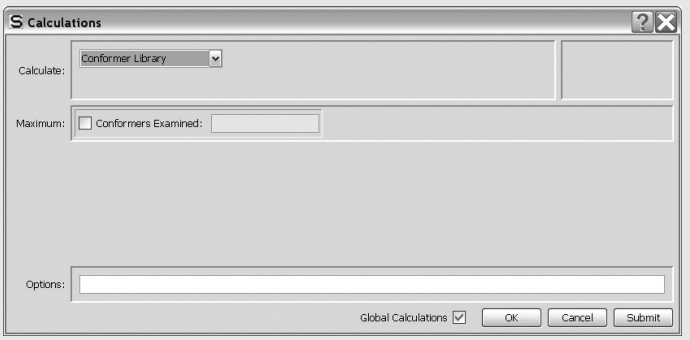
The following recipe is used to generate a conformer library:

1. Step through all possible conformers in a systematic manner and calculate equilibrium structures and energies for each using the MMFF force field. Eliminate only duplicates and very high-energy conformers.
2. Place spheres on each nitrogen, and oxygen atom and at the center of each hydrophobic region for each conformer.

3. Find the smallest set of conformers for which spheres do not overlap, that is, eliminate any conformer for which all of its spheres overlap with the corresponding spheres on another conformer.

This procedure involves a systematic search of all possible conformers. A set of conformer assignment rules that eliminate degrees of freedom that are unlikely to produce significant shape changes are available. See **Set Torsions** under the **Geometry** menu (**Chapter 13**) and **Settings Preferences** under the **Options** menu (**Chapter 18**). A limit may be placed on the total number of conformers actually considered by first calculating the fraction of conformers that are to be kept (total number of conformers to be examined/total number of systematic conformers), and then randomly deciding to examine or skip each conformer in a systematic search in order to enforce this fraction. Such a procedure is different from the Monte-Carlo method, in that it strives to span all possible shapes open to a flexible molecule without consideration of energy.

Selection of **Conformer Library** results in a different dialog.

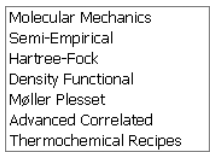


The only control, **Conformers Examined**, sets the maximum number of conformers.

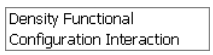
Energy applies to all methods. **Equilibrium Geometry**, **Equilibrium Conformer**, **Conformer Distribution** and **Energy Profile** apply to all methods except advanced correlated methods and thermochemical recipes. **Transition State Geometry** applies to all methods except molecular mechanics methods, advanced correlated methods and thermochemical recipes. **Conformer Library** applies only to MMFF. **Similarity Analysis** does not involve any molecular mechanics or quantum chemical methods.

Specification of **State** (electronic state) is by way of a pull-down menu. The default setting **Ground** may be changed to **First Excited** by *clicking* on ⬍ to the right of the box and back again to **Ground** by *clicking* on ⬍. Different methods are available to handle ground and excited-state species.* Hartree-Fock, Møller-Plesset as well as a variety of advanced correlated methods are available only for ground-state species, while configuration interaction methods are available only for excited-state species. Density functional methods are available for both ground and excited-state species. **State** is ignored for molecular mechanics calculations, and semi-empirical calculations are limited to the ground state.

Specification of **Type of Calculation** is by way of a menu, the selections in which depend on electronic state. For ground states:

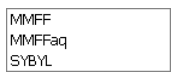


For excited states:

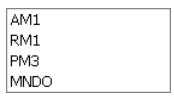


Some calculation types require additional information:

Molecular Mechanics



Semi-Empirical



Density Functional**

---

* Note that gradients for density functional methods for excited states need to be calculated numerically. This means that determining an equilibrium geometry for an excited state of a molecule is significantly slower than determining an equilibrium geometry for the ground state.

** A variety of other functionals, including local and EDF2 functionals are available and may be entered directly into the **Options** box. A full listing is provided in **Appendix A**. Additionally, any combination of supported exchange and correlation functionals may be specified.

Møller-Plesset	MP2 RI-MP2
Advanced Correlated*	MP3 MP4 CCSD CCSD(T)
Thermochemical Recipes**	G3(MP2) T1
Configuration Interaction	CIS CIS(D)

The most important application of **molecular mechanics models** is conformational searching, in particular, for molecules with several degrees of conformational freedom. In these cases, even the simplest quantum chemical models may not be practical. The MMFF force field is known to properly assign equilibrium conformation in a variety of simple molecules for which experimental data are available. It also has been shown to provide a reasonable account of conformational energy differences in larger organic molecules as obtained from high-level quantum chemical calculations. Aside from conformational energy differences, molecular mechanics models are not suitable for thermochemical calculations.

Semi-empirical models are simplest methods based on quantum mechanics. They are applicable to molecules containing 100 - 200 atoms, but not to molecules containing thousands of atoms, for example, proteins. The MNDO, AM1, RM1 and PM3 models generally provide geometries in good accord with experimental structures. This makes them suitable for evaluation of such properties as polar surface area that depend solely on geometry. However, none of the present-generation semi-empirical models is suitable for the calculation of relative energies as might be required to conclude, whether a chemical reaction is weakly or strongly *exothermic*, thermoneutral or weakly or strongly *endothermic*, or whether one isomeric product of a reaction is likely to be more or less stable than another, or to identify the lowest-energy tautomer. Also, none of the methods reliably accounts for conformational energy differences, as would be required to establish whether a particular conformer has a good chance at actually being present or to calculate the average value of a property based on the relative (Boltzmann) populations of different conformers.

Hartree-Fock models follow from the **Schrödinger equation** by requiring that the electrons be independent particles. This is known as the **Hartree-Fock approximation**. Here, the motions of electrons in

---

* A number of other advanced correlated methods are available, and may be entered directly into the **Options** box: CCSD(2), OD, OD(T), QCCD, QCCD(2), QCISD and QCISD (T).

** G2 and G3 recipes are available and may be entered into the **Options** box.

molecules (molecular orbitals) are approximated by a sum of the motions of electrons in atoms (atomic orbitals). A second approximation, termed the **LCAO** or **L**inear **C**ombinations of **A**tomic **O**rbital **approximation**, distinguishes different Hartree-Fock models. Different models use different basis sets, that is, different numbers and kinds of atomic orbitals. For example, the model with the 3-21G split-valence basis set uses one atomic orbital to describe each non-valence electron and two atomic orbitals to describe each valence (core) electron, that is, it comprises one 1s atomic orbital and two sets of 2s and 2p atomic orbitals.

1s
2s, $2p_x$, $2p_y$, $2p_z$
$2s'$, $2p_x'$, $2p_y'$, $2p_z'$

The model with the 6-31G* polarization basis set increases the flexibility by including d type atomic orbitals that, while not occupied in the atom, are used in molecules.

1s
2s, $2p_x$, $2p_y$, $2p_z$
$2s'$, $2p_x'$, $2p_y'$, $2p_z'$
d_{xx}, d_{xy}, d_{xz}, d_{yy}, d_{yz}, d_{zz}

Larger basis sets involving even higher splitting of valence shells, including higher-order (f and g type) functions as well as diffuse functions can be and have been defined and are available in ***Spartan***. Description of their makeup is provided in ***A Guide to Molecular Mechanics and Quantum Chemical Calculations*** which is provided as a PDF under the **Help** menu.

3-21G and 6-31G* and larger basis set Hartree-Fock models generally provide good accounts of the geometries of organic molecules. This extends to unusual or unstable systems for which experimental data may be limited (and hence may be poorly represented in parameterization of the semi-empirical models). The 6-31G* model (and to a lesser extent, the 3-21G model) also properly accounts for thermochemistry, at least insofar as being able to say whether a particular reaction is weakly or strongly *exothermic* or weakly or strongly *endothermic*. It is also able to properly order the stabilities of isomeric products and to reliably identify the lowest-energy tautomer. Finally, 6-31G* calculations provide a qualitatively correct account of the relative stabilities of different conformational arrangements. Hartree-Fock models with larger basis sets show similar behavior.

The most conspicuous and most important flaw of Hartree-Fock models (even with very large basis sets) is their inability to properly account for the energetics of bond dissociation. This can be traced to the Hartree-Fock approximation which in effect replaces interactions between electrons by interactions between one electron and the average field created by all other electrons. As a consequence, electrons "get in each other's way" to a greater extent than they should. This leads to overestimation of the electron-electron repulsion energy, most significantly for electrons that are paired. Because bond dissociation reduces the number of electron-pairs (by one), the overestimation will be greater for reactants than for products and the calculated dissociation energy will be too small.*

Models that do away with the Hartree-Fock approximation, or at least lessen its effect, are termed **correlated models**. These divide into two broad categories, **density functional models** and **wavefunction-based models**. Density functional models explicitely introduce an approximate correlation term. They are not much more costly in terms of computation time than Hartree-Fock models, although it is not apparent how to improve on a particular choice of correlation correction. Wavefunction-based models, typified by **Møller-Plesset models** and **configuration interaction models** extend the flexibility of Hartree-Fock models by mixing ground-state and excited-state wavefunctions. They are significantly more costly than both Hartree-Fock and density functional models, but offer the advantage over the latter of a clear path to improvement. In the limit of complete mixing, so-called full configuration interaction and infinite order Møller-Plesset models, wavefunction-based models lead to the exact result, although in practice this limit cannot be reached.

Density functional and wavefunction-based correlated models make use of the same basis sets as Hartree-Fock models, except that 3-21G (and smaller) basis sets do not yield satisfactory results. Among the most popular models are the B3LYP/6-31G* density functional model and the MP2/6-31G* (second-order) Møller-Plesset model. The primary use of these models (and other correlated models) is to obtain thermochemistry, including for reactions in which bonds are made or broken and including activation energies.

While geometry is well described with simple quantum chemical models, energy (or more precisely relative energy) has proven to be more difficult

---

* The limiting case is H_2. Here, the Hartree-Fock energy of the products (hydrogen atoms each with only one electron) is "exact", while the energy of the reactant (H_2 with two electrons) is too large. Therefore, the Hartree-Fock bond dissociation energy will be too small.

to calculate accurately. Hartree-Fock models properly describe the energies of ***isodesmic*** reactions, where formal bond count is maintained but are less successful for reactions in which reactants and products have markedly different bonding. Indeed, high-level correlated calculations with large underlying basis sets are needed to properly account for the energetics of bond breaking or to relate the energies of reactants to that of a transition state.

A number of thermochemical recipes or combinations of different quantum chemical methods have been formulated for the purpose of providing accurate thermochemical data. The goal has been to be able to reproduce experimental heats of formation to within 4-8 kJ/mol. (Except for very simple molecules, experimental thermochemical data are seldom determined more accurately than this.) G3(MP2) is the simplest and perhaps the most widely used of the recipes that have been formulated. It reproduces experimental heats of formation for a wide variety of small molecules to within 6 kJ/mol (mean absolute error), with only a small number of systems showing errors twice this amount. The problem with G3(MP2) and other schemes of this type are their large computational cost, and closely associated, the severe limits on the size of systems that can be dealt with. In practice, G3(MP2) may easily be applied to molecules comprising ten heavy (non-hydrogen) atoms and only with great difficulty to molecules comprising fifteen or more heavy atoms.

Three different energies are required in the G3(MP2) recipe:

E_A MP2/6-31G* using the MP2/6-31G* equilibrium geometry
E_B MP2/6-311++G(2df, 2p) using the MP2/6-31G* equilibrium geometry
E_C QCISD(T)/6-31G* using the MP2/6-31G* equilibrium geometry

These are combined in order to take into account both the effects of increased basis set size and of electron correlation beyond the MP2 level. Note, that the two effects are treated independently in order to avoid having to perform a QCISD(T) calculation using a large basis set.

$$E = E_A + (E_B - E_A) + (E_C - E_A)$$

Two corrections are needed to bring the calculated energy in line with what is actually measured experimentally: a correction for zero-point energy; $E_{zero\text{-}point}$, and a correction to the enthalpy (energy) to account for finite temperature (298K), H(T).

$$E_{zero\text{-}point} = \frac{1}{2} h \sum_{i}^{vibrations} \nu_i$$

$$\Delta H(T) = H_{trans}(T) + H_{rot}(T) + \Delta H_{vib}(T) + RT$$

$$H_{trans}(T) = \frac{3}{2} RT$$

$$H_{rot}(T) = \frac{3}{2} RT \text{ (RT for a linear molecule)}$$

$$\Delta H_{vib}(T)^* = H_{vib}(T) - H_{vib}(T) - H_{vib}(0)$$

$$= Nh \sum_{i}^{vibrations} \frac{\nu_i}{(e^{h\nu_i/kT} - 1)}$$

h is Planck's constant, k is the Boltzmann constant, R is the gas constant and N is Avogadro's number. Both zero-point and temperature corrections require the full set of vibrational frequencies. In the G3(MP2) recipe, these are obtained from Hartree-Fock/6-31G* calculations and require the corresponding equilibrium geometry.

To summarize, the G3(MP2) recipe requires two equilibrium geometry calculations (at HF/6-31G* and MP2/6-31G*), a HF/6-31G* frequency calculation, a MP2/6-311++G (2df, 2p) energy calculation and a QCISD(T)/6-31G* energy calculation. The QCISD(T) calculation dominates with increasing molecular size and also requires the most temporary memory and disk storage. The next most costly term is likely to be the MP2/6-31G* geometry optimization, although both the large basis set MP2 energy and the HF/6-31G* frequency calculation may also be significant.

T1 is a simplified recipe based on G3(MP2) and intended to closely reproduce G3(MP2) heats of formation has been developed. It eliminates the QCISD(T)/6-31G* energy calculation (E_C), the zero-point energy calculation ($E_{zero\text{-}point}$) and the temperature correction to the enthalpy ($\Delta H(T)$), and replaces the large basis set MP2 energy calculation (E_B) with a dual-basis-set RI-MP2 calculation, where G3MP2large is the large basis set (the same basis set that is used in the G3(MP2) recipe) and 6-311G* is the small basis set. To compensate, an empirical correction based on the Hartree-Fock and RI-MP2 energies and Mulliken bond orders is introduced. This follows from a linear regression analysis involving G3(MP2) heats of formation for ≈1,050 diverse molecules.

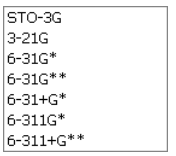
---

* Note, that the expression for $\Delta H_{vib}(T)$ goes to RT instead of 1/2 RT as ν_i goes to zero. This requires that contributions that would exceed 1/2 RT be set to 1/2 RT.

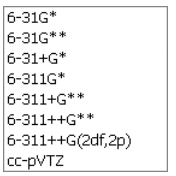
T1 uses HF/6-31G* equilibrium geometries, instead of MP2/6-31G* geometries in G3(MP2).

Except for semi-empirical methods, either **restricted** or **unrestricted scf methods** may be used. The former restricts paired electrons to the same spacial orbital, whereas the latter allows spin up and spin down electrons to occupy different orbitals. The default (singlets are described using restricted methods and non-singlets are described using unrestricted methods) may be overridden using the **Scf** keyword (**Appendix C**).

Except for molecular mechanics and semi-empirical methods and for thermochemical recipes, a basis set needs to be specified. Menus provide a number of choices.* For Hartree-Fock methods:



For all other methods:



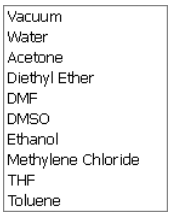
The 3-21G basis set for second-row and heavier main-group elements incorporates a set of d-type functions. While these are not occupied in the atomic ground state, they have been found necessary for proper description of bonding in molecules. 6-311++G (2df,2p) uses the 6-311++G (3d2f,2p) basis set for second-row elements.

---

* A number of other basis sets are available, and may be entered directly into the **Options** box: These include the Pople 6-31G or 6-311G basis sets followed by: (*i*d*j*f) or (*i*d*j*f, *k*p*l*d) where *i* is the number of sets of d functions on non-hydrogen atoms, j is the number of sets of f functions on non-hydrogen atoms, *k* is the number of s functions on hydrogen atoms and *l* is the number of sets of p functions on hydrogen atoms. If 0 eliminate p, d, f; if 1 eliminate the number; 6-31G or 6-311G with "+" or "++" inserted before G; cc-pVDZ and cc-pVQZ.

The combination of a method, for example, B3LYP, and a basis set, for example, 6-31G*, constitutes a **theoretical model** or more simply a **model**. The nomenclature is to separate method and basis set by a "/", for example, B3LYP/6-31G*. By convention, specification of Hartree-Fock (HF) as a method is optional. That is, specification of basis set alone refers to Hartree-Fock models, for example, both HF/6-31G* and 6-31G* refer to combination of the Hartree-Fock method and the 6-31G* basis set. Methods without basis sets, for example, MMFF (molecular mechanics) and PM3 (semi-empirical) are themselves models.

Finally, a solvent can be specified for Hartree-Fock and density functional calculations. This is done by replacing ***vacuum*** in the menu to the right of **in** either by water or by one of a selection of common organic solvents.*



The effect of the solvent on the energy and wavefunction (and hence on any properties derived from the wavefunction) is approximated by way of the SM8 model. This has been parameterized to reproduce experimental energies of solvation and related quantities. It has not been parameterized to reproduce equilibrium geometries and properties other than the energy, and there is no solid calibration of its success or failure for anything other than energy. It should be used with caution.

The cost of SM8 in terms of computation time is roughly equal to that of an energy calculation. Therefore, solvated Hartree-Fock and density functional energy calculations are twice as costly as the corresponding calculation in vacuum while geometry calculations are roughly 30% more costly.

If *checked*, **Pseudopotential** signifies that a pseudopotential is to replace the all-electron basis set for elements for which all-electron

* Other solvents are available and can be entered via the **Solvent** keyword in the **Options** box.

basis sets are not available. In effect, this limits description of heavy elements to valence electrons only. Pseudopotential pertains to Hartree-Fock, density functional, MP2 and advanced correlated calculations only, and only for third-row and heavier elements. Does not apply to RI-MP2 calculations and to calculations using the STO-3G and 3-21G basis sets. For availability and description of pseudopotentials in ***Spartan***, consult the help facility (**Chapter 19**).

If *checked*, **Dual Basis** signifies that a Hartree-Fock, density functional, MP2 or RI-MP2 energy equilibrium and transition-state geometry calculations with the 6-311+G**, 6-311++G**, 6-311++G (2df, 2p), or cc-pVTZ basis set will be approximated by a calculation in which scf convergence is first achieved using the smaller 6-311G* basis set and then corrected pertubatively for the effects of basis set extension. Other combinations of small and large basis sets may be specified under **Options** (see discussion following). Use of dual basis has only a small effect on calculated relative energies, but can reduce the cost of large-basis-set energy calculations by two to ten times (depending on the choice of the large basis set).

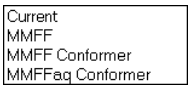
An additional control, **IRC** (**I**ntrinsic **R**eaction **C**oordinate), appears when a transition-state geometry has been requested. If *checked*, **IRC** signifies that the resulting transition state will be used to generate a pathway (an **intrinsic reaction coordinate**) leading first to reactant and then to product. (It is advisable, although not necessary, to compute the IR spectrum prior to calculation of an intrinsic reaction coordinate.) The full sequence of steps: reactant → transition state → product, will be placed in a new file ***document.IRC.identifier*** where ***document*** is the name of the document submitted, ***IRC*** designates the origin as an IRC calculation and ***identifier*** is the molecule identifier within the document. The default number of steps in the sequence (40) may be changed using the keyword **IRCSteps** (see **Appendix C**). **IRC** is not available for semi-empirical models.

Start from (Energy, Equilibrium Geometry)

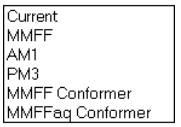
Energy and equilibrium geometry calculations (only) from one model may be preceded by equilibrium geometry calculations at another (simpler) model. Choices depend on the method selected:

Molecular Mechanics not applicable

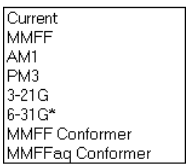
Semi-Empirical



Hartree-Fock



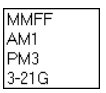
Density Functional, Møller-Plesset, Advanced Correlated, Thermochemical Recipes



Selection of **MMFF conformer** or **MMFFaq conformer** signifies that geometry optimization will follow a search for the lowest-energy conformer. These lead to an additional menu specifying the level of geometry optimization.

Molecular Mechanics, Semi-Empirical not applicable

Hartree-Fock



Density Functional, Møller-Plesset, Advanced Correlated, Thermochemical Recipes

Maximum (Equilibrium Conformer, Conformer Distribution)

Maximum Conformers Examined sets a maximum for the number of possible conformers that will be considered both for **Equilibrium Conformer** and **Conformer Distribution** tasks. Note, that if a systematic search has been requested

(**SearchMethod** keyword; **Appendix B**) this invokes an algorithm whereby systematic conformers are randomly eliminated in order to enforce the conformer limit.

Maximum Conformers Kept sets a maximum for the number of conformers returned at the end of a **Conformer Distribution** task. It must be smaller or equal to the number of conformers examined.

Subject to

Spartan allows calculations to be carried out in the presence of geometrical constraints or with atoms which have been frozen in place. In addition, ***Spartan*** attempts to make use of symmetry.

If *checked*, **Constraints** signifies use of any previously defined constraints on distances, angles and dihedral angles into equilibrium and transition-state geometry optimization, conformation searching and generation of energy profiles. Does not apply to energy calculation, conformer library generation or similarity analysis. See **Constrain Distance**, **Constrain Angle** and **Constrain Dihedral** under the **Geometry** menu (**Chapter 13**) for information about constraining geometrical parameters.

NOE data are not treated as constraints, but rather as post-calculation filters in establishing conformer distributions.

If *checked*, **Frozen Atoms** signifies that the coordinates of any atoms that have previously been frozen will not be moved during equilibrium and transition-state geometry optimization, conformation searching and generation of energy profiles. Does not apply to energy calculation, conformer library generation or similarity analysis. See **Freeze Center** under the **Geometry** menu (**Chapter 13**) for information about freezing atoms. See also discussion in **Chapter 17** regarding the freezing of heavy atoms in structures resulting from a search of the Cambridge Structural Database.

If *checked*, **Symmetry** signifies that molecular symmetry is to be employed wherever possible to simplify the calculation. By default **Symmetry** is checked.

Compute

Entries under this section request calculation of a number of important spectral quantities. Other quantities such as the dipole moment, the estimated aqueous solvation energy* and LogP** that are calculated automatically, whereas calculation of still others may be requested from the **Options** box (see discussion following).

IR

If *checked*, calculates vibrational frequencies and intensities together with the corresponding vibrational modes. These are available in the output (**Output** under the **Display** menu; **Chapter 16**) along with selected thermodynamic properties (entropies*** and Gibbs energies). Vibrational modes may be animated and an IR spectrum displayed from the **IR** dialog accessible from **Spectra** under the **Display** menu (**Chapter 16**). Second derivatives required for frequency calculation are evaluated analytically for molecular mechanics, semi-empirical, Hartree-Fock and density functional models, but involves numerical differentiation of analytical gradients for RI-MP2 and MP2 models, this is much more costly in terms of computation. Frequency calculations for higher-order correlated models are not available.

---

* The aqueous solvation energy is estimated using the SM5.4 procedure of Cramer, Truhlar and co-workers (C.C. Chambers, G.D. Hawkins, C.J. Cramer and D.G. Truhlar, *J. Phys. Chem.,* **100**, 16385 (1996)). This is added to the gas phase total energy and written to the output. This sum is also available in the spreadsheet (**Chapter 16**). The SM8 solvation model (A.V. Marenich, R.M. Olsen, C.P. Kelly, C.J. Cramer and D.G. Truhlar, *J. Chem. Theory Comput.*, **3**, 2011 (2007)) is also available for Hartree-Fock and density functional calculations (see previous discussion). Unlike SM5.4, this affects the wavefunction (and any properties derived from the wavefunction). It can be used for equilibrium and transition-state geometry optimization as well as for calculation of vibrational frequencies. Parameters have been provided for water as well as a variety of organic solvents. Unlike SM5.4, the computational cost of the SM8 model is significant, and it is not invoked unless specifically requested.

** LogP is estimated both according to the method of Ghose, Pritchett and Crippen (*J. Computational Chem*., **9**, 80 (1988)), and of Villar (*J. Computational Chem*., **6**, 681 (1991); *Int. J. Quantum Chem*., **44**, 203 (1992)), and written to the output. Both Ghose-Crippen and Villar LogP values may also be brought into the spreadsheet (**Chapter 16**).

*** Note that the linear harmonic oscillator approximation incorrectly leads to a vibrational contribution to the entropy that goes to ∞ as frequency goes to 0. Vibrational entropy contributions from frequencies below 300 cm^{-1} should be treated with caution.

NMR

If *checked*, specifies that NMR chemical shifts will be calculated. These are then available in the output (**Output** under the **Display** menu; **Chapter 16**) as well as from the **Atom Properties** dialog (**Display** menu) and as atom labels (**Configure...** under the **Model** menu; **Chapter 12**). ^{13}C (proton decoupled) and ^{1}H (shifts only) spectra from Hartree-Fock or density functional calculations may be displayed from the **NMR Spectra** dialog accessible from **Spectra** under the **Display** menu (**Chapter 16**)*. ^{13}C chemical shifts from the B3LYP/6-31G* model (only) that have been empirically corrected for local environment (numbers and kinds of directly-bonded atoms) are also available in the printed output, as atom labels, from the **Atom Properties** dialog and in spectral plots. Line intensities are assumed to be proportional to the number of equivalent carbons or hydrogens. Coupling constants associated with ^{1}H spectra are presently not calculated, but three-bond HH coupling constants are estimated empirically and may be displayed either in terms of spectral splittings or as COSY plots.

UV/vis

If *checked*, specifies that a series of excited-state energy calculations will be performed (following a ground-state calculation) and that state-to-state energy differences will be reported. A UV/visible spectrum may be displayed from the **UV/vis Spectra** dialog accessible from **Spectra** under the **Display** menu (**Chapter 16**). The default number of excited states is 6, but may be changed using the **UV States** keyword (**Appendix C**). UV/visible spectra calculations are restricted to Hartree-Fock models (CIS for the excited state) and density functional models (time-dependent density functional theory for the excited state).

---

* Chemical shifts for other nuclei are available in the **Output** dialog (**Output** under the **Display** menu) and may also be attached as labels (**Configure...** under the **Model** menu).

Total Charge

Total charge. The default setting **Neutral** may be changed either by *clicking* on ⇕ to the right of the box, and selecting **Anion**, **Dianion**, **-3**, etc. from the menu which appears, or by replacing any menu entry by a numerical value. **Total Charge** is ignored for molecular mechanics calculations.

Multiplicity

Spin multiplicity. The default setting **Singlet** may be changed either by *clicking* on ⇕ to the right of the box, and selecting instead **Doublet** or **Triplet** from the menu which appears, or by replacing any menu entry by a numerical value. Multiplicity is 1 for singlets, 2 for doublets, 3 for triplets, 4 for quartets, etc. **Multiplicity** is ignored for molecular mechanics calculations.

Print

If *checked*, writes the quantity to the output window. Text output may be seen by selecting **Output** from the **Display** menu (**Chapter 16**) and printed by first selecting the output dialog and then selecting **Print Output** from the **File** menu (**Chapter 10**). Additional printing may be requested from the **Options** box (see discussion following).

Orbitals & Energies

If *checked*, writes the orbitals and energies to the output. Orbital energies are also available in the spreadsheet (**Chapter 16**).

Thermodynamics

If *checked*, writes thermodynamic quantities to the output. Requires that vibrational frequencies (**IR** under **Compute**) have been calculated. Selected thermodynamic quantities are also available in the spreadsheet (**Chapter 16**).

Vibrational Modes

If *checked*, writes vibrational frequencies and modes to the output. Requires that **IR** under **Compute** have been calculated.

Atomic Charges

If *checked*, writes Mulliken, natural and electrostatic-fit charges to the output. Note that Mulliken and natural charges from semi-empirical models are identical. Atomic charges are also available from the **Atom Properties** dialog (**Display** menu; **Chapter 16**), in the spreadsheet (**Chapter 16**) and as atom labels (**Configure...** under the **Model** menu; **Chapter 12**). Atomic charges are not available for calculations using advanced correlated models or for thermochemical recipes.

Options

Program options may be specified using keywords entered into the **Options** box. Keywords may be either single words or expressions. Keyword=N indicates an integer argument, and keyword=F indicates a real argument. **Appendix C** contains a listing of commonly-used keywords.

Converge

If *checked*, invokes procedures to assist in convergence of quantum chemical calculations. May lead to an increase in computation time.

Similarity

Spartan provides a procedure to assess and quantify the similarity of one or more molecules in a document to one or more molecules in one or more different documents. Similarity may either be based on molecular structure (geometry) or on chemical function descriptors (CFD's). ***Spartan*** also assesses the similarity of one or more pharmacophores in a document to one or more molecules in one or more different documents.

Three different permutations of similarity analysis are available. Here, ***template*** refers to the selected document in the graphical user interface and ***library*** refers to one or more different documents specified in the **Calculations** dialog.

template	*library*
pharmacophores	molecular structures for which CFD's are assigned on-the-fly
molecular structures	molecular structures
CFD's from molecular structures	molecular structures for which CFD's are assigned on-the-fly

Two types of libraries are supported for molecular structures (and CFD's generated on-the-fly from molecular structures). In the first, each entry represents a single molecule. In the second, each entry represents a selection of different conformers of a single molecule and spanning all possible shapes that the molecule may take on. The latter are generated using **Conformer Library** under the **Calculate** menu (see discussion earlier in this chapter), and ***may only be used as library entries in similarity analysis and not as templates***.

Of the three cases for similarity analysis, only the second makes direct use of molecular structures (selected atomic centers for the molecules in the template) as a similarity measure. The other two libraries make use of CFD's generated and (optionally) edited for the template and generated on-the-fly in the library. Selection of structure or CFD as the basis of similarity measure must be explicitely stated (see discussion following).

Seven different chemical function descriptors (CFD's) are recognized by the similarity analysis procedure, the first six of which are attributes of a molecule while the seventh derives from knowledge of a molecule bound to a protein or nucleic acid.

hydrophobe (sterically crowded region)
aromatic hydrophobe (aromatic ring)
hydrogen-bond donor
hydrogen-bond acceptor
positive ionizable site
negative ionizable site
excluded volume

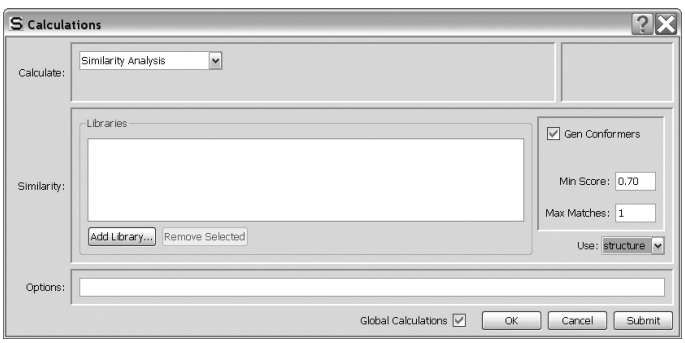
Discussion of CFD types has already been provided (see **CFD's** under the **Model** menu; **Chapter 12**).

The individual CFD's for molecules in the template (only) may either be used or ignored in the similarity procedure. This is specified from **Set Similarity Centers** under the **Geometry** menu (**Chapter 13**). The **CFD Properties** dialog (**Properties** under the **Display** menu; **Chapter 16**) allows for changing or extending the definition of a CFD For example, the nitrogen in a dimethylamino group might initially be assigned as a hydrogen bond acceptor CFD, but could be changed to a positive ionizable site (as such a center could become positively-charged at biological pH) or could be extended to cover both possibilities.

Spartan assumes that any chiral molecules in a library are unresolved (they comprise an equal mixture of enantiomers). By default, similarity comparisons are carried out on both the library molecules as they are stored and (if the molecules are chiral) on their mirror images. The default procedure may be overridden (providing comparisons only to the stored molecules) by way of the **SingleEnantiomer** keyword entered in **Options** in the **Calculations** dialog. See discussion later in this chapter and also **Appendix C**.

Note that only a single enantiomer of a chiral template molecule is considered. If comparisons involving both enantiomers are desired, the second enantiomer must be added to the query.

Selection of **Similarity** from the **Calculate** menu with either a molecule or pharmacophore (list of molecules/pharmacophores) leads to a dialog for setting up a similarity analysis.



One or more libraries need to be identified. This is accomplished by *clicking* on the **Add Library...** button, and gives rise to a file browser. Once selected, *clicking* on **Open** adds the library to the list of templates. A library may be deleted from the list by *clicking* on its name in the list and *clicking* on the **Remove Selected** button at the bottom of the dialog. **Shift** and **Ctrl** keys function in the usual manner in both library addition and removal.

Prior to starting a similarity analysis, you need to select either **structure** or **CFD** from the **Use** menu near the bottom of the **Calculations** dialog.

