Theor Chem Acc (2003) 109:285–297
DOI 10.1007/s00214-003-0432-8

Theoretical
Chemistry Accounts

Regular article

Implementation of a quantum mechanics/molecular mechanics approach in the parallel density functional program ParaGauss and applications to model copper thiolate clusters

Teerakiat Kerdcharoen[1,2], Uwe Birkenheuer[1,3], Sven Krüger[1], André Woiterski[1], Notker Rösch[1]

[1]Institut für Physikalische und Theoretische Chemie, Technische Universität München, 85747, Garching, Germany
[2]Department of Physics, Faculty of Science, Mahidol University, 10400, Bangkok, Thailand
[3]Max-Planck-Institut für Physik komplexer Systeme, 01187, Dresden, Germany

Received: 25 May 2002 / Accepted: 4 November 2002 / Published online: 16 May 2003

Abstract. Based on the integrated molecular orbital and molecular mechanics approach, we constructed a combined quantum mechanics (QM) and molecular mechanics (MM) method, by combining the parallel density functional program ParaGauss with the MM3 force field. We examined different ways to describe link atoms at the boundary between the QM and MM regions which we tested successfully on various organic molecules. We applied the new tool, intended for the description of metal–ligand interaction, to model Cu thiolate clusters. Separation at the first C–C bond was favored over a transition at the S–C bond, closer to the metal particle. We successfully checked the effect of different ligand orientations on the cluster geometry by comparison with QM calculations. Hybrid calculations correctly described geometric rearrangements to avoid steric stress of larger ligands. Thus, the QM/MM approach is also applicable when direct metal–metal interactions have to be treated.

Keywords: QM/MM – Cluster compounds – Copper thiolate – Density functional calculation

Introduction

Ligand-stabilized transition-metal clusters have attracted considerable interest owing to their unique properties that can be useful for a wide spectrum of applications, ranging from industrial catalysts to nanoelectronics to molecular biology [1, 2, 3]. Among these clusters, alkane thiolate covered gold clusters are very popular owing to their excellent stability and usability [1, 4]. Sophisticated techniques, such as transmission electron microscopy [5], scanning tunneling microscopy [6], and small-angle X-ray scattering [4, 6], were applied to characterize these gold thiolate clusters. Nevertheless, detailed information on structure and bonding as well as electronic or magnetic properties is still scarce. For this purpose, accurate quantum chemical calculations are very promising: for instance, bare metal clusters [7, 8] as large as Au_{147} or Pd_{309} were treated by a scalar-relativistic all-electron density functional (DF) method using the parallel computer program ParaGauss [8, 9]. Nevertheless, when covered with a stabilizing ligand shell, such large clusters still present a strong challenge to precise quantum mechanical calculations. Besides the size of the combined metal–ligand system, also the symmetry lowering due to large ligands renders this type of calculation prohibitively demanding. Consequently, shells of simple model ligands like PH_3 [10, 11], SH [11, 12], and SCH_3 [13] were introduced to represent larger phosphines or thiolate groups, respectively. However, such an oversimplification fails to describe steric effects of bulky ligands which act back on the metal cluster.

A hybrid approach provides an attractive strategy for describing cluster–ligand interactions. Thereby, one combines a highly accurate quantum mechanical method (e.g. a Kohn–Sham procedure) of the metal particle with an inexpensive molecular mechanics (MM) description of the outer part of the ligand shell. Such combined quantum mechanics (QM) and MM methods were formulated to study reactions of small organic molecules [14, 15]. Since then, new techniques have been developed to render such strategies applicable to a large number of problems, for example, to enzyme reactions [16], molecular cavities [17, 18], structural molecular biology [19, 20], surface phenomena [21, 22], liquids [23], and solutions [24, 25, 26]. To the best of our knowledge, the present work is the first attempt to tailor a QM/MM method for describing ligand-stabilized transition-metal

Correspondence to: N. Rösch
e-mail: roesch@ch.tum.de

clusters. We present a series of test calculations on organic molecules to validate our QM/MM implementation which combines DF and force field (FF) methodologies. As a first feasibility study on ligated metal clusters, we present results for small Cu thiolate species, $Cu_4(S(CH_2)_nCH_3)_2$ ($n = 1$–3). These model systems incorporate all essential interactions present in thiolate-coated transition-metal clusters.

Although QM/MM strategies are available in some widely distributed quantum chemistry packages [27, 28], they are often implemented for specialized purposes only, restricting choices of the electronic structure method as well as the computational accuracy and efficiency. For instance, the quantum chemistry package Gaussian98 [27] does not provide the MM3 FF which is very popular with organic chemists. As a result, many research groups have chosen to implement their own QM/MM modules [19, 24, 29, 30, 31, 32, 33, 34, 35]. In the present work, we describe a flexible QM/MM code that is able to combine essentially any existing QM and MM package for QM/MM calculations. As QM code, we use our own parallel DF program PARAGAUSS [8, 9]. Our design adapts the integrated molecular orbital and MM (IMOMM) method [36] for the treatment of metal–ligand systems, including direct metal–metal interactions, which are not accessible with many common FFs.

QM/MM method

In this section, we first discuss pertinent details of the IMOMM method as well as our modifications. This section also covers the formal treatment of link atoms between the regions described by QM and MM. Then we describe our implementation of the QM/MM interface; subsequently we provide computational details of the test systems. Finally, we examine different schemes for treating link atoms, different choices of the subsystem treated quantum mechanically, and steric effects. We also present results of tests on various functionalized organic molecules.

Integrated molecular orbital and MM

The central idea of the QM/MM approach is to separate a large and complex system XY into a "central" part X and its surrounding Y, where the central part is to be treated by an accurate quantum mechanical method, whereas the surrounding is described at a lower level of theory, for example, by a FF approach. Typical systems are active sites of transition-metal complexes used in homogeneous catalysis and defects of surfaces as well as reaction centers in zeolites or enzymes. The topic of this work is another class of systems, namely metal clusters in shells of bulky ligands.

In general, the energy expression for such a combined approach is written as

$$E(\mathrm{XY}) = E_{\mathrm{high}}(\mathrm{X}) + E_{\mathrm{low}}(\mathrm{Y}) + E_{\mathrm{int}}(\mathrm{X},\mathrm{Y}), \tag{1}$$

where E is the total energy of the complete system XY, $E_{\mathrm{high}}(\mathrm{X})$ and $E_{\mathrm{low}}(\mathrm{Y})$ are the energies of the subsystems X and Y, and $E_{\mathrm{int}}(\mathrm{X},\mathrm{Y})$ provides a correction for the interface between the two subsystems. Most QM/MM methodologies have applied this energy expression in which the self-consistency of the QM electronic structure is not separable from the MM environment [14, 15, 37, 38, 39, 40]. This interdependent combination can be referred to as a "connection scheme".

Alternatively, if the total energy is to be "extrapolated" from a lower level of theory, one writes

$$E(\mathrm{XY}) \approx E_{\mathrm{app}}(\mathrm{XY}) = E_{\mathrm{high}}(\mathrm{X}) + E_{\mathrm{low}}(\mathrm{XY}) - E_{\mathrm{low}}(\mathrm{X}), \tag{2}$$

avoiding the explicit treatment of an interface term as in Eq. (1). This so-called extrapolation scheme which is realized in the IMMOM approach of Maseras and Morokuma [36] is equivalent to the connection scheme if

$$E_{\mathrm{int}}(\mathrm{X},\mathrm{Y}) = E_{\mathrm{low}}(\mathrm{XY}) - E_{\mathrm{low}}(\mathrm{X}) - E_{\mathrm{low}}(\mathrm{Y}). \tag{3}$$

The term $E_{\mathrm{app}}(\mathrm{XY})$ on the left-hand side of Eq. (2) is an approximation to the total energy of the combined system computed at a high level of theory,

$$E_{\mathrm{high}}(\mathrm{XY}) = E_{\mathrm{app}}(\mathrm{XY}) + \Delta. \tag{4}$$

The target energy $E_{\mathrm{high}}(\mathrm{XY})$ can safely be substituted by the extrapolated energy if the error Δ of extrapolation remains essentially constant with respect to structural variations of the system XY so that the potential-energy surfaces of the exact and the approximated systems run parallel to each other. This would render results of a geometry optimization or a molecular dynamics simulation consistent with the true energy surface.

