ONIOM and Its Applications to Material Chemistry and Catalyses

Keiji Morokuma

Cherry L. Emerson Center for Scientific Computation and Department of Chemistry,
Emory University, Atlanta, Georgia 30322, USA

One of the largest challenges for quantum chemistry today is to obtain accurate results for large complex molecular systems, and a variety of approaches have been proposed recently toward this goal. We have developed the ONIOM method, an onion skin-like multi-level method, combining different levels of quantum chemical methods as well as molecular mechanics method. We have been applying the method to many different large systems, including thermochemistry, homogeneous catalysis, stereoselectivity in organic synthesis, solution chemistry, fullerenes and nanochemistry, and biomolecular systems. The method has recently been combined with the polarizable continuum model (ONIOM-PCM), and was also extended for molecular dynamics simulation of solution (ONIOM-XS). In the present article the recent progress in various applications of ONIOM and other electronic structure methods to problems of homogeneous catalyses and nanochemistry is reviewed. Topics include 1. bond energies in large molecular systems, 2. organometallic reactions and homogeneous catalysis, 3. structure, reactivity and bond energies of large organic molecules including fullerenes and nanotubes, and 4. biomolecular structure and enzymatic reaction mechanisms.

Key Words : ONIOM, Hybrid method, Catalyses, Material chemistry

Introduction to ONIOM

The challenge to quantum chemists in the new millennium is to perform accurate and reliable quantum mechanical calculations for large molecular systems. As quantum chemistry plays the more and more central role in many complicated and complex real chemical problems, it has become necessary to obtain accurate results for large molecular systems.

The ONIOM (*Our N-layered Integrated molecular Orbital + molecular Mechanics*) method developed by the author and his colleagues is an onion skin-like extrapolation method, as shown in Figure 1(A) that allows to combine a variety of quantum mechanical methods as well as a molecular mechanics method in multiple layers.[1-6]

The concept of the ONIOM method is extremely simple. The target calculation is the *high* level calculation for a large *real* system, E(*high, real*), which is prohibitively expensive. You can perform for the *real* system an inexpensive *low* level calculation E(*low, real*) that may not be sufficiently accurate for the property you want to calculate. You can also carry out an accurate *high* level calculation for a smaller *model* system E(*high, model*). Starting from E(*low, model*), if one assumes that the correction for the *high* level, E(*high, model*) − E(*low, model*), and the correction for the *real* system, E(*low, real*) − E(*low, model*), to be additive, as illustrated in Figure 1(B), the energy of the *real* system at the *high* level can be estimated extrapolatively from three independent calculations as:

$$E(\text{ONIOM, } real) = E(high, model) + E(low, real)$$
$$- E(low, model). \tag{1}$$

The scheme can easily be extended to thee- or multiple-layers, as shown in Figure 1(C). As one see here, there is no restriction on the methods used at various levels and any MO method as well as MM method can be combined. Combination of different theoretical levels of MO methods is a unique feature of the ONIOM method.

The definition of the model system is rather straightforward when there is no covalent bond between the layers. If one considers a solute molecule complexed with one solvent molecule as the real system, the solute will be the model system, and the solute-solvent interaction is naturally included at the low level calculation of the real system. The situation becomes more complicated when a covalent bond exists between the layers. The ONIOM method adopts the most straightforward way to ensure that the model system is representative for the real system, *i.e.*, saturating the dangling bonds at the end of the model layer with link atoms. For example, if a methyl group is in the low level layer, this group is substituted by a hydrogen atom in the model system.