Similarity analysis is most valuable where the template molecule is conformationally flexible, and it is not apparent which (if any) of the different shapes it adopts might match a particular target in a library. It becomes indispensable where the target is also flexible and its conformation is not established. Here, the process must be repeated for every template conformer-target conformer pair, that is, thousands or tens of thousands of times.

This process can be shortened if the purpose is simply to identify which target molecules a particular template molecule might match, irrespective of conformer. In this case, the first good match concludes the analysis, and the user is left with the knowledge that template and target can be made to match (together with an example of a match).

The similarity algorithm incorporated in ***Spartan*** attempts to examine template-target pairs that are most likely to lead to matches before examining pairs that are less likely. It also allows early termination upon finding a single good match, or a user-specified number of good matches*. Taken together, these can lead to one or two orders of magnitude speedup in the analysis.

---

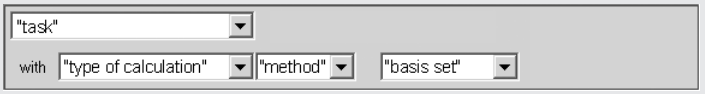
* Criteria for early termination are specified using program options. See **Similarity Analysis Options** in **Appendix C**.

Global Calculations

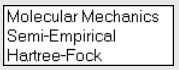
If *checked*, signifies that settings in the **Calculations** dialog are to be applied to all molecules in the document.

The **Calculations** dialog may be exited by clicking on **Submit**, **Cancel** or **OK** at the bottom right of the dialog, or on ☒ at the top. (**Submit** and **OK** are not available if the job is already executing.) Clicking on **OK** or on **Submit** overwrites any previous information. Additionally, **Submit** enters the job in the execution queue (see discussion later this chapter). Clicking on **Cancel** or on ☒ exits the **Calculations** dialog without saving any changes.

The **Calculations** dialog in the Essential Edition of ***Spartan'08*** (**NAL**) differs primarily in the **Calculate** section.



The **Task** menu is the same. The **Type of Calculation** menu is restricted to molecular mechanics, semi-empirical and Hartree-Fock models under the **Method** menu.



The **Basis Set** menu (Hartree-Fock models only) is identical. Pseudopotentials are not available.

Additionally, intrinsic reaction coordinate (**IRC**) calculations and NMR spectra and UV/visible (**NMR** and **UV/vis** under **Compute**) are not available in the essential edition. IR spectra (**IR** under **Compute**) are available, but the required second-derivative calculations are performed numerically rather than analytically and are significantly slower than in the full edition.

Surfaces

Spartan allows graphical display of the HOMO and LUMO among other molecular orbitals, the electron density, the spin density for molecules with unpaired electrons, the electrostatic potential and the local ionization potential.

The ***electron density*** is the number of electrons found at a point in space. It is the quantity measured in an X-ray diffraction experiment that is then used to locate atomic positions, that is, most electrons are closely associated with atoms. While the electron density is non-zero everywhere, it is possible to define surfaces of constant density. The most important of these contains most of a molecule's electrons and that roughly corresponds to a space-filling model, that is, a van der Waals surface. This is generally referred to simply as an ***electron density***, and is interesting because it reveals overall molecular size and shape and demarks the steric barrier seen by encroaching molecules. Another important surface, referred to as a ***bond density***, contains fewer electrons in total and demarks atomic connectivity.

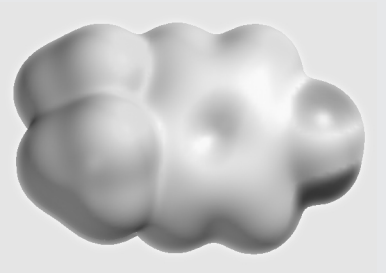
The ***spin density*** is the difference in the number of electrons of α and β spin at a point in space. It indicates the location of the unpaired electron in a radical or unpaired electrons in a triplet or higher multiplicity state.

The ***electrostatic potential*** is the energy of interaction of a positive charge with a molecule. It represents a balance between repulsive interactions involving the positively-charged nuclei and attractive interactions involving the negatively-charged electrons. Regions where the balance tips toward attraction are said to be electron rich (basic) and subject to attack by electrophiles, while regions where the balance tips toward repulsion are said to be electron poor (acidic) and subject to attack by nucleophiles. Electron-rich regions such as lone pairs are typically located outside the van der Waals surface. As such, they may be easily identified by constructing a surface of negative (attractive) electrostatic potential. While interesting electron-poor areas such as acidic hydrogens also lie outside the van der Waals surface, the electrostatic potential is also positive (repulsive) throughout the region inside this surface.

The ***local ionization potential*** indicates the ease or difficulty of electron removal (ionization). Like the negative regions of the electrostatic potential, regions of low local ionization potential are likely to be subject to attack by electrophiles.

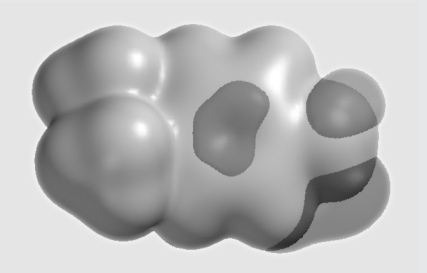
Additionally, any one of the quantities listed above (except the electron density) may be mapped onto any surface (except a molecular orbital surface). In practice, the only maps to have received widespread attention are those based on the electron density surface (depicting overall molecular size and shape). Most common are the electrostatic potential map, the local ionization potential map and the LUMO map.

The ***electrostatic potential map*** paints the value of the electrostatic potential onto an electron density surface. By convention, colors toward red depict negative potential, while colors toward blue depict positive potential, and colors in between (orange, yellow, green) depict intermediate values of the potential. Thus, an electrostatic potential map for *p*-*tert*-butylphenol will show oxygen to be red, its attached (acidic) hydrogen to be blue, the π faces of benzene to be orange or yellow and the *tert*-butyl group to be green.



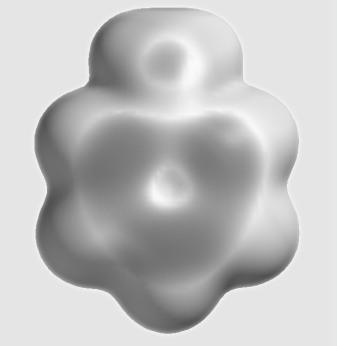
The main advantages of this presentation relative to separate electron density and electrostatic potential surfaces are its clarity and its compactness. The main disadvantage is that it provides information only about the contact surface and does not reveal how far electron-rich and electron-poor areas extend beyond the surface.

An alternative to an electrostatic potential map, referred to as an ***exposed electrostatic potential surface***, is a composite of three different surfaces: an electron density surface depicting overall molecular size and shape, a negative electrostatic potential surface identifying electron-rich regions and a positive electrostatic potential surface identifying electron-poor regions. These surfaces need to be generated and then displayed simultaneously. The electron density may either be displayed as an opaque solid or as a transparent solid (in order that the molecular skeleton may be seen inside). The two potential surfaces are best represented as transparent solids, the negative surface colored red and the positive surface colored blue. The exposed electrostatic potential surface for *p*-*tert*-butyl phenol is shown on the next page.



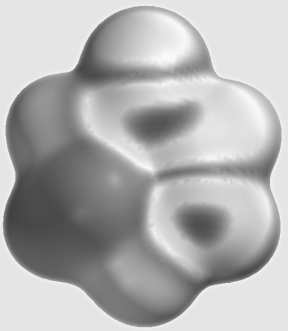
Note that the exposed electrostatic potential surface provides the same information as the electrostatic potential map. Red areas in the map correspond to regions when the negative electrostatic potential surface is likely to protrude from the electron density while blue areas correspond to regions where the positive electrostatic potential surface is likely to stick out. There is broad similarity of this graphic to CFD representations discussed previously (see **CFD's** under the **Model** menu in **Chapter 12**). Hydrogen-bond acceptors and/or positive ionizable sites are associated with protruding negative potential surfaces, whereas hydrogen-bond donors and/or negative ionizable sites are associated with protruding positive potential surfaces.

The ***local ionization potential map*** paints the value of the local ionization potential onto an electron density surface. By convention, colors toward red indicate low ionization potential, while colors toward blue indicate high ionization potential. Thus, the local ionization potential map for aniline shows that the *ortho* and *para* ring positions have a lower ionization potential than the *meta* positions, consistent with the known directing ability of an amino group in electrophilic aromatic substitution.

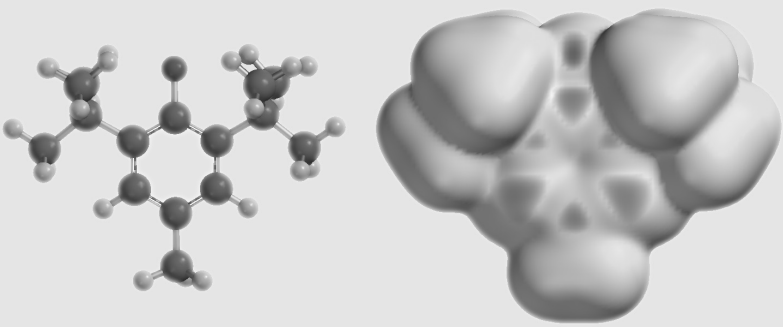


The ***LUMO map*** paints the absolute value of the lowest-unoccupied molecular orbital (the LUMO) onto an electron density surface. By convention, colors near blue indicate high concentration of the LUMO, while colors near red indicate low concentration. Given that the LUMO designates space available for a pair of electrons, a LUMO map indicates where nucleophilic attack would likely occur. For example, a LUMO

map for cyclohexenone shows concentration in two regions, one over the carbonyl carbon and the other over the β carbon, consistent with both carbonyl addition and Michael (conjugate) addition.



The ***spin density map*** paints the value of the spin density onto an electron density surface. By convention, colors near blue indicate high concentration of spin density, while colors near red indicate low concentration. For example, a spin density map for the radical resulting from loss of hydrogen from 3,5-di-*tert*-butylhydroxytoluene (BHT) shows that the spin has delocalized from oxygen onto the *ortho* and *para* ring positions.



This radical would be expected to be particularly stable, which explains why BHT acts as an antioxidant (scavenging less favorable localized radicals).

All of these maps use an electron density surface and to delineate overall molecular size and shape. Not all regions on this surface are accessible and therefore available for interaction with their environment (or with an incoming reagent). ***Spartan*** allows these regions to be identified.*

* A region on a density surface is designated as inaccessible if a sphere of radius 1.0 Å centered on a line normal to the surface and touching a point in the middle of the region, impinges on any other regions of the density surface. The sphere radius may be changed in the **Settings Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 18**).

Surfaces (including those underlying maps) connect points of equal value (they are isosurfaces), and may be displayed as an arrangement of dots, a mesh, or an opaque or translucent solid. Examples of graphical output in orthogonal projection are provided in **Figure 15-1**. Surfaces (and maps) may also be rendered in perspective (see **Chapter 12**) and in stereo (see **Chapter 2**).

Calculated quantities may also be displayed as two dimensional contour plots (slices). Unlike surfaces and maps, these can be translated, rotated and zoomed independently of the molecular skeleton. An example of a slice display is provided in **Figure 15-1**.

Several different surfaces, maps and slices may be simultaneously displayed. In addition, any of the usual structure models may be displayed along with the graphic. The total display can become very complex, and selective use of meshes and/or translucent solids (as opposed to opaque solids) may facilitate visualization.

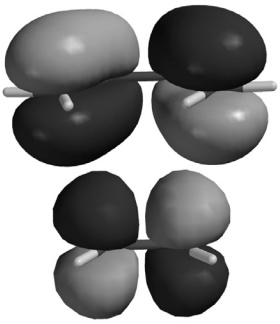
Discussion of the utility of graphical models for describing molecular structure and chemical reactivity is provided in ***A Guide to Molecular Mechanics and Quantum Chemical Calculations***, provided as a PDF under the **Help** menu.

Selection of **Surfaces** from the **Setup** menu results in display of the **Surfaces** dialog.

The box at the top of the dialog identifies graphical surfaces for later calculation and display. The **Add...** button at the bottom of the dialog is used to specify new graphical surfaces. *Clicking* on it leads to the **Add Surfaces** dialog that contains three menus and a check box:

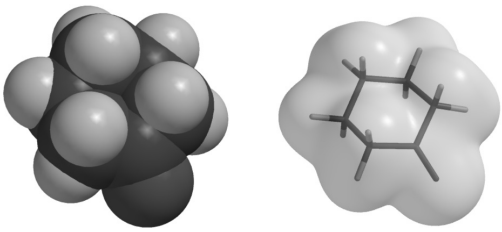
Figure 15-1: Examples of Graphical Displays Available in *Spartan*

Frontier orbitals for a symmetry-allowed Diels-Alder reaction,



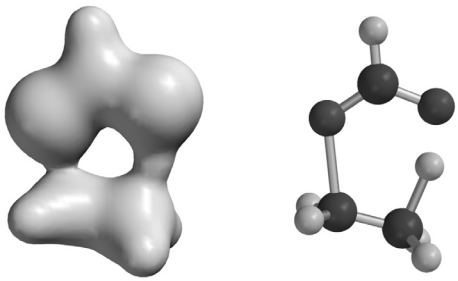
showing interaction of the HOMO of 1,3-butadiene and the LUMO of ethylene.

Space-filling model and electron density surface of cyclohexanone,



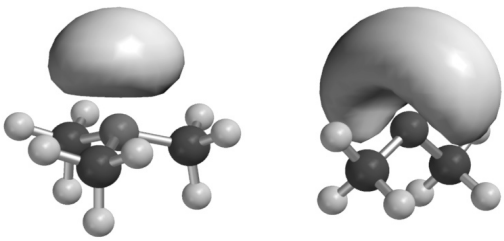
showing overall molecular size and shape.

Electron density surface (0.08 electrons/ au^3) of transition structure for pyrolysis of ethyl formate,



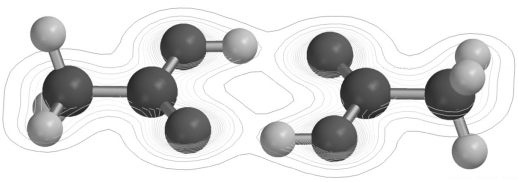
showing bonding in the transition state.

Electrostatic potential surfaces (-40 kJ/mol) of trimethylamine (left) and dimethyl ether (right),



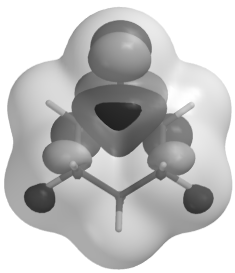
showing the lone pairs on nitrogen and oxygen, respectively.

Electron density slice for acetic acid dimer,

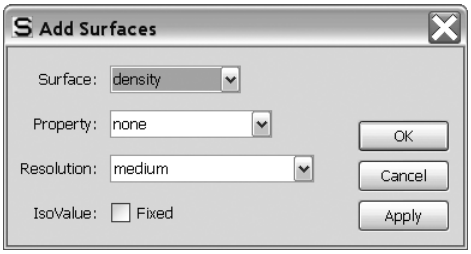


showing hydrogen bonding.

Simultaneous display of the LUMO and the electron density surfaces of cyclohexanone,

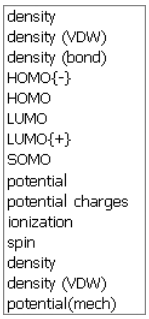


showing accessibility for nucleophilic attack.



Surface

Available surface types appear under the **Surface** menu.



Density portrays overall molecular size, ***density (bond)*** locates bonds, ***density (VDW)*** approximates ***density*** with van der Waals spheres, ***HOMO{-}, HOMO, LUMO, LUMO{+}, SOMO**** are molecular orbitals, ***potential*** is the electrostatic potential, ***ionization*** is the local ionization potential and ***spin**** is the spin density.

> Unlike the molecular orbitals, the electron density and the electrostatic potential, the local ionization potential does not go to zero with increasing distance. This makes its use as a surface problematic.

Selection of **HOMO{-}** and **LUMO{+}** results in display of a box to decrement the HOMO and increment the LUMO. This allows any molecular orbital to be specified..

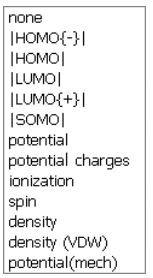
Slice designates that a plane will cut through the graphic defined by **Property** (see discussion following).

---

* These menu entries appear only for molecules with unpaired electrons.

Property

Properties for maps appear in the **Property** menu.

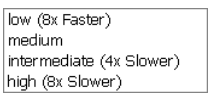


Available properties are the molecular orbitals (**HOMO{-}**, **HOMO**, **LUMO**, **LUMO{+}**, **SOMO***), the electrostatic potential (**potential**), an approximate electrostatic potential based on (electrostatic) charges (**potential (charges)**), the local ionization potential (**ionization**) and the spin density (**spin**)*. **none** indicates that no property is to be mapped onto the surface). As with **Surface** above, selection of **HOMO{-}** and **LUMO{+}** leads to a decrement (increment) box.

A **Spin** button will be displayed if **Multiplicity** (in the **Calculations** dialog) is set to a value other than **Singlet**, or if **Scf=unrestricted** has been specified in the **Options** box (**Calculations** dialog; see also **Appendix C**) and if **HOMO{-}**, **HOMO**, **LUMO** or **LUMO{+}** has been selected for **Surface** or for **Property**. *Clicking* on **Spin** toggles it between **Alpha** and **Beta**. **Alpha** designates that the molecular orbital either to be displayed as a surface or mapped as a property corresponds to α spin; **Beta** designates that the molecular orbital corresponds to β spin.

Resolution

Selection of surface resolution is from the **Resolution** menu.



Medium resolution generally is intended for routine work,

* These menu entries appear only for molecules with unpaired electrons.

low resolution is used to get rough images very quickly, while intermediate and high resolution may be employed to obtain publication quality graphics. Both calculation time and disk storage increase significantly with increasing resolution.

IsoValue

This allows a fixed isosurface to be generated in place of a three-dimensional grid of points (from which different isosurfaces, and/or slices can be constructed). Fixed isosurfaces are particularly attractive for property maps (that are generally based on a density surface representing molecular size and shape) as they require less computer time and disk storage. Note that a fixed density surface needs to be based on a specific isovalue (see below) and not on percentage enclosure of the total number of electrons.

Checking the box to the left of **Fixed** leads to display of the isovalue (0.002 electrons/au^3 in the case of density). This value may be changed. **Fixed** may not be specified in conjunction with **Slice** as a **Surface**.

Following **Surface**, **Property**, **Resolution** and (optionally) spin selection, *clicking* on **OK** adds the requested graphic to the list and removes the (**Add Surfaces**) dialog. *Clicking* on **Apply** adds the requested graphic to the list but leaves the dialog on screen. *Clicking* on **Cancel** does not add a graphic to the list but removes the (**Add Surfaces**) dialog.

The process (*clicking* on **Add...**, surface and property selection, *clicking* on **OK** or **Apply**) may be repeated as required.

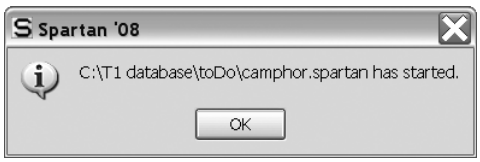
An existing graphic may be deleted from the list by first highlighting (*clicking* on) it and then *clicking* on **Delete**.

Global Surfaces, if *checked*, signifies that the requested surfaces will be calculated for all members of the list.

Only one copy of the **Surfaces** dialog may appear on screen, and any actions relate to the currently selected molecule. The dialog may be removed by *clicking* on ☒.

Submit

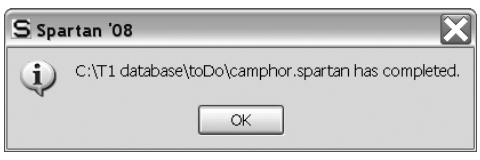
Following setup of a molecular mechanics or quantum chemical calculation, or a similarity analysis, as well as any requests for properties, spectra and/or graphical displays, the required calculations will begin when **Submit** is selected. If the job has not previously been named, selection of **Submit** triggers a request for a name. If the document being submitted comprises a single molecule and the molecule exists in the Spartan Molecular Database, the name in the database will be presented as a default name. Otherwise, the default name presented will be ***spartan*** for the first job and ***spartanX*** (X an integer starting with 1 for all successive jobs). After a name has been provided (or if a name already exists or if the default name is accepted) a dialog appears indicating that the job has actually been submitted.*



Click on **OK** to remove it. After a job has started, and until it has completed, all associated files will be designated read only.

In addition to the local (Windows, Macintosh or Linux) machine, ***Spartan's*** computational codes may run on a remote (licensed) Linux server. The only exception is the similarity analysis code which must be run on the local machine. If remote submission has been set up (see **Hosts** from the **Preferences** dialog under the **Options** menu; **Chapter 18**), selection of **Submit** will also request a name for the destination machine.

Another dialog appears following completion of a calculation.



Click on **OK** to remove it.

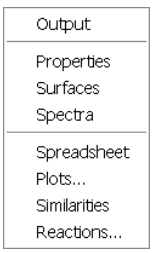
* The job is submitted to a job queue and will be submitted for execution only when released from this queue. See **Monitor** under the **Options** menu (**Chapter 18**) for discussion.

Upon completion, conformer distribution and energy profile jobs, as well as jobs that request generation of an intrinsic reaction coordinate, lead to an additional document being created for each molecule in the original document. These new documents are named ***document.task.identifier*** where ***document*** is the name given to the original document, ***Conf*** for a conformer distribution, ***Prof*** for energy profile and ***IRC*** for an intrinsic reaction coordinate and ***identifier*** identifies the molecule inside the original document.

A query dialog is provided asking whether the resulting ***Spartan*** documents are to be opened.

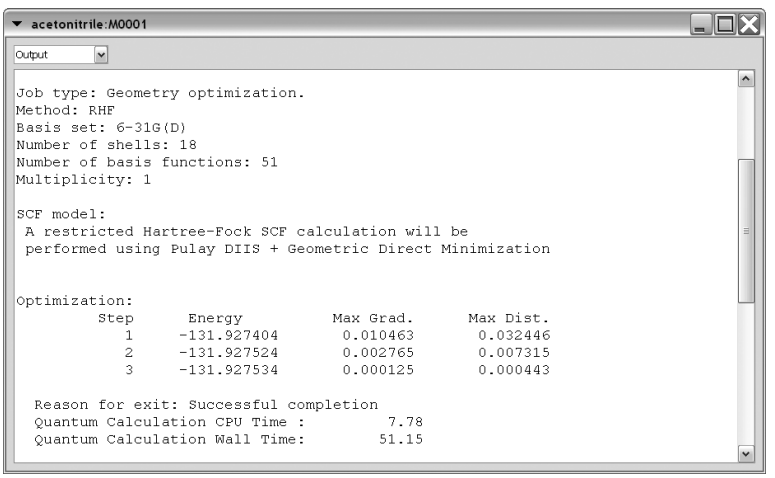
Chapter 16
The Display Menu

*Functions available under the **Display** menu provide for text, dialog, spreadsheet and graphical displays following structure/ property calculation, conformational analysis or similarity analysis. Functions are also available to query a variety of on-screen objects, display and compare IR, NMR and UV/visible spectra, animate vibrational motions, perform linear regression and prepare plots from spreadsheet data, display similarities between molecular structures and between molecular structures and pharmacophores and calculate reaction energies.*

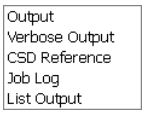


Output

Selection of **Output** opens a window.



A menu at the top left of the window selects the type of output for display.



Output provides normal output. **Verbose Output** contains more detailed output, but will be eliminated upon normal completion unless **Keep Verbose** has been checked in the **Setting Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 18**). **CSD Reference** provides the literature reference for data in the Cambridge Structural Database (CSD). **Job Log** and **List Output** contain diagnostic information in the event of a job failure.

The contents of the output window may be scrolled in the usual manner and may be paged up or down by *clicking* above or below the scroll bar. The contents may be printed by selecting **Print Output...** from the **File** menu (this replaces **Print...** when an output window is selected), or copied to the clipboard by selecting **Copy** from the **Edit** menu when an output is selected. **Find...** and **Find Next** functions from the **Edit** menu are also available.

A single output window is associated with each document, and changes focus as different molecules from the document are selected. As many output windows as desired may be simultaneously open on screen. An output window may be closed by *clicking* on ⊠ at the top right.

> Except for CSD data, output from jobs that are currently executing or are in the execution queue, is unavailable. Output for jobs that execute locally may be viewed using the **Monitor** under the **Options** menu (**Chapter 18**).

Properties

Spartan provides specialized dialogs for reporting and (in some cases) changing the properties of molecules, atoms, bonds, graphical surfaces and geometrical constraints. They also allow for changing default plot styles and limits, as well as for altering the nature of fitting functions. Finally, they allow editing CFD's and substituent libraries. Only one

Properties dialog is permitted on screen, and this dialog refers either to the selected molecule (**Molecule Properties**), or to the selected component (atom, bond, etc.) or attribute (spectra, graphical surface, substituent, etc.) of the selected molecule (**Atom Properties**, **Bond Properties**, **Surface Properties**, etc.), or to a plot constructed from data in a spreadsheet (**Plot Properties**, **Curve Properties**, etc.), or to a fitting function (**Regression Properties**). Selection of a different molecule (with a **Properties** dialog on screen) leads to the **Molecule Properties** dialog for the newly selected molecule. Dialogs that refer to components/attributes of the (newly selected) molecule follow by *clicking* on the component/attribute.

Dialog selection operates in toggle mode, with the **Molecule Properties** dialog being in the "neutral" position. In this position (with the **Molecule Properties** dialog on screen), *clicking* on a component/attribute (of the currently selected molecule) brings up the appropriate **Properties** dialog. For example, *clicking* on an atom brings up the **Atom Properties** dialog. *Clicking* on a different component or attribute brings up the **Properties** dialog for this component or attribute. However, *clicking* a second time on the same component reverts back to the **Molecule Properties** dialog.*

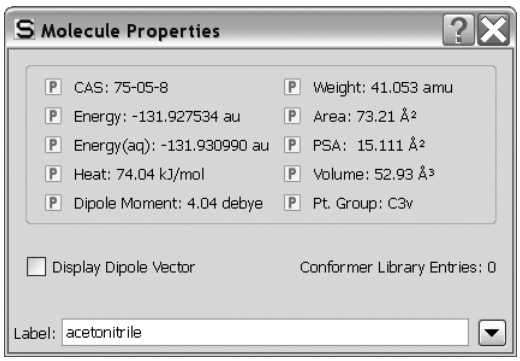
Most **Properties** dialogs have an associated **Utilities** or **Style** dialog. For example, associated with the **Molecule Properties** dialog is a **Molecule Utilities** dialog. These access additional information about the molecule and its components/attributes, or provide style and color controls unique to the selected molecule, component or attribute. This is useful for highlighting (or dehighlighting) a particular molecule, component or attribute. **Utilities**/**Style** dialogs are reached by *clicking* on ▼ at the bottom right of the appropriate **Properties** dialog. Return to the **Properties** dialog follows from *clicking* on ▲ at the bottom right of the associated **Utilities**/**Style** dialog. *Clicking* on a new molecule with a **Utilities**/**Style** dialog on screen leads to the **Molecule Properties** dialog for that new molecule.

---

* The exception to this involves *clicking* on a property map to obtain the value of the property at a particular surface location. *Clicking* a second time on a new location will report a new value of the property. *Clicking* on the background leads to the **Molecule Properties** dialog.

The **Properties** (or **Utilities/Style**) dialog may be removed from the screen by *clicking* on ☒.

Molecule Properties



The **Molecule Properties** dialog displays the energy (in au) of a Hartree-Fock, density functional, Møller-Plesset or other correlated calculation, the heat of formation (in kJ/mol) of a semi-empirical calculation or thermochemical recipe, or the sum of the strain energy and non-bonded energy (in kJ/mol) of a molecular mechanics calculation, the energy (heat of formation, strain energy) corrected for an aqueous environment (in the same units as the non-corrected quantity), dipole moment (in debyes), molecular weight (in amu), surface area, polar surface area* and volume of a space-filling model** (in $Å^2$ and $Å^3$, respectively), the point group and (if available) the experimental heat of formation (in kJ/mol) and the CAS number. These may be posted to the spreadsheet using the P buttons to the left of their values. Also provided is the number of conformers generated for the molecule using **Conformer Library**. **Label** (also referred to as identifier) identifies the molecule in a document. It appears in the first column of the spreadsheet (see **Spreadsheet** later in this chapter) and is the name given to the molecule in the Spartan Molecular Database.***

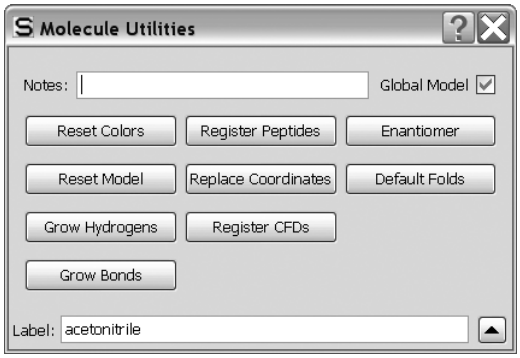
* By default, polar surface area is defined as the area due to nitrogen and oxygen and any hydrogens attached to nitrogen and oxygen. Polar surface areas corresponding to arbitrary alternative definitions are available for posting into the spreadsheet using the **PAREA** function. See **Table 16-3**.

** Area and volume calculations are not performed for molecules with more than 500 atoms.

*** This assumes that the molecule exists in the Spartan Molecular Database. If it does not, the name will be ***M000X*** (X=1,2,...) unless altered by the user.

Finally, the dialog permits display of the dipole moment vector by *checking* the box to the right of **Display Dipole Vector**.

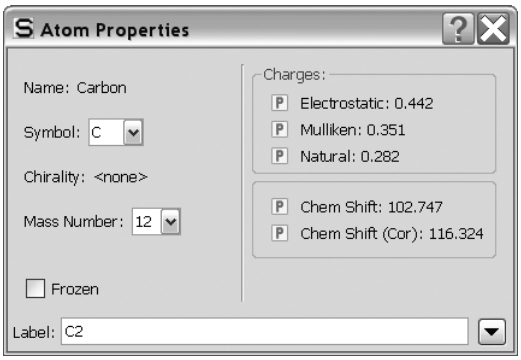
Clicking on ▼ at the bottom right of the **Molecule Properties** dialog brings up the **Molecule Utilities** dialog (*clicking* on ▲ at the bottom right of this dialog returns to the **Molecule Properties** dialog).



Notes is user-supplied text that will be reproduced in the output. Controls allow resetting model color and style, adding missing hydrogens and bonds, providing information about amino acids in polypeptides and about CFD's, replacing coordinates by standards based only on atomic connectivities, changing enantiomers and resetting default conformer selections.

Atom Properties

Selection of an atom with a **Properties** dialog on screen, or selection of **Properties** following selection of an atom, leads to the **Atom Properties** dialog.



This displays the element name (and allows changing the element), R/S chirality, Mulliken, electrostatic-fit and natural charges (in electrons), calculated and (for the B3LYP/6-31G* model) corrected NMR ^{13}C chemical shifts (in ppm relative to tetramethylsilane). It also allows freezing the atom (see **Freeze Center** in **Chapter 13**), changing its mass number and the default label, and posting atomic charges and calculated and corrected chemical shifts to the spreadsheet.

Clicking on ▼ at the bottom right of the **Atom Properties** dialog brings up the **Atom Style** dialog (not shown). This contains controls for selecting atom color and changing model style. *Clicking* on ▲ at the bottom right of this dialog returns to the **Atom Properties** dialog.

Bond Properties

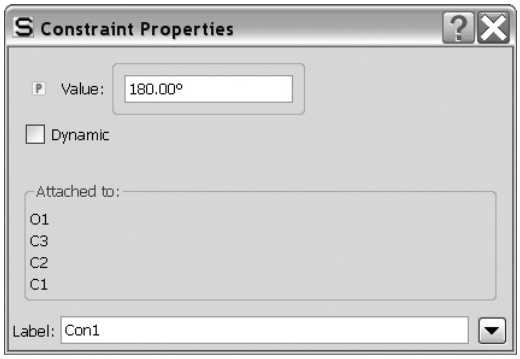
Selection of a bond with a **Properties** dialog on screen, or selection of **Properties** following selection of a bond leads to the **Bond Properties** dialog.

This displays the bond length and bond type (and allows changing the bond type).

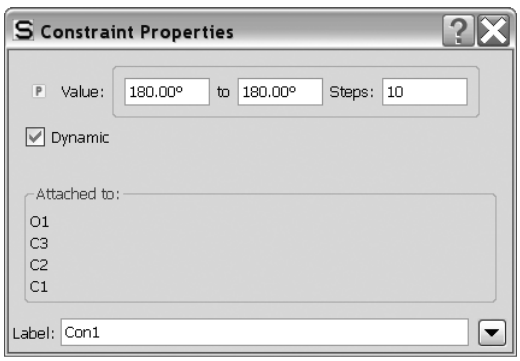
Clicking on ▼ at the bottom right of the **Bond Properties** dialog brings up the **Bond Style** dialog (not shown). This contains controls for changing bond color. *Clicking* on ▲ at the bottom right of this dialog returns to the **Bond Properties** dialog.