The energy expression (Eq. 2) is a formal one as long as the energy terms for X and XY are not defined in detail. This is relatively easy to do if there are no (strong) chemical bonds between subsystems X and Y, for example, for a solvated molecule in a simple solution. In the case of (covalent) bonds connecting the two subsystems X and Y, an appropriate method is needed to saturate the dangling bonds which result from the subdivision of the composite system XY. For this purpose, two types of techniques have been proposed: one is to cap the dangling bonds with atoms (atom capping or link atom approach) [14, 15, 36] and the other is to saturate the dangling bonds with pseudo orbitals (orbital capping) [37, 38, 39, 41, 42]. Although the extrapolation scheme of Maseras and Morokuma is not exclusive to one type of bond saturation, most applications to date have applied the link atom approach because it is more straightforward and simpler to implement. Also, when dividing a system XY into subsystems, Coulomb interactions deserve special attention owing to their long-range nature. In a general and accurate QM/MM scheme, electrostatic fields due to charges in the outer part Y have to be

accounted for in the energy terms of the central subsystem, $E_{\mathrm{low}}(\mathrm{X})$ and $E_{\mathrm{high}}(\mathrm{X})$, Eq. (2). When modeling neutral systems without strongly polar bonds, like ligated metal clusters (see later), a simpler yet sufficiently accurate procedure results if one neglects the electrostatic coupling of the QM central part with its surrounding; this strategy is also common in sophisticated QM/MM schemes [26].

In our implementation we investigated various treatments of link atoms. The basic idea is to cap a covalent bond that is cut at the border of the QM subsystem by a hydrogen atom, called the link atom. The situation is illustrated in Fig. 1, where R_1 and R_3 represent the coordinates of atoms involved in the frontier bonds that connect the two subsystems of the QM and MM regions, respectively. R_2 denotes the coordinates of the link atom which caps a frontier bond. R_0 and R_4 represent coordinates of atoms in the QM and MM subsystems, respectively, that do not participate in frontier bonds.

Using this classification of atoms, the IMMOM energy expression (Eq. 2) can be explicitly written as follows:

$$\begin{aligned} E(R_0,R_1,R_3,R_4) = {} & E_{\mathrm{QM}}(R_0,R_1,R_2) \\ & + E_{\mathrm{MM}}(R_0,R_1,R_3,R_4) \\ & - E_{\mathrm{MM}}(R_0,R_1,R_2). \end{aligned} \tag{5}$$

The link atom coordinates R_2 are assumed to be defined by the coordinates of the QM frontier atom R_1 and MM frontier atom R_3,

$$R_2 = R_2(R_1,R_3). \tag{6}$$

Otherwise, if the link atom degrees of freedom were optimized without constraints, artificial degrees of freedom, not present in the original undivided system, would be introduced. The gradients of the energy expression (Eq. 5) are given as

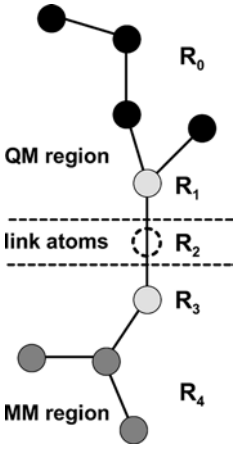

Fig. 1. Classification of atomic coordinates in a quantum mechanics (*QM*)/molecular mechanics (*MM*) calculational scheme: R_0 atoms inside the QM region, R_1 QM frontier atoms, R_2 link atoms, R_3 MM frontier atoms, R_4 atoms inside the MM region

$$\begin{aligned} \nabla E_{\mathrm{total}}(R_0,R_1,R_3,R_4) = {} & \nabla E_{\mathrm{QM}}(R_0,R_1,R_2)J(R_2;R_1,R_3) \\ & - \nabla E_{\mathrm{MM}}(R_0,R_1,R_2)J(R_2;R_1,R_3) \\ & + \nabla E_{\mathrm{MM}}(R_0,R_1,R_3,R_4), \end{aligned} \tag{7}$$

where $J(R_2;R_1,R_3)$ is the Jacobian matrix associated with transforming gradients with respect to coordinate R_2 to R_1 and R_3. Note that the original IMOMM scheme [36] does not include coordinates R_3 in the energy expression; rather, the link atom coordinates R_2 are treated as explicit geometric variables,

$$\begin{aligned} E(R_0,R_1,R_2,R_4) &= E(R_0,R_1,R_3,R_4) \\ &\text{with } R_3 = R_3(R_1,R_2), \end{aligned} \tag{8}$$

and the corresponding gradient expression reads

$$\begin{aligned} & \nabla E_{\mathrm{total}}(R_0,R_1,R_2,R_4) \\ & \quad = \nabla E_{\mathrm{QM}}(R_0,R_1,R_2) - \nabla E_{\mathrm{MM}}(R_0,R_1,R_2) \\ & \qquad + \nabla E_{\mathrm{MM}}(R_0,R_1,R_3,R_4)J(R_3;R_2,R_1). \end{aligned} \tag{9}$$

The reason for transforming gradients from coordinate R_3 to R_2 and R_1 in the original IMOMM scheme is that a geometry optimization on the coordinate set (R_0,R_1,R_2) for the QM subsystem can be performed using existing quantum chemistry programs while geometry optimization on R_4 is done separately by a MM program where all other coordinates R_1, R_2, and R_3 are fixed at their QM values. Since our design is based on an independent geometry optimization program, the "true" coordinate set (R_0,R_1,R_3,R_4) can be used in the context of Eqs. (5), (6), and (7). Employing this true coordinate set and an independent geometry optimization module has several advantages [35]: it avoids the unwanted explicit treatment of the degrees of freedom of the link atoms, and, even more importantly, it admits a modular design of the QM/MM interface which allows flexible use of various QM and MM codes.

To facilitate the discussion of the various ways for treating the bond saturating link atoms, we introduce the terms "link bond" as the bond between a QM frontier atom and a link atom, R_1–R_2, and "frontier bond" as the bond between a QM frontier atom and a MM frontier atom, R_1–R_3. Our implementation includes three different options to determine the link atom coordinates subject to the constraint that the link atom remains located on the frontier bond vector (Fig. 1):

1. Fixed link-bond length R_{12} and fixed frontier-bond length R_{13} [36]. If the link-bond length and the frontier-bond length are fixed, no transformation of gradients on the link atom to the frontier atom coordinates is required. This option is the simplest and was widely employed in early applications of the IMOMM method [43, 44, 45, 46]. Since it freezes the frontier bonds, it restricts the geometry of the combined system XY in a rather artificial fashion. This

option was introduced into our code for comparison only and thus will not be discussed further.

2. Flexible link-bond length and flexible frontier-bond length [47]. This option allows a change of the link-bond length proportional to the length of the frontier-bond length, shown in the following where g is the constant scaling factor.

$$R_2 := R_1 + g(R_3 - R_1), \tag{10}$$

3. Fixed link-bond length and flexible frontier-bond length [35]. With this option, the frontier-bond length is flexible, but the link-bond length is fixed at a user-defined value R_{12}^0,

$$R_2 := R_1 + \frac{R_{12}^0}{|R_3 - R_1|}(R_3 - R_1). \tag{11}$$

By expanding the Jacobian matrix in Eq. (7), the explicit gradients for each type of coordinate in a QM/MM calculation can be written as follows:

$$\begin{aligned}
\frac{\partial E}{\partial R_0} &= \frac{\partial E_{\mathrm{QM}}(R_0,R_1,R_2)}{\partial R_0} + \frac{\partial E_{\mathrm{MM}}(R_0,R_1,R_3,R_4)}{\partial R_0} \\
&\quad - \frac{\partial E_{\mathrm{MM}}(R_0,R_1,R_2)}{\partial R_0} \\
\frac{\partial E}{\partial R_1} &= \frac{\partial E_{\mathrm{QM}}(R_0,R_1,R_2)}{\partial R_1} + \frac{\partial E_{\mathrm{QM}}(R_0,R_1,R_2)}{\partial R_2}\frac{\partial R_2(R_1,R_3)}{\partial R_1} \\
&\quad + \frac{\partial E_{\mathrm{MM}}(R_0,R_1,R_3,R_4)}{\partial R_1} - \frac{\partial E_{\mathrm{MM}}(R_0,R_1,R_2)}{\partial R_1} \\
&\quad - \frac{\partial E_{\mathrm{MM}}(R_0,R_1,R_2)}{\partial R_2}\frac{\partial R_2(R_1,R_3)}{\partial R_1} \\
\frac{\partial E}{\partial R_3} &= \frac{\partial E_{\mathrm{QM}}(R_0,R_1,R_2)}{\partial R_2}\frac{\partial R_2(R_1,R_3)}{\partial R_3} \\
&\quad + \frac{\partial E_{\mathrm{MM}}(R_0,R_1,R_3,R_4)}{\partial R_3} \\
&\quad - \frac{\partial E_{\mathrm{MM}}(R_0,R_1,R_2)}{\partial R_2}\frac{\partial R_2(R_1,R_3)}{\partial R_3} \\
\frac{\partial E}{\partial R_4} &= \frac{\partial E_{\mathrm{MM}}(R_0,R_1,R_3,R_4)}{\partial R_4}.
\end{aligned} \tag{12}$$

For option 2, the Jacobian transformations in Cartesian coordinates is

$$\begin{aligned}
\frac{\partial R_{2,i}}{\partial R_{1,j}} &= (1-g)\delta_{ij}, \\
\frac{\partial R_{2,i}}{\partial R_{3,j}} &= g\delta_{ij},
\end{aligned} \tag{13}$$

where i and j denote Cartesian components x, y, and z, and δ is the Kronecker symbol. For option 3, the Jacobian transformations can be written as

$$\begin{aligned}
\frac{\partial R_{2,i}}{\partial R_{1,j}} &= \delta_{ij} + \frac{R_{12}^0}{|R_3 - R_1|}e_i e_j - \frac{R_{12}^0}{|R_3 - R_1|}\delta_{ij}, \\
\frac{\partial R_{2,i}}{\partial R_{3,j}} &= -\frac{R_{12}^0}{|R_3 - R_1|}e_i e_j + \frac{R_{12}^0}{|R_3 - R_1|}\delta_{ij},
\end{aligned} \tag{14}$$

where e_i and e_j are the components i and j of the unit vector $e = (R_3 – R_1)/|R_3 – R_1|$.

