LUND UNIVERSITY

On the Convergence of QM/MM Energies

Hu, LiHong; Soederhjelm, Paer; Ryde, Ulf

Link to publication

On the convergence of QM/MM energies

LiHong Hu [ab], Pär Söderhjelm [c], and Ulf Ryde [a] *

[a] Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124, SE-221 00 Lund, Sweden

[b] Faculty of Chemistry, North-east Normal University, Changchun, 130024, P. R. China

[c] Department of Chemistry and Applied biosciences - Computational Science, ETH Zürich, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland

Correspondence to Ulf Ryde, E-mail: Ulf.Ryde@teokem.lu.se,

Tel: +46 – 46 2224502, Fax: +46 – 46 2228648

2012-02-03

Abstract

We have studied the convergence of QM/MM calculations with respect to the size of the QM system. We study a proton transfer between a first-sphere cysteine ligand and a second-sphere histidine group in [Ni,Fe] hydrogenase and use a 446-atom model of the protein, treated purely with QM methods as reference. We have tested 12 different ways to redistribute charges close to the junctions (to avoid overpolarisation of the QM system), but once the junctions are moved away from the active site, there is little need to redistribute the charges. We have tested 13 different variants of QM/MM approaches, including two schemes to correct errors caused by the truncation of the QM system. However, we see little gain from such correction schemes; on the contrary they are sensitive to the charge-redistribution scheme and may cause large errors if charges are close to the junctions. In fact, the best results were obtained with a mechanical embedding approach that does not employ any correction scheme and ignores polarisation. It gives a mean unsigned error for 40 QM systems of different sizes of 7 kJ/mol with a maximum error of 28 kJ/mol. The errors can be significantly decreased if bonds between the QM and MM system (junctions) are moved one residue away from all active-site residues. Then, most QM/MM variants give mean unsigned errors of 5–9 kJ/mol, maximum errors of 16–35 kJ/mol, and only 5–7 residues give an error of over 5 kJ/mol. In general, QM/MM calculations converge faster with system size than pure QM calculations.

Key Words: QM/MM, hydrogen link-atom, electrostatic embedding, mechanical embedding.

Introduction

During the last two decades quantum mechanical (QM) calculations have been established as an attractive and competitive complement to experiments to study biochemical reactions.[1,2,3,4] However, there is still no consensus how such calculations are best performed. In principle, two schools have arisen. In the first, which we will call QM-only in the following, a small part (20–200 atoms) is cut out of the macromolecule of interest, typically the active site and a few nearby residues.[5,6] This system is studied with QM methods, whereas the rest of the macromolecule is either ignored, or more commonly, modelled as a featureless continuum, characterised by a dielectric constant of ~4. It is typically necessary to fix a number of atoms at the periphery of the QM system to model steric restrictions of the macromolecule. Entropic effects can be modelled by a harmonic model, based on calculated vibrational frequencies.

The alternative approach is to include the whole macromolecule in the calculation by the use of combined QM and molecular mechanics (MM) methods, the QM/MM approach.[7,8] In this approach, a central system of a similar size as for the QM-only approach is treated by QM methods, whereas the rest of the macromolecule, as well as some explicit solvent molecules are modelled by MM methods. The advantage with this approach is of course that the whole macromolecule is explicitly modelled and that free energies can be calculated by free-energy perturbations or related approaches.[9,10,11] On the other hand, the size of the system makes the method more expensive and it becomes hard to control the conformation of the MM system.

Unfortunately, there are few direct comparisons of the two approaches. Ochsenfeld and coworkers have studied proton transfer within a 32-residue polypeptide and a 1637-atom model of triose isomerase and shown that the QM/MM approach converges appreciably faster than the QM-only approach with respect to the size of the QM system, although the convergence is quite slow for both approaches.[12] For example, with 299 QM atoms, the errors for the QM-only and QM/MM approaches were 43 and 12 kJ/mol, respectively, and even with 1092 QM atoms, the error in the QM-only approach was still 6 kJ/mol. We have observed a similar slow convergence of the QM-only approach for QM systems up to 696 atoms.[13] Even worse, different ways to select what residues to include in the QM system gave widely different results. In fact, after the addition of 40 residues, there was still a difference of 60 kJ/mol if the residues were added according to their distance to the QM system or if they were added according to their energy components in a QM/MM free-energy perturbation approach.[14] This gives a quite pessimistic view of the use of QM-only methods for the study of biochemical systems.