Constraint Properties

Selection of a constraint marker with a **Properties** dialog on screen, or selection of **Properties** following selection of a constraint marker, leads to the **Constraint Properties** dialog.



This allows setting the value of a constraint, posting it to the spreadsheet and changing the default constraint label. *Checking* **Dynamic** leads to an expanded dialog.



This allows specifying a sequence of constraints for an energy profile (see **Calculations...** under the **Setup** menu; **Chapter 15**). The value of the starting constraint is given in the lefthand box to the right of **Value**, and the value of the ending constraint is given in the righthand box to the right of **Value**. The number of steps in the profile is given in the box to the right of **Steps**. Initially, the numbers in both boxes to the right of **Value** will be the same, and **Steps** will be set to **1**. These may be altered by typing the desired numbers into the appropriate boxes and then *pressing* the **Enter** key (**return** key on Mac). An energy profile may involve

constraints on more than one geometrical variable (changed in lock-step with one another). In this case, the constraint ranges for the individual variables need to be selected, but the number of steps must and will be the same for each variable.

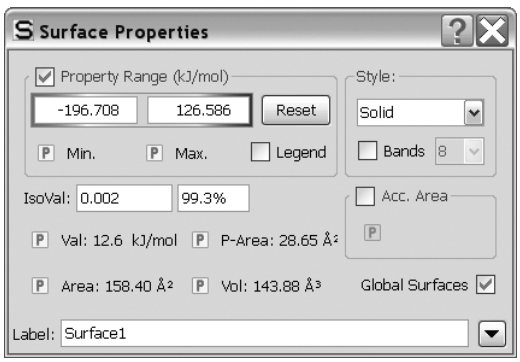
Clicking on ▼ at the bottom right of the **Constraint Properties** dialog brings up the **Constraint Style** dialog (not shown). This contains controls for selecting the color of the constraint marker. *Clicking* on ▲ at the bottom right of this dialog returns to the **Constraint Properties** dialog.

Point and Plane Properties

Selection of a user-defined point or plane with a **Properties** dialog on screen, or selection of **Properties** from the **Display** menu following selection of a point or plane, leads to the **Point Properties** or **Plane Properties** dialog (not shown). These allow changing point or plane labels. *Clicking* on ▼ at the bottom right of the **Point** (**Plane**) **Properties** dialog gives rise to the **Point** (**Plane**) **Style** dialogs (not shown). These contain controls for adjusting point and plane colors, respectively. *Clicking* on ▲ at the bottom right of this dialog returns to the **Point** or **Plane Properties** dialog.

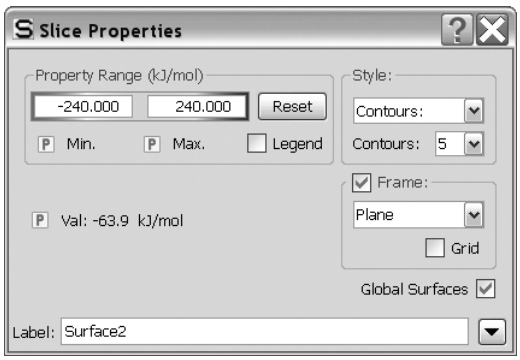
Surface Properties

Selection of a graphical display with a **Properties** dialog on screen, or selection of **Properties** following selection of a graphical display, leads to the **Surface Properties** dialog.



This allows changing display style, isovalue (and in the case of electron density surfaces, percentage of the electrons contained inside the surface), turning "on" mapped properties, selecting between continuous and banded displays and setting the range of the property in property maps, displaying accessible area of surfaces and maps and changing the default surface labels. The dialog also reports (and optionally posts to the spreadsheet) the area and volume of the graphic, the accessible area*, the polar area of an electrostatic potential map**, maximum and minimum value of the mapped property and its value at the cursor position***. If *checked*, **Legend** displays a color scale. If *checked*, **Global Surfaces** designates that the settings apply to all molecules in the document.

In the event that the selected graphical surface is a slice, a different dialog **Slice Properties** appears.



This contains similar controls to that found in the previous dialog, with variations to account for the number of contours displayed and to the type of slice (plane, cylinder or sphere). Check boxes allow for a frame around the slice and for a grid.

---

* A region on a density surface is designated as inaccessible if a sphere of radius 1.0 Å centered on a line normal to the surface and touching a point in the middle of the region, impinges on any other regions of the density surface. The default radius (**Accessible Area Radius**) may be changed in the **Settings Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 18**).

** This is defined as that part of the surface area for which the absolute value of the electrostatic potential is > 100 kJ/mol. The cutoff (**Polar Area Range**) may be changed in the **Settings Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 18**).

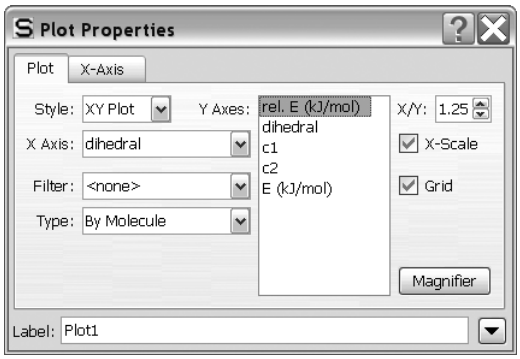
*** To determine property value at another position *click* on it. To bring up the **Molecule Properties** dialog, *click* on the background.

Clicking on ▼ at the bottom right of the **Surface Properties** dialog brings up the **Surface Style** dialog (not shown). This contains controls for selecting surface color. *Clicking* on ▲ at the bottom right of this dialog returns to the **Surface Properties** dialog.

Curve and Plot Properties

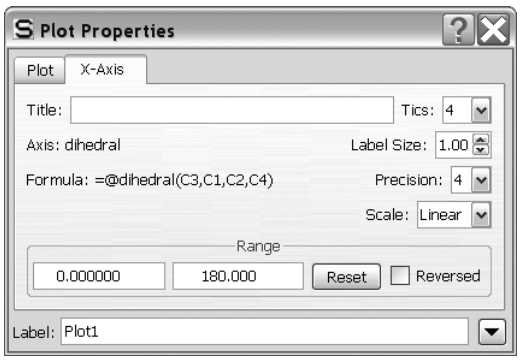
Plots created either using data from the spreadsheet or from spectra may be edited using a set of closely-related dialogs. These refer to "curves" (a relationship between variables), and "plots" (a set of relationships that share variables).

Selection of a plot frame with a **Properties** dialog on screen, or selection of **Properties** following selection of a plot frame, leads to the **Plot Properties** dialog (XY plot case shown).



The **X Axis** menu and the **Y Axes** list duplicate functionality under the **Plots** dialog (see **Plots...** later in this chapter) and allow specification of a molecular property to serve as the X axis of the plot and one or more properties to serve as the Y axes. The contents of the **Y-Axes** list depends on selection from the **Type** menu: **By Molecule** and **Free Form** reference explicit entries in the spreadsheet (see discussion later in this chapter) while **By Atom** references ***Spartan's*** file of atom-based properties (the ***Property Archive***). Initial selections may be altered as desired. In addition, the aspect ratio of the plot may be changed, the numerical scale of the X (horizontal) axis turned "on" or "off", a grid added to the plot, and the default label changed. **Magnifier** brings up a "magnifying glass" that may be moved independently.

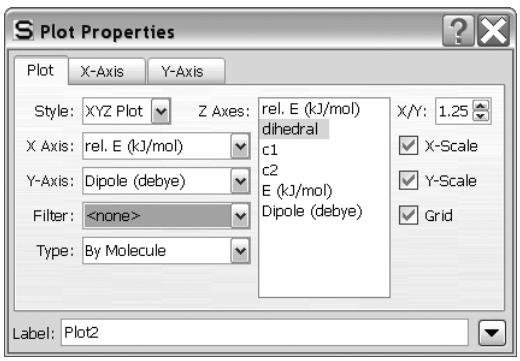
The scale along the X (horizontal) axis may be altered from default settings by *clicking* on the **X-Axis**. tab at the left of the dialog. This brings up the **Plot X-Axis Properties** dialog (*clicking* on the **Plot** tab returns to the **Plot Properties** dialog).



To alter the scale, edit the values inside the **Range** boxes, and *press* the **Enter** key (**return** key on Mac). Controls are also available to alter the number of tic marks on the X axis, the label size and precision and change the scale from linear to logarithmic.

Clicking on ▼ at the bottom right of either the **Plot Properties** or **Plot X-Axis Properties** dialog brings up the **Plot Style** dialog (not shown). This contains controls for selecting the color of the plot frame, as well as the label associated with the X axis. *Clicking* on ▲ at the bottom right of this dialog returns to either the **Plot Properties** or **Plot X-Axis Properties** dialog.

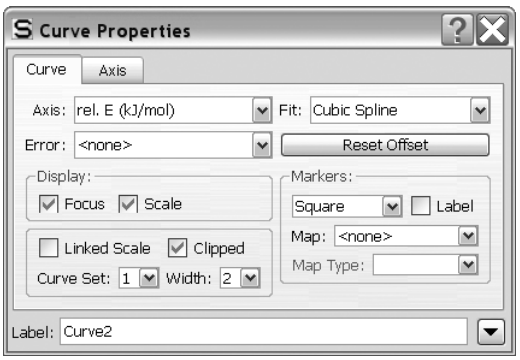
The **Style** menu in the **Plot Properties** dialog allows shifting between 2D (XY) and 3D (XYZ) plots. Selection of **XYZ Plot** leads to an extended **XYZ Plot Properties** dialog.



This is identical to the **XY Plot Properties** dialog, except that it includes an additional menu (**Y Axis**) from which one property needs to be selected. In addition, the **Y Axes** list in the **XY Plot Properties** dialog has been renamed to **Z Axes**.

Scales along X and Y axes may be adjusted by *clicking* on **X-Axis** and **Y-Axis** tabs to the left of the dialog. This leads to the **Plot X-Axis** (**Plot Y-Axis**) **Properties** dialog (*clicking* on the **Plot** tab at the left of this dialog returns to the **Plot Properties** dialog). These are analogous to the (**Plot X-Axis Properties**) dialog discussed for XY plots.

Selection of an individual curve in a plot with a **Properties** dialog on screen, or selection of **Properties** from the **Display** menu following selection of an individual curve, leads to the **Curve Properties** dialog.

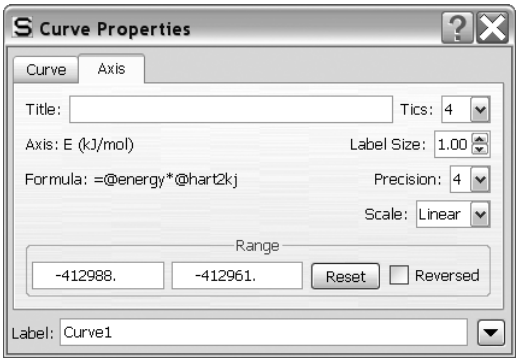


This allows changing the quantity that is plotted, providing a scale and error bars (down from the column in the spreadsheet designated in the **Error** menu) and turning "on" clipping. Controls are available to designate the kind of markers for the data points, request marker labels, and designate the kind of fit to the data points and the line width of the fitting curve*. If *checked*, **Focus** marks the selected molecule on the curve.

The scale along the Y axis (vertical axis) in the case of an XY plot (or Z axis in the case of a XYZ plot) may be altered from

* Fitting functions include linear, quadratic and cubic least square, a Fourier series, a Lorentzian or a Gaussian. The data points may also be connected by straight lines (point-to-point) or by a smooth curve (cubic spline) or be presented as a skyline. Finally, the fitting curve can be removed.

default settings (from the minimum to maximum value of the Y or Z variable) by *clicking* on the **Axis** tab at the left of the dialog. This brings up the **Curve Axis Properties** dialog.

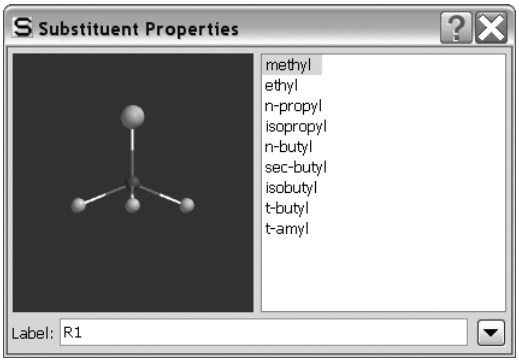


To alter the scale, type in new values inside the boxes underneath **Range**. You need to *press* the **Enter** key (**return** key on Mac) following each entry. Controls are also available to alter characteristics of labels associated with the Y (XY plot) or Z axis (XYZ plot): tics the number of tic marks on the axis label size and precision.

Clicking on ▼ at the bottom right of either the **Curve Properties** or **Curve Axis Properties** dialog brings up the **Curve Style** dialog (not shown). This contains controls for selecting the color of the curve (and the label associated with the Y (or Z) axis). *Clicking* on ▲ at the bottom right returns to the original dialog.

Substituent Properties

Selection of a substituent icon with a **Properties** dialog on screen, or selection of **Properties** following selection of a substituent icon, leads to the **Substituent Properties** dialog.



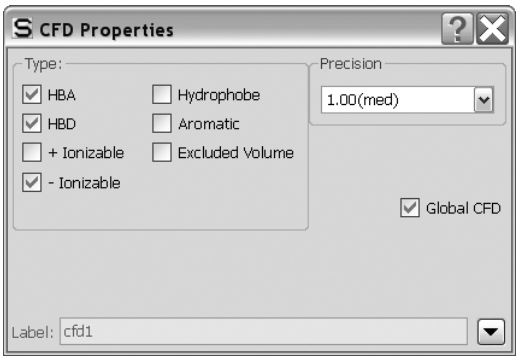
This allows substituent libraries already attached to a molecule to be edited, and the label on the substituent icon to be changed. It does not allow for attachment points to be reassigned. The dialog comprises a scroll box (listing the names of the substituents) on the right, a viewing window (showing a model of the substituent) on the left and a label field at the bottom.

To delete a substituent from the library, *click* on its name in the scroll box, then *right click* and select **Delete** from the menu that appears. **Ctrl** and **Shift** keys may be used in the usual way to remove several substituents at once. To add a new substituent that has been copied to the clipboard to the library, *right click* inside the scroll box and select **Paste** from the menu that appears. Note, that if the clipboard contains several molecules, all will be added to the substituent library. A substituent (or list of substituents) in an existing ***Spartan*** document can also be added by *right clicking* inside the scroll box, selecting **Append** from the menu that appears and choosing a document from the browser. Alternatively, *drag* the document that contains the substituent(s) from an external window into the scroll box.

Clicking on ▼ at the bottom right of the **Substituent Properties** dialog brings up the **Substituent Style** dialog (not shown). This contains controls for selecting substituent (icon) color. *Clicking* on ▲ at the bottom right of this dialog returns to the **Subtituent Properties** dialog.

CFD Properties

Selection of a chemical function descriptor (CFD) or of an element of a pharmacophore (which is functionally the same as a CFD) with a **Properties** dialog on screen, or selection of **Properties** (**Display** menu) following selection of a CFD or pharmacophore element, leads to the **CFD Properties** dialog.



This allows changing the definition of the CFD, changing the label on the CFD and the precision of the fit (by changing the radius of the CFD; the bigger the radius, the less precise the fit). If *checked*, **Global CFD** specifies that CFD definitions will apply to all molecules in the document.*

Clicking on ▼ at the bottom right of the **CFD Properties** dialog brings up the **CFD Style** dialog (not shown). This contains controls for selecting CFD color. *Clicking* on ▲ at the bottom right of this dialog returns to the **CFD Properties** dialog.

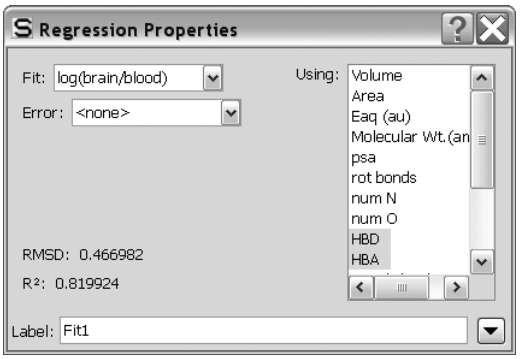
Regression Properties

Following a linear regression analysis (see discussion later in this chapter), a new row, labeled **Fit1****, appears near the bottom of the spreadsheet. This contains information about the fit. *Clicking* on this line with a **Properties** dialog on screen, or selecting **Properties** (**Display** menu) after *clicking* on the line, leads to the **Regression Properties** dialog.

---

* This is generally applicable only where the molecules are conformers, or the CFD labels have been adjusted to associate positions on different molecules. Label reassignment is accomplished using the **Atom Properties** dialog (see discussion earlier in this chapter).

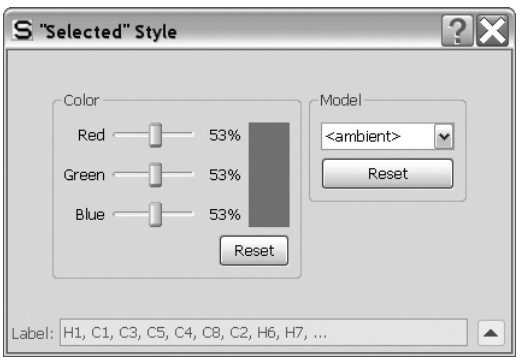
** More precisely, a row will be written for each fit, and labelled **Fit1**, **Fit2**, ...



This reports RMSD and R^2 error statistics related to the fit, as well as allows for changing what is to be fit (**Fit** menu) and what it is to be fit to (list to the right of **Using**). Following selection of either or both, the error statistics will update, allowing the user immediate feedback.

"Selected" Style

In addition to **Style** dialogs associated with several **Properties** dialogs, there is **"Selected" Style** dialog. This is accessed either by defining a selection box (see discussion under **Mouse and Keyboard Operations** in **Chapter 2**) with a **Properties** dialog on screen, or by selecting **Properties** from the **Display** menu after defining a selection box.



This dialog contains controls for changing the color and model type of whatever is included in the selection box.

Surfaces

This accesses the same dialog already described in **Chapter 15**.

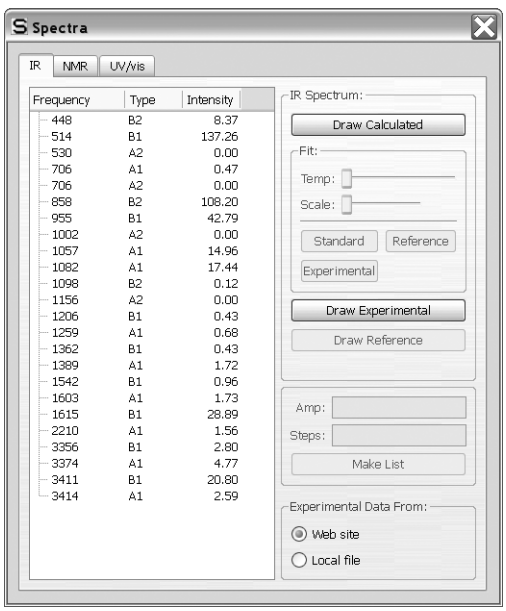
Spectra

Spartan provides for display of calculated IR, NMR and UV/visible spectra. (All of these need to have been previously requested from the **Calculations** dialog under the **Setup** menu; **Chapter 15**.) In addition, it allows on-line access and display of experimental spectra from publicly available collections. This allows direct visual comparison of calculated and experimental spectra. Finally, the **IR Spectra** dialog allows fitting of a calculated spectrum to a measured spectrum, using a single linear scaling parameter and a peak width parameter to adjust the calculated spectrum.

Access to spectra display is via a series of tabbed dialogs (**IR**, **NMR** and **UV/vis**).

IR

Clicking on the **IR** tab leads to the **IR Spectra** dialog that not only provides for display of a calculated and (if available) experimental infrared spectra, but also allows for animated display of the vibrational modes as well as for generation of a sequence of structures along a vibrational coordinate.



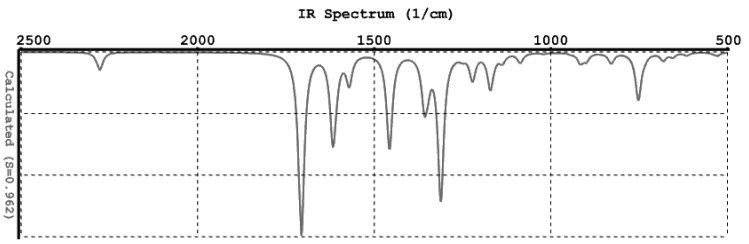
At the left-hand side of the dialog is an ascending list of frequencies (in cm^{-1})* together with (infrared) intensities and symmetry labels. A frequency is selected for display by *clicking* on it. This results in display (animation) of the vibrational motion. *Clicking* again deselects it (stops the animation).

The amplitude of vibrational motion (the maximum displacement away from equilibrium of any pair of atoms) may be changed from the default amplitude (0.5Å which is much larger than appropriate for vibrational motion at room temperature) by altering the contents of the box to the right of **Amp** at the lower right of the dialog. The default number of steps that make up the display (11) may be changed by altering the contents of the box to the right of **Steps**** at the lower right of the dialog. The greater the number of frames, the slower will be the vibration.

Clicking on **Make List** creates an unnamed ***Spartan*** document containing a sequence of structures corresponding to motion along the selected vibrational mode. Sequences (corresponding to different vibrations) can be created.

Note that the information required to produce vibrational motions (animations) and to construct lists of structures along different vibrational coordinates ***is not available*** as part of an entry in the Spartan Molecular Database (see discussion in **Chapter 17**). To conserve space, only frequencies and intensities are provided.

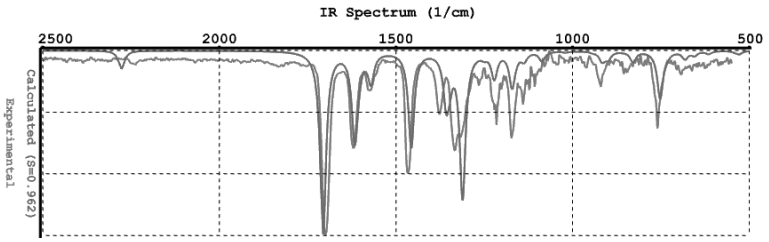
The calculated spectrum may be drawn by *clicking* on **Draw Calculated** at the top right of the dialog.



* Imaginary frequencies, for example, corresponding to the reaction coordinate for a transition state, will appear first in the list, and will be designated by the letter **i** in front of the frequency value.

** It is recommended that the number of steps be odd ensuring that the center point corresponds to the actual equilibrium or transition-state structure.

The experimental spectrum (if available) may be drawn with or without the calculated spectrum by *clicking* on **Draw Experimental** at the center right of the dialog.



Finally, a reference spectrum if available (see discussion of the Spartan Infrared Database in **Chapter 17**) may be displayed along with the calculated and/or experimental spectra or both.

Draw operations for both calculated and experimental IR spectra, apply to all molecules in the document (not just the selected molecule). Note, however, that the reference spectrum is common to all molecules in the document.

Experimental IR spectra are accessed from the NIST (National Institute of Standards and Technology) website (http://webbook.nist.gov/chemistry). This comprises approximately 7,000 IR spectra, primarily for organic molecules. ***Spartan*** keeps an index to facilitate retrieval but the actual data are on the website. Alternatively, the user may supply a spectra file (see **Appendix I**). Selection (website vs. file) is made at the bottom right of the dialog.

The default range for calculated, experimental and reference IR spectra is from 4000 cm^{-1} to 500 cm^{-1}. This corresponds to the range commonly measured and reported. This range may be changed by selecting **Properties** from the **Display** menu, *clicking* on the horizontal axis to bring up the **Plot Properties** dialog, *clicking* on the **X Axis.** tab at the left of the dialog and editing the values under **Range**. The spectrum is treated like any other graphical object, and may be translated and scaled using the appropriate mouse operations. It may not be rotated.

The calculated spectrum may be removed by *clicking* on **Delete Calculated** (that has replaced **Draw Calculated**) in the dialog.

Similarly, the experimental spectrum may be removed by *clicking* on **Delete Experimental** (that has replaced **Draw Experimental**) and the reference spectrum may be removed by *clicking* on **Delete Reference** (that has replaced **Draw Reference**). The last spectrum (calculated, experimental or reference) removed will also cause removal of the plot axes. These operations apply to all molecules in the list (not just the selected molecule). An individual (calculated or experimental) spectrum or entire plot may also be removed by selecting **Delete** from the **Build** menu and then *clicking* on a spectrum or on the plot axes, respectively. Holding down the **Delete** key while *clicking* on the spectrum or plot axes has the same effect.

Calculated IR frequencies often exhibit systematic errors. For example, frequencies obtained from (limiting) Hartree-Fock models are typically 10-12% larger than measured frequencies, while frequencies obtained from density functional and MP2 models are typically 5-6% larger than measured frequencies. Systematic errors may be revealed (and quantified) by first simultaneously displaying calculated and experimental spectra and then scaling one of the spectra to provide a best visual fit. To scale the calculated spectrum, use the slider bar to the right of **Scale** near the top right of the dialog. This is limited in range (from 0.8 to 1.0). If it is necessary to scale outside this range, *click* on the spectra (not on an axis) to select and move the mouse up and down while simultaneously depressing the **Shift** and **Alt** keys in addition to the right button. The scale factor (multiplying the frequency) will be displayed on the vertical axis. Scaling applies uniformly to all molecules in the list (not just the selected molecule).

A scaling factor may be applied to frequencies, and to all properties that make use of frequencies, prior to reporting using the **FreqScale** keyword typed into the **Options** box in the **Calculations** dialog (**Calculations...** under the **Setup** menu, **Chapter 15**).

A second slider bar marked **Temp** controls peak width. (This is loosely connected with the temperature at which the experimental measurement is carried out.) Low "temperature" (slider to the left) will produce sharp spikes, whereas high "temperatures" (slider to the right) will produce broad bands. The default setting is intended to give calculated spectra that broadly resemble experimental infrared spectra.

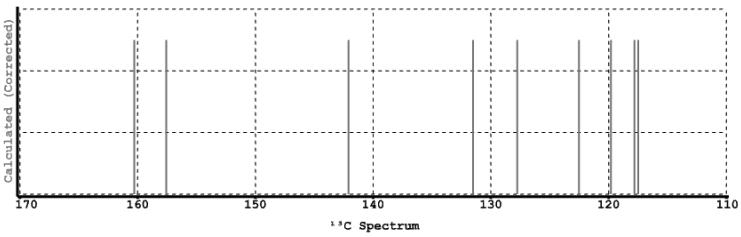
NMR

Clicking on the **NMR** tab leads to the **NMR Spectra** dialog that provides for display of both calculated and (if available) experimental ^{1}H and ^{13}C NMR spectra.*

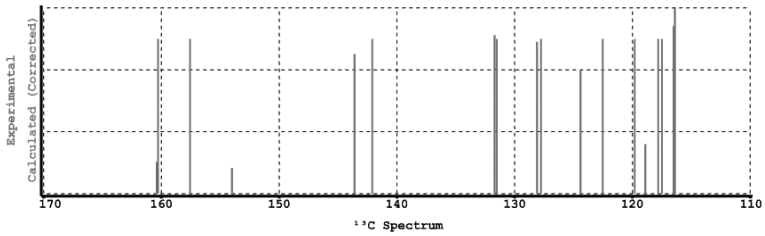
Calculated proton and ^{13}C spectra may be drawn by *clicking* on their respective **Draw Calculated** buttons. If **Use Corrected Shifts** is checked, ^{13}C NMR spectra will be based on chemical shifts from the B3LYP/6-31G* model (only) that have been empirically corrected to account for local environment.

---

* At present, only chemical shifts are calculated. Coupling constants are estimated empirically and relative intensities are assumed to be in proportion to the number of equivalent nuclei. Chemical shifts for nuclei other than hydrogen and carbon may be displayed as labels (see **Configure...** under the **Model** menu; **Chapter 12**).



The corresponding experimental ^{13}C spectrum (if available) may be drawn together with or independently of the calculated spectra by *clicking* on **Draw Experimental**. Draw operations for both calculated and experimental NMR spectra apply to all molecules in the list.



Experimental ^{13}C spectra are accessed from http://nmrshiftdb.ice.mpg.de. This comprises approximately 15,000 experimental ^{13}C spectra, all for organic molecules. ***Spartan*** keeps an index to facilitate retrieval but the actual data are on the website. A suitable collection of experimental proton spectra is not presently available on-line. Alternatively, the user may supply a spectra file (**Appendix I**). Selection (website vs. file) is made at the bottom right of the dialog.

Proton spectra may be drawn to reflect spin-spin coupling due to neighboring nuclei. Coupling constants are not calculated, but are estimated empirically. Note that the resulting spectra may become very complicated, and a magnifier is available under the **Plot Properties** dialog (see discussion earlier in this chapter). Finally, a COSY plot may be constructed.

The default range for both calculated and experimental proton spectra is from 0 ppm to 10 ppm (relative to TMS), whereas the default range for ^{13}C spectra is from 0 ppm to 225 ppm (again relative to TMS). These correspond to ranges typically measured

and reported. These ranges may be changed by selecting **Properties** from the **Display** menu, *clicking* on the horizontal axis to bring up the **Plot Properties** dialog, *clicking* on the **X Axis** tab at the left of the dialog and editing the values under **Range**. ^{1}H and ^{13}C spectral plots are treated like any other graphical object, and may be translated and scaled using the appropriate mouse operations. They may not be rotated.

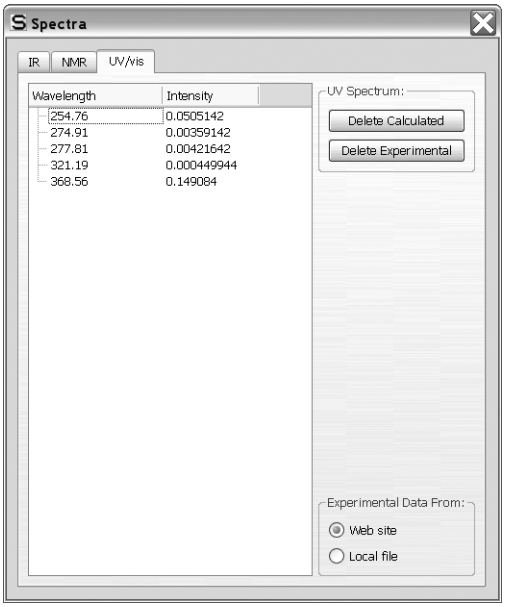
The calculated spectrum may be removed by *clicking* on **Delete Calculated** (that has replaced **Draw Calculated**) in the appropriate section of the dialog. Similarly, the experimental (^{13}C) spectrum may be removed by *clicking* on **Delete Experimental**). The last spectrum removed will also cause removal of the plot axes. These operations apply to all molecules in the document (not just the selected molecule). An individual (calculated or experimental) spectrum or entire plot may also be removed by selecting **Delete** from the **Build** menu and then *clicking* on the spectrum or on the plot axes, respectively. Holding down the **Delete** key while *clicking* on the spectrum or plot axes has the same effect.

To reveal and quantify systematic errors in chemical shifts, it may be useful to scale either the calculated and experimental spectra to provide a best visual fit. To scale, *click* on the spectra (not on an axis) to select and move the mouse up and down while simultaneously depressing the **Shift** and **Alt** keys in addition to the right button. The scale factor (multiplying the chemical shifts) will be displayed on the vertical axis. Scaling applies uniformly to all molecules in the document (not just the selected molecule).

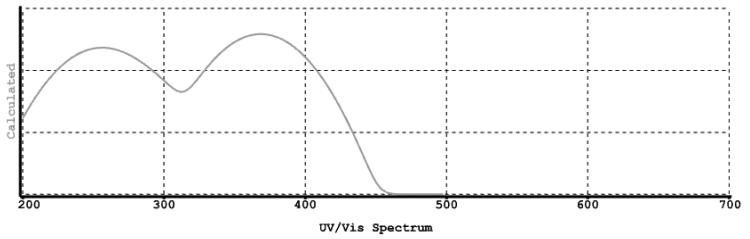
Only a single copy of one of the dialogs under **Spectra** may appear on screen; scaling will relate to the currently selected molecule. The dialog may be removed by *clicking* on ☒.

UV/vis

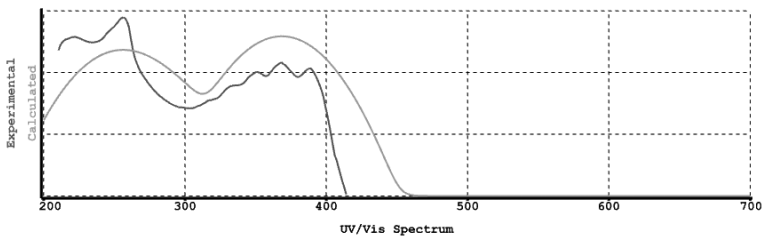
Clicking on the **UV/vis** tab leads to the **UV/vis Spectra** dialog that provides for display of both calculated and (if available) experimental UV/visible spectra.