Implementation of the QM/MM approach

Our QM/MM suite includes three independent programs which are combined by interfaces and drivers. QM and MM energies and gradients are computed by the parallel DF program ParaGauss [8, 9] and the FF program TINKER [48], respectively. An interface layer then pipes results from the two programs to the separate geometry-optimizing module OPTIMIZER [49] which performs the overall structure optimization. As stated earlier, this modular structure hinges crucially on using the true coordinates (R_0,R_1,R_3,R_4) as opposed to the original IMOMM approach [36], which employed the coordinate set (R_0,R_1,R_2,R_4). All modules can be utilized as ''black boxes''. At the level of the QM/MM interface, no special knowledge of the internal structure of the various modules is required; the interface suite of modules therefore is designed as a series of command shell scripts (and programs) that coordinate the data exchange between existing programs. Pertinent tasks of the interface and driver layer are the input preparation for the QM and MM programs derived from a generalized master input, the summation of various gradient contributions to a complete set of gradients for the system XY, which is then fed to OPTIMIZER, and, finally, the calculation of the total energy. A special feature of ParaGauss is a symmetrization procedure which allows application of local symmetry constraints during the QM calculation so that this local symmetry can be favorably exploited to speed up the QM calculation which by far is the most expensive computational step.

Computational details

To test our approach, we carried out DF calculations [50] using the linear combination of Gaussian-type orbitals fitting-functions DF [51] method as implemented in the parallel program ParaGauss [8, 9]. Since almost all of our tests focused on geometric aspects, we applied the local density approximation (LDA) [52] for the exchange–correlation potential [53]. For our test calculations, we chose economic basis sets, which were contracted in generalized form, derived from atomic eigenfunctions: (9s,5p,1d) → [4s,3p,1d] for carbon, oxygen and nitrogen [54], (12s,9p,2d) → [5s,4p,1d] for sulfur [55], (6s,1p) → [3s,1p] for hydrogen [54], and (15s,11p,6d) → [6s,4p,3d] for copper [56, 57]. We constructed auxiliary basis sets according to a standard

procedure [51] to represent the electron charge density for approximate evaluation of the classical interelectronic Coulomb energy. We employed the MM3 FF [58, 59, 60, 61] for the MM calculations since its quality for describing organic compounds (e.g. ligands with long alkyl chains) is well established [62]. We discuss later how we treated link atoms and related parameters. In all geometry optimizations, we required Cartesian components of the gradient vectors to be converged to at least 10^{-5} au.

Tests on organic molecules

The results of QM/MM calculations on small organic molecules to be presented in the following sections served a twofold purpose. First we examined the influence of different treatments of link atoms and we studied the sensitivity of pertinent parameters with respect to molecular geometries. Then, these examples validated our combined DF and FF QM/MM implementation and they allowed an estimate of the achievable accuracy of calculated molecular structures.

From the three options of link atom treatment described before, option 1 has been widely adopted and thus its validity has been tested thoroughly [43, 44, 45, 46]. For option 2, the proper choice of the scaling factor g was recently addressed [47]; here, we report a more extensive elaboration of how to select the pertinent parameters for this treatment of link atoms. To the best of our knowledge, option 3 was not previously validated. Besides hydrocarbons, our tests also included functionalized organic molecules.

Choices of scaling factor value g and fixed link-bond length R_{12}^0

Dapprich et al. [47] proposed the value $g = 0.709$ for the scaling factor which determines the link-bond length that substitutes a C–C bond at the border of the QM to MM regimes. This value is the ratio of a standard C–H bond length (1.084 Å) and a standard C–C bond length (1.528 Å) [47]. To determine the parameter R_{12}^0 for option 3, it was proposed to use the C–H bond length obtained by optimization of the isolated QM fragment [35]. We selected ethane as a simple test molecule to study in detail the sensitivity of link-bond parameters. We chose a methyl moiety as the QM fragment; then the boundary between the QM and MM regions falls on a C–C bond (Fig. 2). We varied the value of g between 0.5 and 0.9 and the link-bond lengths from 0.8 to 1.4 Å. We compare results of geometry optimizations by QM, MM and QM/MM methods in Tables 1 (g variation) and 2 (R_{12}^0 variation).

Inspection of Table 1 shows that variation of g does not affect the C–H bond length, in neither the QM nor the MM subsystems; changes are below 0.002 Å. On the other hand, the value of g affects, to some extent, the H–C–H bond angle of the QM subsystem; this angle increases with increasing value of g, varying by about 5° in the parameter range examined. The same trend was previously observed for acetaldehyde [47]. On the other hand, for not too extreme values of g the calculated C–C bond length varies in an apparently random fashion in a range of 0.02 Å. Variations of the H–C–H angle are

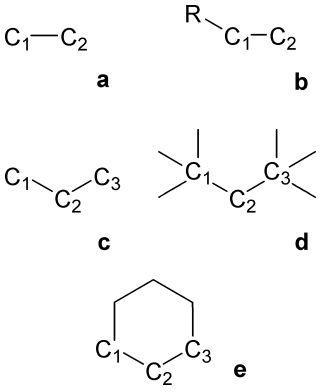

Fig. 2. Sketches of organic molecules used as test systems

Table 1. Influence of the scaling factor g (which determines the length of the link atom bond; see text) on the optimized geometry of ethane. Values pertaining to the quantum mechanics (*QM*) subsystem are given in *boldface*. In the right-most column, we give the corresponding fixed link bond length parameter R_{12}^0 (see text). For comparison, we list results of full QM and molecular mechanics (*MM*) calculations. Distances in angstrom, angles in degree

g	C1–H1	C2–H2	C1–C2	H–C1–H	H–C2–H	R_{12}^0
0.50	**1.106**	1.113	1.513	**104.7**	107.4	0.756
0.60	**1.108**	1.113	1.500	**105.1**	107.3	0.900
0.65	**1.108**	1.113	1.513	**105.8**	107.4	0.983
0.70	**1.108**	1.113	1.522	**106.4**	107.4	1.065
0.75	**1.108**	1.113	1.522	**107.1**	107.4	1.142
0.80	**1.108**	1.113	1.517	**107.8**	107.4	1.214
0.90	**1.107**	1.113	1.530	**109.4**	107.4	1.376
QM	**1.110**	**1.110**	**1.513**	**107.2**	**107.2**	
MM	1.113	1.113	1.531	107.5	107.5	

Table 2. Influence of the fixed link bond length R_{12}^0 ($C–H_{link}$) on the optimized geometry of ethane. Values pertaining to the QM subsystem are given in *boldface*. In the right-most column, we give the corresponding proportional link bond length parameter g. For comparison, we also list results of full QM and MM calculations. Distances in angstrom, angles in degree

R_{12}^0	C1–H1	C2–H2	C1–C2	H–C1–H	H–C2–H	g
0.800	**1.106**	1.113	1.528	**104.9**	107.5	0.52
0.900	**1.108**	1.113	1.529	**105.4**	107.5	0.59
1.000	**1.108**	1.113	1.530	**106.0**	107.5	0.65
1.100	**1.108**	1.113	1.531	**106.8**	107.5	0.72
1.200	**1.108**	1.113	1.532	**107.7**	107.5	0.78
1.300	**1.107**	1.113	1.533	**108.7**	107.4	0.85
1.400	**1.107**	1.113	1.535	**109.7**	107.4	0.91
QM	**1.110**	**1.110**	**1.513**	**107.2**	**107.2**	
MM	1.113	1.113	1.531	107.5	107.5	

below 3° for g values between 0.6 and 0.8, which is quite a wide range around the suggested value of $g = 0.709$. These findings permit a relatively large margin for selecting the scaling factor. We observed similar trends for option 3 with a fixed link-bond length, also entailing a wide safety margin for selecting values of R_{12}^0 between 0.9 and 1.3 Å. In comparison to the proportional link-bond approach, option 2, the C–C bond varies regularly for option 3, increasing with the link-bond length. Whereas option 2 (Table 1) exhibits an irregular variation of this bond length between 1.50 and 1.53 Å, option 3 (Table 2) features a notably smaller variation over a range of 0.007 Å for the parameter range scanned; also, the C–C bond length remains close to the MM result. Note that acceptable values of R_{12}^0 and g correspond to each other (Tables 1, 2).