On the other hand, QM/MM calculations also show a quite slow convergence with respect to the size of the QM system. For example, it has been shown that the error in QM/MM forces are sizeable for QM regions with a radius of up to 9 Å and that QM/MM free energies change by 12 kJ/mol when increasing the size of the QM system from 3 to 5 Å.[15]

In this paper, we supplement our QM-only study by an investigation of the corresponding convergence of the QM/MM approach. Thus, we use a 446-atom model of [Ni,Fe] hydrogenase from our previous investigation[14] as a reference and investigate how the QM/MM energy converges towards the QM energy of the full system as the QM system in the QM/MM treatment is enlarged. We concentrate on the energy of a reaction that takes place inside the QM system, in accordance with the common use of QM/MM methods. Therefore, errors arise primarily from two sources. The first is the treatment of the surroundings by an MM potential, rather than by a more accurate QM method.

The second problem with the QM/MM approach is that it is typically necessary to truncate the QM system by cutting some (normally C–C) bonds in the macromolecule. This is a well-known problem in QM/MM that has been much discussed.[17,13,16,17,18,19] There are several ways to solve it, but the most simple and common one is to truncate the QM system with hydrogen atoms. If the surrounding macromolecule is included in the QM calculations as a point-charge model, there is a risk of overpolarisation, owing to the fact that some point

charges are located close to the hydrogen junction atoms. This has been much discussed and several solutions have been tested and compared.[12,19,20,21,22,23] In general, the results are varying and large errors (60–100 kJ/mol) are occasionally found, especially for charged systems.

We also address another related question that has received much less attention, viz. whether the errors caused by the junctions can be corrected. We test two such correction schemes, one available in the ONIOM approach[24] and another used in the QTCP approach[7] and also some variants of them. Moreover, we test several variants of mechanical embedding (i.e. when electrostatic interactions between the QM and MM systems are treated at the MM level).

We study the energy of a simple proton-transfer reaction between a bridging Cys ligand and a second-sphere His residue in [Ni,Fe] hydrogenase. Although the proton moves only 0.97 Å and the structure of the surroundings hardly change, it has been shown that this reaction is very sensitive to the surroundings, changing the reaction energy from ~0 kJ/mol in vacuum to ~80 kJ/mol in the protein.[26] Therefore, it provides a sensitive test case for the modelling of effects of the surroundings in biochemical reactions.

Methods

QM/MM calculations with electrostatic embedding
In the QM/MM approach,[17,13] a small but important part of the total system (called system 1 or the QM system) is treated by quantum mechanics (QM), whereas the rest (called system 2 or the MM system) is treated by molecular mechanics (MM). The QM/MM calculations have been performed with the program ComQum,[25,26,27] which is a modular combination of the QM software Turbomole 5.10[28] and the MM software Amber 9.[29]

Special attention is needed when there are covalent bonds between the QM and MM systems. Many approaches have been suggested to treat such junctions, e.g. by truncating the QM system by certain link atoms or by using localised orbitals at the junctions.[17,13] We have employed the simplest and most widely used approach, the hydrogen link-atom approach, in which the QM system simply is truncated with hydrogen atoms. To simplify the discussion, we will use the following nomenclature, illustrated in Figure 1:[13,22] The hydrogen link-atom is called HL, whereas the MM atom it replaces is called M1. The QM atom directly connected to HL is Q1, QM atoms directly bound to Q1 are called Q2, and those directly bound to Q2 are called Q3, and so on. Likewise, MM atoms directly connected to M1 are called M2, those directly bound to the M2 atoms are called M3, and so on. Sometimes, we divide all atoms into three systems, viz. those in the MM system, excluding M1 (M), the HL atoms (J), and the rest of the QM atoms (Q).

In principle, the HL atoms will introduce additional degrees of freedom. To avoid this, the HL atoms are placed along the Q1–M1 bond, with a Q1–HL bond length ($r_{Q1\text{-}HL}$) that is proportional to the Q1–M1 bond length ($r_{Q1\text{-}M1}$) according to:

$$r_{Q1-HL} = r_{Q1-M1} \frac{r_{Q1-HL}^{QM0}}{r_{Q1-M1}^{MM0}} \tag{1}$$

where r_{Q1-M1}^{MM0} is the equilibrium Q1–M1 bond length in the MM force field used and r_{Q1-HL}^{QM0} is optimum length of the Q1–HL bond optimised with the QM method and basis sets used.[8] Thereby, the HL and M1 atoms can be considered to be the same atom, albeit with different positions (and sometimes also different charges) in the calculations with system 1 alone or with both system 1 and 2. The QM/MM forces are calculated with the help of the chain rule.[27,30]

