

Computational Quantum Chemistry

$$H\Psi = E\Psi$$

C. David Sherrill



Quantum Chemistry is Useful For...

- ◆ Understanding molecules at the subatomic level of detail
- ◆ Simulation and modeling: make predictions. Particularly useful for systems hard to study by experiment.

Quantum Mechanical Models

- ◆ Molecules are small enough that classical mech doesn't always provide a good description. May need Quantum Mechanics.
- ◆ Quantum effects large for proton or electron transfers (e.g., biochemistry).
- ◆ Quantum mechanics required for electronic processes (e.g., spectra).

Comparison to Classical Methods

- ◆ Quantum models don't necessarily need empirical parameters: applicable in principle to *any* molecule
- ◆ Quantum mechanics provides *all information that can be knowable* about a system (QM postulate).
- ◆ Often much more accurate and reliable.
- ◆ Computations can be vastly more time-consuming.

Quantum Theory of Chemistry

- ◆ For non-relativistic atoms, the Schrödinger equation is all we need!
- ◆ Time dependent form: $H \Psi = i \hbar \frac{d \Psi}{d t}$
- ◆ Time independent-form: $H \Psi = E \Psi$
- ◆ For heavier atoms (2nd transition row and beyond) need corrections for relativity or the full Dirac equation.

Good News / Bad News

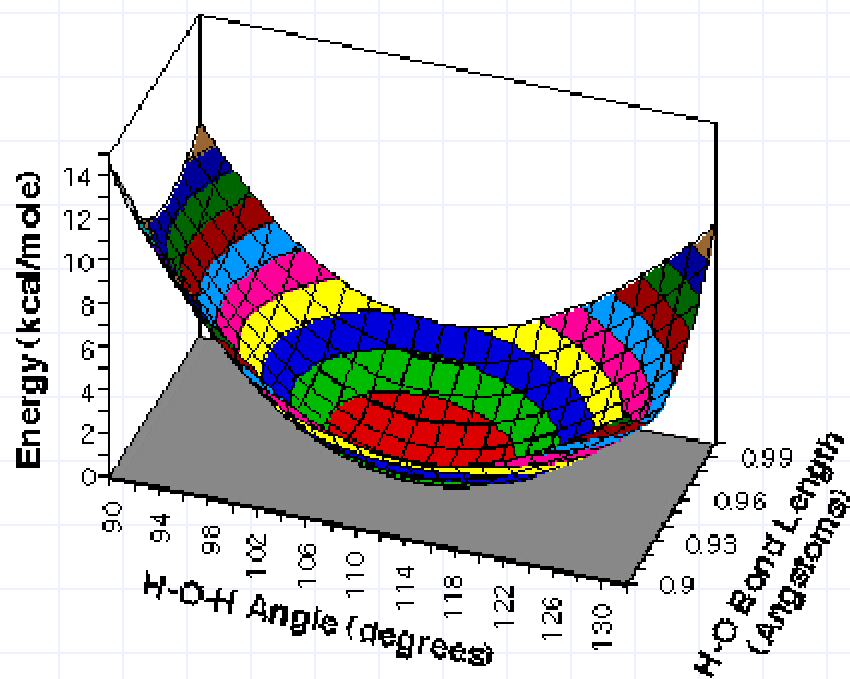
“The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.” ---P.A.M. Dirac, *Proc. Roy. Soc. (London)*, **123** 714 (1929).

Born-Oppenheimer Approximation

- ◆ The time-independent Schrödinger equation, $H\Psi = E\Psi$, approximately factors into electronic and nuclear parts.
- ◆ The electronic Schrödinger equation $H_e(r;R)\Psi_e(r;R) = E_e(R)\Psi_e(r;R)$ is solved for all nuclear coords R to map out the potential energy surface.

Potential Energy Surfaces

- ◆ Potential energy at each geometry is the electronic energy $E_e(R)$.



Quantum Simulation and Modeling

- ◆ Electronic Structure Theory: Solve for electronic Schrödinger Equation, get electronic properties and potential energy surfaces.
- ◆ Dynamics: Solve for motion of nuclei on the potential energy surfaces.
Important for detailed understanding of mechanisms.

Properties

Electronic wavefunction and its derivatives give...

- ◆ Dipole moment, polarizability, ...
- ◆ Molecular structures
- ◆ Spectra: electronic, photoelectron, vibrational, rotational, NMR, ...