Effects of the QM/MM partitioning scheme: a case study on propane

A propane molecule (Fig. 2) can be divided in various ways for a QM/MM treatment. With one CH_n moiety ($n = 2$ or 3) treated by QM and the other two CH_n moieties by MM (symbolized by Q and M, respectively), three types of model strategies are possible: M-Q-M, Q-M-M, and Q-Q-M. The first two variants allow a comparison of the QM treatment of an end and an inner group, while the third scheme can be regarded as an improved description of an end group (cf. the last two variants). As one can see from Table 3, the QM/MM method yields acceptable results even with the smallest QM model Q-M-M. It predicts the bond length C1–H and the angles H–C1–H of the QM subsystem within an error of 0.003 Å and 1°, respectively. These errors are reduced to 0.001 Å and 0.3° when the QM subsystem is enlarged, in Q-Q-M. Deviations of the link atom approaches used here are comparable to those obtained previously in a study with an orbital capping method [41]. For the geometric parameters characterizing the environment of the central carbon atom C2, the partitioning Q-Q-M yields slightly better results than the partitioning M-Q-M (improvement by 0.001 Å and 0.4–0.7° for C2–H and H–C2–H, respectively, Table 3); note that the two methyl end groups (head and tail groups) are treated at different accuracy. Overall, the QM results are reproduced in a very satisfactory fashion, but one should keep in mind that deviations of the QM/MM approach may be larger than differences between QM and MM calculations; cf. the angle H–C2–H in Table 3.

Table 3 also affords a comparison of the fixed and flexible link atom approaches 2 and 3. We used the parameters $g = 0.709$ for 2 and $R_{12}^0 = 1.106$ Å for 3. For the fixed link atom scheme (3), we obtained slightly better results for bond angles (by up to 0.5°). In addition, the C–C bond length at the QM/MM boundary, which must be represented by the MM potential owing to the QM/MM energy expression, is better reproduced by option 3, as already noted for ethane (Tables 1, 2).

Table 3. Influence of the QM and MM partitioning on the optimized geometry of propane. For the QM/MM calculations, we compare values obtained with the proportional (option 2) and fixed (option 3) link bond length approach (see text). Values pertinent to the QM subsystem are given in *boldface*. Results of full QM and MM calculations are listed for comparison. Distances in angstrom, angles in degree

	QM	MM	QM/MM		
			Q-M-M (2/3)	M-Q-M (2/3)	Q-Q-M (2/3)
C1–C2	**1.512**	1.534	1.525/1.533	1.527/1.533	**1.517**/**1.516**
C2–C3	**1.512**	1.534	1.534/1.534	1.527/1.533	1.529/1.534
C1–H	**1.111**	1.113	**1.108**/**1.108**	1.113/1.113	**1.110**/**1.110**
C2–H	**1.113**	1.115	1.115/1.115	**1.110**/**1.110**	**1.111**/**1.111**
C3–H	**1.111**	1.113	1.113/1.113	1.113/1.113	1.113/1.113
C1–C2–C3	**112.4**	112.4	112.4/112.4	113.0/112.6	112.7/112.7
H–C1–H	**107.5**	107.4	**106.5**/**106.8**	107.4/107.4	**107.2**/**107.2**
H–C2–H	**107.7**	106.7	106.7/106.7	**105.5**/**106.0**	**106.2**/**106.4**
H–C3–H	**107.5**	107.4	107.4/107.4	107.4/107.4	107.4/107.4

Steric effects in bulky and ring systems: branched nonane (2,2,4,4-tetramethylpentane) and cyclohexane

One of the reasons to apply a QM/MM method is that steric effects of bulky ligands can be incorporated in an accurate treatment of a smaller model system via an economic MM approach. For systems with large steric effects, a QM/MM approach may even become the method of choice since DF calculations on the whole system, employing one of the common exchange–correlation approximations, are unable to reproduce dispersion interactions correctly [53]. Thus, with an adequate parameterization of weak interactions, a QM/MM calculation may be even superior to a full DF calculation. Here we consider a branched conformer of nonane, 2,2,4,4-tetramethylpentane $(CH_3)_3C(CH_2)C(CH_3)_3$ (Fig. 2), where only the central CH_2 group is treated by QM. In the QM/MM calculation, this unit is capped with two link atoms that generate a methane model molecule. The results of QM, MM, and QM/MM calculations are compared in Table 4.

In particular the bond angles at the central carbon atom C2 deviate from the ideal tetrahedral value of

Table 4. Comparison of QM, MM, and QM/MM results for the optimized geometry of 2,2,4,4-tetramethylpentane, $(Me)_3CCH_2C(Me)_3$, a branched nonane. Values pertinent to the QM subsystem are given in *boldface*. Distances in angstrom, angles in degree

	C1–C2	C2–C3	C2–H	H–C2–H	C1–C2-C3
QMmethane			**1.106**	**109.5**	
QM	**1.536**	**1.536**	**1.118**	**105.2**	**125.7**
MM	1.555	1.555	1.115	104.5	124.1
QM/MM proportional link bond	1.545	1.545	**1.114**	**102.7**	124.5
QM/MM fixed link bond	1.554	1.554	**1.115**	**102.3**	124.5

109.5°. The angle H–C2–H is reduced to 105° in the QM calculation, while the angle C1–C2–C3 opens to 126°. Both QM/MM approaches underestimate this angle by only 1°, whereas the angle H–C2–H deviates by up to 3°. Nevertheless, the essential features of the central CH_2 moiety of 2,2,4,4-tetramethylpentane are correctly reproduced (Table 4). The C2–H bond lengths of the QM subsystem as obtained by the QM/MM calculations, 1.114 and 1.115 Å, are in good agreement with the QM result, 1.118 Å.

Steric effects of the MM part of a system onto the QM part can also be demonstrated for cyclohexane (Table 5, Fig. 2); we chose the QM subsystem to comprise three (adjacent) CH_2 units that generate propane when capped by two link atoms. The QM/MM method was found to predict bond lengths and bond angles with deviations of at most 0.003 Å and 0.7° in comparison with a full QM calculation. The structural parameters of the propane subunit in hexane and of free propane are not too different, but for instance bond length C1–C2 and angle C1–C2–C3 illustrate that the QM/MM approach is able to reproduce even relatively small differences at least qualitatively (Table 5).

For the two examples just discussed, the fixed link bond option 3 yields almost the same results for the QM subsystem as the flexible link bond approach (option 2). As for the examples discussed previously, the C–C bond length at the QM/MM boundary calculated with the fixed link bond approach, 1.554 Å, is closer to the MM result of 1.555 Å than the flexible link bond variant, which yields 1.545 Å; this latter value falls between the MM value and the QM value, 1.536 Å.

Substituted hydrocarbons

Finally, we tested organic molecules H_3CCH_2R (Fig. 2) that contain an ethyl moiety H_3CCH_2 and a functional group R = OH, NH_2, SH, CHO, or COOH. In the QM/MM calculations, the CH_2R moiety constitutes the QM subsystem and the methyl group the MM subsystem. We compare results of the QM/MM approach with QM and MM calculations on the substituted system and with QM calculations on the isolated HCH_2R subunits in Table 6. We applied only the slightly more accurate fixed link bond approach because we found results of the flexible link bond variant to be very close.

Inspection of Table 6 reveals that the QM/MM method reproduces the QM geometries of all molecules very accurately. This is particularly remarkable because the differences between analogous methyl and ethyl species are quite small. On average, they amount to 0.0036 Å for the 17 bonds displayed and to 0.3° for the five angles (Table 6). The deviations of the QM/MM from the QM result are on average 0.0016 Å and 0.5°, respectively, where the largest deviation of an angle is due to a single unfavorable case (ethyl aldehyde). When geometrical differences between substituted methyl and ethyl compounds are somewhat larger, the QM/MM approach yields results close to the full QM calculation. Examples are the bond lengths C1–O in ethanol, C1–H in ethyl amine, and C1–S in ethyl thiol (Table 6). Errors obtained with our QM/MM approach are comparable or even smaller than those obtained by an orbital capping method [42]. That approach for saturating dangling bonds can lead to larger errors for geometric parameters involving the frontier atoms, although it accurately predicts bond properties inside the QM subsystem. For instance, the error in bond lengths [C1–O in ethanol, C1–N in ethyl amine, C1–C(O) in propionic acid] obtained with the bond capping method are 0.014, 0.030, and 0.028 Å, respectively, whereas our results are an order of magnitude more accurate, with errors of at most 0.002 Å. This success provides a pragmatic justification as to why, despite criticism, the link atom approach is very popular [37, 38, 39, 40, 41, 42].