Most calculations in this paper have been calculated with electrostatic embedding (EE),[17,13] meaning that a point-charge model of the MM system is included in the QM calculations, so that the QM system is polarised by the MM system. Thus, the total QM/MM energy is calculated from

$$E_{QM/MM}^{EE} = E_{1+ptch2}^{QM,HL} + E_{12,no1el}^{MM,M1} - E_{1,no1el}^{MM,HL}$$ (2),

where the three terms on the right-hand side are the QM energy of the QM system with HL atoms, including the point-charge model of the MM system, the MM energy of all atoms (with M1, rather than HL atoms), but with the charges of the QM system zeroed, and the MM energy of the QM system (again with HL atoms), with zeroed charges. The latter term is needed to cancel the MM term of the QM system from the second term, to avoid double counting. Likewise, the charges of the QM atoms are zeroed in order to avoid double counting of the electrostatic interactions between the QM and MM systems. The self-energy of the point-charge model is excluded from the QM term (this energy is instead included in the $E_{12,no1el}^{MM,M1}$ term).

Charge-redistribution schemes

When using electrostatic embedding, it is not fully clear what atoms should be included in the point-charge model of the MM system. In particular, it is unclear whether the charges of the M1 atoms should be included or not. If HL and M1 are considered to be the same atom, it is evident that M1 should not be included. This becomes even clearer if you do not use a hydrogen link atom, but rather a reparametrised atom that behaves like a carbon atom[19,31,32,33] – then the HL and M1 atoms will overlap.

However, most discussions and developments have started from the assumption that also the M1 charge should be included.[22,23] The reason for this is probably that many QM/MM methods are based on force fields with charge groups (e.g. OPLS and CHARMM[34,35]), i.e. where small chemical groups, like a CH_2 unit, have a neutral charge. Then, it is natural to include also the M1 charge, to keep the charge-neutrality.[22]

On the other hand, the distance between HL and M1 is quite short, ~0.5 Å, which can lead to a significant overpolarisation of the QM system. In fact, even the distance between HL and M2 is quite short, 1.3–1.7 Å. Therefore, it is common to exclude or redistribute some point charges in the QM calculations.[17,13,22,23] In this paper, we have tested six different approaches, which were chosen among those that have given the best results in previous tests:[19,22,23]

- Z0: All charges are included, including those on the M1 atoms.
- Z1: The charges of the M1 atoms are excluded.
- Z2: The charges of the M1 and M2 atoms are excluded.
- Z3: The charges of the M1, M2, and M3 atoms are excluded.
- RCD (the redistributed charge and dipole method[22]): The charges of the M1 atoms are redistributed over all the M2 atoms, keeping the bond dipole constant by adding a compensating charge at the bond midpoint between each M1 and M2 atom. If we let q_0 be the original charge on M1 divided by the number of M2 atoms, then the charge on the bond midpoint will be $2q_0$, whereas q_0 will subtracted from the charge on each M2 atom.[22]
- CS (the charge shift scheme[36,37]): Similar to RCD, in that the M1 charge is redistributed, keeping the bond dipole by compensating charges. However, q_0 is added to the charges on M2 atoms and two point charges are placed on each side of M2 along the M1–M2 bond. We place the two point charges at 6% of the M1–M2 bond length from the M2 atom with charges $\pm 50/6\ q_0$. This is slightly different from the implementation of this approach in ChemShell, where the distance is only approximately 6% to allow the charges to be truncated after four decimals (to ensure

numerical stability in geometry optimisations; P. Sherwood, pers. commun.; in this paper we do not change the geometries).

As mentioned above, the Z0, RCD, and CS schemes were designed for MM force fields that use charge groups, so that the part of each junction residue that is not in the QM system (including the M1 atom) have a net integer charge. This is not the case with the Amber force field[38] we are using, for which only the full residue has an integer charge. In order to test the importance of charge groups, we used two different approaches: In the first, we simply used the original charges, although they do not sum up to an integer. This approach is denoted by the six abbreviations noted above (Z0, Z1, Z2, Z3, RCD, and CS).

Alternatively, we changed the charge on the M1 atom so that the sum of the charges of the atoms in the junction residue that are not in the QM system becomes an integer (zero in all cases tested here, except for the carboxy-terminal His-591).[19] Then, each of the charge-distribution approaches was performed as described above. These balanced approaches will be called BZ0, BRCD and BCS in the following.

The Z1, Z2, and Z3 approaches are not affected by this redistribution of the M1 charge. For these, we instead tested to redistribute the sum of the deleted charges evenly on the other MM atoms in that residue. These approaches will be called DZ1, DZ2, and DZ3. The DZ2 approach is default in ComQum.[8] Finally, for Z2, we also tested to set the sum of the remaining MM charges in the residue to zero by adding the same increment to all charges. This approach will be called NZ2 below. All these approaches are implemented in our local software changeparm, which generates the point-charge file from the MM topology file.