Electronic Structure Methods

Method	Accuracy	Max atoms
Semiempirical	Low	~2000
Hartree-Fock & Density Functional	Medium	~500
Perturbation & Variation Methods	High	~50
Coupled-Cluster	Very High	~20

Hartree-Fock Method

- ◆ Assumes electronic wavefunction can be written as an antisymmetrized product of molecular orbitals (a Slater Determinant).
- ◆ Same thing as assuming that each electron only feels an *average* charge distribution due to other electrons.
- ◆ An approximation! Orbitals aren't real...

Slater Determinants

Ψ is the Hartree-Fock wavefunction, and ϕ_i is molecular orbital i .

$$\Psi(1,2,\dots,n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \mathbf{f}_1(1) & \mathbf{f}_2(1) & \dots & \mathbf{f}_n(1) \\ \mathbf{f}_1(2) & \mathbf{f}_2(2) & \dots & \mathbf{f}_n(2) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{f}_1(n) & \mathbf{f}_2(n) & \dots & \mathbf{f}_n(n) \end{vmatrix}$$

LCAO-MO

- ◆ The molecular orbitals in the wavefunction are determined as a Linear Combination of Atomic Orbitals (LCAO). Coefficients found by minimizing energy (Variational Theorem).

$$f_i(i) = \sum_m C_{mi} \mathbf{C}_m(i)$$

Hartree-Fock Energy

$$E = \sum_{i=1}^n h_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (J_{ij} - K_{ij})$$

Sums run over occupied orbitals. h_i is the one-electron integral, J_{ij} is the Coulomb integral, and K_{ij} is the exchange integral. For example:

$$J_{ij} = \iint dr_1 dr_2 \mathbf{f}_i^*(r_1) \mathbf{f}_i(r_1) \frac{1}{r_{12}} \mathbf{f}_j^*(r_2) \mathbf{f}_j(r_2)$$

A Catch-22?

- ◆ The orbitals ($C_{\mu i}$ coefficients) are found by minimizing the HF energy.
- ◆ But....the energy depends on the orbitals!
- ◆ Solution: start with some guess orbitals and solve *iteratively*.
- ◆ Self-consistent-field (SCF) model.

Two-electron integrals are hard to compute

- ◆ The two-electron integrals (J and K) are usually the time-consuming part. Recall each MO is a LCAO...substitute the AO's into the integral equation to get something evil like:

$$(\mathbf{m} | \mathbf{rs}) = \iint dr_1 dr_2 \mathbf{c}_m^*(r_1) \mathbf{c}_n(r_1) \frac{1}{r_{12}} \mathbf{c}_r^*(r_2) \mathbf{c}_s(r_2)$$

Each AO could be on a different atom ...
4-center integral!

Computational Cost of Hartree-Fock

- ◆ The 4-center integrals lead to a formal scaling of $O(N^4)$... things get *much* worse for larger molecules!
- ◆ The integral $(\mu\nu|\rho\sigma)$ is small unless the pairs (μ,ν) and (ρ,σ) have non-negligible overlap: leads to $O(N^2)$ scaling. (Advances in 1980's).
- ◆ Recent work (1990's) using multipole expansions is bringing the cost down to linear!

Basis Sets

- ◆ How do we pick the AO's used in Hartree-Fock? There are "standard" sets of AO's for each atom called **Basis Sets**.
- ◆ Usually the AO's (or basis functions) are atom-centered Gaussian functions *or* combinations of such functions.

Gaussian Basis Functions

- ◆ Hydrogen orbitals are Slater functions, but quantum chemists use Gaussian functions instead (easier to compute).
- ◆ A primitive Gaussian Type orbital (GTO)

$$\mathbf{C}_{nlm}^{GTO}(x, y, z) = x^n y^l z^m e^{-Vr^2}$$

- ◆ A contracted Gaussian orbital (CGTO)

$$\mathbf{C}_{nlm}^{CGTO}(x, y, z) = \sum_i C_i \mathbf{C}_{nlm,i}^{GTO}(x, y, z)$$

STO-3G Minimal Basis

- ◆ A minimal basis has one basis function for each AO in the real atom (e.g., 1 for H, 5 for C, etc).
- ◆ In STO-3G, each basis function is a CGTO of 3 Gaussian orbitals, with coefficients C_i picked to fit the corresponding Slater type orbitals (STO).
- ◆ This basis is pretty bad ... it's too small!