This lists absorption wavelengths (from lowest to highest) together with intensities. The calculated UV/visible spectrum may be drawn by *clicking* on **Draw Calculated**.



The corresponding experimental spectrum (if available) may be drawn together with or independently of the calculated spectrum by *clicking* on **Draw Experimental**. Draw operations for both calculated and experimental UV/visible spectra and apply to all molecules in the list.



Experimental UV/visible spectra are accessed from the NIST (National Institute of Standards and Technology) website (http://webbook.nist.gov/chemistry). This comprises approximately 1,500 UV/visible spectra, primarily for organic molecules. ***Spartan*** keeps an index to facilitate retrieval but the actual data are on the website. Alternatively, the user may supply a spectra file (see **Appendix I**). Selection (website vs. file) is made at the bottom right of the dialog.

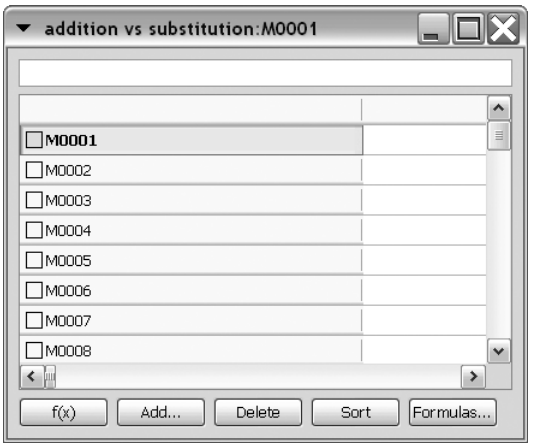
The default range for both calculated and experimental UV/visible spectra is from 200 nm to 700 nm. This corresponds to the range commonly measured and reported. This range may be changed by selecting **Properties** from the **Display** menu, *clicking* on the horizontal axis to bring up the **Plot Properties** dialog, *clicking* on the **X Axis** tab at the left of the dialog and editing the values under **Range**. The spectrum is treated like any other graphical object, and may be translated and scaled using the appropriate mouse operations.

The calculated spectrum may be removed by *clicking* on **Delete Calculated** (that has replaced **Draw Calculated**) in the dialog. Similarly, the experimental spectrum may be removed by *clicking* on **Delete Experimental** (that has replaced **Draw Experimental**). The last spectrum removed will also cause removal of the plot axes. These operations apply to all molecules in the document (not just the selected molecule). An individual (calculated or experimental) spectrum or entire plot may also be removed by selecting **Delete** from the **Build** menu (or holding down the **Delete** key) and then *clicking* on the spectrum or on the plot axes, respectively.

To help reveal similarities and differences between calculated and experimental UV/visible spectra, one or the other may be scaled. To scale, *click* on the spectra (not on an axis) to select and move the mouse up and down while simultaneously depressing the **Shift** and **Alt** keys in addition to the right button. The scale factor will be displayed on the vertical axis. Scaling applies uniformly to all molecules in the document (not just the selected molecule).

Spreadsheet

Associated with each ***Spartan*** document (including documents comprising a single molecule) is a spreadsheet. This may be displayed by selecting **Spreadsheet**.



The spreadsheet comprises a series of rows (corresponding to different molecules in the document) and columns (corresponding to different molecular properties). This gives rise to cells, the number of which is the product of the number of rows (molecules) and the number of columns (molecular properties). The spreadsheet may be expanded or contracted by positioning the cursor at one of the corners, *pressing* the left mouse button and *dragging* the mouse.

Only one molecule from one document may be selected (although several molecules may be simultaneously displayed). Molecule selection follows either by *clicking* on the spreadsheet cell containing the molecule label or identifier (leftmost column), or by using the ◂| and |▸ buttons or the scroll bar at the bottom left of the screen. Molecules may be ***animated*** (stepped through in succession) using the ▸ button at the bottom left of the screen. Animation speed may be adjusted from the **Settings Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 18**). Selection of a new molecule in the document results in deselection of the previously selected molecule. A molecule may be designated for permanent display by *checking* the box to the left of its identifier in the spreadsheet. The molecules in a document may either be translated and rotated in concert or

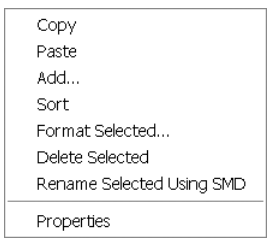
manipulated independently. This is controlled by **Coupled** under the **Model** menu (**Chapter 12**). By default (**Coupled** checked) molecules move in concert. *Uncheck* **Coupled** to move them independently.

Upon initial entry, all columns of the spreadsheet except the leftmost column, are blank. The leftmost column contains a label that may be changed either by directly typing a new label into the spreadsheet or into the **Label** box in the **Molecule Properties** dialog (see discussion earlier in this chapter). Additionally, default identifiers (***M0001, ...***) may be replaced by chemical names if the molecule exists in the Spartan Molecular Database (SMD) by *right clicking* inside the cell that contains the identifier and selecting **Rename Selected Using SMD** from the menu that results. To replace all identifiers, *right click* inside the header cell of the leftmost column and again select **Rename Selected Using SMD**.

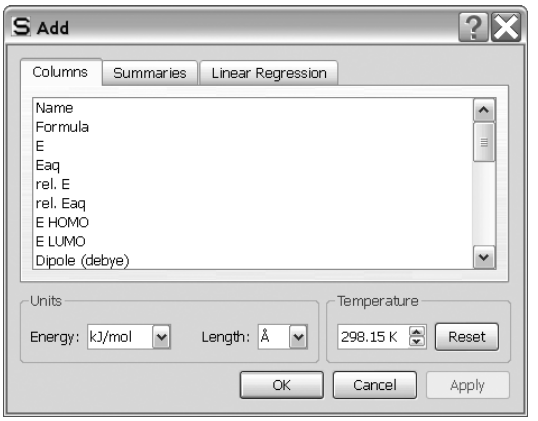
Information may be added to the spreadsheet in several different ways:

From the Add Dialog

A selection of molecular properties may be entered into the spreadsheet by first *clicking* on the header cell of an empty column, and then *clicking* on **Add...** at the bottom of the spreadsheet. Alternatively, *right click* inside the header cell and then to select **Add...** from the menu that results.



This leads to the **Add Columns** dialog.



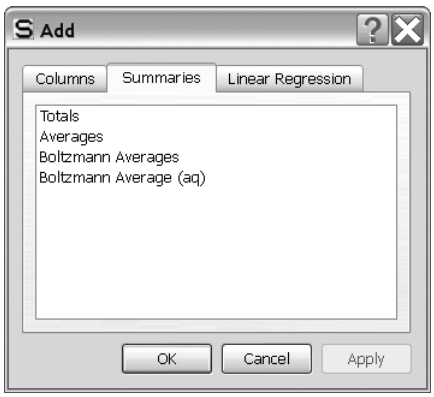
Name	molecule name as it appears in SMD
Formula	molecular formula
E	energy (heat of formation, strain energy)
E_{aq}	aqueous energy (heat of formation, strain energy) based on the SM5.4 model
rel. E	energy (heat of formation, strain energy) relative to selected molecule
rel. E_{aq}	aqueous energy (heat of formation, strain energy) relative to selected molecule based on the SM5.4 model
E HOMO	energy of highest-occupied molecular orbital
E LUMO	energy of lowest-occupied molecular orbital
Dipole	dipole moment (in debyes)
CPK Area	surface area of a space-filling model
CPK Volume	volume of a space-filling model
PSA	polar surface area (N,O + attached hydrogens)
Molecular Weight	molecular weight (in amu)
HBD Count	number of hydrogen-bond donors*
HBA Count	number of hydrogen-bond acceptors*
+ Ionizable Count	number of positive ionizable centers*
- Ionizable Count	number of negative ionizable centers*

* Counts of hydrogen-bond donors and acceptors and positive and negative ionizable sites derive from CFD assignments. These in turn follow from atomic connectivity together with special case designations for common organic functional groups. CFD's assignments may be modified on a case-by-case basis using the **CFD Properties** dialog (**Properties** under the **Display** menu; **Chapter 16**).

H°	enthalpy at room temperature
S°	entropy at room temperature
G°	Gibbs energy at room temperature
ZPE	zero-point energy
Boltzmann Distribution	Boltzmann distribution based on energy
Boltzmann Distribution (aq)	Boltzmann distribution based on aqueous energy as estimated from SM5.4 model
Alignment Scores	$1\text{-}R^2/N$, where R^2 is the root mean square distance and N is the number of alignment centers. 1 is a perfect score

One or more properties may be added to the spreadsheet by *clicking* on their entries, then specifying appropriate units from the **Energy** and **Length** menus, varying temperature from the **Temp** menu (as applied to the entropy and to the Boltzmann distributions), and finally *clicking* on **OK** or on **Apply**. In the former case, the dialog is dismissed and in the latter it is left on screen. *Clicking* on **Cancel** or ☒ removes the dialog without further affecting the contents of the spreadsheet. (Any previous additions to the spreadsheet as a result of *clicking* on **Apply** are maintained.)

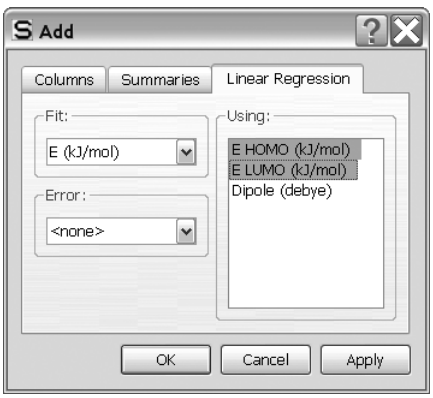
Column totals, averages and Boltzmann-weighted averages may be accessed by *clicking* on the **Summaries** tab. This leads to the **Add Summaries** dialog.



Totals	sum of column values
Averages	average of column values
Boltzmann Averages	Boltzmann weighted average of column values
Boltzmann Averages (aq)	Boltzmann weighted average of column values using energies corrected for aqueous environment based on the SM5.4 model

Clicking on one or more of these followed by *clicking* on **OK** or **Apply**, leads to the requested summaries as rows at the bottom of the spreadsheet, identified as **Totals**, **Averages**, etc. (one row for each summary). Note, that this dialog can be brought up directly by *clicking* on an empty row header (instead of an empty column header) inside the spreadsheet prior to *clicking* on **Add...**.

Finally, a linear regression analysis may be performed on the data in the spreadsheet. First *click* on the **Linear Regression** tab to bring up the **Add Linear Regression** dialog.



Select one entry from the **Fit** menu and one or more entries from the list under **Using**. *Clicking* on **OK** or **Apply** performs the linear regression analysis and places the results in a row at the bottom of the spreadsheet identified by **Fit**. As many regression analyses as desired may be performed on the data in the spreadsheet. The individual results will be entered as separate rows in the spreadsheet, with names **Fit1**, **Fit2**, etc. Additional information about the regression analyses is available from the **Regression Properties** dialog (see discussion earlier in this chapter).

Numerical Data

Numerical data may be entered by typing directly into the spreadsheet. A column header first needs to be specified. *Double click* on an empty column header cell, type in a name and *press* the **Enter** key (**return** key on Mac). Then, type the data into individual cells of the new column (*press* the **Enter** key following each data entry).

User-Defined Expressions

An expression may be entered either into a header cell (in which case it refers to all entries in a column) or into an individual cell (in which case it refers only to a single entry). Expressions in the column header take the form ***name=formula***, where ***formula*** is made up of arithmetic operations (**Table 16-1**), specialty functions (**Table 16-2**), calculated quantities (**Table 16-3**), conversion factors and constants (**Table 16-4**) in addition to numerical values. References to specialty functions, molecular mechanics and quantum chemical quantities and conversion factors and constants must be preceded by **@**. For example, ***mu = @DIPOLE*** typed into a header cell gives the dipole moment. Some functions have arguments, for example, ***c1*** and ***c2*** in the expression ***c12= @DISTANCE (c1,c2)*** refer to atoms c1 and c2, while ***3*** in the expression ***orbitalE=@HOMO (-3)*** designates the energy of the molecular orbital three orbitals below the HOMO. It is necessary to *press* the **Enter** key (**return** key on Mac) following entry of the expression into a cell. The leading ***name=*** is optional for entries in an individual (non-header) cell. Examples of user-defined expressions are provided in **Table 16-5**.

From Post (P) Buttons

Post buttons **(P)** found in a number of properties dialogs provide an alternative method to the **Add** dialog for entering calculated properties into the spreadsheet. Note that some properties may require user specification. These include individual bond

Table 16-1: Arithmetic and Boolean Operations and Mathematical Functions

arithemetic operations		boolean operations	
+	addition	>	greater than
_	subtraction	>=	greater than or equal to
*	multiplication	<	less than
/	division	<=	less than or equal to
^	raise to a power	==	equal to
		!=	not equal to
		|	or
		&	and

mathematical functions

ABS(x)	absolute value	LN(x)	natural logarithm
ACOS(x)	inverse cosine	LOG(x)	log (base 10)
ASIN(x)	inverse sine	SIN(x)	sine
ATAN(x)	inverse tangent	SQRT(x)	square root
COS(x)	cosine	TAN(x)	tangent
EXP(x)	exponential		

Table 16-2: Specialty Functions

AVG (column name)	average of values in column
FITVAL (fit name)	column of fit values from regression analysis
MIN (column name)	minimum of values in column
MAX (column name)	maximum of values in column
NUM (column name)	number of defined entries in column
ROW	the number of the row in the spreadsheet
ROW(molecule name)	the number of the row of molecule
REF(i, x)	the value of the x referenced to row i
STDEV (column name)	standard deviation of values in column
SUM (column name)	sum of values in column

Table 16-3: Calculated Quantities

ANGLE(i, j, k)	angle involving atoms i, j, k (degrees)
AREA	area of a user-defined plane ($Å^2$)
DIHEDRAL(i, j, k, l)	dihedral angle involving atoms i, j, k, l (degrees)
DISTANCE(i, j)	distance involving atoms i, j (Å)
ELECTRONEGATIVITY	QSAR descriptor -(HOMO+LUMO)/2 (eV)
ELECTROSTATIC (i)	electrostatic charge on atom i (electrons)
HARDNESS	QSAR descriptor -(HOMO-LUMO)/2 (eV)
HOMO(-n)	energy of n^{th} orbital below the HOMO (eV)
HOMOBETA(-n)	energy of the n^{th} orbital below the β HOMO (eV)
HYPERPOLARIZABILITY	beta hyperpolarizability parameter
INERTIA(i)	principle movements of inertia from largest (i=1) to smallest (i=3)
ISOTOPE(i)	mass number of atom i
LENGTH (i)	length of bond i (Å)
LOGPC	LogP from Crippen model
LOGPV	LogP from Villar model
LUMO(+n)	energy of the n^{th} orbital above the LUMO (eV)
LUMOBETA(+n)	energy of the n^{th} orbital above the β LUMO (eV)
MULLIKEN(i)	Mulliken charge on atom i (electrons)
NATURAL(i)	natural charge on atom i (electrons)
OVALITY	QSAR descriptor area/$[4\pi(3\bullet volume/4\pi)^{2/3}]$
PAREA (i, j, . . .)	partial surface area of a space-filling model due to all atoms of type i, j, ..., where i, j, ... are either atomic numbers or elemental symbols. i:, j:, ... signifies that attached hydrogens are also included (the surface area due to nitrogen, oxygen and any bonded hydrogen is commonly referred to as the polar surface area) ($Å^2$)
POLARIZABILITY	alpha polarizability parameter

Table 16-4: Conversion Factors and Constants

ANGS2AU	Ångstroms to atomic units
AU2ANGS	atomic units to Ångstroms
EV2HART	eV to atomic units (hartrees)
EV2KCAL	eV to kcal/mol
EV2KJ	eV to kJ/mol
HART2KCAL	atomic units (hartrees) to kcal/mol
HART2EV	atomic units (hartrees) to eV
HART2KJ	atomic units (hartrees) to kJ/mol
KCAL2EV	kcal/mol to eV
KCAL2HART	kcal/mol to atomic units (hartrees)
KCAL2KJ	kcal/mol to kJ/mol
KJ2EV	kJ/mol to eV
KJ2HART	kJ/mol to atomic units (hartrees)
KJ2KCAL	kJ/mol to kcal/mol
PI	π

Table 16-5: Examples of User Defined Expressions

E/area = @ENERGY/@AREA	energy divided by surface area
RelE = @ENERGY-@REF (6,@ENERGY)	energy relative to energy of molecule in row 6
Eq = @EXP (-@ENERGY/592.1)	equilibrium constant at room temperature
EnergyFilter = @ENERGY<-99.43	"true" (≠0) for all energies <-99.43
RowFilter = @ROW>10	"true" (≠0) all entries past row 10

distances, angles and dihedral angles (available from **Measure Distance**, **Measure Angle** and **Measure Dihedral** under the **Geometry** menu; **Chapter 13**), bond distance, angle and dihedral angle constraints (available from **Constrain Distance**, **Constrain Angle** and **Constrain Dihedral** under the **Geometry** menu; see discussion in **Chapter 13**), atomic charges, chemical shifts and corrected chemical shifts (available from the **Atom Properties** dialog; this chapter), the accessible area of an electron density surface, the polar area of an electrostatic potential map, minimum and maximum property values on a map and the value of the property at a specific location on a property map (available from the **Surfaces Properties** dialog; this chapter). With the exception of the property value on a map, post generates an entire column. Where atom labels are involved, for example, in defining a specific distance, post can be expected to yield consistent results for all molecules in a document only where the molecules are closely related, for example, molecules resulting from a conformational search, or where labels have been explicitly reassigned*. The property value on a map is posted only for the selected molecule. Post buttons are also available for CAS numbers and experimental heats of formation contained in ***Spartan's*** internal database.

From the Clipboard

Properties of one or more molecules in a document may be copied into the clipboard and then pasted into individual spreadsheet cells. These include (but are not restricted to) bond distances, angles and dihedral angles (**Measure Distance**, **Measure Angle** and **Measure Dihedral** under the **Geometry** menu), bond distance, angle and dihedral angle constraints (**Constrain Distance**, **Constrain Angle** and **Constrain Dihedral** under the **Geometry** menu), atomic charges and chemical shifts (**Atom Properties** dialog), infrared frequencies and chemical shifts (**IR Spectra** and **NMR Spectra** dialogs, respectively) and the value of a property on a property map (**Surface Properties** dialog). To

* Label reassignment is accomplished using the **Atom Properties** dialog (see discussion earlier in this chapter).

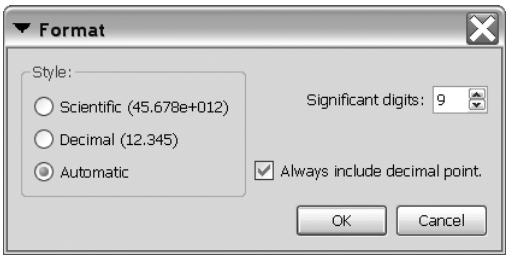
copy the spreadsheet, first highlight the numerical value of the property in the appropriate screen location (distances, etc.) or dialog (charges, etc.), then select **Copy** from the **Edit** menu, then *click* on the appropriate (destination) cell in the spreadsheet, and finally select **Paste** from the **Edit** menu.

Each row in a spreadsheet corresponds to a molecule in a document, and new rows are automatically added in response to adding new molecules to the document. New molecules are added by building (**New Molecule** under the **File** menu; **Chapter 10**), by appending one or more existing documents each containing one or more molecules using either **Append Molecule(s)...** under the **File** menu (**Chapter 10**), or by *right clicking* inside the header cell of the first available row and selecting **Append** from the menu that appears, by pasting from the clipboard, or by *dragging* from the file system. To copy a molecule into the clipboard, first select (*click* on) it, and then select **Copy** from the **Edit** menu, or *click* on its identifier (left most column) in its spreadsheet, and then select **Copy** from the **Edit** menu. Alternatively *right click* either on the molecule or on its identifier in the spreadsheet and select **Copy** from the menu that appears. Use of the clipboard permits several molecules to be selected (and copied) at once using the **Shift** and **Ctrl** keys in the usual manner. To copy the contents of the clipboard to its destination, *click* on an empty row header in the spreadsheet (for the destination document), and then select **Paste** from the **Edit** menu. An alternative to the two-step **Copy-Paste** procedure is to *drag* the molecule or set of molecules from one spreadsheet to another.

A row (molecule) may be deleted from a spreadsheet, either by first selecting the molecule and then selecting **Delete Molecule** from the **File** menu, or by first *clicking* on its identifier in the spreadsheet (leftmost column) and then either *clicking* on the **Delete** button at the bottom of the spreadsheet, or by *right clicking* on its identifier in the spreadsheet and then selecting **Delete Selected** from the menu that appears. In all cases, a warning is provided prior to deletion. An entire column in the spreadsheet may be deleted by first *clicking* inside its header cell and then *clicking* on the **Delete** button (or **Delete Selected** from the menu).

Rows in the spreadsheet may be sorted according to the numerical values in any column either by first *clicking* inside the header cell and then *clicking* on the **Sort** button at the bottom of the spreadsheet or by *right clicking* inside the header cell and selecting **Sort** from the menu that appears. The rows are placed in ascending order, the smallest (least positive) value of the selected property at the top, largest (most positive) value at the bottom. To sort in descending order, hold down the **Shift** key before *clicking* on the **Sort** button or selecting **Sort** from the menu.

Information in one or more columns of the spreadsheet may be formated by *right clicking* inside the header cell(s) and selecting **Format Selected** from the menu that appears.



Format as desired and *click* on **OK** to remove the dialog. The full contents of the spreadsheet may be formated by *right clicking* inside the header cell for the left most column and then selecting **Format Selected** from the menu.

A button at the bottom right of the spreadsheet toggles between numerical representation of data, **f(x)**, and formula presentation, **=?**.

The spreadsheet may be printed by selecting **Print Spreadsheet...** from the **File** menu (that replaces **Print...** when a spreadsheet is selected).

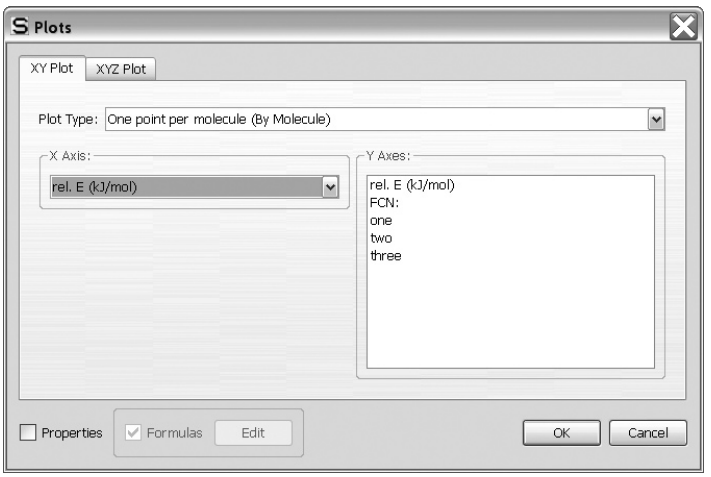
As many spreadsheets as desired (corresponding to the same or to different documents) may be open on screen. A spreadsheet is removed when the associated document is closed and may also be removed by *clicking* on ⊠.

The contents of the spreadsheet may be brought into Excel™ using the clipboard. Select whatever cells are to be copied, select **Copy** from the **Edit** menu. Alternatively, *right click* with the proper cells selected and select **Copy** from the menu that appears. **Paste** into Excel.

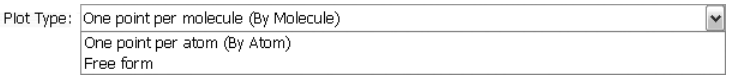
The contents of an Excel spreadsheet may be brought into ***Spartan***. Copy whatever information is to be transferred to the clipboard, move into ***Spartan***, then *click* on the appropriate cell and select **Paste** from the **Edit** menu (or *right click* on the appropriate cell and select **Paste** from the menu that appears). Note, that information on the clipboard that goes beyond the number of rows in ***Spartan's*** spreadsheet will be ignored.

Plots...

Plots may be constructed from data in a spreadsheet and a variety of simple curves fit to these data. Selection of **Plots...** from the **Display** menu leads to the **Plots** dialog.



A tab at the top of the dialog selects between **XY Plot** (two dimensions) or **XYZ Plot** (three dimensions). Common to both is the **Plot Type** menu.



This distinguishes between plots based on the entire set of molecules in a document (**By Molecule**), to those based on the atoms in a single molecule (**By Atom**). The latter is the mode used to plot calculated IR

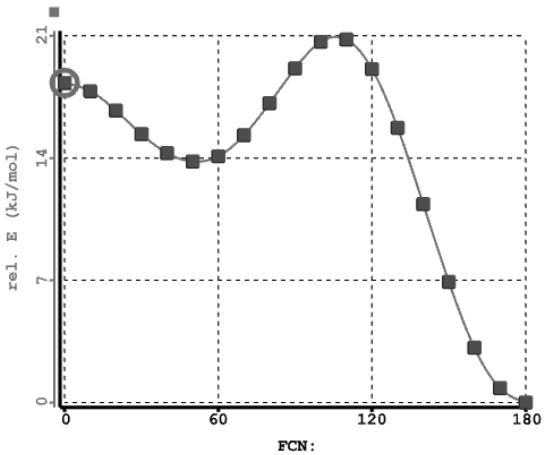
and NMR spectra. The third choice (**Free Form**) is used to construct single molecule plots calculated UV/vis spectra as well as plots of experimental IR, NMR and UV/visible spectra.

"By Atom" and "Free Form" plots make use of vector capabilities in ***Spartan's*** spreadsheet. See the FAQ ***Vector Spreadsheet Operations*** available under **Help** from the **Help** menu (**Chapter 19**).

Also common to both is the **Properties** checkbox at the bottom left of the dialog. Unchecked, this provides access for the **X-Axis** (X-Axis and Y-Axis for an XYZ plot) menu and **Y-Axes** (**Z-Axes** for an XYZ plot) only to quantities that have been entered in the spreadsheet. Checked, it provides access to all quantities in the ***Spartan's Property Archive***.

XY Plots

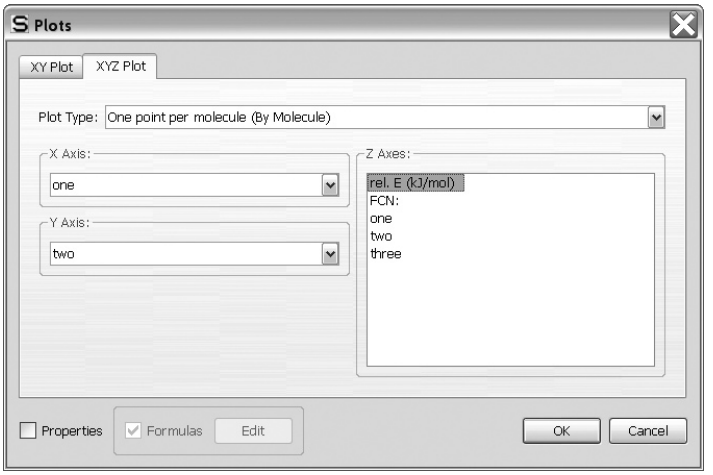
The XY Plot incorporates an **X Axis** menu designating the molecular property to be displayed among the X (horizontal) axis, and a list of properties to be displayed along the Y (vertical) axis. These properties correspond one to one to the columns in the associated spreadsheet. To construct a 2D plot, select an item from the **X Axis** menu, then *click* on one or more items from the **Y Axes** list and finally *click* on **OK**. (Repeated *clicking* on a property in the **Y Axes** list turns it "on" and "off".) The dialog is removed from the screen and a plot appears.



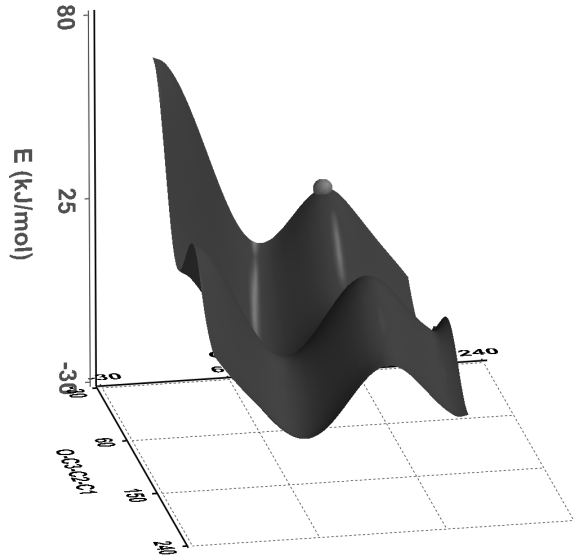
The plot may be moved about the screen. Select (*click* on) either the frame or on one of the curves (it will turn gold in response), then hold down the right mouse button and *drag* the mouse. The plot may also be scaled (expanded or shrunk) by first selecting either the frame or one of the curves and *dragging* the mouse while holding down both the right button and the **Shift** key. The plot may not be rotated.

XYZ Plots

The XYZ Plot is reached by clicking on the **XYZ Plot** tab at the top of the dialog. This leads to a new dialog.



This is very similar to the XY Plot, except that it incorporates **X Axis** and **Y Axes** menus to designate quantities to be displayed along the X and Y axes, respectively, as well as a list of properties to be displayed along the Z axis. To construct a 3D plot, select one item from each of the **X Axis** and **Y Axis** menus, then *click* on one or more items from the **Z Axes** list and finally *click* on **OK**. A plot appears.



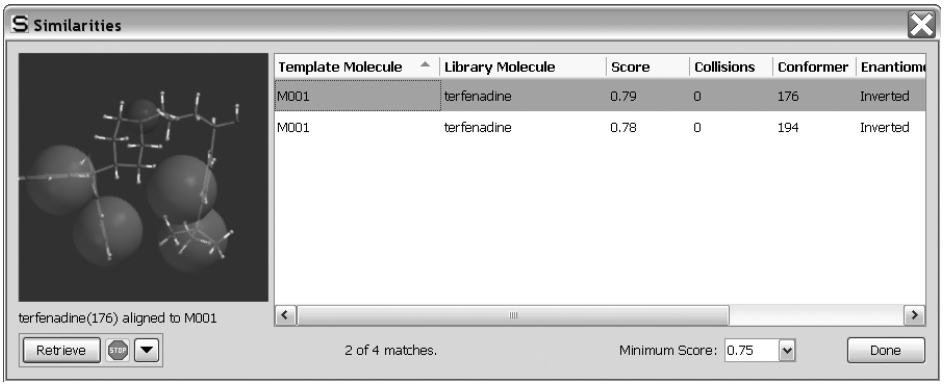
The plot can be moved about the screen. Select (*click* on) either the frame or on one of the curves (it will turn gold in response), then hold down the right mouse button and *drag* the mouse. The plot may also be scaled (expanded or shrunk) by first selecting either the frame or one of the curves and *dragging* the mouse while holding down both the right button and the **Shift** key. Finally, 3D plots (unlike 2D plots discussed earlier) may be rotated, by first selecting either the frame or one of the curves and *dragging* the mouse while holding down the left button.

Plots initially presented connect the data points with cubic splines. A variety of curve fits (linear, least squares, Fourier, etc.) as well as a variety of different presentation formats are available under the **Plot Properties** and associated dialogs (see **Properties** earlier in this chapter).

A plot may be deleted by first selecting (*clicking* on) **Delete** from the **Build** menu, or *clicking* on , or holding down the **Delete** key, and then *clicking* on either the plot frame or on an individual curve. *Clicking* on the frame removes it and all associated curves, while *clicking* on a curve removes only this curve. However, the last curve deleted also deletes the frame. IR, NMR and UV/vis spectral plots may also be removed from the corresponding **Spectra** dialog.

Similarities...

The results of a similarity analysis based either on selected atoms in a structure or on a set of CFD's as a template may be viewed using the **Similarities** dialog.



The box at the right of the dialog identifies each hit in terms of the identity of the Template, the Library molecule, and in the case where a library molecule has been expanded in terms of a set of diverse conformers, the identity of the conformer. It also provides a score* and, in the case where excluded volume CFD's are present in the template, the number of collisions between these elements and the library molecule. Finally, it indicates whether the enantiomer for the molecule included in the library has been inverted. Entries in the dialog may be sorted by score or number of collisions (as well as by the names of the template and library molecules) by *clicking* on the appropriate column header at the top of the box.

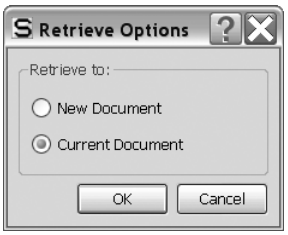
A minimum score can be set from the **Minimum Score** menu at the bottom right of the dialog. The number of matches that satisfy this minimum is reported.