Table 5. Comparison of QM, MM and QM/MM results (proportional link bond as well as fixed link bond length approach) for the optimized geometry of cyclohexane. The QM subunit is propane. Distances in angstrom, angles in degree

	QM	MM	QM/MM		QM propane
			Proportional bond length	Fixed bond length	
C1–C2	1.518	1.536	1.519	1.518	1.512
C1–H	1.114	1.114	1.112	1.113	1.113
C2–H	1.114	1.114	1.111	1.112	1.111
H–C1–H	106.7	106.7	106.0	106.0	107.7
H–C2–H	106.7	106.7	107.0	107.1	107.5
C2–C1–C3	111.2	111.3	111.5	111.6	112.4

Table 6. Comparison of QM, QM/MM (fixed link bond lengths) and MM optimized geometries of various substituted ethane species CH_3CH_2R. For comparison, we list the results of QM calculations on the QM subsystems HCH_2R (QM_{sub}). Distances in angstrom, angles in degree

		QM	QM/MM	QM_{sub}	MM
OH	O–H	0.980	0.979	0.979	0.948
	C1–O	1.410	1.409	1.403	1.418
	C1–H	1.120	1.119	1.118	1.114
	C1–O–H	108.5	108.2	108.4	108.4
NH_2	N–H	1.032	1.030	1.031	1.016
	C1–N	1.448	1.450	1.445	1.456
	C1–H	1.118	1.118	1.111	1.115
	H–N–H	107.1	107.0	106.8	106.4
SH	S–H	1.371	1.370	1.370	1.343
	C1–S	1.850	1.844	1.838	1.813
	C1–H	1.109	1.110	1.106	1.114
	C1–S–H	97.6	96.7	96.9	96.8
CHO	C=O	1.209	1.207	1.208	1.209
	C–H	1.134	1.136	1.135	1.119
	C1–C(O)	1.491	1.490	1.486	1.522
	C1–H	1.117	1.114	1.112	1.114
	C1–C=O	124.0	126.0	124.6	124.8
COOH	C=O	1.211	1.211	1.210	1.207
	O–H	0.991	0.992	0.991	0.974
	C1–C(O)	1.494	1.494	1.488	1.523
	C1–H	1.113	1.110	1.107	1.114
	O=C–O	122.0	122.0	122.3	121.7

Application to Cu thiolate species

In the following we apply our QM/MM method to small metal thiolate species which can serve as models for evaluating our approach with regard to metal cluster compounds. First we describe the model clusters and some computational details, then we compare and discuss the results of various QM and QM/MM calculations.

Details of models and methods

Models of ligated metal clusters

The best studied metal cluster compounds with thiolate ligands are gold thiolate clusters [1, 4, 5, 6]. They are synthesized in analogy to the well-known self-assembled monolayers of thiolates on gold surfaces [63]. However, for the present evaluation we chose the first-row transition-metal copper of the triad of coinage metals, to minimize the computational effort. While we are not aware of a synthesis of Cu thiolate clusters, analogous self-assembled monolayers on Cu surfaces are known [63, 64].

As the simplest test systems for the QM/MM approach we chose a Cu atom with an alkane thiolate ligand. Next, we extended this oversimplified model to mono- and biligated Cu_4 species. The model cluster Cu_4L_2 with $L = S(CH_2)_nCH_3$ comprises all essential interactions present in metal thiolate clusters, namely metal–metal bonds, metal–ligand bonds, and for some conformations, if suitable symmetry constraints are applied, also ligand–ligand interactions.

MM parameters

Organometallic compounds are difficult to treat by means of a FF approach because transition-metal species demand specific FF parameters which may depend on oxidation state and coordination. Parameters have been established for molecules including single-centered transition-metal complexes [65, 66]. However, the situation is rather difficult, if not hopeless, for metal clusters stabilized by ligand shells owing to the metal–metal interaction which is not well described by a conventional FF approach that relies on directional bonds and angular forces. Also, experience from coordination compounds [67, 68] with a single metal center as well as bioinorganic molecules [69] shows that metal parameters are hardly transferable from one group of compounds to the next.

The problem of parameterizing metal–metal interactions can be circumvented by partitioning the metal–ligand system under consideration in such a way that all metal atoms belong to the QM region. Since in the QM/MM expressions for the energy (Eq. 2) and the forces (Eq. 7), the MM contributions due to atoms in the QM region cancel exactly, these contributions can be completely omitted or treated by any suitable approximation. In the following we illustrate this approach for the metal alkane thiolate species examined. Different requirements arise regarding the FF parameters, depending on where the border between the QM and MM regions is drawn.

All terms of the FF referring to atoms in the MM subsystem have to be treated explicitly, whereas we will apply approximations in the QM part. For metal thiolate species we considered two partitioning schemes that assign the metal moiety to the QM region. In the first scheme, the ligands are terminated directly after the sulfur head group (SC scheme), whereas in the second one the carbon center bound to sulfur is also included in the QM region (CC scheme). Thus, either S–C or C–C dangling bonds have to be saturated. The link atoms were treated according to the fixed link bond length scheme (3). We set the parameter R_{12}^0 to 1.375 Å for the S–H bond (SC scheme) and to 1.108 Å for the C–H bond (CC scheme); these values were adopted from DF-LDA optimizations of the moieties CuSH and $CuSCH_3$, respectively.

For all metal atoms M of the QM region we reduced the FF to the van der Waals terms only. This approach guaranties cancellation of the contributions of these terms in the QM/MM expressions of energy and forces and is easy to realize in a conventional MM code. Nevertheless, depending on the partitioning scheme, different FF terms involving metal centers are still required. For the MM3 FF which we applied, the following terms (including link atoms) occur:

SC scheme : bending $M-S-H$, $M-S-C$
torsion $M-M-S-H$, $M-M-S-C$,
$M-S-C-H$, $M-S-C-C$

CC scheme : torsion $M-S-C-H$, $M-S-C-C$

If we neglect the small torsional parameters involving metal atoms M, we are left without additional parameters in the CC scheme, and we have to introduce only two metal-related bending terms in the SC scheme.

Although transferring missing parameters from other FFs in general is not recommended [70], there are some successful examples in the literature [41, 71, 72]. In such cases, van der Waals parameters for metals of a universal FF (UFF) were used within an MM3 approach [71]. To determine the missing parameters, we adopted a procedure which was designed specifically for the MM3 FF [70]. This strategy allows construction of FF parameters from atomic parameters similar to the approach taken in rule-based FFs like DREIDING [73] and UFF [74]. We chose the van der Waals parameters of Cu ($r_0 = 2.26$ Å, $\varepsilon = 0.296$ kcal/mol) in accord with this recipe. We also determined the force constants for the bending terms Cu–S–H and Cu–S–C along these lines: 0.54 and 0.70 mdynÅ/rad, respectively [69]. We adopted the equilibrium bond angles (Cu–S–H 94.0°, Cu–S–C 104.9°) from DF-LDA calculations on the

complexes CuSH and $CuSCH_3$, respectively. To all other degrees of freedom, we applied the MM3 FF [58, 59, 60, 61] without modifications.

Results and discussions

Cu*L* and Cu_4L, $L = SCH_3$, SCH_2CH_3

The complexes CuL with $L = S(CH_2)_nCH_3$ comprise the simplest system for testing the QM/MM scheme and the parameterization for thiolate-ligated transition-metal species introduced earlier. We used the species with $n = 0$ as well as CuSH to determine MM parameters; therefore, the first real test case contains a ligand with $n = 1$. We compare the QM/MM results for $CuSCH_3$ and $CuSCH_2CH_3$ to QM results obtained from DF-LDA calculations in Table 7.

The QM reference results exhibit only small differences between geometric parameters of different Cu thiolate species. The Cu–S distance is the same (2.094 Å) for CuSH and $CuSCH_3$; it shortens by only 0.003 Å for $CuSCH_2CH_3$. Also the S–C distances of the two larger species are very similar; for $CuSCH_2CH_3$, the distance is longer by about 0.01 Å. For angles, one notes more pronounced effects. The Cu–S–H angle of CuSH is 94°; this bond angle opens to 105° and 106° for the two larger complexes. The binding energy of the S-*R* ligand to a copper atom decreases with increasing size of the substituent *R*, from 3.7 eV for $R = H$ to 3.2 eV for $R = CH_2CH_3$ (Table 7).

We carried out QM/MM calculations for both partitioning schemes described earlier. Whereas one can apply the SC scheme to $CuSCH_3$, taking CuSH as the QM subsystem, both schemes can be compared for $CuSCH_2CH_3$. With deviations of up to 0.002 Å, all QM/MM calculations reproduce the Cu–S distance very well. For the S–C distance, on the other hand, relatively large deviations, about 0.04 Å, are obtained with the SC scheme, whereas the CC scheme yields a S–C bond length of 1.855 Å, only 0.005 Å shorter than the QM result for $CuSCH_2CH_3$. Also, all QM/MM calculations describe the bond angle Cu–S–C well, with deviations from the QM results being at most 3° in the SC scheme and about 1° in the CC scheme. We were able to trace the apparent low quality of the SC results for the S–C distance to the MM3 bond equilibrium parameter of 1.807 Å; this value is considerably shorter than the typical value of 1.85 Å obtained in our QM calculations. When we recalculated the QM/MM(SC) results with the S–C bond length parameter set to 1.848 Å according to our QM result for $CuSCH_3$, we obtained a S–C bond distance of 1.856 Å in the QM/MM calculations, in very good agreement with the CC partitioning scheme (Table 7).