Larger Basis Sets

- ◆ Atomic orbitals deform based on their environment. This is impossible with a minimal basis set: Need to add extra basis functions.
- ◆ Bigger/smaller basis functions allow the orbitals to expand/contract
- ◆ Higher angular momentum functions allow orbitals to *polarize*.

Larger Basis Example: 6-31G*

- ◆ Split-valence basis: the core orbitals described by a single CGTO made of 6 primitives, the valence orbitals described by *two* functions, one a CGTO made of 3 primitives, the other 1 primitive.
- ◆ The * denotes a set of polarization functions on “heavy” (non-H) atoms.

Other Basis Set Items

- ◆ cc-pVXZ: Dunning's "correlation consistent" polarized X-zeta basis sets. cc-pVDZ roughly equivalent to 6-31G**.
- ◆ + denotes diffuse functions: *critical* for anions (called aug-cc-pVXZ by Dunning).
- ◆ ANO: Atomic Natural Orbitals.

Accuracy of Hartree-Fock

Property	Accuracy
Bond lengths	$\pm 0.02 \text{ \AA}$
Bond angles	$\pm 2^\circ$
Vibrational frequencies	$\pm 11\%$
Dipole moments	$\pm 0.3 \text{ D}$
Relative energy	$\pm 25\text{-}40 \text{ kcal/mol}$ for dissociation energies

Semiempirical Methods

- ◆ Attempt to speed up Hartree-Fock by replacing some of the two-electron integrals by empirical parameters.
- ◆ Examples: MNDO, AM1, PM3.
- ◆ Rotational barriers around partial double bonds too low, weakly bound complexes poorly predicted, parameters not always available (e.g., metals).

Semi-empirical Performance

Error in...	MNDO	AM1	PM3
ΔH_f (kcal/mol)	46.2	27.6	11.6
Bond to C (Å)	0.002	0.002	0.002
Bond to H (Å)	0.015	0.006	0.005
Bond to N (Å)	0.015	0.014	0.012
Bond to O (Å)	0.017	0.011	0.006

Electron Correlation Methods

- ◆ Hartree-Fock is an approximation. It replaces the instantaneous electron-electron repulsion ("electron correlation") by an *average* repulsion term.
- ◆ Models which explicitly treat electron correlation are more accurate.

Many-Body Perturbation Theory

- ◆ Also called Møller-Plesset Perturbation Theory
- ◆ Treat electron correlation as a “small” perturbation to the Hartree-Fock description.
- ◆ More accurate but cost goes way up; scales with system size as $O(N^5)$. Efforts underway to cut the cost.

Configuration Interaction

- ◆ Express the wavefunction as the Hartree-Fock determinant *plus* many other determinants which put electrons in *different* orbitals. CISD is $O(N^6)$.
- ◆ Use the Variational Theorem to find best coefficients.

$$\Psi(r_1, r_2, \dots, r_n) = \sum_i C_i \Phi_i(r_1, r_2, \dots, r_n)$$

Coupled-Cluster Methods

- ◆ Expresses the wavefunction as an exponential product: has higher-order corrections “built-in” as products of lower-order terms. Cost is not much more than comparable CI treatment but much more accurate. CCSD, CCSD(T).

$$\Psi(r_1, r_2, \dots, r_n) = e^T \Phi_{HF}$$

Convergent Methods

Method	STO-3G	6-31G*	cc-pVTZ	...	CBS
HF		X			
MP2		X			
CISD		X			
CCSD			X		
CCSD(T)			X		
...				X	
Exact					"Truth"

Coupled-Cluster Performance

Typical accuracy of coupled-cluster methods

Property	Accuracy
Bond lengths	$\pm 0.004 \text{ \AA}$
Bond angles	$\pm 0.03^\circ$
Vibrational frequencies	$\pm 2\%$
Dipole moments	$\pm 0.05 \text{ D}$
Relative energy	$\pm 1.5 \text{ kcal/mol}$ for dissoc/ioniz energies

Density Functional Theory

- ◆ A large part of the 1998 Nobel Prize in Chemistry (Kohn and Pople) recognized work in this area.
- ◆ Use the density instead of complicated many-electron wavefunctions.
- ◆ Basic idea: minimize the energy with respect to the *density*. Relationship of energy to density is the “functional” $E[\rho]$ (true form of this functional is unknown: use approx.)