---

* The score is defined as $[(1-R^2)/N]$-penalty, where R^2 is the root mean square distance between template and library molecule centers and N is the number of similarity centers. Whereas molecule alignment as a prelude to similarity scoring is based on atomic, CFD and pharmacophore centers ***only***, penalties are given for alignments that lead to unfavorable steric interactions between template and library entries or (in the case of molecule alignment by CFD's) to incorrect orientation of hydrogen bond donor or acceptor CFD's. These are subtracted from the score based on R^2. 1 is a perfect score.

As hits are selected by *clicking* on them in the box, the library entry superimposed onto the template will be displayed in the window at the left of the dialog. This graphic may be manipulated (rotated, translated, scaled) with the usual mouse/keyboard commands (the cursor needs to be positioned inside the window). The display style cannot be altered, nor can any measurements be made.

The library entry associated with a particular hit may be retrieved (brought into ***Spartan's*** main window) by first selecting it and then *clicking* on the **Retrieve** button at the bottom of the **Similarities** dialog. The template cannot be retrieved. Retrieval can either be to a New Document or to the Current Document depending on the setting in the **Retrieve Options** dialog accessed by *clicking* on ▾ to the right of the **Retrieve** button.



Retrieval to the Current Document (that with the template) allows both the template and one or more hits from the library to be displayed simultaneously, with a full range of model styles and display options.

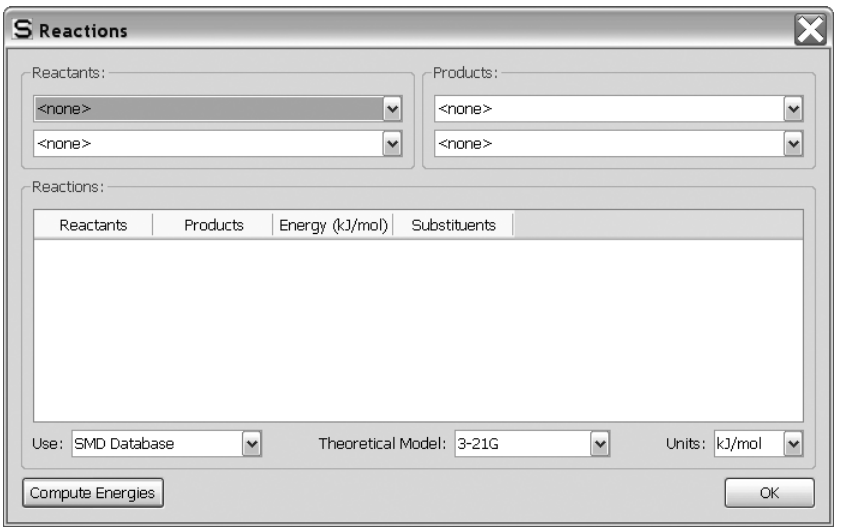
The **Similarities** dialog is closed by *clicking* on the **Done** button.

Reactions...

Data entered in a ***Spartan*** document may be used to caculate reaction energies including activation energies.

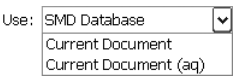
$$\Delta E = E_{product1} + E_{product2} - E_{reactant1} + E_{reactant2}$$

Selection of **Reactions...** from the **Display** menu leads to the **Reactions...** dialog.



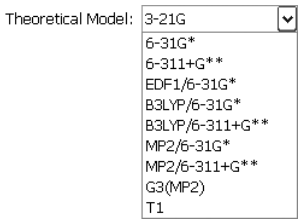
Two menus under **Reactants:** and two menus under **Products:** identify the reactants and products of reaction. They correspond to the labels (identifiers) of the molecules in the document, plus a null entry **<none>**. Selections must be made such that the overall reaction is mass balanced.

The **Use** menu identifies the source of the energies to be used in the reaction energy calculation.



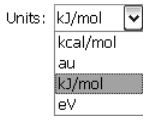
SMD Database refers to energies drawn from entries in the Spartan Molecular Database (SMD), **Current Document** refers to energies in document and **Current Document (aq)** refers to energies in the document that have been empirically corrected for aqueous solvent using the SM5.4 model.

Selection of **SMD Database** (from the **Use** menu) requires specification of theoretical models. This is made from the **Theoretical Model** menu.

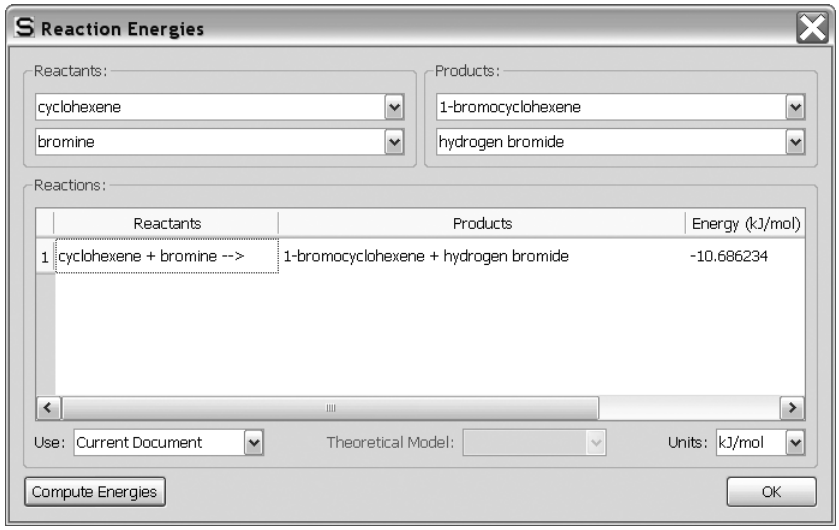


All 10 models covered by SMD are included.

The **Units** menu allows for selection of units.



A reaction energy is computed by *clicking* on **Compute Energies** at the bottom left of the dialog.



Substituted molecules can be used in place of "real molecules" as reactants and products. By accessing SMD, this means that the

energies for a series of related (by substitution) reactions can be computed at once.*

The results of a reaction energy calculation may be printed by *right clicking* inside the display area of the **Reactions** dialog and selecting **Print** from the menu that results.

The **Reactions** dialog is closed by either *clicking* on **OK** or on ☒.

* An example is provided in the tutorial ***Positional Selectivity of Substituted Naphthalenes*** in **Chapter 5**.

Chapter 17
The Search Menu

*Functions available under the **Search** menu provide substructure searching of the Cambridge Structural Database (CSD) of experimental X-ray structures, replacement from and searching by name, formula, molecular weight, isomer and substructure of the Spartan Molecular Database (SMD) of calculated structures, energies, spectra and properties, substructure searching of the Spartan Reaction Database (SRD) of calculated transition states, searching the Spartan Infrared Database (SIRD) based on matching an input infrared spectrum, on-line access to the Protein Data Bank (PDB) and extraction of ligands from PDB files. The menu also accesses a facility for automatically guessing transition states for reactions based on their similarity to transition states in SRD and a procedure for identifying tautomers.*

Databases ( )

This provides access to the Cambridge Structural Database (CSD) of experimental X-ray crystal structures, the Spartan Molecular Database (SMD) of calculated structures, energies, spectra and molecular properties, the Spartan Reaction Database (SRD) of calculated transition states, the Spartan Infrared Database (SIRD) of calculated infrared spectra and the Protein Data Bank (PDB).* CSD, SMD and SRD may be searched based on substructure. SMD may also be searched by name, formula, molecular weight and isomer. If a selected molecule on screen exists in SMD, it may be replaced by a database entry and, even if it is not replaced, the molecule name in

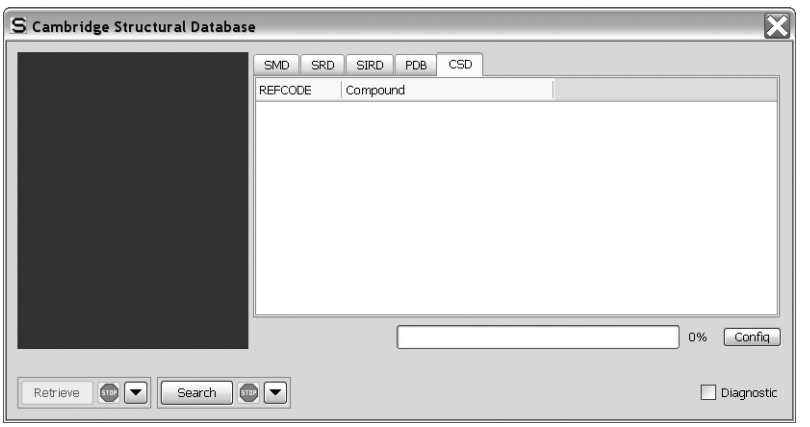
* CSD, SMD and SRD may be located on the local machine or on a remote server. Setup details are provided in **Appendix F** for CSD and in **Preferences** under the **Options** menu in **Chapter 18** for SMD, SRD and SIRD. PDB is available only through web access.

SMD may be used. SIRD is searched by comparison with an input infrared spectrum. PDB is accessed only by its PDB ID number.

Cambridge Structural Database (CSD)

The **C**ambridge **S**tructural **D**atabase (**CSD**) is a collection of more than 450,000 experimental X-ray crystal structures for organic and organometallic molecules.* It is maintained by the Cambridge Crystallographic Data Centre (CCDC) and grows approximately 10% per year. CSD not only contains information about the molecular geometry, but also about the manner in which molecules pack in the solid state**. In short, the CSD represents a gold mine of detailed experimental structural information, and also serves to identify molecules that can be (and have been) synthesized and purified through crystallization.

Access to the data in CSD is via a substructure searching procedure. A molecule is first constructed and one or more attachment points are identified. *Clicking* on the **CSD** tab at the top of the **Database** dialog leads to the **CSD** dialog.



This contains a window at the left for previewing structures, a box at the right for listing hits on CSD and buttons at the bottom to control

---

* CSD is not included with ***Spartan*** but is available by subscription from CCDC (http://www.ccdc.cam.ac.uk/). For installation instructions, see **Appendix F**.

** ***Spartan***'s interface to CSD accesses only molecular information. Additional information related to packing in the crystal may be accessed using CCDC's program ConQuest, that is supplied as part of the CSD subscription.

aspects of the search and subsequent transfer of CSD data into the file system. Both the preview window and the text box may be scaled independently. A substructure search is set up by specifying one or more of the following*:

(i) **Attachment Points**

A free valence is designated as an attachment point by *clicking* on it. It will be marked with an orange cone. *Clicking* again removes the designation and the cone. Anything may be grown off an attachment point (including hydrogen), although substituents are not permitted. Free valences not designated as attachment points are assumed to be hydrogens.

(ii) **Wild-Card Atoms**

An atom may be designated as a wild-card (meaning that element type is unimportant) by *clicking* on it. An orange ball will surround the atom. *Clicking* again removes the wild-card designation and the ball. Use of wild-card atoms will result in structures that incorporate variants of the original substructure with different atoms at designated positions. For example, wild-card designation of an unsubstituted position of a substituted benzene will allow molecules with aromatic nitrogen (substituted pyridines) to be located.

(iii) **Wild-Card Bonds**

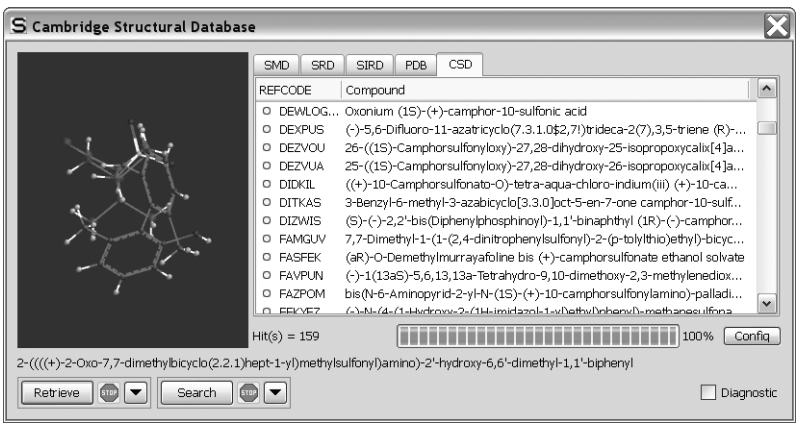
A bond may be designated as a wild-card (meaning that bond type is unimportant) by *clicking* on it. An orange cylinder will be drawn around it. *Clicking* on this cylinder removes the wild-card designation and the cylinder. This may be particularly important in certain heterocycles where bond typing may be ambiguous.

To limit a CSD search either to organic molecules (about half the total collection) or to molecules containing transition metals, lanthanides or actinides, bring up the **Search Options** dialog by *clicking* on ▾ to the right of the **Search** button at the bottom right of the dialog, and select **Organics** or **Inorganics** respectively.

---

* If nothing is specified, a search for an exact match is carried out assuming that all atoms and bonds are exactly as given in the query and that all free valences are assumed to be hydrogens.

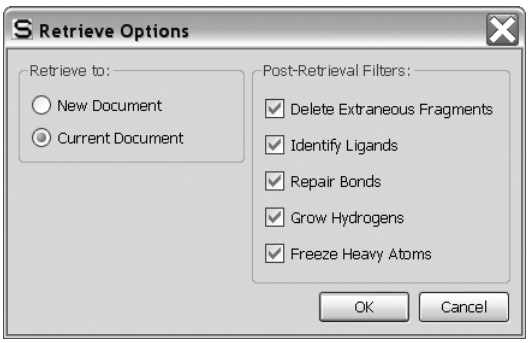
After a substructure has been specified, a search of CSD is initiated by *clicking* on the **Search** button at the bottom right of the dialog. If there are hits to be found, these will appear in the scroll box at the right of the dialog. The search can be terminated at any time by *clicking* on STOP to the right of the **Search** button at the bottom of the dialog. Once completed (or stopped), the CSD reference codes (**REFCODE**) and the names for the hits are displayed in the box and the number of hits is indicated immediately below. Structural data are available only for entries preceded by a filled yellow circle, although experimental references are available for all entries (see **Output** under the **Display** menu; **Chapter 16**). A ball-and-wire model for any entry for which a structure is available may be displayed in the window at the left of the dialog by *clicking* on its name in the box.



The model can be manipulated (rotated, translated, scaled) inside the window with the usual mouse/keyboard commands (you need to position the cursor inside the window).

The structure is ***exactly as it appears in CSD***. In particular, it may contain two or more different molecules, two or more different

conformations of the same molecule and/or extraneous molecules (most commonly solvent molecules and counterions). Hydrogens will be present only where they have been assigned in the experimental structure. The structure may also contain errors in bonding due to uncertainties in the original data. Modifications to the CSD entry to address these issues may be made prior to transfer to the file system using the **Retrieve Options** dialog. This is accessed by *clicking* on ▾ to the right of the **Retrieve** button at the bottom of the **CSD** dialog.



(i) **Delete Extraneous Fragments**
If *checked*, deletes any detached molecules that do not contain the original substructure.

(ii) **Identify Ligands**
Identifies ligands in structures that are returned from CSD (primarily for organometallic compounds) and replace whatever bonding is supplied by a representation in which a dashed line is drawn between the "center of the ligand" and the atom to which it connects.

(iii) **Repair Bonds**
If *checked*, attempts to fix bonding errors based on the actual geometry in CSD and on normal valence rules.

(iv) **Grow Hydrogens**
If *checked*, "grows" hydrogens wherever they appear to be missing, based on the actual geometry and on normal valence rules. This function is also available under **Grow Hydrogens** in the **Molecule Utilities** dialog (**Chapter 16**).

(v) **Freeze Heavy Atoms**
If *checked*, freezes all atoms except for hydrogens. This allows experimental X-ray hydrogen positions that are often poorly determined, to be refined using molecular mechanics or quantum chemical methods, while maintaining the heavy atom skeleton.*

(vi) **Retrieve to:**
Selection of **New Document** means that the search results are to be stored in a new document and leads to a request for a document name, whereas selection of **Current Document** results in the retrieved molecules being appended to the end of the current document.

The **Retrieve Options** dialog is dismissed with all changes to existing settings maintained by *clicking* on **OK**. *Clicking* on **Cancel** or on ☒ also dismissed the dialog but any changes to settings are lost.

Retrieval is accomplished by first selecting (*clicking* on) one or more hits in the scroll box and then pressing the **Retrieve** button at the bottom of the **CSD** dialog. The **Shift** and **Ctrl** keys may be used in the usual way to specify retrieval of multiple hits. Retrieval can be stopped at any time by *clicking* on  to the right of the **Retrieve** button at the bottom of the **CSD** dialog.

---

* Hydrogen positions may be refined using molecular mechanics using **Minimize** under the **Build** menu; **Chapter 14**. Refinement using quantum chemical models is specified in the **Calculations** dialog (**Calculations...** under the **Setup** menu; **Chapter 15**) by way of **Frozen Atoms** (**Freeze Center** under the **Geometry** menu; **Chapter 13**). Additional discussions provided in **Chapter 15**.

Spartan Molecular Database (SMD)

The **S**partan **M**olecular **D**atabase (**SMD**) comprises three collections.* ***Small molecules*** comprises approximately 150,000 molecules, the structures, energies, and selected properties of most of which have been obtained from five different theoretical models: Hartree-Fock models with 3-21G and 6-31G* basis sets, EDF1/6-31G* and B3LYP/6-31G* density functional models and the MP2/6-31G* model. Infrared spectra have been provided for approximately 50,000 molecules at one or more levels of calculation, UV/visible spectra primarily from B3LYP models for approximately 2,000 molecules and NMR spectra from the HF/6-31G* and B3LYP/6-31G* models for approximately 20,000 molecules. In addition, Hartree-Fock, B3LYP and MP2 calculations with larger basis sets (primarily 6-311+G**) are provided for approximately 2,000 molecules and G3(MP2) calculations have been provided for approximately 1,200 molecules. In total, the ***small molecules*** collection in SMD contains about 500,000 data records. It is expected to grow at 25%/year.

The ***T1*** collection comprises approximately 40,000 molecules obtained from the T1 thermochemical recipe. Most of the entries are also found in the ***small molecules*** collection. T1 heats of formation closely mimic those from G3(MP2) calculations, but because T1 requires 2-3 orders of magnitude less computation time than G3(MP2) is available for much larger systems. T1 results for conformationally-flexible molecules are based on the best (lowest heat of formation) T1 conformer, instead of the best MMFF conformer for entries in the ***small molecules*** collection.

The third collection comprises molecules that have been found to bind to proteins or nucleotides (***ligands***) as abstracted from the Protein Data Bank. Each molecule is represented by one or more conformers found in the protein or nucleotide complex. This collection is intended to be used as a library in similarity analysis.

---

* A subset of approximately 5,000 molecules primarily from the ***small molecules*** collection provided with ***Spartan***, each with a full selection of theoretical models. The full SMD database is supplied without charge under a ***Spartan*** maintenance contract or may be licensed from Wavefunction.

An SMD entry is referenced by a one-dimensional (SMILES like) character strung together with a theoretical model. Structural isomers, for example, *cis-trans* isomers are distinguished as are diastereomers, but enantiomers are not distinguished.

SMD is open-ended and users may construct their own data collections with additional molecules and/or results from additional theoretical models on existing molecules (**Appendix G**). Additional collections of organic molecules, each molecule represented in terms of a set of diverse conformers as obtained from MMFF molecular mechanics and intended for use in similarity analysis, may be licensed from Wavefunction. Finally, the raw data that makes up the ***T1*** collection, that is, a selection of up to 100 diverse conformers for each flexible molecule may be licensed.

While the majority of entries in the ***small molecules*** collection are represented in terms of a single (lowest-energy) conformer (see text box below), some entries are represented in terms of two or more conformers. In these cases, the available conformers will be ranked according to energy, and the user needs to choose among them.

The individual entries in the ***small molecules*** collection of SMD have been obtained according to the following recipe:

1. Identify the lowest-energy conformer using molecular mechanics and the MMFF (Merck) force field. Systems with only a few degrees of conformational freedom are searched systematically guaranteeing location of the global minimum, while a Monte-Carlo search is applied to more complicated systems. The latter does not guarantee location of the global minimum.
2. Starting from the best conformer located, perform a geometry optimization using the appropriate quantum chemical model. This generally means that all quantum chemical calculations for a particular molecule will refer to the same conformer (unless this conformer turns out to be unstable at a particular quantum chemical level).

The individual entries in the **T1** collection of SMD have been obtained according to the following recipe:

1. Identify a set of up to 100 diverse conformers using the PM3 semi-empirical model. This has been done using a so-called systematic Monte-Carlo search (see description in **Chapter 15**).

2. Perform T1 calculations on all conformers and select the one with the lowest (T1) heat of formation.

The ***ligands*** collection in SMD comprises approximately 1250 small molecules bound to proteins/nucleotides contained in the January 6, 2005 snapshot of PDB. It contains approximately 1650 entries in total, and has been constructed according to the following procedure:

1. Restrictions

   H, C, N, O, S and halogens only
   200 ≤ molecular weight ≤ 700
   LogP ≤ 7
   number of hydrogen-bond acceptor sites ≤ 14
   number of hydrogen-bond donor sites ≤ 7

   In addition, ligands that are incomplete (according to the compound that is supposed to be incorporated) or are severely distorted have been excluded.

2. Refinement

   Allow atoms to float as much as 0.05A from experimentally assigned positions
   Add hydrogens and correct obvious bonding errors
   No molecular mechanics or quantum chemical calculations are performed

3. Eliminate duplicates, that is, ligands in the same conformation.

A listing of PDB identification codes for all proteins/nucleotides that incorporate the ligand in a particular conformation may be added to the spreadsheet by typing ***PDBids*** in an empty column header (see discussion in **Chapter 16**). A maximum of ten codes are supplied with additional codes signaled by the word ***more*** at the end of the list.

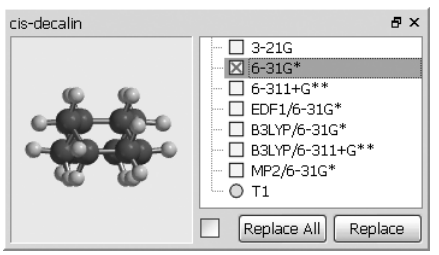
Collections in SMD may be accessed in several ways. Most important is access based on structure or "structural root" (substructure). For the ***small molecules*** and ***T1*** collections, this involves either replacement of a query structure with a database entry corresponding to a particular quantum chemical model, or location of all entries in the database (obtained from one or more quantum chemical models) that contain a particular substructure. Substructure searching of SMD is strictly analogous to the way ***Spartan*** accesses the Cambridge

Structural Database (see previous discussion). In the case of the ***ligands*** collection where there are no quantum chemical calculations, replacement of the query results in the structure that is found in the protein. The ***small molecules*** collections may also be searched by molecule name, chemical formula, molecular weight and isomer.

Replacement

The existence of one or more entries in the ***small molecules*** or ***T1*** collections (corresponding to different quantum chemical models) for the selected molecule is signaled by its name being displayed at the bottom of the screen.

Details are provided by *clicking* on ▲ to the immediate left of the molecule name (it then changes to ▼). This brings up a dialog that comprises a viewing screen on the left and a list of quantum chemical models for which database entries exist on the left.



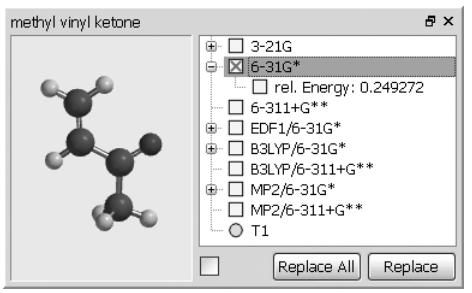
By default, the entry in the list corresponding to a Hartree-Fock 6-31G* calculation is *checked* and a ball-and-spoke model corresponding to its structure appears in the viewing screen. (If a Hartree-Fock 6-31G* calculation does not exist in SMD, the topmost entry in the list is *checked*.) This model may be rotated, translated and scaled using the usual mouse/keyboard commands (you need to position the cursor inside the viewing area). The model may be made to tumble automatically by *checking* the box to the left of **Replace All** at the bottom of the dialog.* Model style

* Molecule tumbling in the main window is controlled through the **Settings Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 18**).

may not be changed. A different database entry (corresponding to a different quantum chemical model) may be selected by *checking* the box to the left of the name of the quantum chemical model in the dialog.

The selected (on-screen) molecule may be replaced by the selected database entry by *clicking* on **Replace** at the bottom of the dialog. (Replacement can be undone by selecting **Undo** from the **Edit** menu; **Chapter 11**). Note that replacement destroys all information except the molecule identifier (leftmost column of the spreadsheet)associated with the molecule being replaced. If desired, the molecule identifier can be changed to the molecule name by first bringing up a spreadsheet (**Spreadsheet** under the **Display** menu; **Chapter 16**), *right clicking* inside the cell corresponding to the leftmost column and selecting **Rename Selected Using SMD** from the menu that results.

Some database entries from the ***small molecules*** collection (only) also comprise information on higher-energy conformers (in addition to that on the presumed ground-state conformer). These entries are designated by a ⊞ icon to the left of the check box. *Checking* this icon leads to an expanded entry that allows access to these higher-energy conformers, as well as provides their relative energies (in kJ/mol).



Entries are selected as before, after which they may be examined in the viewing screen or used to replace the on-screen molecule.

In the event that the selected (on-screen) molecule belongs to a multi-molecule document, it is possible to replace all molecules in the document for which database entries for the specified level

of calculation are available. This is accomplished by *clicking* on **Replace All** instead of **Replace**. A warning message is provided prior to replacement. Note that the lowest-energy conformer for all other (non-selected) molecules in the document is used even though a higher-energy conformer may be indicated for the selected molecule. (Replacement of document entries with non lowest-energy conformations needs to be done one by one using **Replace** instead of **Replace All**.) Molecules in the document for which there are no database entries at the specified level of calculation are not affected. As with single molecule replacement, all information except the molecule identifier is replaced by that in the database. Molecule names from SMD for all molecules in a document may be substituted for the original molecule identifiers by *right clicking* on the header cell of the leftmost column in the spreadsheet and selecting **Replaced Selected Using SMD** from the menu that results.

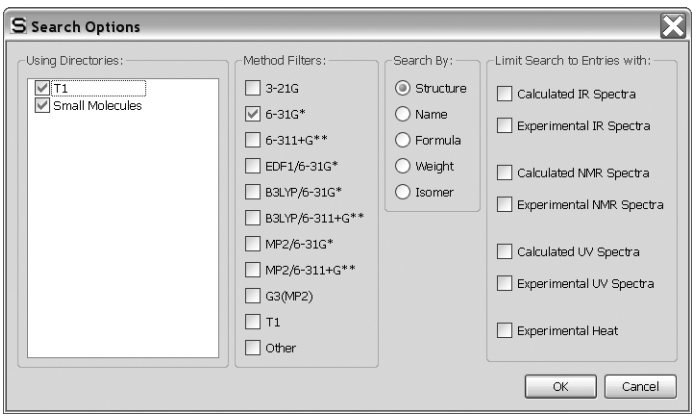
Substructure Searching

Substructure searching from SMD is similar to that from CSD and requires a three-dimensional structure and (optionally) one or more growth points specified. As SMD contains results corresponding to different theoretical models, it is also necessary to specify which model (or set of models) is requested.

Access to SMD is provided by *clicking* on the **SMD** tab at the top of the **Database** dialog. This gives rise to the **SMD** dialog.

This contains a box on the right and a window on the left for listing and examining hits, and buttons at the bottom to define search parameters and transfer of data into ***Spartan***'s file system. Both the window and box can be sized independently.

A substructure search is set up by specifying one or more attachment points, in the same manner as CSD. Wild-card atoms and bonds are not available. However, substituents may be attached to broaden the search beyond a single starting molecule (see **Substituent Model Kit** under the **Build** menu; **Chapter 14**). Prior to starting the search, it is necessary to verify that a structure search is desired and to specify which theoretical model or set of models are to be included. Optionally, the search can be limited to database entries that contain calculated and/or experimental IR, NMR and/or UV/visible spectra and/or an experimental heat of formation. Model specification and search limitations are set in the **Search Options** dialog that is reached by *clicking* on ▾ to the right of the **Search** button.



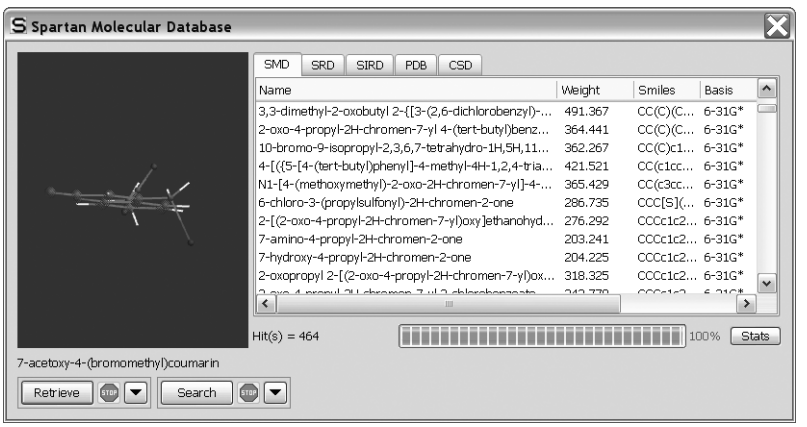
Using Directories lists the collections that make up SMD. Note that the user may add collections (**Appendix G**). **Method Filters** specify one or more models to be included in the search. **Other** accesses data with models that are not explicitly mentioned, and is needed to access the *ligands* collection as well as user-defined collections.

Search by structure (substructure) is indicated by **Structure** being selected under **Search By**.

Checking one or more entries under the **Limit Search to Entries with** restricts the search to database entries that include IR, NMR and/or UV/visible spectra, and/or for entries where the corresponding experimental spectra exist in on-line databases accessible to ***Spartan***, and/or where experimental heats of formation are available. Note, that only entries in the ***small molecules*** collection (may) contain calculated spectra. In particular, no entries in the ***T1*** collection contain spectra.

Clicking on **OK** removes the dialog. (*Clicking* on **Cancel** or ☒ also removes the dialog, but selections are not made.)

Following specification of attachment points and designation of theory level(s) and additional requirements, a search on SMD is carried out by *clicking* on the **Search** button at the bottom right of the dialog. Any hits will appear in the box at the right of the dialog.* The search may be terminated at any time by *clicking* on STOP to the right of the **Search** button at the bottom right of the dialog. A 3D structure for a particular hit may be previewed by *clicking* on its name in the box. The ball-and-wire model is displayed in the window at the left of the dialog, and its name appears immediately below the window.



The model may be manipulated (rotated, translated, scaled) inside the window in the usual way, but the model style cannot

* Normally, only the name of the molecule is visible, but molecular weight and theoretical model may also be displayed by appropriately *sliding* the header cells at the top of the box.

be changed, nor can geometrical measurements be made. Hits may be sorted by name, molecular weight or theoretical model by *clicking* on the appropriate tab.

Searching Name, Formula, Weight, Isomer

SMD may also be searched by name, formula, weight, or isomer. Specification is provided under **Search by** in **Search Options** dialog.

(i) **Name**
Selection of **Name** followed by *clicking* on **OK** leads to a box **By Name** at the bottom right of the **SMD** dialog. The search will return all entries that include whatever text string is entered into this box. For example, typing in ***toluene***, will not only result in toluene, but also molecules like para-toluenesulfonic acid and 4-chloro-2-fluorotoluene.

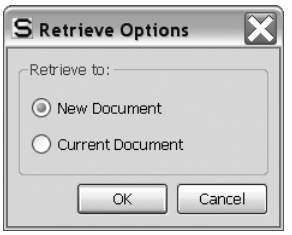
(ii) **Formula**
Selection of **Formula** followed by *clicking* on **OK** leads to a box **By Formula** at the bottom right of the **SMD** dialog. A formula of the form $C_xH_yHet_z$... where C and H stand for carbon and hydrogen respectively and Het stands for other elements (heteroatoms). x, y, z are integers and are required, for example, the formula for nitrobenzene needs to be written as ***C6H5N1O2***. The order of atoms is irrelevant. The search will return all entries with this formula. For example, the formula ***C6H5N1O2*** will return not only nitrobenzene but also molecules like phenyl nitrite and pyridine-4-carboxylic acid.

(iii) **Weight**
Selection of **Weight** followed by *clicking* on **OK** leads to a text box **By Weight** at the bottom right of the **SMD** dialog. The search will return all entries with the specified molecular weight ± 0.05 amu.

(iv) **Isomer**
Selection of **Isomer** followed by *clicking* on **OK** specifies that all isomers of the selected molecule will be returned.

Retrieval

Prior to retrieval of one or more database hits, it is necessary to specify a destination. This is accomplished with the **Retrieve Options** dialog, brought up by *clicking* on ▾ to the right of the **Retrieve** button at the bottom left of the **SMD** dialog.