Comparing binding energies of ligands from QM and QM/MM calculations reveals that this quantity is much more sensitive to the ligand approximation than structural parameters. As expected, the QM/MM binding energy of a ligand is always similar to the binding energy of the corresponding QM subsystem. In a pure QM approach, we calculate a ligand binding energy of 3.2 eV for Cu–SCH_2CH_3. The QM/MM SC scheme yields a value of 3.77 eV, which is similar to the result obtained for the isolated QM subsytem CuSH of 3.74 eV (Table 7). Correspondingly, the value for the QM/MM CC scheme of 3.29 eV is close to the QM result for $CuSCH_3$ of 3.31 eV. Thus, ligands, simplified in the QM/MM approach, are characterized by binding energies which are about as accurate as QM binding energies of those simplified ligands (see later). This result—not unexpectedly—reflects the fact that electronic structure affects the ligand binding more than small geometric variations accounted for in the QM/MM approach.

Now, we turn to Cu_4 as substrate with one thiolate ligand which, for simplicity, is attached in singly coordinated fashion at a corner of the distorted tetrahedral Cu_4 cluster; Cu–Cu distances vary from 2.19 to 2.26 Å. Assuming C_s symmetry, two orientations of the alkane thiolate chain are conceivable: toward a face or an edge of the trigonal Cu_4 pyramid. Our QM optimizations showed that both orientations of the thiolate chain yield very similar results. Thus, for evaluating the QM/MM implementation we focused on the face orientation. We present the same geometric parameters that we discussed already for the Cu_1 species in Table 7; we compare results of the QM calculations to those of various QM/MM procedures. In addition, we list the average Cu–Cu distance of the Cu_4 moiety.

Compared to the Cu_1 thiolates (Table 7), the QM results for Cu_4SH show an elongated Cu–S distance of 2.111 Å, which shortens by 0.010 and 0.014 Å for the ligands SCH_3 and SCH_2CH_3, respectively. The S–C distance is considerably less affected when going from the mononuclear Cu species to the Cu_4 substrate; it decreases by less than 0.003 Å. We observed small effects on the angle Cu–S–(H)C; it opens by 2.2° for the ligand SH and by about 3° for the larger thiolates. All these very small effects are correctly reproduced by both QM/MM partitioning schemes. Only the S–C distance

Table 7. Comparison of QM and QM/MM (SC and CC partitioning, see text) calculations on monoligated Cu and Cu_4 species. Bond lengths in angstrom, angles in degree, ligand binding energy (*BE*) in electron volt

		Cu–Cu	Cu–S	S–C	Cu–S–(H)C	BE
CuSH	QM		2.094		94.0	3.74
$CuSCH_3$	QM		2.094	1.848	104.9	3.31
	QM/MM(SC)		2.093	1.807	107.7	3.89
$CuSCH_2CH_3$	QM		2.091	1.860	105.6	3.20
	QM/MM(SC)		2.093	1.815	107.6	3.77
	QM/MM(CC)		2.094	1.855	104.5	3.29
Cu_4SH	QM	2.269	2.111		96.2	4.19
Cu_4SCH_3	QM	2.263	2.101	1.847	107.5	3.77
	QM/MM(SC)	2.265	2.104	1.806	109.6	4.37
$Cu_4SCH_2CH_3$	QM	2.262	2.097	1.857	108.4	3.43
	QM/MM(SC)	2.265	2.103	1.814	109.6	4.25
	QM/MM(CC)	2.263	2.100	1.852	107.6	3.76

of the SC scheme deviates again strongly from the QM results (see earlier). As expected, the values obtained with the CC scheme agree slightly better with the QM reference than the values determined with the SC scheme. In general, the deviations of the QM/MM results from the QM geometries are of the same order as in the case of the simpler Cu_1 species (Table 7). The binding energies per S*R* ligand exhibit the same trend as for the Cu_1 compounds; energies are always close to the values of the corresponding free QM subsystem (see earlier).

The Cu–Cu distances of the QM/MM and QM calculations agree very well with each other. For $Cu_4SCH_2CH_3$, the SC result deviates from the QM value by 0.003 Å, whereas the CC scheme agrees up to 0.001 Å or better. In part, this high accuracy results from error compensation because we are discussing an average of four different Cu–Cu bonds. The average absolute deviations of individual Cu–Cu bond lengths from the QM results are 0.007 Å for the SC scheme and 0.002 Å for the CC scheme.

In summary, for both substrate models of monoligated species, Cu_1L and Cu_4L, our QM/MM implementation exhibits very satisfactory performance, with the CC scheme being slightly favored over the SC scheme.

Cu_4L_2, $L = S(CH_2)_nCH_3$, $n = 1$–3

To probe the effect of ligand–ligand interaction on a Cu cluster, we chose the model Cu_4L_2 with the two thiolate ligands attached to faces of the distorted tetrahedral Cu_4 cluster (Fig. 3). To reduce the number of possible conformations, we restricted the cluster symmetry to C_{2v}. It was not our intention to identify the minimum energy conformation; rather, we were interested in controlling the ligand orientation and in testing the QM/MM approach for various situations. For this purpose, we selected three ligand conformations of the cluster $Cu_4(SCH_2CH_3)_2$ (Fig. 4). Conformation A avoids direct interligand interaction. The S–C bonds of the alkane thiolate chains are oriented "downward" (Fig. 4); in this way, the two ligands assume an almost "equatorial" orientation, pointing essentially in opposite directions. Upward orientation of the S–C bond leads to conformation B, with the ligands almost forming a right angle from the center of the substrate cluster. Searching for stronger interligand interaction, we also inspected conformation C, which is derived from conformation B by flipping the first C–C bond of cluster B "inward" (Fig. 4). We applied the more accurate QM/MM(CC) approach; also, for comparison, the less accurate SC scheme was applied to conformation A.

We compare QM and QM/MM results for the three conformations of the cluster $Cu_4(SCH_2CH_3)_2$ in Table 8. Because we are mainly interested in the response of the cluster to different orientations of the ligands, we list only average Cu–Cu bond lengths, the Cu–S distances to the upper (Cu1) and lower (Cu2) copper atoms of the cluster (see Fig. 3) as well as S–C bond lengths and angles Cu1–S–C.

From the QM results we note an effect of the different ligand orientations on the structure of the Cu cluster. When two ethyl thiolate ligands are attached to Cu_4, the average Cu–Cu bond elongates compared to the Cu–Cu bond length of 2.28 Å which was calculated for the isolated tetrahedral Cu_4 species. This bond elongation amounts to 0.01, 0.03, and 0.08 Å for conformations B, A, and C, respectively. As expected, it is strongest for conformation C which is assumed to imply the strongest steric interactions between the ligands. Nevertheless, as

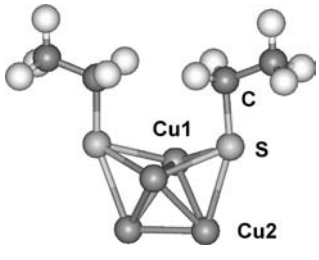

Fig. 3. Structure of the copper thiolate cluster $Cu_4(SCH_2CH_3)_2$ (conformation B, see Fig. 4)

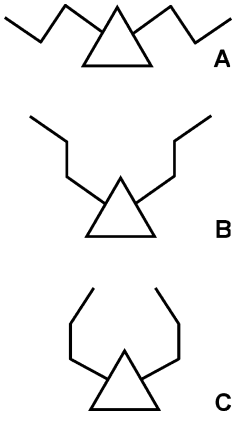

Fig. 4. Sketches of various conformations (ligand orientations) of the copper thiolate cluster $Cu_4(SCH_2CH_3)_2$

Table 8. Comparisons of QM and QM/MM results (SC and CC partitioning) for three conformations A–C of $Cu_4(SCH_2CH_3)_2$ (see Figs. 3,4). Distances in angstrom, bond angles in degree, BE per ligand in electron volt, energy difference, *E*, relative to conformation A in electron volt

		Cu–Cu	Cu1–S	Cu2–S	S–C	Cu–S–C	BE	*E*
A	QM	2.314	2.298	2.683	1.882	107.2	3.56	–
	QM/MM(SC)	2.279	2.330	2.995	1.819	118.8	4.25	–
	QM/MM(CC)	2.319	2.302	2.700	1.872	106.5	3.87	–
B	QM	2.291	2.328	2.548	1.887	105.8	3.44	0.24
	QM/MM(CC)	2.292	2.332	2.559	1.877	104.6	3.76	0.23
C	QM	2.360	2.306	2.253	1.883	133.4	3.56	0.03
	QM/MM(CC)	2.366	2.303	2.254	1.876	133.9	3.83	0.09

can be seen from the total energies (Table 8), conformation C is by 0.2 eV more stable than conformation B and thus close in energy to the most stable conformation A. The ligand binding energies are in line with total energy differences; for conformations A and C, we obtained very similar results, whereas the ligand binding of the less stable structure, conformation B is 0.12 eV smaller (Table 8). These small energy differences suggest weak steric interactions at most.