Performance of DFT

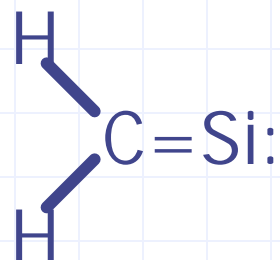
- ◆ Formulation is very similar to Hartree-Fock and cost is only slightly more, but DFT includes electron correlation.
- ◆ Examples: S-VWN, BLYP, B3LYP
- ◆ Often very high accuracy (comparable to coupled-cluster), particularly for B3LYP.
- ◆ Sometimes empirical parameters go into functionals.
- ◆ Problem: not a convergent family of methods.

Example Applications

- ◆ Predicting/confirming spectra
- ◆ Structures/energies of highly reactive molecules
- ◆ Interaction between possible drugs and enzyme active sites
- ◆ Computational materials science

Interstellar Molecule Spectra

- ◆ 1-Silavinylidene has been predicted to be abundant in interstellar space



- ◆ Could search for it if we knew what its microwave/infrared spectra looked like
- ◆ Bengali and Leopold performed tricky experiments and requested theoretical confirmation.

Theory Confirms Assignment

Mode	Theory	Experiment
ω_1 (CH sym str)	3084	2980 ± 20
ω_2 (CH ₂ scissor)	1345	1250 ± 30
ω_3 (Si-C str)	927	930 ± 20
ω_4 (Si oop bend)	690	
ω_5 (CH asym str)	3165	
ω_6 (CH ₂ rock)	305	~ 265

Using TZ2Pf CCSD(T) theoretical method.

Sherrill and Schaefer, *J. Phys. Chem.* **99**, 1949 (1995).