The Choice of Model and Level, and the S-value Test

The most important question when using the ONIOM scheme is how to select the methods that will be combined, and the partitioning of the system into high and low level layers. These two factors are closely related. An arbitrary choice of model and method combination usually does not work. The first rule of thumb is that if one knows by intuition who are the major players and who are the minor players, include the major players in the model system to be treated at high level, while the minor players can be treated only at low level. For instance, in chemical reactions in
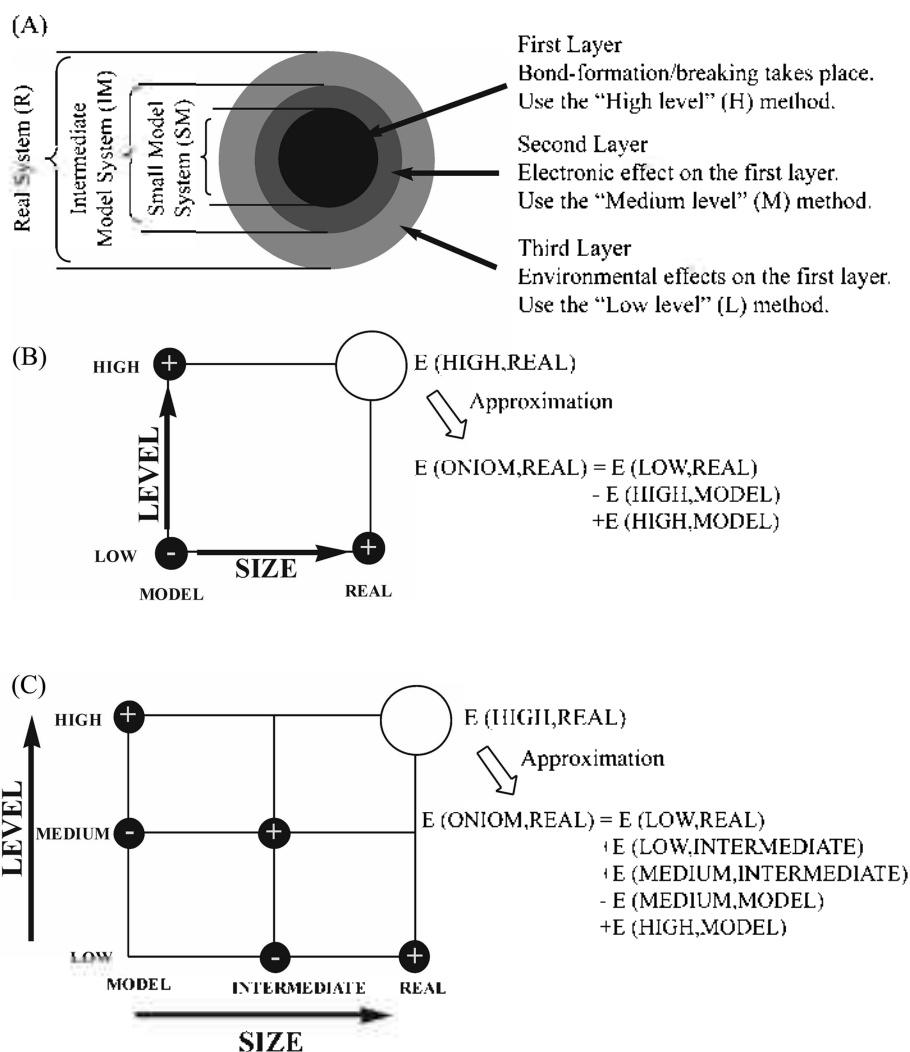
Figure 1. The ONIOM method. (A) The onion skin-like layers and models. (B) The two-layer ONIOM method. (C) Three-layer ONIOM method.

solvated cluster (the real system), the reactants (and possibly a few solvent molecules directly involved in the reaction) will constitute the model system, and the rest of the system will be treated only at the low level. There are cases where the ONIOM method does not work. In such cases, one cannot identify major players or minor players: all the parts of a large system cooperate strongly and one cannot find a workable partition of the system.

A systematic procedure to test the reliability of the level/ model combination has been devised. In the ONIOM method, E(*high, model*) is the target calculation that E(ONIOM, *real*) is trying to approximate. Let us define the S-value at a given level as:

$$S(level) = E(level, real) - E(level, model). \tag{2}$$

The S-value, S(*level*), is the difference between the *real* and the *model* system, or the effect of the substituent (second layer) evaluated at a given level. From Eq. 1 and Figure 1, it is obvious that

if $S(low) \rightarrow S(high)$, E(ONIOM, *real*) $\rightarrow$ E(*high, real*).