Selection of destination is accomplished by *toggling* between **New Document** and **Current Document**. Selection of the former creates a new document that will need to be named later, while selection of the latter results in the retrieved molecules being appended to the end of the current document.* The **Retrieve Options** dialog is dismissed by *clicking* on **OK**. *Clicking* on **Cancel** or ☒ also dismisses the dialog but any changes are lost.

Retrieval is accomplished by first selecting (*clicking* on) one or more hits in the box and then pressing the **Retrieve** button at the bottom of the **SMD** dialog. The **Shift** and **Ctrl** keys may be used to specify retrieval of multiple hits. A block of entries may be specified by *clicking* on the top (or bottom) entry, then holding down the **Shift** key and finally *clicking* on the bottom (or top) entry. Multiple entries may be specified by *clicking* on each in turn while holding down the **Ctrl** key. Retrieval may be stopped at any time by *clicking* on STOP to the right of the **Retrieve** button at the bottom of the **SMD** dialog.

---

* Note, that the SMD dialog and other dialogs under **Databases** in the **Search** menu can be entered without a document open on screen (structure and isomer searching are not available). In this case, selection of **Current Document** creates a new document.

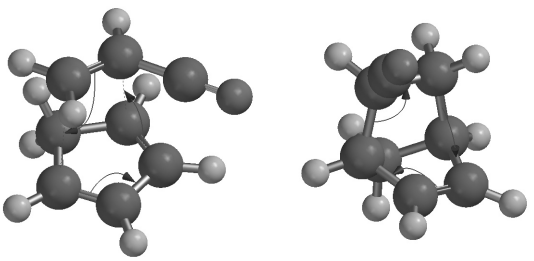
Spartan Reaction Databases (SRD)

The **S**partan **R**eaction **D**atabase (**SRD**) is a collection of transition states for approximately 1,500 organic and organometallic reactions obtained from one or more of the following theoretical models: AM1, PM3, HF/3-21G and HF/6-31G*. In addition to the (transition-state) geometry and the energy (heat of formation for semi-empirical models), each data entry contains the infrared spectrum. This should have a single imaginary frequency, the associated vibrational motion for which corresponds to the reaction coordinate.

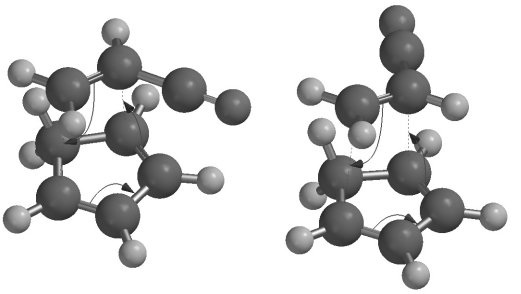
An entry in SRD is in terms of a three-dimensional structure (corresponding to the reactants or the product), together with an appropriate set of reaction arrows (connecting the reactants to the product or vice versa) and designation of theoretical model.

> Additional information about drawing reaction arrows is provided later in this chapter under **Transition States**.

For example, an entry for the Diels-Alder reaction of cyclopentadiene and acrylonitrile is given on the left, while that for the corresponding retro Diels-Alder is given on the right.



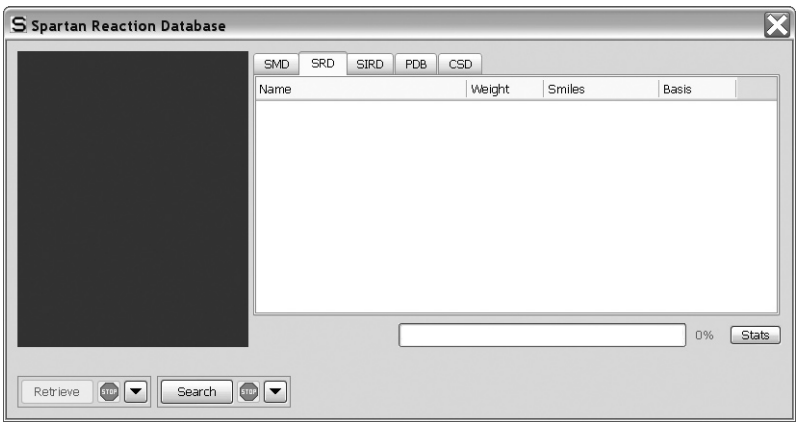
Note, that stereoisomers are distinguished by the way the reactants are positioned. For example, reactions leading to *endo* and *exo* Diels-Alder adducts are distinct.



SRD is open-ended and users may construct their own data collections with additional reactions and/or results from additional theoretical models on existing reactions (**Appendix G**).

Substructure Searching

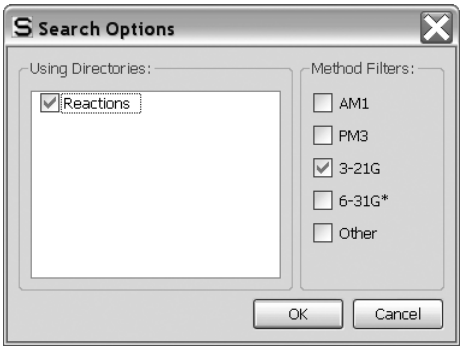
Substructure searching from SRD is very similar to that from SMD. Access to SRD is provided by *clicking* on the **SRD** tab at the top of the **Database** dialog. This gives rise to the **SRD** dialog.



Like the **SMD** dialog, the **SRD** dialog contains a window on the left for previewing structures and a box on the right for listing hits on SRD, as well as buttons at the bottom to control various aspects of the search and subsequent transfer of data into ***Spartan***'s file system. Both the preview window and the text box can be sized independently.

A substructure search is set up by specifying one or more attachment points on the reactant structure that has been augmented with reaction arrows (see discussion under **Transition States** later in this chapter). Prior to starting the actual search, it is necessary to verify

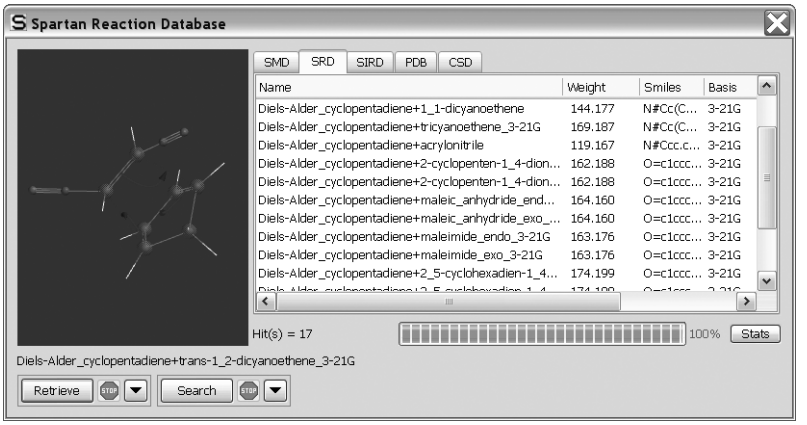
that a structure search is desired and to specify which theoretical model or set of models are to be included. Model specification and type of search are set in the **Search Options** dialog that is reached by *clicking* on ▼ to the right of the **Search** button at the bottom right of the dialog.



Using Directories indicates the libraries to be searched. One is supplied with ***Spartan*** and the users are free to replace or supplement this with their own collections. *Checking* one or more boxes under the **Method Filters** specifies the theoretical model(s) to be searched.

Clicking on **OK** removes the dialog. (*Clicking* on **Cancel** or ☒ also removes the dialog, but selections are not made.)

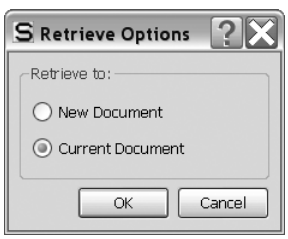
Following specification of structure/reaction arrows, attachment points and designation of theory level(s), a search on SRD is carried out by *clicking* on the **Search** button at the bottom right of the dialog. If there are hits, they will begin to appear in the box at the right of the dialog after a few seconds. The search may be terminated at any time by *clicking* on STOP to the right of the **Search** button at the bottom right of the dialog.



Database hits are displayed in the box at the right of the dialog*. They may be sorted by name, molecular weight and theoretical model by *clicking* on the appropriate tab. A 3D structure for a particular hit may be previewed by *clicking* on its name in the box. A ball-and-wire model of the transition state with bonding consistent with the reactant and associated reaction arrows is displayed in the window at the left of the dialog, and its name appears immediately below the window. The model may be manipulated inside the window in the usual way, but the model style cannot be changed, nor can geometrical measurements be made.

Retrieval

Prior to retrieval of one or more database hits, it is necessary to specify a destination. This is accomplished with the **Retrieve Options** dialog, brought up by *clicking* on ▾ to the right of the **Retrieve** button at the bottom left of the **SRD** dialog.



Selection of **New Document** results in request for a file name;

* Normally, only the name of the reaction is visible, but molecular weight and theoretical model may also be displayed by appropriately *sliding* the header cells at the top of the box.

selection of **Current Document** results in the retrieved molecules being appended to the end of the current document. The **Retrieve Options** dialog is dismissed by *clicking* on **OK**. *Clicking* on **Cancel** or on ☒ also dismisses the dialog but any changes to settings are lost.

Retrieval is accomplished by first selecting (*clicking* on) one or more hits in the box and then pressing the **Retrieve** button at the bottom of the **SMD** dialog. The **Shift** and **Ctrl** keys may be used to specify retrieval of multiple hits. A block of entries may be specified by *clicking* on the top (or bottom) entry, then holding down the **Shift** key and finally *clicking* on the bottom (or top) entry. Multiple entries may be specified by *clicking* on each in turn while holding down the **Ctrl** key. Retrieval may be stopped at any time by *clicking* on STOP to the right of the **Retrieve** button at the bottom of the **SRD** dialog.

Spartan Infrared Database (SIRD)

The **S**partan **I**nfrared **D**atabase (**SIRD**) presently comprising ~40,000 entries* is different from Spartan Molecular Database and the Spartan Reaction Database in that it is searched not based on structure (substructure) but rather on a measured (or calculated) infrared spectrum. This is input in JCAMP format (.dx).** At present the search is carried out in two steps, a "fast" search, the computational cost of which depends on the overall size of the database, and a second more thorough and significantly slower search that is independent of the size of the database.

Infrared spectra for SIRD have been obtained from the EDF2/6-31G* model*** based on best conformations from the T1 model.

---

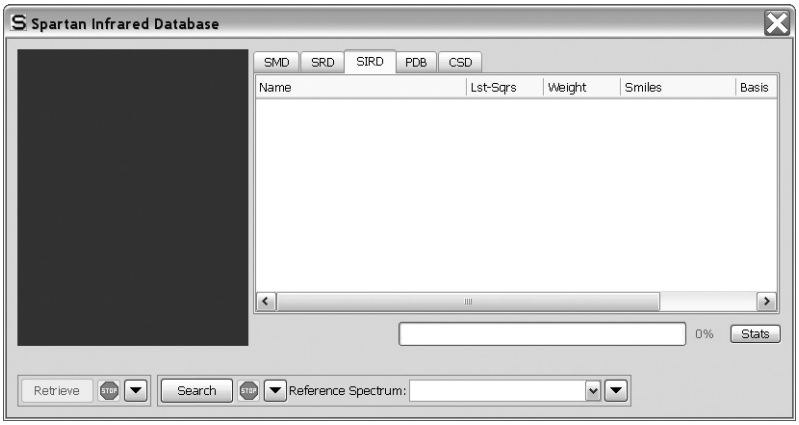
* As with SMD, a small subset of ~5000 molecules is provided with ***Spartan***. The full SIRD database is provided without charge as part of a maintenance contract or may be separately licensed.

** A sample .dx file is provided in **Appendix I**.

*** Infrared spectra obtained from the Hartree-Fock 3-21G and 6-31G* models and the EDF1/6-31G* and B3LYP/6-31G* density functional models are also available as alternative (or supplementary) libraries in SIRD. Each collection presently comprises ~40,000 molecules (nearly identical to the number in the EDF2/6-31G* collection). Note, however, that the molecules in these collections are represented by the best conformation obtained by MMFF molecular mechanics (rather than by T1). Infrared spectra produced by these models are not as successful for spectral searches as those resulting from the EDF2/6-31G* model.

Spectra Searching

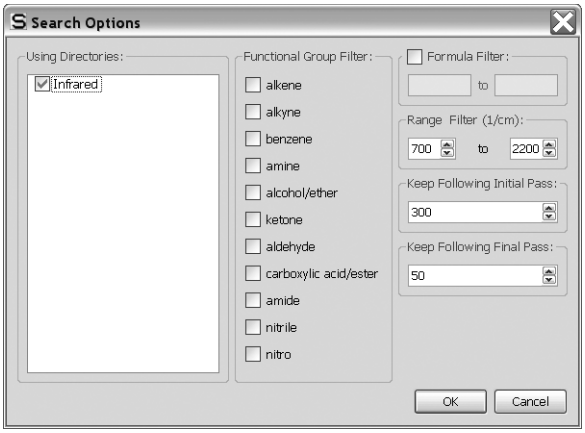
Access to SIRD is provided by *clicking* on the **SIRD** tab at the top of the **Database** dialog. This gives rise to the **SIRD** dialog.



Like the **CSD**, **SMD** and **SRD** dialogs, this comprises a window on the left for previewing structures and a box on the right for listing hits, as well as buttons at the bottom to control various aspects of the search and subsequent of data into ***Spartan's*** file system. Both the preview window and the box can be sized independently.

A search on an infrared spectrum is initiated by *clicking* on (▼) to the extreme right of ***Reference Spectrum*** at the bottom right of the dialog. This leads to a file browser that has been set up to recognize only .dx files. Selection of a file in the usual manner followed by *clicking* on **Open** (or *double clicking* on the name of the .dx file) leads to the spectrum being selected as a reference for the search and the file name displayed in the box to the right of ***Reference Spectrum***.

The details of the search are specified from the **Search Options** dialog which is accessed by *clicking* on (▼) to the right of the **Search** button.



The window at the left lists the directories that make up the SIRD database, and if checked will be used in the search. Normally, the only entry will be EDF2/6-31G* (either the subset or the full SIRD database). Users may add their own (EDF2/6-31G* entries) or infrared spectra from other calculation models.* Changes are made using the **Path Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 18**).

The remaining entries in the dialog serve to filter database entries prior to the search and to control the number of hits returned.

Functional Group Filters allow the search to be restricted to molecules that incorporate one or more of the functional groups checked. Presently these include alkene, alkyne, benzene, amine, alcohol and ether, ketone, aldehyde, carboxylic acid and ester, amide, nitrile and nitro functionality.

Formula Filter allows selection of the "minimum" and "maximum" molecular formula (or both) to be considered in the search. The format of the input is $C_cN_nO_o$ ..., where C, N, O, etc. are atomic symbols and c, n, o are the number of each of the designated atoms. Hydrogens are not considered as are any elements that are not explicitely designated.

Range Filter provides the range (in cm^{-1}) that the search will be considered. Frequencies below 500-600 cm^{-1} (which typically

* It is not advisable to mix spectra obtained from different theoretical models.

correspond to torsional motions and may depend on conformation) require special instrumentation to measure and are seldom reported. While the experimental spectrum is usually recorded and reported to 4500 cm^{-1}, the region beyond ~2800 cm^{-1} is dominated by CH stretching vibrations and may be too crowded to be of value in spectral searching. A good range is from 700-2200 cm^{-1}.

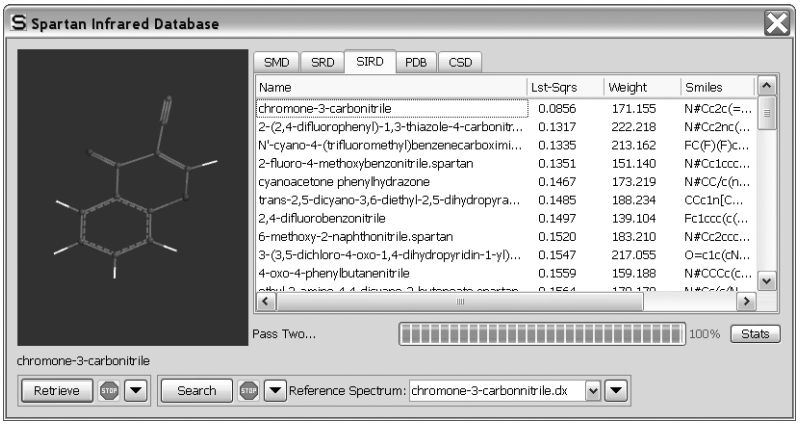
Keep Following Initial Pass controls the number of entries that are kept following the initial pass of ***Spartan's*** spectra matching algorithm. The larger the number, the more computer time will be needed and the more likely it will be to produce a "good" hit (assuming that a closely-related molecule is actually in the database).

Keep Following Final Pass controls the number of hits that are returned to the user. Does not affect the time required for the search.

Clicking on **OK** exits the dialog. *Clicking* on **Cancel** or ☒ also dismisses the dialog but any changes to the settings are lost.

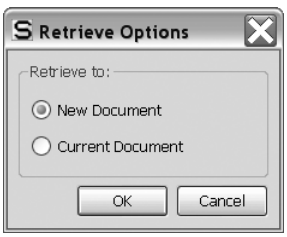
A search on the reference spectrum restricted by whatever conditions are specified in the **Search Options** dialog is carried out by *clicking* on the **Search** button at the bottom of the **SIRD** dialog. The search is carried out in two passes. The first pass may require several tens of seconds depending on the size of the database and the filters that have been set. The second pass may require several minutes (depending on the number of hits kept following the initial pass as specified in the **Search Options** dialog). Searching may be stopped at any time by *clicking* on (STOP) to the right of the **Search** button at the bottom of the **SIRD** dialog. However no results will be returned if the search is stopped prematurely.

When the search has completed a listing of hits (the number of which is specified by **Keep Following Final Pass** in the **Search Options** dialog) ordered by their least squares errors (lowest is best) appears in the box at the right of the **SIRD** dialog. Hits may be sorted by least squares error as well as name and molecular weight by *clicking* on the appropriate tab.



Retrieval

Prior to retrieval on one or more database hits, it is necessary to specify a destination. This is accomplished with the **Retrieve Options** dialog, brought up by *clicking* on ▾ to the right of the **Retrieve** button at the bottom left of the **SIRD** dialog.



Selection of destination is accomplished by *toggling* between **New Document** and **Current Document**. Selection of the former creates a new document which will need to be named later, while selection of the latter results in the retrieved molecules being appended to the end of the current document. The dialog is dismissed by *clicking* on **OK**. *Clicking* on **Cancel** or ☒ also dismisses the dialog but any changes are lost.

Retrieval is accomplished by first selecting (*clicking* on) one or more hits in the box and then *pressing* the **Retrieve** button at the bottom of the **SIRD** dialog. The **Shift** and **Ctrl** keys may be used to specify retrieval of multiple hits. A block of entries may be specified by *clicking* on the top (or bottom) entry, then holding down the **Shift** key and *clicking* on the bottom (or top)

entry. Multiple entries may be specified by *clicking* on each in turn while holding down the **Ctrl** key. Retrieval may be stopped at any time by *clicking* on () to the right of the **Retrieve** button at the bottom of the **SIRD** dialog.

Comparison of Reference Spectrum and Database Hits

This is accomplished using the **IR Spectra** dialog (**Spectra** under the **Display** menu; **Chapter 16**).

Protein Data Bank (PDB)

Spartan provides on-line access to the **P**rotein **D**ata **B**ank (**PDB**)* comprising upwards of 37,000 protein and nucleotide structures.

Access to PDB is provided by *clicking* on the **PDB** tab at the top of the **Database** dialog. This gives rise to the **PDB** dialog.

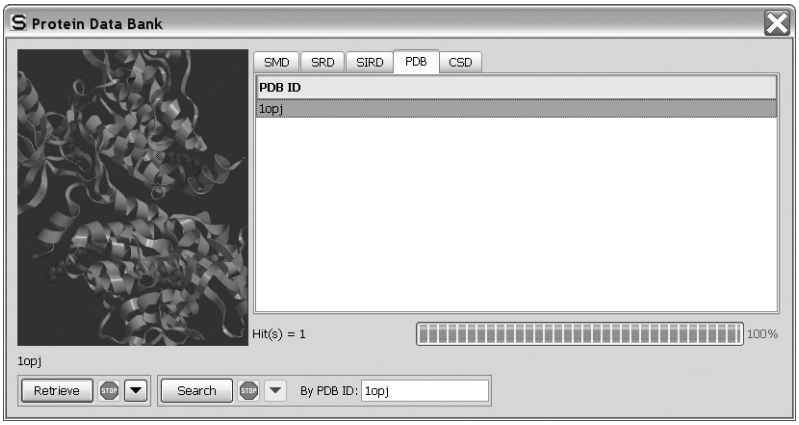
This contains a window on the left for previewing structures and a scroll box on the right for listing hits on PDB, as well as buttons at the bottom to control transfer of data into the file system.

To access a PDB structure, enter the four character identification code in the box to the right of **By PDB ID** at the bottom of the dialog and *click* on the **Search** button. If the PDB entry contains more than one structure (common with protein structures obtained from NMR

* The web address is http://www.rcsb.org.

spectroscopy) and/or the PDB ID yields more than one entry, all structures will be returned.

PDB access will typically require a few seconds. The PDB ID will appear at the right and a ribbon model of the protein or nucleotide will appear at the left.



To retrieve a structure, *click* on the PDB ID (or, using the **Ctrl** and **Shift** keys, set of ID's) inside the box in the dialog and then *click* on the **Retrieve** button. The **Retrieve Options** dialog (not shown) designates whether the structure will go into a **New Document** or be appended onto the **Current Document**.

Ligands ( )

Spartan provides a facility for extracting ligands from PDB files* either as molecular structures or as footprints (pharmacophores) that these structures leave behind. The latter is separated into three parts, the first two of which follow from the molecular structure of the ligand and the third which follows from the immediate (protein or nucleotide) environment:

(i) hydrogen-bond donor/acceptor sites and positive/negative ionzable sites

(ii) hydrophobes

(iii) excluded volumes

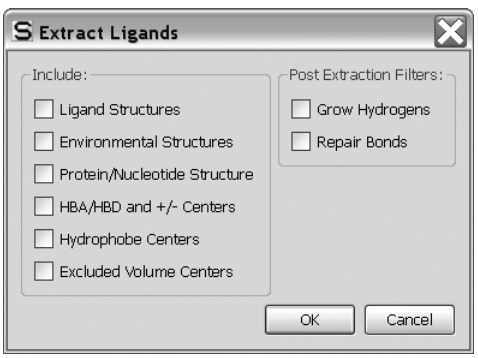
* This capability is not restricted to files in PDB, but extends to files written in PDB format.

A pharmacophore can be constructed by requiring (i) to account for hydrogen-bonded and electrostatic interactions and (optionally) ***either*** (ii) or (iii) to account for steric interactions. Selection of (ii) leads to a pharmacophore that is visually simpler, but selection of (iii) may provide a more realistic picture.

Selection of **Ligands** leads to a message at the bottom of the screen.

Select Ligands.

Ligands in PDB files are represented as a series of transparent spheres much like a space-filling model. *Clicking* on a ligand selects it and displays the PDB **HET name*** at the bottom of the screen (if a HET name is available). Multiple ligands may be selected if desired by holding down on the **Ctrl** key. All ligands may be selected by *clicking* on the **Select All Ligands** button at the bottom right of the screen. After one or more ligands have been selected, *clicking* on the **Extract Ligands** button at the bottom right of the screen brings up the **Extract Ligands** dialog.



This allows selection of whether **Ligand Structures** or one or more pharmacophore elements (**HBA/HBD and +/- Centers**, **Hydrophobe Centers** including aromatic centers, **Excluded Volume Centers**) or ***both*** structure and one or more pharmacophore elements are to be abstracted from the PDB file, and whether amino acid and/or nucleotide residues in close hydrogen-bonding or charge-charge contact with the ligand (**Environmental Structures**) are to be

* This is the code given to small molecules that are associated with proteins/nucleotides in PDB files.

identified. It also allows the **Protein/Nucleotide Structure** (minus all ligands) to be extracted.

In the case of extraction of **Ligand Structures**, the **Extract Ligands** dialog also provides a set of utilities:

(i) **Grow Hydrogens**
If *checked*, "grows" hydrogens wherever they appear to be missing, based on the actual geometry and on normal valence rules. This function is also available under **Grow Hydrogens** in the **Molecule Utilities** dialog (**Properties** under the **Display** menu; **Chapter 16**).

(ii) **Repair Bonds**
If *checked*, attempts to fix bonding errors based on the actual geometry in the PDB entry and on normal valence rules.

Clicking on **OK** at the bottom of the **Extract Ligands** dialog extracts the ligand(s) or pharmacophore(s), placing them in a single new (unnamed) document. *Clicking* on **Cancel** dismisses the dialog without ligand (pharmacophore) extraction.

Transition States ()

Spartan provides a facility for automatically guessing transition state geometries based on the similarity of the reaction of interest with one or more entries in ***Spartan***'s reaction database*. An exact hit is possible, in which case the guess will be the exact transition state for the particular theoretical model entered in the database**. This will not commonly be the case, and ***Spartan*** will attempt to provide as close a match as possible with a database entry. This will generally involve a less substituted system or one in which substituents differ. Here, ***Spartan*** will use those parts of the structure of the transition state in the database that are common, and will optimize the remaining parts (using molecular mechanics). No conformational searching is performed, and it may be essential for the reactants to

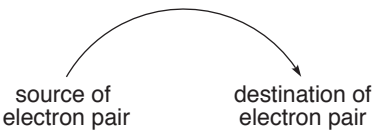
* The Spartan Reaction Database (SRD) may also be directly searched by substructure. See discussion earlier in this chapter.

** At the present time, the majority of database entries derive from very simple (semi-empirical and Hartree-Fock) quantum chemical models.

be properly oriented to reflect the desired stereochemical outcome of the reaction.

Where a reaction is completely unknown to the database, a fallback technique similar to the linear synchronous transit method is automatically invoked.

Input to ***Spartan***'s transition-state guessing procedure will be very familiar to (organic) chemists, in that it is based on reaction arrows. The difference is that arrows are superimposed onto a three-dimensional structure rather than a two-dimensional drawing. The reaction is specified using curved arrows, where each arrow identifies the movement of one electron pair. The direction of electron flow follows customary practice:



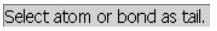
There are two possible sources of an electron pair and three possible destinations, leading to six combinations:

lone pair → lone pair	move lone pair
lone pair → bond	use lone pair to increase bond order
lone pair → space between atoms	use lone pair to create new (single) bond
bond → lone pair	decrease bond order to make lone pair
bond → bond	decrease order of one bond to increase order of another bond
bond → space between atoms	decrease order of one bond to make a new (single) bond

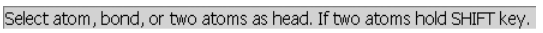
The first of these is a null operation, and its inclusion has no effect.

Spartan's transition-state guessing procedure presently does not distinguish between a lone pair and a radical center. Examples are provided in the next section.

Selecting **Transition States** results in a message at the bottom left of the screen.



The tail of the arrow corresponds to the source of the electron pair. If the source is a lone pair, then select the atom that holds the lone pair. If the source is a bond, then select the bond. *Clicking* on an atom or bond highlights (colors gold) the atom or bond and leads to a new message at the bottom left of the screen.



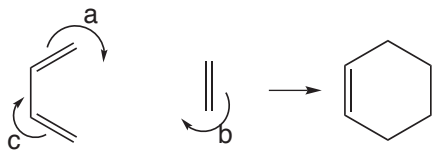
The head of the arrow corresponds to the destination of the electron pair. If the destination is an atom (leading to a lone pair), then select the atom that will hold the lone pair. If the destination is an existing bond (leading to an increase in bond order from single → double → or double → triple), then select the bond. If no bond presently exists, *press* the **Shift** key and select the two atoms that will become bonded upon reaction. These operations result in a curved arrow being drawn on the reactant structure. This extends from an atom, or the center of a bond to an atom, or the center of a bond, or the center of a dotted line that has been drawn between atoms that are to be bonded. The original message returns to the bottom left of the screen.

The process (tail selection followed by head selection) is repeated as necessary to fully define the reaction. Incorrect reaction arrows may be removed by selecting **Delete** from the **Build** menu () followed by *clicking* on the arrow to be deleted. In the latter case, you need to again select **Transition States** () in order to continue arrow selection. Alternatively, *click* on the arrow(s) to be deleted while holding down the **Delete** key

After all reaction arrows have been properly designated, *click* on  at the bottom right of the screen to replace the reactant with a guess at the transition state. In the event that the guess is unreasonable, this operation may be "undone" (select **Undo** from the **Edit** menu). This allows you to review your assignment of arrows and make changes as needed.

Examples

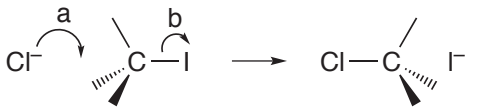
Diels-Alder reaction of 1,3-butadiene and ethylene



a, b. double bond to empty space leading to single bond and single bond

c. double bond to single bond leading to single bond and double bond

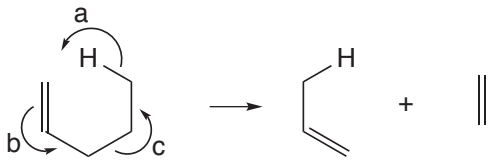
S_N2 reaction of chloride and methyl iodide



a. lone pair on Cl^- to empty space leading to ClC bond

c. CI bond leading to lone pair on I

Ene reaction of 1-pentene

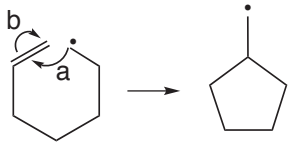


a. CH bond to empty space leading to CH bond

b. double bond to single bond leading to single bond and double bond

c. single bond to single bond leading to empty space and double bond

Ring closure of 1-hexenyl radical to methylcyclopentyl radical

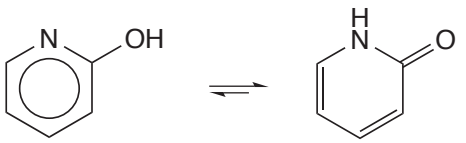


a. radical center to empty space (between C_2 and C_6) leading to single bond

b. double bond to terminal carbon leading to single bond and radical center

Tautomers ( **)**

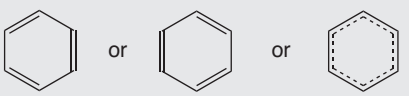
Spartan provides a facility to automatically identify tautomers, that is, isomers arising from transfer of hydrogens among heteroatoms. Tautomer identification is especially important in dealing with heterocyclic compounds where two or more different tautomers may exist in equilibrium and the identity of the dominant tautomer may not be obvious, for example, while 2-hydroxypyridine is an aromatic molecule and would be expected to be unusually stable, the non-aromatic 2-pyridone molecule is actually thermodynamically favored.



2-hydroxypyridine 2-pyridone

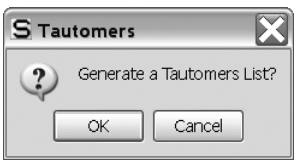
This, of course, has consequences in trying to locate a particular molecule in a database where it may be represented by a different tautomer.

Note, that different tautomers are not the same as different ways to draw the same structure. For example, benzene can be represented as one of two structures with alternating double and single bonds or a single "delocalized" structure in which all bonds are the same and intermediate in length between single and double bonds.



The procedure incorporated into ***Spartan*** is limited to tautomers involving nitrogen, oxygen, phosphorous and sulfur. Tautomers involving carbon, in particular, have been excluded. In addition, zwitterion tautomers, for example, $H_3N^+CH_2CO_2^-$ as a tautomer of glycine, are not considered. Within these limits, all possible tautomers will be identified. Note, however, that there may be duplicates as well as tautomers that are likely to be highly unfavorable.*

The existence of tautomers is signaled by the word **Tautomer** displayed at the bottom right of the screen. If tautomers exist, individual structures may be examined or a complete list of structures generated by selecting **Tautomers** from the **Search** menu (or *clicking* on the icon on the top of the screen). Following this, *clicking* on the step buttons at the bottom right of the screen moves through the list of tautomers. Any one of these may be used in place of the original. The full list of tautomers may be generated by *clicking* on (at the bottom right of the screen). This leads to a dialog.



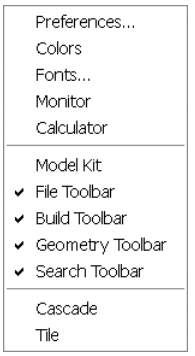
Clicking on **OK** dismisses the dialog and leads to a new (unnamed) document containing the full list of tautomers. (The original document is unaffected.) *Clicking* on either **Cancel** or dismisses the dialog without generating a tautomer list.

As tautomers are generated without consideration of steric crowding, it is important to minimize each in turn (**Minimize** under the **Build** menu; **Chapter 14**) prior to using them for quantum chemical calculations.