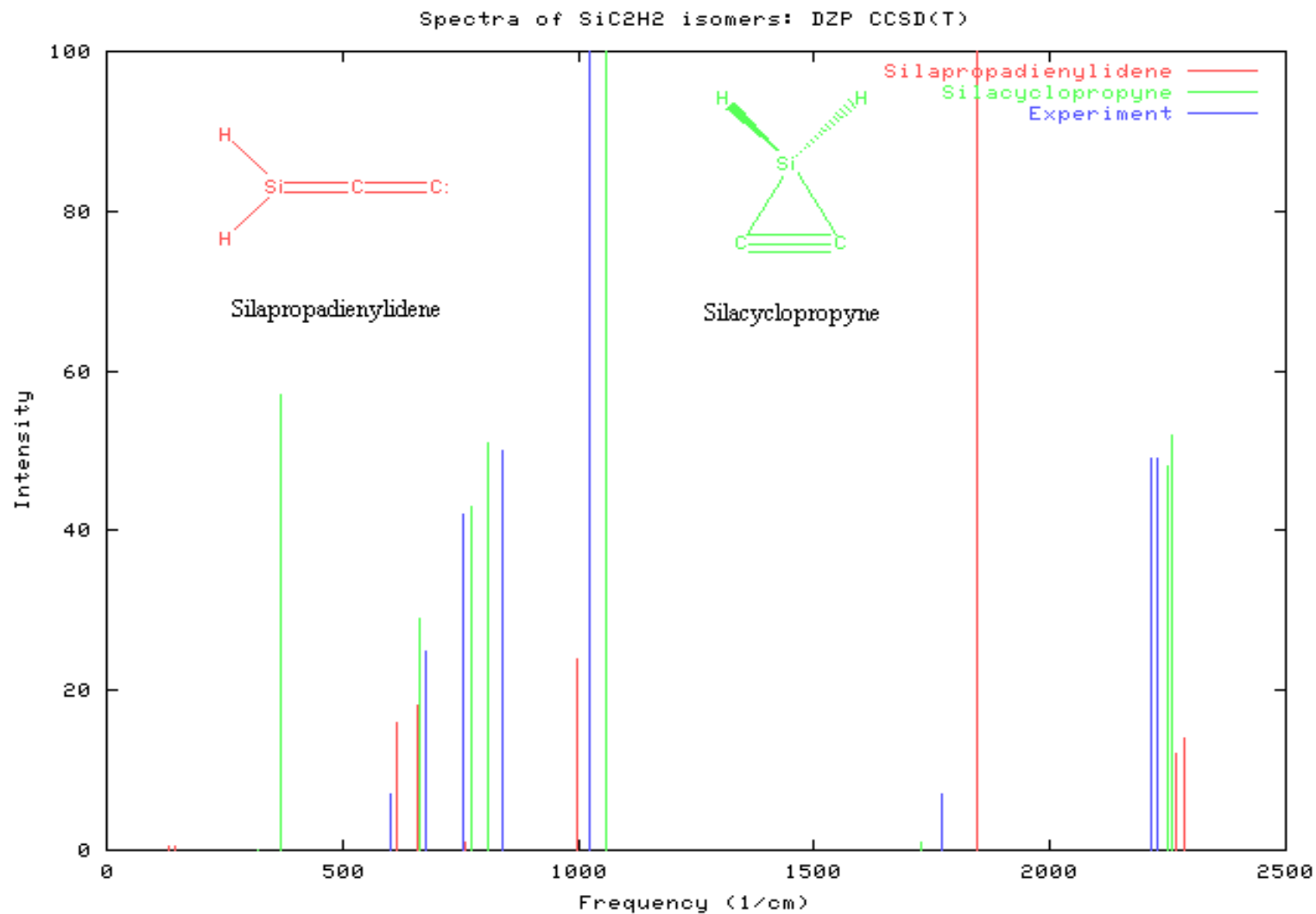
Vibrational Spectra of SiC₂H₂ Species

- ◆ Maier et al. isolated a new, unknown SiC₂H₂ species
- ◆ The matrix IR spectrum indicated an SiH₂ group (absorption at 2229 and 2214 cm⁻¹).
- ◆ Theory shows which species was seen.

Maier et al. *Angew. Chem. Int. Ed. Eng.* **33**, 1248 (1994).

Sherrill et al. *J. Am. Chem. Soc.* **118**, 7158 (1996).

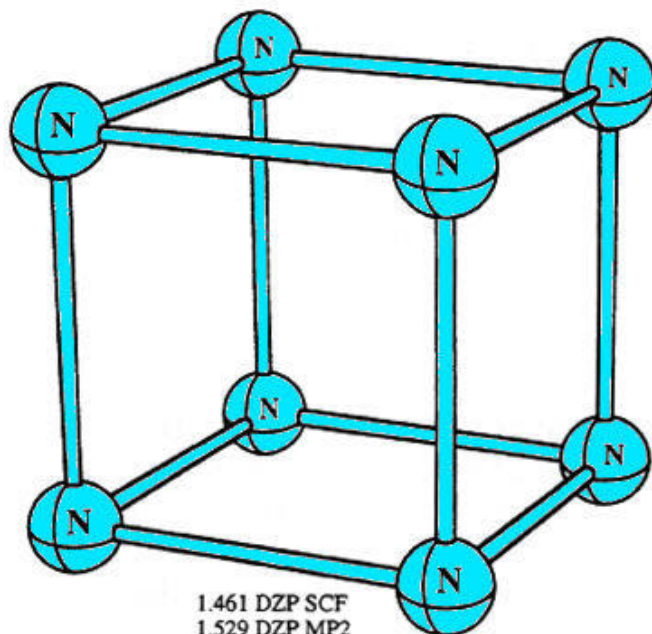
Theory Shows Silacyclopropyne Seen



Highly Reactive Systems

- ◆ Many molecules are hard to study experimentally: e.g., radicals, diradicals, highly strained molecules.
- ◆ Theory can be helpful in understanding such systems.

N_8 : Possible Rocket Fuel



1.461 DZP SCF
1.529 DZP MP2
1.479 DZP CISD
1.517 DZP CCSD

(1)

Figure 1. Theoretical equilibrium geometries of the O_h symmetry octahedral structure of N_8 . Bond lengths are in angstroms, and bond angles are in degrees.

❖ First studied theoretically.

❖ Energy for decomposition $N_8 \rightarrow 4 N_2$ is computed as 423 kcal/mol !

❖ Efforts underway to synthesize it.