Also the Cu–S distances exhibit distinct differences between the various conformations (Table 8). The distances Cu1–S to the "upper" Cu atoms are 2.30–2.33 Å and thus do not depend significantly on the ligand orientation (Fig. 4). On the other hand, the bond distance Cu2–S varies significantly among the three conformations. In conformation A, a large value of 2.68 Å shows that the ligands are attached closer to atoms Cu1, thus featuring a quasi-twofold "bridge" coordination. Ligand orientation B entails a slight shift of the ligands toward the triangular faces of the Cu cluster; this is reflected in the shorter Cu2–S distance of 2.55 Å. In conformation C, the ligands are attached close to the center of the triangular faces; the Cu2–S distance, 2.25 Å, is even shorter (by 0.05 Å) than the bond length Cu1–S (Table 8).

For the S–C bond lengths, we found only small effects for different confirmations; compared to conformation A, the bond distance elongates by 0.005 and 0.001 Å for conformations B and C, respectively. For conformations A and B, an angle Cu–S–C of about 106° is found, comparable to that calculated for the monoligated cluster (Table 7), whereas the sterically strained conformation C yields a larger Cu–S–C angle, 133° (Table 8). While other geometric parameters reflect interligand repulsion, H–H distances between ligands are found to be larger than 2.46 Å (conformation B); this shows that no steric stress is present in the equilibrium structures obtained for the various isomers.

Taking into account all these results, we can rationalize the effects of different ligand orientations of conformations B and C as a consequence of avoiding steric repulsion. As demonstrated by the shift of the S atom across the triangular face of the Cu_4 cluster, the cluster itself and the ligand–metal bonds react in flexible fashion and they rearrange. As a result, the final geometries are due to the interplay between steric effects and the flexibility of the S_2Cu_4 core of the system. Ultimately, one has to expect further low-energy local minima on this overall shallow potential-energy surface.

We start the discussion of the QM/MM results by examining geometry A as obtained by means of the SC scheme. Here, the quasi-twofold coordination of the S atom poses a particular problem. The situation demands an adjustment of the Cu–S–C and Cu–S–H angular FF parameters because these angular potentials were parameterized for a singly coordinated S atom. As a first approximation, we chose the same equilibrium angles and force constants as determined for the singly coordinated system; yet, inspection of the results reveals serious deficiencies: The Cu2–S distance is overestimated by 0.3 Å, a large value even if one takes into account that this distance does not correspond to a chemical bond. In addition, the average Cu–Cu distance is underestimated by 0.035 Å, the Cu1–S bond length is overestimated by 0.032 Å, and the Cu–S–C angle is relatively large, 119° compared to 107° as obtained by the QM approach. Because these results were considerably less accurate than the geometry obtained with the CC approach, we refrained from pursuing further the SC approach. Clearly a careful parameterization of the FF is required if the QM region is to terminate close to the metal species.

On the other hand, with the QM/MM CC procedure we obtained quite accurate results for all three conformations. The average deviation of bond lengths amounts to 0.004 Å, while angles are reproduced with deviations of up to 1°. The largest differences in bond lengths are about 0.01 Å. Here we take the Cu2–S distance as a bond when it is comparable to the length of Cu1–S (about 2.4 Å or less, see conformation C). More important than these absolute values is the inspection of trends. All variations of bond distances and angles due to different ligand conformations are reproduced by the CC scheme. Even relatively small elongations of the Cu1–S bond from A to C to B by 0.03 Å are in line for both methods (Table 8). Also the relative QM/MM energies calculated with the CC scheme are in reasonable agreement with the corresponding QM results. The order of stability is A > C > B with about 0.1 eV difference between two conformations. For the ligand binding energies, we note the same trends as observed earlier for monoligated species. The SC scheme yields considerably larger values than the QM results (conformation A), in agreement with the larger binding energy of SH compared to SCH_2CH_3. Because SCH_3 is a better approximation for an ethyl thiolate ligand, ligand binding energies obtained by the CC scheme are closer to the QM reference results (Table 8). Moreover, the variation of the ligand bond strength for the different conformations, A∼C > B, is satisfactorily reproduced by the QM/MM CC approach.

To examine further how different ligand orientations affect the structure of the cluster Cu_4, we also investigated longer alkane chains of the thiolate ligands $S(CH_2)_nCH_3$, extended by one ($n=2$) and two ($n=3$) CH_2 groups; we calculated conformations B and C as examples. Since no short H–H contacts were present in the equilibrium structures with ethylthiolate ligands (see earlier), no strong effects are expected from ligand elongation. QM/MM CC results for pertinent parameters are collected in Table 9. We found no structural effects for either conformation B or C, irrespective of the number of CH_2 groups added. Changes of bond distances remain below 0.002 Å, angles vary by less than 0.5°. This structural stability of the conformation of the cluster core with regard to different ligand tail groups corroborates that the computational approach correctly implements the coupling of the QM and MM regions.

Table 9. Effect of the alkane chain length in $Cu_4(S(CH_2)_nCH_3)_2$, $n=2$–4, for conformations B and C (Fig. 4). Distances in angstrom, angles in degree

	n	Cu–Cu	Cu1–S	Cu2–S	S–C	Cu–S–C
B	2	2.292	2.332	2.559	1.877	104.5
	3	2.292	2.331	2.559	1.877	104.4
	4	2.292	2.331	2.559	1.877	104.4
C	2	2.366	2.303	2.254	1.876	133.9
	3	2.366	2.303	2.254	1.877	133.8
	4	2.366	2.302	2.255	1.877	133.6

These examples demonstrate that our DF-MM3 implementation of the QM/MM approach is able to reproduce QM calculations with sufficient accuracy as well as to transfer the effect of larger ligands successfully to a smaller accurately treated model system. Thus, we are confident that this approach will provide a sufficiently accurate treatment of larger clusters covered by extended ligand shells—systems which, at least for the time being, are outside the scope of conventional "first principles" quantum mechanical methods.

Summary

To model ligated metal clusters, we implemented a QM/MM approach based on the IMOMM method. The approach was realized in a modular way, combining our parallel DF program ParaGauss and the MM software TINKER via a separate geometry-optimization module. We implemented three options for treating link atoms that saturate dangling bonds at the border of QM and MM regions. We evaluated two of them which provide a flexible "frontier bond"; we found them to be essentially equivalent and relatively insensitive with respect to the link bond parameters chosen. Test calculations on bulky compounds and ring systems revealed that the QM/MM method properly takes steric effects into account. We have examined organic molecules that contain functional groups, namely alcohol, thiol, amine, aldehyde, and carboxylic substituents. We showed that the link atom approach reproduces QM optimized geometries with an accuracy comparable to the orbital capping method.

To demonstrate the applicability of the QM/MM approach to ligated metal clusters, we treated various small systems modeling metal–ligand interaction. As test systems we chose copper thiolate species which are analogous to the well-known thiolate-stabilized gold nanoparticles; with copper instead of gold, we were able to keep the computational effort small even in all-electron calculations on ligated metal clusters. The computational strategy chosen relied on a partitioning where the metal species as well as the metal–ligand interaction is treated by QM and the ligand shell is modeled by MM. Exploiting the cancellation of FF terms for the QM subsystem, we approximated the FF terms referring to metal atoms by van der Waals terms only. We probed two schemes for partitioning the metal thiolate species where the QM part comprised the metal core and the thiolate ligands either up to the S–C bond or the first C–C bond. We demonstrated that the second approach (CC) performs better; however, with a specially adjusted parameterization, the first procedure (SC) may also yield satisfactory results.

For simple test systems, a Cu atom coordinated by a single ethyl thiolate and a Cu_4 cluster coordinated by a single thiolate ligand, QM/MM results for geometric parameters agreed very well with those of the corresponding system treated at the QM level. As expected, ligand binding energies determined with the QM/MM scheme correspond to results for the isolated QM subsystems because differences in electronic structure between different ligands examined are more important for energtic aspects than (small) structure variations. Also, the QM/MM approach is able to describe subtle structural differences that result when two ethyl thiolate ligands are attached to a Cu_4 cluster in different conformations. These tests were especially convincing because the geometry of the QM subsystem is identical for two of the conformations (B, C) compared. When the ligands are orientated in "diverging" directions (Fig. 3), extending the alkane thiolate chains to propyl and butyl groups does not have any significant effect on the cluster and the attached head groups of the ligands. We were able to observe rearrangements of the cluster core (e.g., opening of the Cu–S–C angle and elongation of Cu–Cu bonds) for conformations where ligand chains were oriented parallel or pointing toward each other.