* Of course, once a complete list of tautomers has been generated it is straightforward to calculate their relative energies using quantum chemical models.

Chapter 18
The Options Menu

*Functions under the **Options** menu* set default colors, fonts, user preferences and van der Waals radii, locate databases, identify remote servers, engage parallel computing, set program queue's and monitor jobs.*

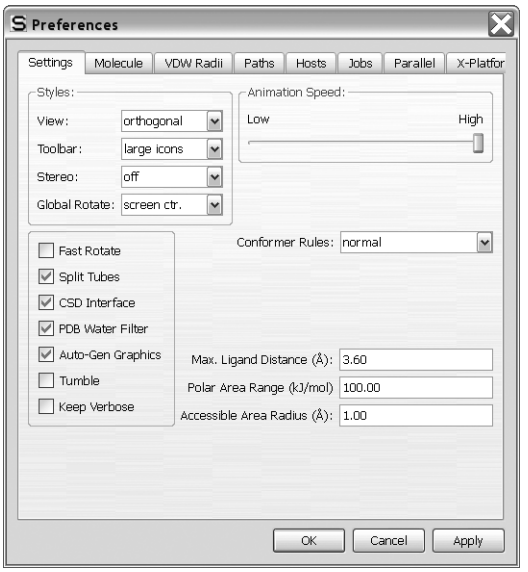


Preferences...

This sets up preferences relating to the graphical user interface (**Settings**), and to molecule displays (**Molecules**). It permits changes to default van der Waals radii used for space-filling models as well as for calculating molecular surface areas and volumes (**VDW Radii**). It also specifies the locations of databases (**Paths**), identifies remote servers (**Hosts**), sets job queues (**Jobs**), (optionally) specifies parallel operations (**Parallel**) and specifies platform dependent features (**X-Platform**). Selection results in one of eight dialogs, depending on which tab has been selected in the previous entry. *Clicking* on a tab brings up the associated dialog. To exit a **Preferences** dialog *click* on **OK**. *Clicking* on **Cancel** or ☒ exits the dialog without instituting any changes.

* **Preferences** are also under the **Spartan'08** menu on the Macintosh version.

Settings



(i) **View: Orthogonal/Perspective**
Controls the view of structural models and graphics.

(ii) **Toolbar Style: text/small icons/medium icons/large icons**
Controls presentation style (text vs. icons) and size of the icons.

(iii) **Stereo: off/red-cyan**
Turns "on" and "off" **red-cyan** stereographic display.

(iv) **Global Rotate: screen ctr./molecule ctr**.
Screen ctr. rotates all molecules about a common center; while **molecule ctr.** rotates each molecule about its own center.

(v) **Animation Speed**
A slider bar controls the maximum speed for animations. This has become an important control as the performance of graphics cards has increased.

(vi) **Fast Rotate**
If *checked*, lowers the resolution of structure models during rotation to improve graphics performance. This should not be necessary except for machines with slow graphics cards.

(vii) **Split-Tubes**
If *checked*, the tubes in tube and ball-and-spoke models are split to designate multiple bonds.

(viii) **CSD Interface**
If *checked*, displays the **CSD** tab in the **Database** dialog (**Database** under the **Search** menu; **Chapter 17**). *Uncheck* if CSD is not installed.

(ix) **PDB Water Filter**
If *checked*, removes water molecules from imported PDB files.

(x) **Auto-Gen Graphics**
If *checked*, performs graphics calculations in the interface.

(xi) **Tumble**
If *checked*, allows automatic tumbling of molecule. To tumble a molecule, select it, *press* the left mouse button, move the mouse and release the button.

(xii) **Keep Verbose**
If *checked*, keeps verbose output. Normally, this is discarded upon successful completion. Verbose output may be useful for identifying the source of problems for calculations that have not successfully completed or have led to suspicious results.

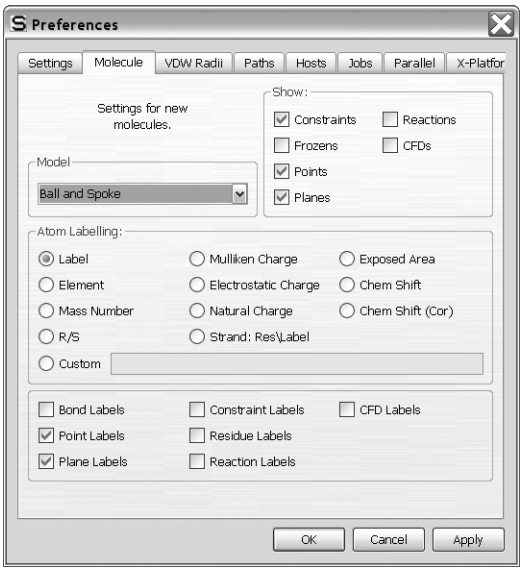
(xiii) **Conformer Rules**
Sets rules for conformational searching. **Normal** should be used for **Equilibrium Conformer** and **Conformer Distribution** where Monte Carlo methods are involved. **Trimmed** eliminates less important degrees of conformational freedom and should be used for **Conformer Library** where systematic searches are carried out.

(xiv) **Max-Ligand Distance**
Set maximum distance (in Å) for an amino acid in a protein to be recognized as bonded to a ligand.

(xv) **Polar Area Range**
Set potential (in kJ/mol) for calculating polar area from elecrostatic potential map. Values > | range | contribute to polar area. Default is 100 kJ/mol.

(xvi) **Accessible Area Radius**
Set sphere radius (in Å) for determining accessible area. Default is 1.0.

Molecule



This specifies default settings for model appearance. These settings may be overridden for a specific molecule (or list of molecules) using entries under the **Model** menu and for specific portions of a molecule using **Utilities/Style** dialogs associated with **Properties** dialogs (**Properties** under the **Display** menu; **Chapter 16**).

(i) **Model: Wire/Ball and Wire/Tube/Ball/Spoke/Space Filling/Line**
Controls default model style.

(ii) **Show: Constraints/Frozens/Points/Planes/Reactions/ CFD's**
If *checked*, constraints and frozen markers, points and planes, reaction arrows and CFD's will always be shown as part

of the model. Otherwise, they will be shown only in the apprpriate mode.

(iii) **Atom Labeling: Label/Element/Mass Number/Mulliken Charge/Electrostatic Charge/Natural Charge/Strand: Residue/Label/R/S/Exposed Area/Chem Shift/Chem Shift (Cor)/Custom**
Controls default label type. **Custom** allows the user to enter an expression designating a label. These are of the same form as used in the spreadsheet. See the section on **User-Defined Expressions** under **Spreadsheet** in the **Display** menu (Chapter 16).

(iv) **Bond Labels**
If *checked*, bond labels will be shown.

(v) **Point Labels**
If *checked*, point labels will be shown.

(vi) **Plane Labels**
If *checked*, plane labels will be shown.

(vii) **Constraint Labels**
If *checked*, constraint labels will be shown.

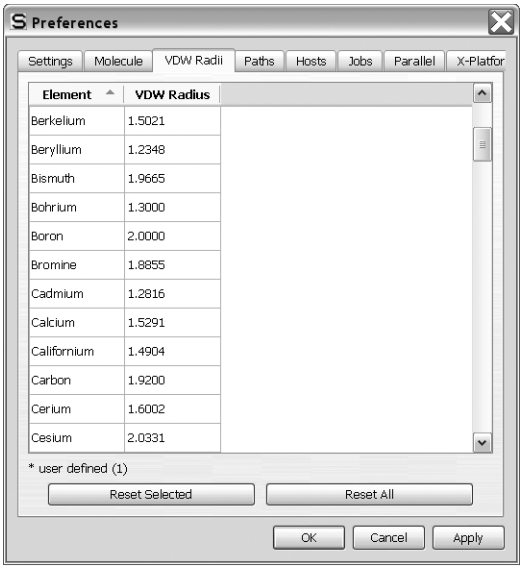
(viii) **Residue Labels**
If *checked*, residue labels will be shown.

(ix) **Reaction Labels**
If *checked*, reaction arrow labels will be shown.

(x) **CFD Labels**
If *checked*, CFD labels will be shown.

VDW Radii

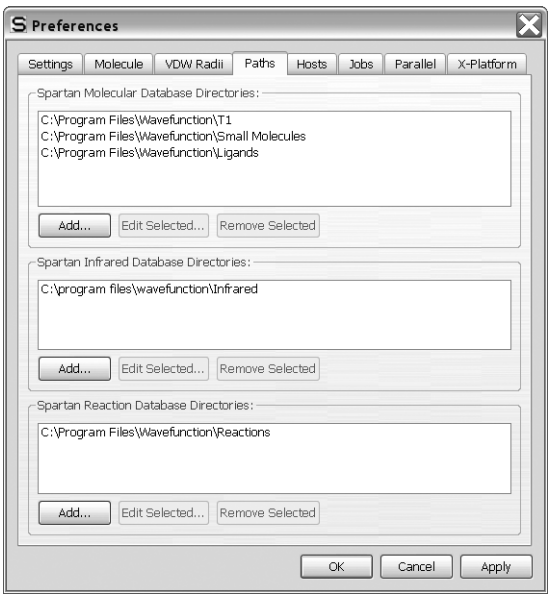
This provides a list of van der Waals radii



To order the list by element name *click* on **Element**, and by atomic radius *click* on **VDW Radius**. Individual entries may be changed from default values by first *clicking* on the entry and then entering a new value. The currently selected entry may be returned to its default radius by *clicking* on **Reset Selected** at the bottom of the dialog, and the full set of radii may be returned to their default values by *clicking* on **Reset All** at the bottom of the dialog.

Paths

This allows setting up of paths for SMD, SIRD and SRD database directories.



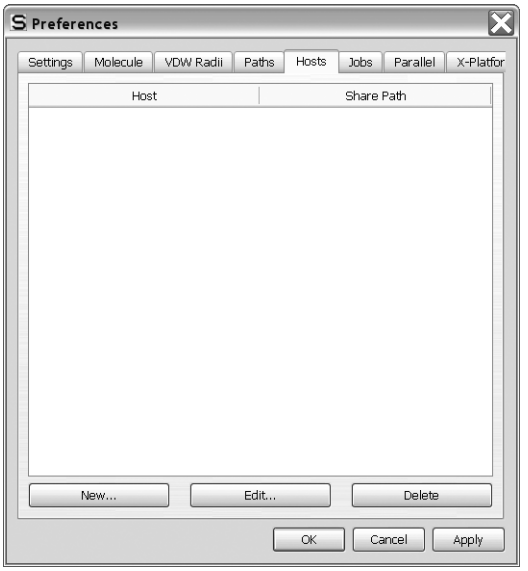
Molecular Database Directories defines the content of SMD, SIRD and SRD in terms of one or more directories. Multiple directories can be accessed at once. Specification of which directory (or directories) is to be searched in the **Search Options** dialogs associated with SMD, SIRD and SRD (**Databases** under the **Search** menu; **Chapter 16**). The default settings include:

Small Molecules, ***T1*** and ***Ligands***	for SMD
Infrared	for SIRD
Reactions	for SRD

meaning that SMD refers to the collection of small molecules, the T1 dataset and a collection of ligands, SIRD refers to a collection of infrared spectra calculations using the EDF2/6-31G* model and assuming conformations assigned by T1, and SRD refers to collections of transition states. Other collections may be added, for example, collections of diverse conformations for use in similarity analyses, or user-supplied collections. **Appendix G** discusses creation of user-defined databases.

Hosts

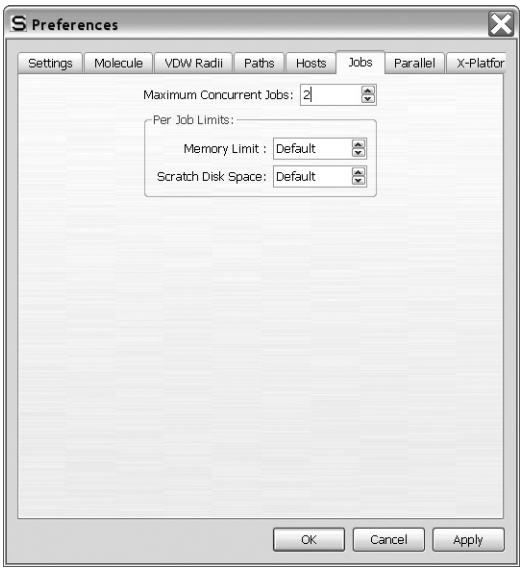
This allows specification of one or more computers that may act as servers.



To add a new host, *click* on the **New...** button and follow the directions provided in the **Host** dialog that appears.

Jobs

This allows setting of job limits.

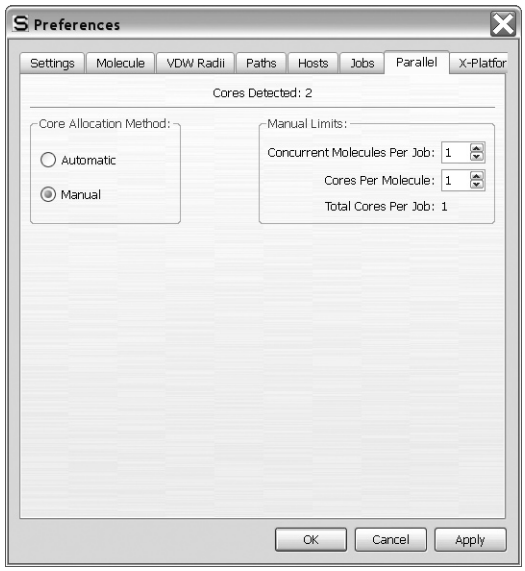


Maximum Concurrent Jobs designates the maximum number of jobs that can be run at one time and **Memory** and **Scratch Disk Space** maximum amounts of memory and disk for a job.

> You need to use caution that the maximum number of jobs does not significantly exceed the number of processors (cores) on your computer and also that you have sufficient memory to support this maximum number.

Parallel

This controls the use of a multi-core machine. The number of available cores is indicated at the top of the dialog.



(i) **Core Allocation Method**
Toggles between **Automatic** and **Manual**. **Automatic** will assign all available cores to a single molecule (single molecule inside a ***Spartan*** document) if there is only one molecule in the document and will divide the cores among the molecules in the document if there is more than one. **Manual** will follow settings under **Manual Limits**.

(ii) **Manual Limits**
Concurrent Molecules Per Job sets the number of jobs (molecules inside a ***Spartan*** document) that will run

simultaneously and **Cores Per Molecule** sets the number of cores used on a single molecule.

> If you are simultaneously working with two or more documents, you need to use caution in specifying more than one core/job. It is very easy (and very inefficient) to demand more cores than are actually available. This is not an issue if you are working only with a single document even if it contains more than one molecule.

X-Platform

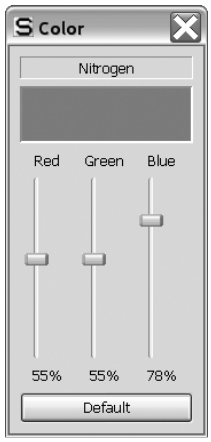
Items here refer to cross-platform applications and should not be of concern to users of a single platform.

(i) **Document Style**
Toggles between **File Based** and **Directory Based**. The latter is for compatibility between Windows and previous Macintosh and Unix/Linux versions of ***Spartan***.

(ii) **Dialog Style**
Menu selects between several dialog formats: **Windows**, **Macintosh**, and **KDE** and **Gnome** Linux. **Default** uses the style associated with the machine on which ***Spartan*** is installed.

Colors

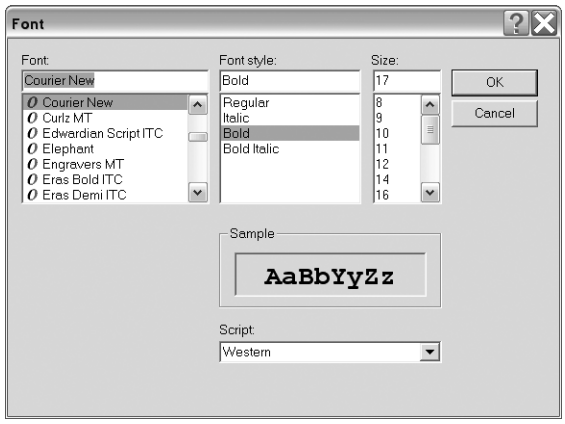
This alters default colors. Selection leads to the **Colors** dialog.



After selecting an object, its color may be set by adjusting the **Red**, **Green** and **Blue** slider bars. The default color may be reset by *clicking* on **Default**. Color selection applies to all objects of the same type, for example, all carbon atoms, and not just to the selected carbon. Further control of colors is available from **Utilities/Style** dialogs associated with **Properties** dialogs (**Properties** under the **Display** menu; **Chapter 16**). *Clicking* on ☒ removes the dialog.

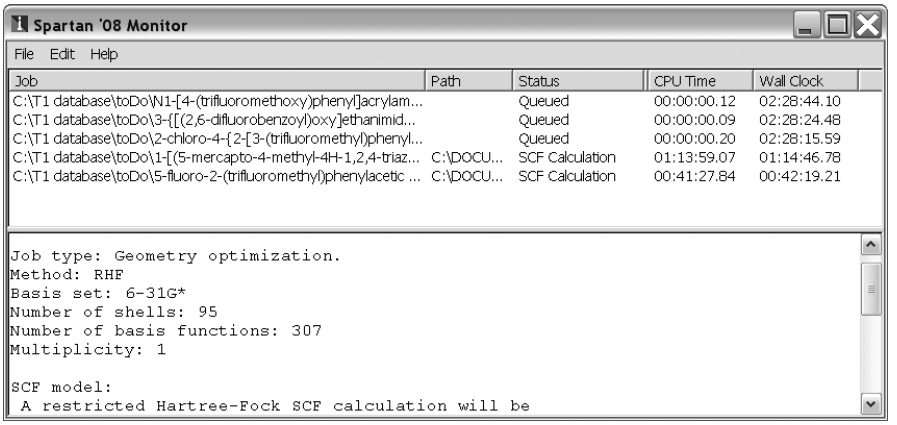
Fonts...

This selects fonts, style and size of labels attached to molecules (**Labels** and **Configure...** under the **Model** menu; **Chapter 12**), and plots (**Plots...** under the **Display** menu; **Chapter 16**). Selection leads to the **Fonts** dialog.



Selections are made from the **Font**, **Font Style** and **Size** menus. *Clicking* on **OK** dismisses the dialog with selections kept. *Clicking* on **Cancel** or on ☒ dismisses the dialog but selections are lost.

Monitor



This provides a listing of all executing/queued jobs and their status. To see accumulated output for an executing job, *click* on its name. To kill a job, *click* on its name, and then select **Kill Selected** from the **Edit** menu at the top of the dialog. To start a queued job (irrespective of the imposed queue limits; see previous discussions under **Jobs** and **Parallel**), *click* on its name and then select **Start Selected** from the **Edit** menu.

The **Monitor** may be removed either by selecting **Exit** from the **File** menu or by *clicking* on ☒ at the top of the dialog.

Calculator

Selection brings up a **Calculator**.

This functions the same way as a normal pocket calculator. The **Calculator** is removed by *clicking* on ☒.

Model Kit

If *checked*, this signifies that a model kit (organic, inorganic, peptide, nucleotide or substituent) is to remain on screen. Does not apply to ChemDraw.

File Toolbar, Geometry Toolbar, Build Toolbar, Search Toolbar

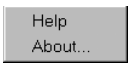
If *checked*, this provides display of toolbars that access functions contained in the **File**, **Build**, **Geometry** and **Search** menus, respectively.

Cascade, Tile

Arranges open windows in a cascade or as tiles on top of ***Spartan***'s main window (**NAM**).

Chapter 19
The Help Menu

This chapter describes help.



Help

This provides information relating to application of computational methods available in ***Spartan***, as well as technical details regarding the program's operation.

A number of topics are dealt with under **Help**; including ***Mouse/ Keyboard Functions***, ***General Operating Features***, ***Units***, ***Differences between Spartan'06 and '08***, ***Selecting a Theoretical Model***, ***Relative Computational Times***, ***Memory Usage***, ***Calculation Options***, and a number of FAQ documents (***Energy Questions***, ***Convergence Questions***, ***Questions about Basis Sets***, ***Problems with Memory***, ***Questions about the Property Module***, ***Questions about Conformations and Energy Profiles, MMFF94 Compliance***, ***Units***) related to computational questions. **Help** also provides a link to Wavefunction's website. Help files are HTML documents and require that Internet Explorer is installed to access them.

The ***Spartan Tutorial and User's Guide*** (this document) and ***A Guide to Molecular Mechanics and Quantum Chemical Calculations*** are also available (as PDF files) under **Help**.

Finally, note that several dialogs, in particular, the **Calculations** dialog incorporate imbedded help messages. *Clicking* on ? at the upper right, followed by *clicking* on a menu, button, etc. in the dialog gives rise to a brief informative message about the object queried.

About...*

Provides information about the user's release of ***Spartan***.

* About is located under the **Spartan 08** menu on the Macintosh version.

Appendix A
Capabilities and Limitations

Essential Edition of *Spartan'08* for Windows and Macintosh

Spartan'08 for Windows and Macintosh is available in two versions: a full version, referred to as ***Spartan'08***, and a lower-cost version, referred to as the Essential Edition. The latter shares most of the same graphical user interface as the full version, although it does not support quantum chemical calculations beyond the Hartree-Fock level and does not provide for calculation of UV/visible or NMR spectra. It also does not support infrared spectral searching.

Only the full version of the Linux implementation of ***Spartan'08*** is available.

Molecular Mechanics Models[1]

The molecular mechanics module provides for the calculation of energy (a combination of strain energy and intramolecular interaction energy), equilibrium geometries and vibrational frequencies, as well as for conformational searching. SYBYL (Tripos, Inc.) and MMFF (Merck Pharmaceuticals) force fields are available. Energies from the latter may be corrected for the effects of aqueous solvent (MMFFaq). Structures are unaffected. There are no atom limits for molecular mechanics calculations.

Semi-Empirical Models[1]

The semi-empirical module provides for calculation of heats of formation, wavefunctions, equilibrium and transition-state geometries and vibrational frequencies, as well as for conformational searching.* MNDO, AM1, RM1 and PM3 models are supported. MNDO/d replaces MNDO for second-row (and heavier) main-group elements,

* Except generation of conformer libraries.

and PM3 has been parameterized for transition metals.* MNDO has been parameterized for H, He, Li-F, Al-Cl, Ca, Zn, Ge, Br, Cd, Sn, I, Hg and Pb; AM1 for H, B-F, Al-Cl, Zn, Ge, Br, Sn and I; RM1 for H, C–F, P, S, Cl, Br and I; PM3 for H-Ne, Mg-Ar, Ca, Ti-Br, Zr, Mo-Pd, Cd-I, Hf-Pt and Hg-Bi and Gd. All semi-empirical models involve a minimal valence basis set.

Preset limits for semi-empirical calculations are 300 atoms.

Hartree-Fock Models[1-3]

The Hartree-Fock module provides for calculation of energies and wavefunctions, equilibrium and transition-state geometries and vibrational frequencies, as well as for conformational searching.** Both closed-shell and open-shell (either ROHF*** or UHF) calculations are supported. A full listing of basis sets is provided later in this appendix.

Preset limits for Hartree-Fock calculations are 200 atoms and 2,000 basis functions.

Density Functional Models[1,2,4]

The density functional module provides for calculation of energies and wavefunctions, equilibrium and transition-state geometries and vibrational frequencies, as well as for conformational searching.** The module supports local density calculations and a wide variety of gradient connected and hybrid functionals. Most of these are ***not*** provided in the menu of functionals, but need to be entered into the **Options** line of the **Calculations** dialog (**Setup** menu; **Chapter 15**). This line may also be used to specify combinations of exchange and correlation components to be used in building a total functional. For example, the line **EXCHANGE = SLATER: 0.9, BECKE: 0.1** indicates a combination of 9 parts Slater exchange and 1 part Becke exchange. Similarly, the line **CORRELATION = VWN: 0.19, LYP: 0.81** indicates a combination of 19 parts VWN and 81 parts LYP correlation.

---

* Parameters for all available transition metals except Zn, Cd and Hg include d-type orbitals in the basis set.

** Except generation of conformer libraries.

*** ROHF calculations are not supported in the Essential Edition of ***Spartan***.

A full listing of available functionals and instructions for constructing custom functionals from available exchange and correlation components is provided in **Table A-1**. A full listing of available basis sets is provided later in this appendix.

Preset limits for density functional calculations are 200 atoms and 2,000 basis functions.

Møller-Plesset Models[1,2,4]

The Møller-Plesset module provides for calculation of energies and wavefunctions and equilibrium and transition-state geometries as well as for conformational searching* using either the MP2 or RI-MP2 models. Vibrational frequencies are also available from both models, but require numerical differentiation of analytical gradients. A full listing of available basis sets is provided later in this appendix.

Preset limits for MP2 and RI-MP2 calculations are 200 atoms and 2,000 basis functions.

MP3 and MP4 calculations are supported under **Advanced Correlated Models**.

Advanced Correlated Models[1,2,4]

The advanced correlated module provides for calculation of energies (only) using the MP3, MP4, CCSD, CCSD(T), CCSD(2), OD, OD(T), QCISD, QCISD(T), QCCD and QCCD(2) correlated models. A full listing of available basis sets is provided later in this appendix.

Depending on basis set, practical considerations limit application of these techniques to molecules comprising 10-20 heavy (non-hydrogen) atoms, although program limits are set much larger.

Thermochemical Recipes

The G2, G3 and G3(MP2) recipes are available for highly-accurate heat of formation calculations. Practical considerations generally limit application of the first two recipes to molecules comprising less than

---

* Except generation of conformer libraries.

Table A1: Available Exchange, Correlation and Combination Functionals

exchange functionals

HF	Fock exchange
Slater	Slater (Dirac 1930)
Becke	Becke (1988)
Gill96	Gill (1996)
GG99	Gilbert, Gill (1999)
B (EDF1)	Becke (uses EDF1 parameters)
PW91	Perdew, Burke, Ernzerhof (1996)

correlation functionals

VWN	Vasko, Wilk, Nusair #5
LYP	Lee, Yang, Parr
PW91	GGA91, Perdew
LYP	(EDF1)
P86	Perdew (1986)
PZ81	Perdew, Zungner
PBE	Perdue, Burke, Ernzerhof (1996)

combination functionals

B3PW91
B3LYP
B3LYP5 (Gaussian's version of B3LYP)
EDF1
EDF2
BMK
M05
M05-2X
M06
M06-2X
M06-HF
M06-L
ωB97
ωB97X
ωB97X-D

10 heavy atoms and the G3(MP2) recipe to molecules comprising less than 15 heavy atoms.

The T1[7] recipe has been developed to closely reproduce G3(MP2) heats of formation. It makes use of an HF/6-31G* geometry, instead of the MP2/6-31G* geometry, replaces the large basis set MP2 calculation in G3(MP2) by a dual-basis set RI-MP2 calculation, eliminates both the QCISD(T) calculation and the vibrational frequency calculation (needed to obtain zero-point energy and to correct the energy for finite temperature) and introduces an empirical correction based on the atom counts and Mulliken bond orders. The result is that T1 requires 2-3 orders of magnitude less computation time than G3(MP2). Calculations on molecules with 30-50 heavy atoms are practical.

T1 is limited to closed-shell, uncharged molecules comprising H, C, N, O, F, S, Cl and Br only, and that can be properly described in terms of a conventional valence structure. Extensions of T1 to B and Si are planned as is a "T1 like" recipe for transition states.

Excited-State Models

Two classes of models are available for calculations of excited states: configuration interaction (CI) models and time-dependent density functional (TDDFT) models. The former comprises three models: CIS (configuration interaction singles), CISD (configuration interaction singles with doubles correction) and RI-CIS(D), while the latter supports the same functionals or combinations of functionals available for ground state density functional calculations. At the present time, only the CIS and RI-CIS(D) models are available for geometry optimization and frequency calculations, all other models being restricted to single-point energy calculations. A full listing of basis sets is provided later in this appendix.

Preset limits for excited-state calculations are 200 atoms and 2,000 basis functions.

Basis Sets*

Available basis sets for Hartree-Fock, density functional, MP2 (RI-MP2) and advanced correlated models for ground states and CIS, CISD (RI-CIS(D)) and time-dependent density functional (TDDFT) models for excited states include the full range of Pople basis sets: minimal STO-3G, split-valence 3-21G, 6-31G, 6-311G and polarization 6-31G*, 6-31G**, 6-311G* and 6-311G**. These may be supplemented with additional polarization functions and/or with diffuse functions. Also available are the cc-pVDZ, cc-pVTZ, cc-pVQZ and aug-cc-pCVQZ basis sets and the G3Large basis set.

STO-3G is available for H–I, 3-21G for H–Cs, 6-31G and basis sets built on it for H–Kr, 6-311G and basis sets built on it for H–Kr and I, cc-pVDZ, cc-pVTZ, cc-pVQZ, and G3 Large for H–Ca and Ge–Kr and aug-cc-pCVQZ for H–Ar.

Dual-basis set techniques are supported for ground-state energy, equilibrium and transition-state geometry calculations for Hartree-Fock, density function and MP2 models. Also available are pseudopotentials for calculations on molecules incorporating heavy atoms (for which all-electron basis sets are not available).[4]

Use of minimal and split-valence basis sets for any but Hartree-Fock calculations generally leads to poor results and is not recommended.

Solvent Models

Spartan'08 supports two different empirical solvation models. The SM5.4 model[5] (available for H, C–F, S–Cl, Br and I) estimates the aqueous solvation energy. This in turn may be added to the gas-phase energy obtained from any molecular mechanics or quantum chemical model. The computational cost of SM5.4 is very low (on the order of that for an energy using a semi-empirical model).

The SM8 model[4,6] (available for H, C–F, S–Cl, Br and I) is applicable only to Hartree-Fock and density functional models. It leads to a

---

* Only the Pople basis sets (STO-3G, 3-21G, 6-31G and 6-311G and extensions) are available in the Essential Edition of ***Spartan***.

change in the wavefunction, meaning that equilibrium or transition-state geometry, properties, infrared spectra and graphical displays may be obtained in the presence of solvent in addition to energy. SM8 has been parameterized for water as well as a variety of organic solvents. A few are available in a menu (acetone, diethyl ether, DMF, DMSO, ethanol, methylene chloride, THF and toluene) while others may be specified using the **Solvent** keyword (**Appendix C**). The computational cost of SM8 is roughly that of an energy with the corresponding Hatree-Fock or density functional model.

Properties and Spectra

The properties module (that is automatically called from the molecular mechanics module or one of the quantum chemical modules) provides for text output printing, population analyses (Mulliken, natural bond orbital and based on fits to electrostatic potentials), evaluation of thermodynamic quantities (enthalpy, entropy and free energy), and calculation of the dipole moment.

The properties module is also responsible for calculating quantities related to infrared/Raman spectra (vibrational frequencies and infrared/Raman intensities), NMR spectra (chemical shifts and ^{13}C chemical shifts correlated for local environment) and UV/visible spectra (excited-state transition energies). IR spectra calculations may be carried out with molecular mechanics models, semi-empirical models, Hartree-Fock models, density functional models and MP2 and RI-MP2 models. NMR spectra calculations to Hartree-Fock and density functional models and UV/visible spectra calculations are limited to Hartree-Fock models (CIS for excited states) and density functional models (TDDFT for excited states).

Graphical Models

The graphics module provides for data preparation associated with the display as surfaces, property maps and slices of molecular orbitals, electron densities, spin densities, electrostatic potentials and local ionization potentials. The sizes of electron density surfaces (and of property maps based on electron density surfaces) may be chosen

either using a specific value of the density or a value that encloses a specific percentage of the total number of electrons. Accessible and inaccessible regions may be distinguished for electron density surfaces and all property maps based on electron density surfaces.

Similarity Analysis

The similarity analysis module evaluates and quantifies the extent to which molecules are similar based either on selected atomic centers or on selected chemical function descriptors (CFD's). In addition, it quantifies the degree of fit of a molecule into a chemical environment (a pharmacophore). Similarity analyses are carried out between one or more molecules or pharmacophores in a document (the template) and one or more molecules or pharmacophores in one or more documents (the library). Molecules in the library (only) may either correspond to single conformers or to collections of diverse conformers.