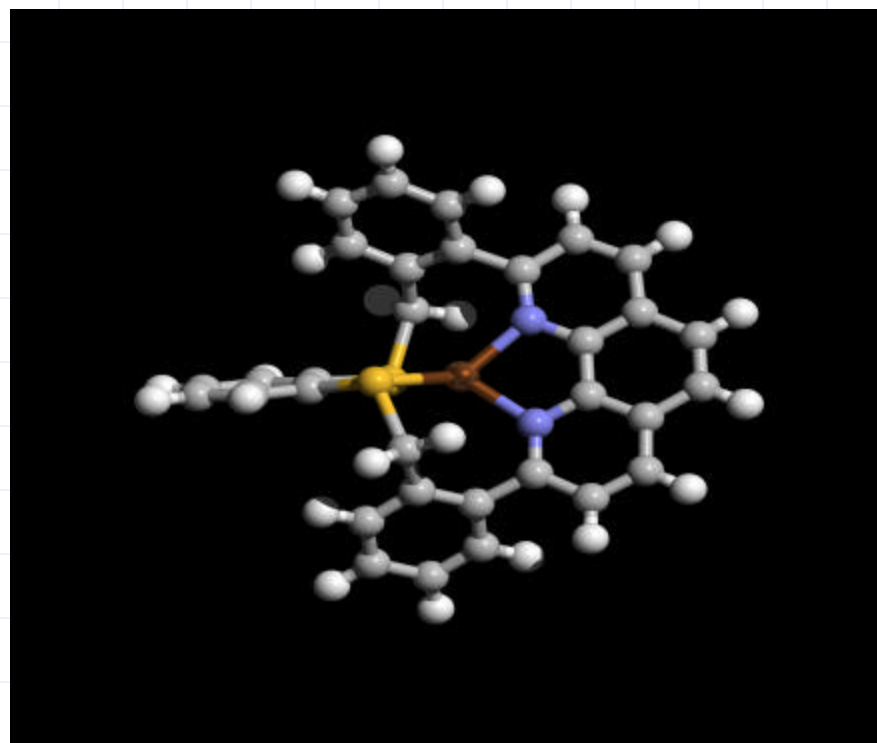
Leininger, Sherrill, and Schaefer, *J. Phys. Chem.* **99**, 2324 (1995).

Design of New Molecules

- ◆ Often need to design a molecule for a specific purpose (e.g., N_2 for rocket fuel).
- ◆ Theory is useful for narrowing down the list of candidate molecules.
- ◆ Ruling out bad candidates early saves time: no need to synthesize something which won't work.
- ◆ This strategy used by many pharmaceutical companies.

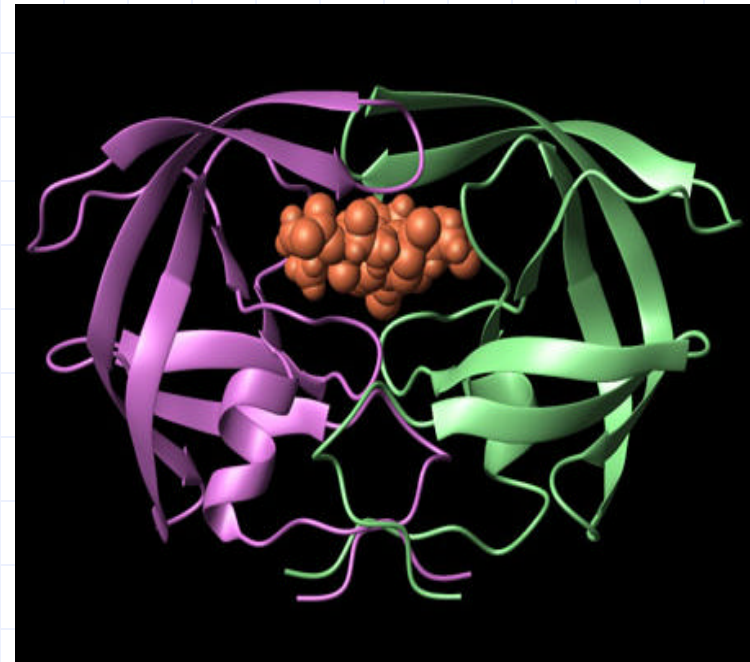
Fluorescent Copper(I) Probes

- ◆ Collaboration with Prof. Christoph Fahrni
- ◆ Need a way to track Cu(I) ions in the body to understand their biochemical role
- ◆ Use theory to predict structures and spectra of possible Cu(I) probes



Rational Drug Design

- ◆ Take structure of enzyme and model interaction with possible drugs
- ◆ Often uses classical mech. models but sometimes refined by Q.M.
- ◆ Promising new drug design approach.



Model of candidate inhibitor for HIV-1 protease.
(Physical Computing Group, Rice U.)

Conclusions

- ◆ Quantum Mechanics is how the world works at small scales: can be vital for understanding physics and chemistry.
- ◆ Can be used to model molecular behavior and speed up experimental work.
- ◆ A wide range of methods are available.