Namely, if the S value evaluated at the low level is the same to that at the high level, the ONIOM energy is the same to the target energy, *i.e.*, the error of the ONIOM extrapolation, E(ONIOM, *real*) − E(*high, real*), is zero. In practice, when one has a series of compounds to which ONIOM is to be applied, one chooses a few of them as the test set, and performs an expensive E(*high, real*) calculation as well as three calculations needed for ONIOM (cf. Eq. 1) with various low level candidates. Table 1 shows an example of the S-value test for C_{sp3}-H bond dissociation energy of a series of hydrocarbons and derivatives. Here the test molecules are *iso*-butane and toluene, and the model compound is H_3C-H, the minimal model for this process. RMP2/6-31G* gives the S-values very close to those at the target high G2MS(R) level, and is an excellent low level method to be combined with G2MS(R) with an estimated ONIOM extrapolation error of less than 2 kcal/mol. On the other hand, the other methods are very poor low level

Table 1. S-values[a] (in kcal/mol) and their differences (in parentheses) for C_{sp3}-H bond dissociation for iso-butane and toluene, employing a minimal (methane) model and B3LYP/6-31G geometries

level	iso-butane	toluene	
G2MS(R)	-7.75(0.00)	-16.21(0.00)	Target (high) level
UHF/6-31G*	-7.29(+0.46)	-26.12(-9.91)	Poor low level
RHF/6-31G*	-6.73(+1.02)	-10.28(+5.93)	Poor low level
UMP2/6-31G*	-7.72(+0.03)	+6.93(+23.14)	Poor low level
RMP2/6-31G*	-8.10(-0.35)	-15.04(+1.17)	Excellent low level

[a]The S-value of -7.75 kcal/mol for iso-butane at G2MS(R) level, for instance, means that the C-H bond energy of this compound is 7.75 kcal/mol lower than that of methane.

method for the present problem.

These analyses, as well as the concept itself, suggests that the ONIOM will work well if the properties to be calculated is principally determined *locally*, i.e., by a small part of a large system, with some corrections from the rest of the system. For instance, the C-H bond dissociation, requiring a very high level theoretical method to describe, is mainly determined by the nature of two bonding atoms but modified by the effects of the rest of molecule, and is a favorable case for ONIOM. The potential energy surface of chemical reactions is another example where ONIOM seems to work well; the transition structure and energy are mainly determined by a few atoms that are directly involved in bond breaking and formation, with the effects of the rest of the system modifying them.

Applications

Thermochemistry. C-H and C-C Bond Dissociation Energies of a Series of Hydrocarbons. ONIOM has successfully been applied to the calculation of varieties of bond dissociation energies.[7-9] Without ONIOM, accurate calculations could be performed only for nine smallest compounds in the table with 7 or less carbon atoms. With the ONIOM method, calculations have been made for all the compounds (with a fraction of time required without ONIOM). The rms error is 2.4 kcal/mol with ONIOM(G2MS:ROMP2/6-31G*), which is not much worse than 1.9 kcal/mol with G2MS and 2 kcal/mol in experiment. A three-layer ONIOM (G2MS: ROMP2/6-31G*:B3LYP/3-21G) calculation predicts the C-C bond dissociation energy of yet to be synthesized Ph_3C-CPh_3 to be 16.4 kcal/mol, indeed a very weak C-C bond.

Homogeneous Catalysts. Computational study of reaction mechanisms catalyzed by metal complexes is one of very popular applications of quantum mechanical calculations in chemistry. Very often, the catalyst is a complicated metal complex with large ligands, and theoreticians are forced to trim off substituents on the ligand to make calculations feasible. With ONIOM, one can treat computationally the real catalyst with large ligands.

We have recently studied the mechanism of copolymerization of CO_2 with cyclohexene oxide catalyzed by a bulky Zn(II) catalyst.[10] Here we used a combination of a DFT

method with a semi-empirical MO method PM3. ONIOM (B3LYP/lanl2dz(d):PM3). Since we are using an MO method as the *low* level, we are including the electronic effects, as well as steric effects, of the substituent, which is essential for the ligand with aromatic rings. For the unactivated catalyst alone, the full DFT calculation took 600 hrs of computer time, while the ONIOM calculation took only 15 hours; without ONIOM, this study would have been impossible.