---

1. For a general discussion and assessment of the techniques and methods available in ***Spartan***, see: W.J. Hehre, ***A Guide to Molecular Mechanics and Quantum Chemical Calculations***, Wavefunction, Inc., Irvine, CA 2003. This is available as a PDF under the **Help** menu.
2. For a review of the quantum chemical methods utilized in ***Spartan*** (except for semi-empirical methods) with emphasis on recent developments, see: Y. Shao, L.F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S.T. Brown, A.T.B. Gilbert, L.V. Slipchenko, S.V. Levchenko, D.P. O'Neill, R.A. DiStasio Jr., R.C. Lochan, T. Wang, G.J.O. Beran, N.A. Besley, J.M. Herbert, C.Y. Lin, T. Van Voorhis, S.H. Chien, A. Sodt, R.P. Steele, V.A. Rassolov, P.E. Maslen, P.P. Korambath, R.D. Adamson, B. Austin, J. Baker, E.F.C. Byrd, H. Dachsel, R.J. Doerksen, A. Dreuw, B.D. Dunietz, A.D. Dutoi, T.R. Furlani, S.R. Gwaltney, A. Heyden, S. Hirata, C-P. Hsu, G. Kedziora, R.Z. Khalliulin, P. Klunzinger, A.M. Lee, M.S. Lee, W.Z. Liang, I. Lotan, N. Nair, B. Peters, E.I. Proynov, P.A. Pieniazek, Y.M. Rhee, J. Ritchie, E. Rosta, C.D. Sherrill, A.C. Simmonett, J.E. Subotnik, H.L. Woodcock III, W. Zhang, A.T. Bell, A.K. Chakraborty, D.M. Chipman, F.J. Keil, A.Warshel, W.J. Hehre, H.F. Schaefer, J. Kong, A.I. Krylov, P.M.W. Gill and M. Head-Gordon, *Phys. Chem. Chem. Phys.,* **8**, 3172 (2006).
3. For an older account see: W.J. Hehre, L. Radom, P.v.R. Schleyer and J.A. Pople, ***Ab Initio* Molecular Orbital Theory**, Wiley, New York, 1986.
4. Not available in essential edition.
5. C.C. Chambers, G.D. Hawkins, C.J. Cramer and D.G. Truhlar, *J. Phys. Chem.*, **100**, 16385 (1996).
6. A.V. Marenich, R.M. Olsen, C.P. Kelly, C.J. Cramer and D.G. Truhlar, *J. Chem. Theory Comput.*, **3**, 2011 (2007).
7. W.S. Ohlinger, P.E. Klunzinger, B.J. Deppmeier, W.J. Hehre; *J. Phys. Chem. A*., **113** 10, 2165 (2009).

Appendix B

Menus

Spartan Screen

File

New	Brings up a model kit for molecule building or substitution
Open...	Opens (imports) a molecule
Close	Closes a molecule
Save	Saves (exports) a molecule
Save As...	Saves a molecule as a ***Spartan*** document under a user-specified name; creates files for use in the Spartan Molecular Database, Spartan Reaction Database and Spartan Infrared Database
New Molecule	Adds a molecule to an existing list; brings up a model kit for molecule building or substitution
Delete Molecule	Deletes a molecule (or molecules) from a list
Append Molecule(s)...	Appends molecules to an existing list
Print...	Prints on-screen display; also prints contents of output window and the **Spreadsheet**
Embedded Data	Associates external files with a ***Spartan*** document, either at molecule or document level (or both)
Start/Stop QuickTime Recording	Starts and stops QuickTime recording of contents of main ***Spartan*** screen (only)
Exit	Exits ***Spartan***

Edit

Undo	Undoes previous operations
Cut	Moves the current molecule or contents of the selection box to the clipboard
Copy	Copies the current molecule or contents of the selection box to the clipboard
Paste	Pastes contents of the clipboard to the screen
Find...	Locates a text string in the output dialog or an on-screen molecular fragment
Find Next	Locates next occurrence of a text string or molecular fragment
Center	Centers the molecule on screen; applies to all molecules in a document
Clear	Clears the selected molecule

Model

Wire	Displays structure as wire-frame model
Ball and Wire	Displays structure as ball-and-wire model
Tube	Displays structure as tube model
Ball and Spoke	Displays structure as ball-and-spoke model
Space Filling	Displays structure as space-filling model
Line	Displays structure as a line drawing
Hide	Hides structure model from view
Global Model	Applies model type and labels of current molecule to all molecules in the document
Coupled	Couples motions of all molecules in the document
Hydrogens	Toggles hydrogens “on” and “off”
Labels	Toggles labels “on” and “off”
Ribbons	Toggles ribbons “on” and “off”
Hydrogen Bonds	Toggles hydrogen bonds “on” and “off”
CFD’s	Toggles chemical function descriptors (hydrophobes, hydrogen-bond donors/

	acceptors, positive/negative centers) "on" and "off"
Configure...	Labels atoms, bonds, etc., provides information about polypeptides/polynucleotides residues and designates ribbon displays

Geometry

Measure Distance	Displays and/or sets bond distance
Measure Angle	Displays and/or sets bond angle
Measure Dihedral	Displays and/or sets dihedral angle
Freeze Center	Freezes selected atomic positions
Set Torsions	Selects bond rings for conformational searching
Set Similarity Centers	Identifies atoms /CFD's for similarity analyses
Constrain Distance	Constrains bond distance
Constrain Angle	Constrains bond angle
Constrain Dihedral	Constrains dihedral angle
Define Point	Defines a point as a geometric mean of a set of atoms; defines ligand point as a position that is perpendicular to the centroid of a plane made by three or more atoms
Define Plane	Defines a plane made by three or more atoms
Define CFD	Defines the location of a CFD
Align	Aligns molecules in a document according to selected atoms, structure, or CFD's

Build

View	Removes the model kit (unless designated for permanent display)
Add Fragment	Brings up a model kit that contains libraries of atomic fragments, functional groups, rings, ligands and substituents; accesses user-defined libraries; provides access to ChemDraw (**NAL**, **NAM**)
Delete	Deletes atoms, bonds, points, planes, etc.

Make Bond	Makes bonds between free valences or atoms
Break Bond	Breaks a bond
Minimize	Performs energy minimization using MMFF molecular mechanics

Setup

Calculations...	Sets up molecular mechanics and quantum chemical calculations and similarity analyses; specifies calculation of IR, NMR and UV/vis spectra
Surfaces	Sets up generation of and displays graphical surfaces
Submit	Submits job to the execution queue

Display

Output	Displays text output
Properties	Displays molecule, bond and atom properties as well as information about CFD's, geometrical constraints, graphical surfaces, plots (including spectral plots) and statistical analyses
Surfaces	Sets up generation of and displays graphical surfaces (same as entry in **Setup** menu)
Spectra	Displays IR, NMR and/or UV/visible spectra, animates vibrational modes, and accesses on-line experimental spectral databases; fits calculated spectra to experimental spectra and (for IR) to reference spectrum
Spreadsheet	Displays spreadsheet
Plots...	Creates plots from the data in the spreadsheet
Similarities	Displays results of similarity analyses
Reactions	Calculates reaction (activation) energies using data either from current document or from the Spartan Molecular Database

Search

Databases	Performs substructure search on Cambridge Structural Database (CSD), Spartan Molecular Database (SMD) and Spartan Reaction Database (SRD); performs spectral search of the Spartan Infrared Database (SIRD); provides on-line access to Protein Data Bank (PDB)
Ligands	Extracts ligands and pharmacophores based on protein-ligand binding from PDB files
Transition States	Provides transition-state guess based on reaction database or, lacking a database entry, based on linear synchronous transit
Tautomers	Identifies tautomers

Options

Preferences...	Sets various run-time and labeling preferences and specifies execution preferences and run-time queue; invokes parallel calculation
Colors	Sets screen and model colors
Fonts...	Sets fonts for labels and plot displays
Monitor	Monitors and allows for killing executing jobs and bypassing queue to start jobs
Calculator	Pocket calculator
Model Kit	Keeps a model kit on screen
File, Build, Geometry, Search Toolbar	Toggles toolbars "on" and "off"
Cascade	Arranges open windows in a cascade
Tile	Arranges open windows as tiles (**NAL**, **NAM**)

Help

Help	Provides information about the performance and timing of computational methods in ***Spartan***; provides information about using

	graphical models in ***Spartan***; bulletin board for FAQs about ***Spartan***
About...	Provides program version information for citation and support

Contextual

Main Screen

Copy	Copies selected molecule to the clipboard
Paste	Pastes the contents of the clipboard into the selected document
Delete Selected	Deletes selected molecule from document
Properties	Brings up the **Molecular Properties** dialog

Spreadsheet

Copy	Copies text of selected cell or cells to the clipboard. If leftmost cell (or cells) selected, copies molecule(s) to the clipboard
Paste	Pastes the contents of the clipboard into selected cells. If leftmost cell (or cells) selected, either pastes text or molecule(s) depending on menu choice
Add	Brings up the **Add** dialog (spreadsheet) for adding calculated quantities into the spreadsheet
Sort	Sorts the column from low to high. *Pressing* the **Shift** key prior to menu selection sorts from high to low
Format Selected	Formats selected cell(s), selected column(s) if selection is in a header cell, or entire spreadsheet if selection is header cell of leftmost column
Delete Selected	Deletes selected molecule(s) from document
Append	Appends the contents of ***Spartan*** document(s)

	to the spreadsheet (corresponding to the selected document)
Rename Selected Using SMD	Rename selected molecule(s) with names in the Spartan Molecular Database; appears only when leftmost cell(s) selected
Properties	Brings up the **Molecular Properties** dialog

Substituent Model Kit

Copy	Copies selected molecule(s) from a custom library onto the clipboard
Paste	Pastes molecule(s) from the clipboard into a custom library
Delete	Deletes selected molecule(s) from the custom library
Append	Appends the contents of ***Spartan*** document(s) to the custom library; brings up a file browser

Reactions

Copy	Copies selected text to the clipboard
Print	Prints selected text

Output Window

Copy	Copies selected text to the clipboard
Print	Prints selected text

Appendix C

Commonly-Used Program Options

This appendix describes a number of commonly-used program options. These are specified using keywords input into the **Options** box in the **Calculations** dialog (**Setup** menu). In addition to the above, any method, functional or combination of functionals and/or basis set supported in ***Spartan*** (and not included explicitly in the menus) may be typed into the **Options** box. For example, typing **EDF2 6-311G(2df,2pd)** will result in density functional calculation using the EDF2 functional and the 6-311G(2df,2pd) basis set. (See **Chapter 15** for further details.)

Keywords are case insensitive. An equals sign (=) separates a keyword from its integer, real or character string value. Keywords may either be single words or expressions. **Keyword=N** indicates an integer argument, **keyword=C** indicates a character argument, and **keyword=F** indicates a floating point argument. Real values can optionally include an **eNN** or **e-NN** (floating point power of ten). Default values for keywords are indicated in **bold**, and alternative values in *italics*.

Conformational Search

keyword	value	description
SearchMethod	**C**	conformational searching method employed; overrides default choice which depends on complexity of system may be overridden with one of the following: *Systematic* use systematic method *MonteCarlo* use Monte-Carlo method *Sparse* use systematic method but randomly eliminate conformers

		to stay within set conformer limits. Used in conformer library generation.
MaxEnergy	**F**	sets the maximum energy (in kJ/mol) at which a conformer will be kept to (minimum energy + F), where minimum energy is the energy of the lowest energy conformer encountered at that point in the search; default is **40** kJ/mol
StartTemperature	**F**	sets initial temperature (in K) for a Monte Carlo search (only) to F; high temperatures will result in the molecule further exploring global space, while low temperatures will result in the molecule further exploring local space; default is **5000**K

Similarity Analysis

keyword	value	description
SingleEnantiomer	**–**	Limits similarity analysis to the enantiomer actually in specified library documents. The default is to consider both enantiomers
QuitSimilarity	**N**	Quit similarity analysis involving a conformer list (in the graphical user interface) and/or one or more conformer libraries when N good matches have been identified; default is **1**
QuitCriterion	**F**	R^2 criterion for a good match; default is **0.9**

Geometry Optimization

keyword	value	description
GeometryCycle	**N**	set maximum number of equilibrium and transition-state geometry optimization cycles to N; default is **number of independent variables +20** for equilibrium geometry optimization and **3x number of independent variables** for transition-state geometry optimization
GradientTolerance	**F**	set convergence criterion in equilibrium and transition-state) geometry optimization for the maximum gradient component (in hartrees/bohr) to F; default is **4.5×10^{-4}**

DistanceTolerance	**F**	set convergence criterion in equilibrium and transition-state geometry optimization for the maximum change in a bond length (in Å) to F; default is **1.8 x 10^{-3}**
Hess	**C**	Choose Hessian (matrix of second derivatives) for start of geometry (transition-state geometry) optimization (in the absence of a Hessian from a previous calculation, the default is **MMFF** molecular mechanics for geometry optimization and **PM3** semi-empirical for transition-state geometry optimization; may be overridden with one of the following:

unit	unit matrix
MMFF	MMFF molecular mechanics
AM1	AM1 semi-empirical
STO-3G	STO-3G Hartree-Fock
3-21G	3-21G Hartree-Fock
*6-31G**	6-31G* Hartree-Fock

Molecular Mechanics Calculations

keyword	description
MMFF94s	Use MMFF94s method insted of MMFF94. The only significant difference between the two is that MMFF94 forces nitrogens attached to aromatic rings to be planar (as in aniline). While this is widely perceived to be the case, in fact the nitrogen in aniline and related molecules are pyramidal not planar.

Quantum Chemical Calculations

keyword	value	description
Diis	**N**	switch on diis all the time in SCF procedure. N is the size of the iterative subspace; it should be an integer between 2 and 10; default is **5**
NoDiis		turn off diis SCF convergence accelerator
ScfCycle	**N**	set maximum number of SCF iterations to N; default is **50**
Guess	**C**	choose initial wavefunction guess; in the absence of a guess from a previous calculation, the

default is either **sad** or **PM3** depending on the calculation type, and may be overridden with one of the following:

core	diagonalize the core Hamiltonian
sad	superposition of atomic densities
PM3	PM3 semi-empirical

Scf **C** SCF procedure; default is **restricted** for closed-shell systems and **unrestricted** for open-shell systems, either of which may be specified for both closed and open-shell systems

SmallBasis **C** specifies the small basis set to be used in dual basis set calculations. The large basis set is that indicated in the menu.

Core **C** use of frozen core approximation in Møller-Plesset and advanced correlated calculations; default is **frozen** which may be changed to *thawed*

ScfTolerance **C** control of all tolerances in SCF procedure and to affect the precision of the algorithm; default is **normal** which may be changed to *high* or *veryhigh*

F set SCF energy convergence criterion (in hartrees) to F; default is **1.0 x 10^{-6}**

BigGrid **–** uses very large grid in density functional calculations

Mix **–** specifies that α and β HOMO's in the guess wavefunction should be constructed according to:

$$HOMO_{\alpha} = \frac{HOMO + LUMO}{\sqrt{2}}$$

$$HOMO_{\beta} = \frac{HOMO - LUMO}{\sqrt{2}}$$

useful for generating a guess wavefunction for singlet diradical

Exchange **E** specifies custom exchange functional or combination of exchange functionals for density functional calculation. See **Appendix A** for details.

Correlation	**C**	specifies custom correlation functional or combination of correlation functionals for density functional calculation. See **Appendix A** for details.
Energy Only		skips calculation of the density matrix for an MP2 energy calculation, reducing the time by 50%

Semi-Empirical Calculations

keyword	value	description
NoAmideFix	**–**	turns "off" molecular mechanics amide correction in PM3 calculations
hh	**C**	turns "on" and "off" hydrogen-hydrogen repulsion term in PM3 calculations; default is **on** only where transition metals are involved and **off** otherwise
Polarizability	**–**	calculates and prints polarizabilities and hyperpolarizabilities

Property and Spectra Calculations

keyword	value	description
NBO	**C**	specify details of NBO population analysis; default is **normal** which may be overridden for *ionic* and *3C* (three center) systems
NoElCharge	**–**	turns "off" calculation of electrostatic charges
ElCharge	**N**	adjusts size of grid used to calculate charge from electrostatic potential; default is **1** point/atomic unit
Temperature	**F**	temperature (in K) used in calculation of thermodynamic properties; default is **298**°K
TEMPRANGE	**Fi,Fj,Fk**	temperature range applied to thermodynamic properties when IR frequencies are calculated, where i=starting temperature, j=ending temperature, and k=increment, all values in K
Pressure	**F**	pressure (in atm) used in calculation of thermodynamic properties; default is **1** atm
Moments	**–**	calculates and prints principal moments of inertia

NoSolvent	**–**	turns "off" calculation of aqueous solvation energy using SM5.4 model
Solvent	**C**	specifies solvent for SM8 model. See **FAQ** (**Properties**) available under **Help** (**Help** menu; **Chapter 19**) for a list of available solvents.
NoAreaVolume	**–**	turns "off" calculation of surface area and volume
MaxPropWeight	**N**	sets limit of molecular weight for electrostatic charge, solvent and area/volume calculations; default is **500** amu
Polarizability	**F**	calculates and prints either static (F not specified) or frequency dependent (frequency F in cm^{-1}) polarizabilities and hyperpolarizabilities
FreqScale	**F**	scales calculated vibrational frequencies by F; default is **1**
NumericalFreq	**–**	uses numerical differentiation (of analytical gradients) for frequency calculation
UVstates	**N**	specifies number of states to be examined in UV/visible spectra calculation; default is **5**
IRCSteps	**N**	maximum number of steps allowed in constructing IRC; default is **40**
Include Triplets		include excited triplet states in a UV spectra calculation where the ground state is a singlet. Ignore where the ground state is a triplet.
IRCStepSize	**N**	maximum step size (in mass-weighted atomic units x 100) for IRC; default is **150**
DropVibs	**N**	ignore all vibrational modes with frequencies below N cm^{-1} in calculating thermodynamic products
ClampThermo		do not allow individual vibrational contributions to the temperature correction to the enthalpy to exceed 1/2 RT, to the entropy to exceed 1/2 R, and for the heat capacity to exceed 1/2 R.
AvgMass		use average mass instead of mass of dominant isotope in calculating thermodynamic properties

Miscellaneous and Printing

keyword	value	description
3DPlot	–	treat dynamic constraints independently (rather than in concert). This allows data for 3D (XYZ) plot to be collected. Two dynamic constraints only.
Execute	**C**	either skip execution (**skip**) or force execution regardless of history (**force**)
PrintVibCoords	–	print coordinates corresponding to normal mode vibrational frequencies
PrintCoords	–	prints Cartesian coordinates
PrintLev	**C**	print level; default is **normal** which may be changed to *verbose*
Print OrbE	–	print energies of molecular orbitals
PruneVirtual	**N,C**	keep N virtual orbitals; default is **10**; **None** keeps all virtual orbitals

Appendix D

Units

Geometries

Cartesian coordinates are given in Ångstroms (Å), and in atomic units (au).

Bond distances are given in Å and in au. Bond angles and dihedral angles are given in degrees (°).

Surface areas, accessible surface areas and polar surface areas are available in $Å^2$ and volumes in $Å^3$, and in au^2 (au^3).

1 Å = 0.1 nm=1.889762 au

Energies, Heats of Formation and Strain Energies, Zero-Point Energies, Enthalpies and Gibbs Energies and Entropies

Total energies from Hartree-Fock calculations are available in au, kcal/mol, kJ/mol and electron volts (eV).

Experimental heats of formation as well as those from semi-empirical calculations and from thermochemical recipes are available in kJ/mol, au, kcal/mol and eV.

Strain energies from molecular mechanics calculations are available in kJ/mol, au, kcal/mol and eV.

Energies, heats of formation and strain energies corrected empirically for the effects of aqueous media are given in the same units as the corresponding gas-phase quantities.

Zero-point energies, enthalpies and Gibbs energies available in kJ/mol, kcal/mol and au/mol. Entropies are available in kJ/mol•degree, kcal/mol•degree and au/mol•degree.

Orbital Energies

Orbital energies are available in eV, kcal/mol, kJ/mol and au.

Energy Conversions

	au	kcal/mol	kJ/mol	eV
1 au	-	627.5	2625	27.21
1 kcal/mol	1.593 (-3)	-	4.184	4.337 (-2)
1 kJ/mol	3.809 (-4)	2.390 (-1)	-	1.036 (-2)
1 eV	3.675 (-2)	23.06	96.49	-

a) exponent follows in parenthesis, e.g., 1.593 (-3) = 1.593×10^{-3}

Electron Densities, Spin Densities, Dipole Moments, Charges, Electrostatic Potentials and Local Ionization Potentials

Electron densities and spin densities are given in electrons/au^3.

Dipole moments are given in debyes.

Atomic charges are given in electrons.

Electrostatic potentials are given in kJ/mol.

Local ionization potentials are given in eV.

Vibrational Frequencies

Vibrational frequencies are given in wavenumbers (cm^{-1}).

Chemical Shifts, Coupling Constants

Chemical shifts are given in parts-per-million (ppm) relative to appropriate standards: hydrogen, tetramethylsilane; carbon, tetramethylsilane; nitrogen, nitromethane; fluorine, fluorotrichloromethane; silicon, tetramethylsilane, phosphorous, phosphoric acid. These are available for Hartree-Fock, BP, BLYP, B3LYP, EDF1 and EDF2 models with the 6-31G*, 6-31+G*, 6-31G**, 6-31+G**, 6-311G, 6-311+G*, 6-311+G** and 6-311++G** basis sets.

Coupling constants are in ppm.

UV/visible Spectra

λ_{max} is given in wavenumbers (cm^{-1}).

Appendix E

Citation

The proper citation for ***Spartan'08*** is as follows:

Spartan'08
Wavefunction, Inc.
Irvine, CA

Except for molecular mechanics and semi-empirical models, the calculation methods used in ***Spartan*** have been documented in: Y. Shao, L.F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S.T. Brown, A.T.B. Gilbert, L.V. Slipchenko, S.V. Levchenko, D.P. O'Neill, R.A. DiStasio Jr., R.C. Lochan, T. Wang, G.J.O. Beran, N.A. Besley, J.M. Herbert, C.Y. Lin, T. Van Voorhis, S.H. Chien, A. Sodt, R.P. Steele, V.A. Rassolov, P.E. Maslen, P.P. Korambath, R.D. Adamson, B. Austin, J. Baker, E.F.C. Byrd, H. Dachsel, R.J. Doerksen, A. Dreuw, B.D. Dunietz, A.D. Dutoi, T.R. Furlani, S.R. Gwaltney, A. Heyden, S. Hirata, C-P. Hsu, G. Kedziora, R.Z. Khalliulin, P. Klunzinger, A.M. Lee, M.S. Lee, W.Z. Liang, I. Lotan, N. Nair, B. Peters, E.I. Proynov, P.A. Pieniazek, Y.M. Rhee, J. Ritchie, E. Rosta, C.D. Sherrill, A.C. Simmonett, J.E. Subotnik, H.L. Woodcock III, W. Zhang, A.T. Bell, A.K. Chakraborty, D.M. Chipman, F.J. Keil, A.Warshel, W.J. Hehre, H.F. Schaefer, J. Kong, A.I. Krylov, P.M.W. Gill and M. Head-Gordon, *Phys. Chem. Chem. Phys.,* **8**, 3172 (2006).

A discussion and assessment of commonly-used calculation methods is found in: W.J. Hehre, ***A Guide to Molecular Mechanics and Quantum Chemical Calculations***, Wavefunction, Inc., Irvine, CA, 2003. This is available as a PDF under the **Help** menu.

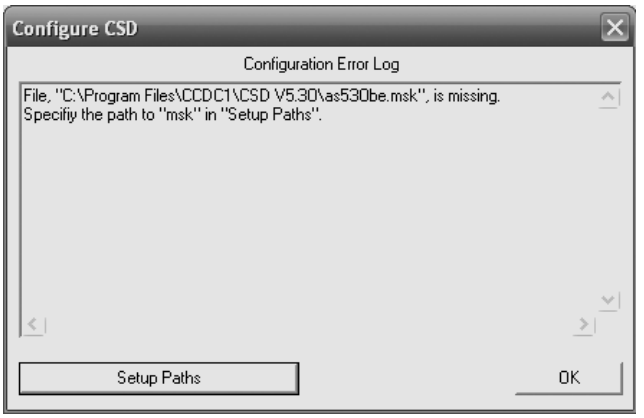
Appendix F

Installing the Cambridge Structural Database

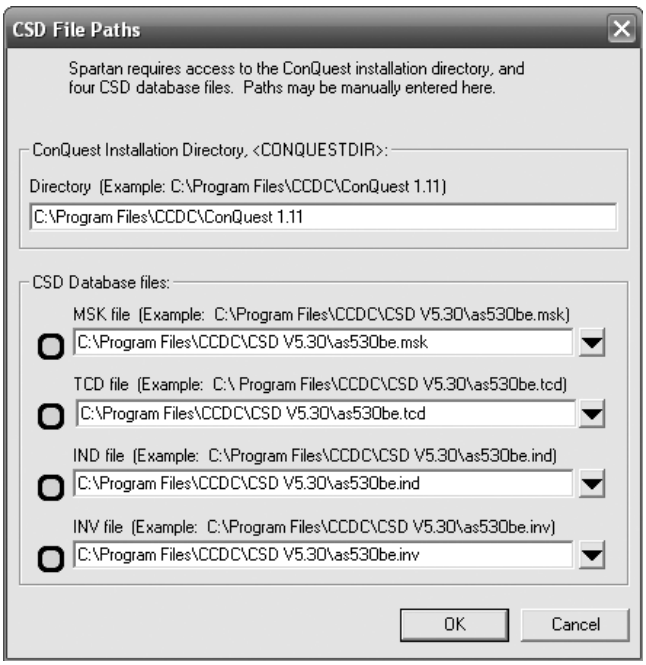
Access to the Cambridge Structural Database (CSD) from ***Spartan*** should be automatic, following successful installation of the ConQuest program (supplied with CSD) according to the detailed instructions provided. If access is denied, the following steps need to be taken:

Windows:

(i) *Click* on the square box immediately to the left of the **Retrieve** button at the bottom of the **CSD** dialog. This brings up the **Configure CSD** dialog,



This will list any files which are missing. *Click* on **Setup Paths** at the bottom of the dialog to bring up the **CSD File Paths** dialog,



This dialog may show one or more red lights.

(ii) Make certain that the path to ConQuest is correctly specified. The most common error will be that the version number is not current. (At the time of writing, the current version is 1.11.) You can get the current version number from your installation of ConQuest.

(iii) Given the proper version number for ConQuest, update the file names in the boxes below **MSK file**, **TCD file**, **IND file** and **INV file**. You can get the current file names from your installation of ConQuest. As you enter the current file names, the red lights turn to green lights. When all lights are green, *click* on **OK** to dismiss the **CSD File Paths** dialog and then *click* on **OK** to dismiss the **Configure CSD** dialog.

Linux:

(i) **CSD HOME** must be defined in order for ***Spartan*** to locate the CSD program.

(ii) Please refer to **Chapter 9** of "Cambridge Structural Database System Release and Installation Notes" for instructions on setting up **CSD HOME**.

Appendix G

Constructing Custom Databases

Custom collections may be added to SMD, SIRD and SRD. These may either include the geometry, energy (heat of formation, strain energy), HOMO and LUMO energies, dipole moment, atomic charges, chemical shifts and vibrational frequencies obtained from quantum chemical calculations, or (for SMD and SIRD only) conformer libraries generated from molecular mechanics calculations. Entries for SRD (transition states) need to derive from ***Spartan's*** transition state guessing facility (**Transition States** under the **Search** menu; **Chapter 17**) in order that they include the original set of reaction arrows.

Molecules (or lists of molecules) that are to be placed in a user database are first saved using **Save As** under the **File** menu with **Save as type** set to **Spartan Database**. This results in two files, a **.spentry** and a **.spindex** file for each molecule (or list). These should then be put in a single directory, the path to which must be entered into the **Paths Preferences** dialog under **Preferences...** in the **Options** menu (**Chapter 18**).

Note that it is important to specify a label for each molecule. Label specification is done either in the **Molecule Properties** dialog (see **Properties** under the **Display** menu in **Chapter 16**) or in the spreadsheet (see **Spreadsheet** under the **Display** menu in **Chapter 16**).

Appendix H

Pharmacophore Input

The following file format is used for input of a pharmacophore (illustrated for a pharmacophore with four hydrophobes and one positive ionizable center) :

```
Blank line
Blank line
Blank line
ENDCART
BEGINCFD
HPHOBE|RING_AROM   8.050  -4.180  -0.400  0  2.000  1.000  1
HPHOBE|RING_AROM   9.630  -0.640  -3.590  1  2.000  1.000  1
HPHOBE|RING_AROM  10.280  -0.380   3.200  2  2.000  1.000  1
HPHOBE|RING_AROM   8.130  -2.300   4.330  3  2.000  1.000  1
POS               14.370  -1.700   0.280  4  3.000  1.000  1
ENDCFD
```

The lines between BEGINCFD and ENDCFD define the individual pharmacophore elements (one line/element):

type1|type2… (| separates types for multiple definitions), Cartesian (X,Y,Z) coordinates (in Angstroms), unique number, radius of element (in Angstroms), weight of element in fit, use the element in fit (1) or ignore the element (5)

The following element types are available:

HPHOBE	hydrophobe
RING_AROM	aromatic ring
POS	positive ionizable center
NEG	negative ionizable center
HBA_CENTER	hydrogen bond acceptor
HBD_CENTER	hydrogen bond donor
EXCL_VOL	excluded volume

Appendix I
Input of Experimental Spectra

The databases of infrared and UV/visible spectra accessible on-line from ***Spartan*** are in JCAMP format. Specification of both and examples can be found at the NIST website (NIST.gov).

The database of ^{13}C spectra is in CML (Chemical Markup Language) format. A description may be found at http://www.ch.ic.ac.uk/omf/cml. A sample file is provided below.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<cml>
  <spectrum type="NMR">
    <metadataList>
      <metadata name="nmr:OBSERVENUCLEUS" content="13C"/>
    </metadataList>
    <peakList>
      <peak xValue="125.1399993896484" peakHeight="0.458" />
      <peak xValue="128.4499969482422" peakHeight="0.558" />
      <peak xValue="129.0099945068359" peakHeight="0.658" />
      <peak xValue="21.64999961853027" peakHeight="0.758" />
      <peak xValue="137.8399963378906" peakHeight="0.858" />
    </peakList>
  </spectrum>
</cml>Blank line
```

This contains five shifts (xValue) and five associated intensities (peakHeight).

Appendix J
Installing a Network HASP

Spartan'08 requires a hardware license key (HASP) which is normally installed on each machine. However, a network hardware key (NetHASP) can be obtained from Wavefunction and installed on a server. This allows ***Spartan'08*** to be run on a specified number of client machines that are on the network. A one-time installation is required.*

On the Server:

1. The License Manager software must be installed. To do so, please run the **lmsetup.exe** file.
   a. (Recommended). The latest release of this file can be obtained by visiting the following link at the Aladdin Web Site:

      http://www.aladdin.com/support/hasp/enduser.asp#lm

   b. The **lmsetup.exe** file may also be obtained directly from a subdirectory in the ***Spartan*** folder if you install the program directly onto your server . . . if you opt to install the program on your server, run through the installation procedure guided via the installation wizard. Upon completion, **lmsetup.exe** can be found in the following subdirectory:

      SPARTAN'08** : C:\...\Wavefunction\Spartan08Vxxx\Network\SERVERS\

2. It is suggested that you choose to configure the program to run as a 'service' during the installation procedure. When the program is finished installing, restart the server and verify that the NetHASP

---

* The network HASP license management option requires a Windows machine to act as the license manager, client machines may be either Windows or Macintosh (**NAL**).

** ***Spartan*** has multiple Windows release versions, the pathname depends on which version you are running (Spartan'08 v.1.0.0 program directory = Spartan08V100) – Each contains the Network and SERVERS folders. Your server must be running Windows 2000 or higher.

License Manager has loaded the necessary protocols for your network.

3. If you are running a Firewall, open ports 475 and 1947. (Used by default, but also configurable).

For additional server options, you can also configure an NHSRV.ini file. Such options include specifying which client machines can access Spartan by IP range.

On Each Client Computer:

1. Install software locally on the network clients that will access the NetHASP through the network. Follow the installation procedure as guided by the installation wizard.

A network configuration file (nethasp.ini) must be added to the directories described below. Several variations of the file can be downloaded from the Aladdin website or you can use the template that is installed with Spartan:

Aladdin: http://www.aladdin.com/support/hasp/enduser.asp#ini

Spartan: The template nethasp.ini can be found in:

C:\...\Wavefunction\Spartan08Vxxx\Network\

2. Locate the NETHASP.ini file and open with a text editor (ex. Notepad). The line that requires modification is the ";;NHSERVER_ADDR = <Addr1>, <Addr2>". Delete the two semicolons (;) and delete the <Addr1>, <Addr2>. Insert the IP address of your server on the right of the =. If you have more than one NETHASP installed, subsequent server IP addresses may be entered by placing a comma after the first address, followed by a space. After these modifications are completed, save the file.

   Next, right Click on this nethasp.ini file and select 'Copy'. Move up one level to the ***Spartan*** folder. Right click on the background and select 'Paste'. Open both the P2 and P4 folder and also right click and 'Paste' the NETHASP.ini file in these folders as well.

Need additional help? Visit the Windows or Macintosh FAQs, www.wavefun.com, support section.

Index

Note: Items in **bold type** refer to entries in ***Spartan*** menus.

D

E

F

G

H

I

M

O

P

Q

R

T

U

V

W

X

Z