In summary, with applications to simple organic as well as organometallic systems, we demonstrated that we successfully implemented a QM/MM approach based on the parallel DF program ParaGauss and a standard FF. QM results for metal–metal, metal–ligand, as well as ligand–ligand interactions showed that the IMOMM strategy is applicable to ligated metal clusters. Next, we will report on applications to transition-metal clusters.

Acknowledgements. We thank J.W. Ponder for providing the source code of the program TINKER. This work was supported by Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie.

References

1. Templeton AC, Wuelfing WP, Murraya RW (2000) Acc Chem Res 33:27
2. Schmid G (1998) J Chem Soc Dalton Trans 1077
3. Schmid G, Maihack V, Lantermann F, Peschel S (1996) J Chem Soc Dalton Trans 589
4. Hostetler MJ, Wingate JE, Zhong C-Z, Harris JE, Vachet RW, Clark MR, Londono JD, Green SJ, Stokes JJ, Wignall GD, Glish GL, Porter MD, Evans ND, Murray RW (1998) Langmuir 14:17
5. Whetten RL, Khoury JT, Alvarez MM, Murthy S, Vezmar I, Wang ZL, Stephen PW, Cleveland CL, Luedtke WD, Landman U (1996) Adv Mater 5: 428

6. Terrill RH, Postlethwaite TA, Chen C-H, Poon C-D, Terzis A, Chen A, Hutchison JE, Clark MR, Wignall G, Londono GD, Superfine R, Falvo M, Johnson CS Jr, Samulski ET, Murray RW (1995) J Am Chem Soc 117:12537
7. Häberlen OD, Chung S-C, Stener M, Rösch N (1997) J Chem Phys 106:5189
8. Belling T, Grauschopf T, Krüger S, Mayer M, Nörtemann F, Staufer M, Zenger C, Rösch N (1999) In: Bungartz H-J, Durst F, Zenger C (eds) High performance scientific and engineering computing. Lecture notes in computational science and engineering, vol 8. Springer, Berlin Heidelberg New York, p 439
9. Belling T, Grauschopf T, Krüger S, Nörtemann F, Staufer M, Mayer M, Nasluzov VA, Birkenheuer U, Hu A, Matveev AV, Shor A, Fuchs-Rohr M, Neyman KM, Ganyushin DI, Kerdcharoen T, Woiterski A, Rösch N (2001) PARAGAUSS version 2.2. Technische Universität München, Munich
10. Häberlen OD, Schmidbaur H, Rösch N (1994) J Am Chem Soc 116:8241
11. Krüger S, Stener M, Mayer M, Nörtemann F, Rösch N (2000) J Mol Struct (THEOCHEM) 527:63
12. Belling T (1998) Dissertation. Technische Universität München, Munich
13. Häkkinen H, Barnett R, Landman U (1999) Phys Rev Lett 82:3264
14. Singh UC, Kollman PA (1986) J Comput Chem 7:718
15. Field MJ, Bash PA, Karplus M (1990) J Comput Chem 11:700
16. Eurenius KP, Chatfield DC, Brooks BR (1996) Int J Quantum Chem 60:1189
17. Clark T, Alex A, Beck B, Gedeck P, Lanig H (1999) J Mol Mod 5:1
18. Treesukol P, Limtrakul J, Truong TN (2001) J Phys Chem B 105:2421
19. Ryde U (1996) J Comput-Aid Mol Des 10:153
20. Eichinger E, Tavan P, Hutter J, Parrinello M (1999) J Chem Phys 110:10452
21. Shoemaker JR, Burggraf LW, Gordon MS (1999) J Phys Chem 103:3245
22. Lopez N, Pacchioni G, Maseras F, Illas F (1998) Chem Phys Lett 294:611
23. Gao J (1998) J Chem Phys 109:2346
24. Kerdcharoen T, Liedl KR, Rode BM (1996) Chem Phys 211:313
25. Bryce RA, Vincent MA, Hillier IH (1999) J Phys Chem A 103:4094
26. Kerdcharoen K, Morokuma K (2002) Chem Phys Lett 355:257
27. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ES, Pople JA (1998) Gaussian 98 (revision A.7). Gaussian, Pittsburgh, PA
28. High Performance Computational Chemistry Group (1998) NWChem version 3.3.1, a computational chemistry package for parallel computers. Pacific Northwest National Laboratory, Richland, WA
29. Truong TN, Stefanovich EV (1997) Chem Phys 218:31
30. Antes I, Thiel W (1999) J Phys Chem A 103:9290
31. Brändle M, Sauer J (1997) J Mol Catal A Chem 119:19
32. Röthlisberger U, Carloni P, Doclo K, Parrinello M (2000) J Biol Inorg Chem 5:236
33. Woo TK, Margl PM, Deng L, Cavallo L, Ziegler T (1999) Catal Today 50:479
34. Murphy RB, Philipp DM, Friesner RA (2000) Chem Phys Lett 321:113
35. Eichler U, Koelmel KM, Sauer J (1996) J Comput Chem 18:463
36. Maseras F, Morokuma K (1995) J Comput Chem 16:1170
37. Thery V, Rinaldi D, Rivail J-L, Maigret B, Ferenczy GG (1994) J Comput Chem 15:269
38. Monard G, Loos M, Thery V, Baka K, Rivail J-L (1996) Int J Quantum Chem 58:153
39. Assfeld S, Rivail J-L (1996) Chem Phys Lett 263:100
40. Nasluzov VA, Rivanenkov VV, Gordienko AB, Neyman KM, Birkenheuer U, Rösch N (2001) J Chem Phys 115:8157
41. Gao J, Amara P, Alhambra C, Field MJ (1998) J Phys Chem A 102:4714
42. Zhang Y, Lee T-S, Yang W (1999) J Chem Phys 110:46
43. Matsubara T, Sieber S, Morokuma K (1996) Int J Quantum Chem 60:1101
44. Matsubara T, Maseras F, Koga N, Morokuma K (1996) J Phys Chem 100:2573
45. Froese RDJ, Musaev DG, Morokuma K (1998) J Am Chem Soc 120:1581
46. Svensson M, Humbel S, Morokuma K (1996) J Phys Chem 105:3654
47. Dapprich S, Komaromi I, Byun KS, Morokuma K, Frisch MJ (1999) J Mol Struc (THEOCHEM) 461:1
48. Ponder JW (1999) TINKER version 3.7. Washington University
49. Nörtemann F (1998) Dissertation. Technische Universität München, Munich
50. Kohn W, Sham LJ (1965) Phys Rev A 140:1133
51. Dunlap BI, Rösch N (1990) Adv Quantum Chem 21:317
52. Vosko SH, Wilk L, Nusair N (1980) Can J Phys 58:1200
53. Görling A, Trickey SB, Gisdakis P, Rösch N (1999) In: Brown J, Hofmann P (eds) Topics in organometallic chemistry, vol 4. Springer, Berlin Heidelberg New York, p 109
54. van Duijnefeldt FB (1971) IBM Res Rep RJ945
55. Veillard A (1968) Theor Chim Acta 12:405
56. Wachters AJH (1970) J Chem Phys 52:1033
57. Hay PJ (1977) J Chem Phys 66:4377
58. Allinger NL, Yuh YH, Lii J-H (1989) J Am Chem Soc 111:8551
59. Schmitz LR, Allinger NL (1990) J Am Chem Soc 112:8307
60. Allinger NL, Chen K, Rahman M, Pathiaseril A (1991) J Am Chem Soc 113:4505
61. Allinger NL, Zhu ZS, Chen K (1992) J Am Chem Soc 114:6120
62. Gundertofte K, Liljefors T, Norrby P-O, Pettersson I (1996) J Comput Chem 17:429
63. Ulman A (1996) Chem Rev 96:1533
64. Driver SM, Woodruff DP (2000) Surf Sci 457:11
65. Landis CR, Root DM, Cleveland T (1995) In: Libkowitz KB, Boyd DB (eds) Reviews in computational chemistry, vol 6. Wiley-VCH, New York, p 73
66. Hagelin H, Åkermark B, Norrby P-O (1999) Organometallics 18:2884
67. Brubaker GR, Johnson DW (1984) Coord Chem Rev 53:1
68. Hay BP (1993) Coord Chem Rev 126:177
69. Zimmer M (1995) Chem Rev 95:2629
70. Allinger NL, Zhou X, Bergsma J (1994) J Mol Struct (THEOCHEM) 312:69
71. Maseras F (1998) New J Chem 327
72. Maseras F, Eisenstein O (1998) New J Chem 5
73. Mayo SL, Olafson BD, Goddard WA III (1990) J Phys Chem 94:8897
74. Rappe AK, Casewit CJ, Colwell KS, Goddard WA III, Skiff WM (1992) J Am Chem Soc 114:10024