With the use of the ONOM, including bulky substituents with an MM method, we have studied many homogenous catalytic reactions and organometallic reactions. We found in many examples that the sterically bulky ligands play the key role in controlling the catalytic activity. For instance, in ethylene polymerization by catalysts of type [2,6-(CMe= N(2,6-^{i}Pr$_2$C$_6$H$_3$))$_2$C$_5$H$_3$N]FeCH$_3^-$.[11] we found that two bulky ortho substituents on the Ar rings results in steric destabilization of the axial positions, and make β-hydride transfer (BHT) chain termination uncompetitive with chain propagation. In ethylene polymerization by a Ni(II) catalyst having with diimine ligands ArN=C(R)C(R)=NAr, Ar=2,6-C_6H_3(i-Pr)$_2$, the bulky substituents on the ligands were found to be necessary to destabilize the Ni-ethylene reactant complex and lower the barrier for chain propagation, which becomes competitive against BHT chain termination process.[12]

Biomolecular Chemistry. Effects of the Protein Environment on the Structure and Energetics of Active Sites of Metalloenzymes. Most theoretical studies to date on structures and reactions in biological systems have been performed using quantum chemical methods for models consisting of up to 50 atoms. With such limited models, the question remains whether inclusion of the surrounding protein in the calculation may change the results. ONIOM, as well as QM/MM, is a method well suited and many applications will be seen in such areas.

We have recently applied the ONIOM(MO:MM) and ONIOM(MO:MO:MM) methods to the structure and energies of both the oxidized forms, MMOH$_{ox}$ and R2$_{met}$ (resting state) and the reduced forms, MMOH$_{red}$ and R2$_{red}$ (active state) of a unit of two non-heme diiron enzymes, methane monooxygenase and ribonucleotide reductase, respectively.[13] The "Active Site Only" model, containing 39-46 MO atoms, and the "Active Site + Four α-Helical Fragments" model, including additional 62 MM amino acid residues were compared. The optimized structures of the active sites with the latter model agree with the X-ray structures (with rms errors of 0.34-0.50 Å) substantially better than the former (errors of 0.623-1.029 Å), indicating that the local structure of active site is in part controlled by the protein environment.

Heterogeneous Catalysis. Chemical reactions on solid surfaces have been studied most commonly with a cluster calculation, in which the solid surface was approximated by a small cluster of atoms. The ONIOM method allows embedding of a small cluster in a larger cluster that is treated at a *low* level. The ONIOM(MO:MM) as well as

ONIOM (MO:MO) method has been used to study chemical reactions on oxide, silicon and metal surfaces as well as in zeolites.[14]

Regio- and Stereo-Selectivity in Organic and Organometallic Reactions. The regio- and stereo-selectivities in organic and organometallic reactions are often controlled by the steric (as well as electronic) interaction of bulky groups, and ONIOM allows the inclusion of these groups explicitly into the calculation. There have been many applications of the ONIOM methodology, in particular the MO:MM version, to this topic.[15]

Chemical Reactions in Cluster and Solution. Theoretical studies of chemical reactions in cluster are hot topics in connection to recent advent of experimental studies. However, including more than a few solvent molecules explicitly into quantum calculation quickly makes the cost prohibitive. The ONIOM will fit in here nicely, if one handles the solute molecule at a *high* level and a few solvent molecules at a less expensive *low* level. Some accurate ONIOM studies of S$_N$2 reactions in cluster have been performed.[16] More applications will be seen in this area.

Material and Nano Chemistry. Usually large molecular systems are involved in material and nano chemistry. Although ONIOM cannot handle truly delocalized properties (such as band structure) efficiently, many interesting properties are local in nature but are influenced by the large size of the system. ONIOM should be able treat such a situation very well, and some applications have been published[17] and more studies are being made in this area.

ONIOM in the Polarizable Continuum, ONIOM-PCM. Possible further developments include the incorporation of the ONIOM method in the polarized continuum or reaction field model for solvation. In this model, the bulk solvent is treated as a polarizable continuum, and the reaction field induced by the charge distribution of the quantum mechanical solute interacts self-consistently with the solute charge distribution. Although the bulk solvent effects on solute can be reproduced qualitatively by this kind of model, the specific interaction of first-shell solvent molecules is not explicitly considered. One may want to include a few solvent molecules explicitly into quantum calculation; however, the cost becomes quickly prohibitive. The ONIOM will fit in here nicely, if one handles the solute molecule at a *high* level and a few solvent molecules at a less expensive *low* level and surround this ONIOM system further by a continuum. We have recently developed a version of this combination, ONIOM-PCM, incorporating Tomasi *et al.*'s PCM method,[18] and it is being tested and applied to a variety of chemical reactions in solution.

Molecular Dynamics with ONIOM, ONIOM-XS/MD. Furthermore, so far we have only explored the ONIOM in a static manner. With its cost effectiveness, direct molecular dynamics (MD) for a large molecular system or assembly may be performed by calculating the ONIOM energy and gradient on the fly. An MD simulation of hydration of simple metal cation using ONIOM(*ab initio* MO:MM) has just recently been carried out[19] considering the cation and

several first-shell water or ammonia molecules quantum mechanically and many other water molecules by MM. In MD simulation, an originally QM water molecule may move away from the first shell, which may be better handled with MM, or vice versa. Therefore, in this simulation, solvent molecules were allowed to move in and out between the MO and the MM region by means of a switching function.

Conclusions

The ONIOM method can combine different levels of MO methods and an MM method into a single integrated calculation of the energy or other electronic properties, a unique feature not available in other hybrid methods. The method is so flexible that the final choice of method/model combination is left to the user, who can tune it to be within the tolerated error (compared to either target calculations or experimental data) for the property under investigation, using an S-value test. However, for a normal problem and an unsophisticated user, one may suggest that if the target (*high*) level is one of "chemical accuracy" methods, such as CCSD(T) with a large basis set or extrapolation scheme such as G2, G3 or CBS, the best choice of *low* level is MP2 with a medium basis set, such as 6-31G*. If MP2 or DFT is the target level, *ab initio* HF with a smaller basis set or semiempirical MO methods are usually good choices of *low* level. These methods can always be combined further with an MM layer for even larger systems.

References

1. Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *16*, 1959.
2. Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357.
3. Frisch, M. J. *et al. Gaussian 98, Revision A.1*; Gaussian, Inc.: Pittsburgh, PA. 1998; Frisch, M. J. *et al. Gaussian 03, Revision A.1*; Pittsburgh, PA. 2003.
4. Maseras, F. *Top. Organomet. Chem.* **1999**, *4*, 165.
5. Froese, R. D. J.; Morokuma, K. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P. A.; Schaefer III, H. F.; Schreiner P. R., Eds.; John Wiley: Chichester, 1998; p 1245.
6. Vreven, T.; Morokuma, K. *J. Comp. Chem.* **2000**, *21*, 1419.
7. Froese, R. D. J.; Morokuma, K. *J. Phys. Chem. A* **1999**, *103*, 4580.
8. Froese, R. D. J.; Morokuma, K. *Chem. Phys. Lett.* **1999**, *305*, 419.
9. Vreven, T.; Morokuma, K. *J. Chem. Phys.* **1999**, *111*, 8799.
10. Liu, Z.; Torrent, M.; Morokuma, K. *Organometallics* **2002**, *21*, 1056.

11. Khoroshun, D. V.; Musaev, D. G.; Vreven, T.; Morokuma, K. *Organometallics* **2001**, *20*, 2007.

12. Froese, R. D. J.; Musaev, D. G.; Morokuma, K. *J. Am. Chem. Soc.* **1998**, *120*, 1581.

13. Torrent, M.; Vreven, T.; Musaev, D. G.; Morokuma, K.; Farkas, O.; Schlegel, H. B. *J. Am. Chem. Soc.* **2002**, *124*, 192.

14. (a) Choi, C. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1999**, *121*, 11311. (b) Sauer, J.; Sierka, M. *J. Comp. Chem.* **2000**, *21*, 1470.

15. (a) Ujaque, G.; Maseras, F.; Lledos, A. *J. Am. Chem. Soc.* **1999**, *121*, 1317. (b) Feldgus, S.; Landis, C. R. *J. Am. Chem. Soc.* **2000**, *122*, 12714.

16. Re, S.; Morokuma, K. *J. Phys. Chem. A* **2001**, *105*, 7185.

17. Irle, S.; Rubin, Y.; Morokuma, K. *J. Phys. Chem. A* **2002**, *106*, 680.

18. Vreven, T.; Mennucci, B.; da Silva, C. O.; Morokuma, K.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 62.

19. Kerdcharoen, T.; Morokuma, K. *Chem. Phys. Lett.* **2002**, *355*, 257; *J. Chem. Phys.* **2003**, *118*, 8856.