Energy correction schemes

Covalent junctions between the QM and MM systems inevitably introduce an unphysical perturbation of the system. The question then naturally arises whether this error can be corrected. Strangely enough, this important question has been much less discussed than possible charge-redistribution schemes. The errors caused by introducing a link atom are of three types:

1. The HL atom is placed in the wrong position compared to the real M1 atom.
2. The HL atom is a hydrogen atom, rather than the correct M1 (typically carbon) atom and thus it will have incorrect MM parameters, in particular an incorrect charge.
3. Electrostatics are treated inconsistently around the junction: In most macromolecular MM force fields, non-bonded interactions between atoms that are directly bonded or that are separated by two covalent bonds (1–2 and 1–3 interactions) are excluded, whereas interactions between atoms connected by three covalent bonds (1-4 interactions) are scaled down. However, the QM software does not know about such exclusion rules and includes all electrostatic interactions between the QM atoms and the point charges. This is illustrated in Figure 2 for a simple model system. consisting of ethanol, in which we use a HOH QM model of the alcohol group, whereas the rest is treated at the MM level. In the QM calculations, the three QM atoms (HO, O, and HL) interact fully with the MM charges of all the seven MM atoms (C1, C2, and H11–H23), giving 21 electrostatic terms. Among these, only the three HO–H21/H22/H23 interactions are present also in the full MM treatment of ethanol. The six HO–C2/H11/H12 and O–H21/H22/H23 interactions should be scaled down (e.g. by a factor of 1.2 for electrostatics and 2.0 for van der Waals interactions in the Amber force field), whereas all the others should be excluded.

In principle, all these errors can be corrected by the general approach of Eqn. 2. For example, if the MM parameters of the HL and M1 atoms are chosen wisely, there will be a cancellation of all bonded and van der Waals terms involving HL (between $E_{1+ptch2}^{QM,HL}$ and

$E_{1,no1el}^{MM,HL}$) so that the total QM/MM energy corresponds to calculations with M1 atoms only. However, for the electrostatic interactions, no such cancellation is obtained with Eqn. 2. This is the case for standard CoMQuM and also in the calculations by Lin and Truhlar.[22] This seems to be the case also in most other QM/MM software, although the details of the implementations are seldom discussed. This approach will simply be called electrostatic embedding (EE) in the following.

However, two approaches have been suggested to correct also the electrostatic interactions: the QM to QM/MM correction in the QTCP (QM/MM thermodynamic cycle perturbation) approach[7] and the implementation of electrostatic embedding in ONIOM.[11] Both approaches assume that the charge distribution of the QM system can be accurately described by a point-charge model. In the QTCP approach, the following correction factor is added to Eqn. 2:

$$E_{corr}^{QTCP} = \sum_{i \in QM\,with\,M1,\,j \in MM} \frac{f_{ij} Q'_i q_j}{4\pi\epsilon_0 r_{ij}} - \sum_{i \in QM\,with\,HL,\,j \in ptch} \frac{Q_i q'_j}{4\pi\epsilon_0 r_{ij}} \qquad (3)$$

where Q_i are charges fitted to the QM electrostatic potential (ESP charges[7,39,40]) for the QM system, including the HL atoms (these charges change when the QM system or the charge-redistribution scheme changes), Q'_i are the same ESP charges, except that the charge on the M1 atom has been modified to be a charge typical for a carbon atom and to give an integer net charge of all atoms (see below), q_j are the standard MM charges for the MM atoms (always the same charges), and q'_j are the point charges, i.e. the q_j charges, but possibly modified by a charge-redistribution scheme; f_{ij} is a scaling factor for MM exclusion rules (for the Amber force field, used in the present calculations, $f_{ij} = 0$ for atoms separated by one or two bonds, $f_{ij} = 0.5$ for atoms separated by two bonds, and $f_{ij} = 1$ otherwise; note that this factor is present only in the first term, not in the second), and r_{ij} is the distance between atoms i and j. Note that the coordinates of the junction atoms are those of M1 in the first term, but those of HL in the second term. The philosophy behind this correction is that the second sum should remove the effect of the wrong positions and charges of the HL atoms, as well as remove the 1–2, 1–3, and 1–4 interactions in the $E_{1+ptch2}^{QM,HL}$ term in Eqn. 2, using an ESP-charge description of the QM system. Then, the first sum should introduce these terms again, but with the correct positions and charges of the M1 atoms, and with correct exclusion rules. All other interactions between QM and MM, which do not involve HL and are more than three bonds apart, are identical in the two sums and therefore cancel in Eqn. 3. If we use the division of the total system into three parts (Q, J, and M), this provides corrected energies for the Q–M and J–M interactions at a MM approximation. In the following, we will call this the QTCP correction.

In the implementation of electrostatic embedding with ONIOM, Morokuma and coworkers took this idea one step further by also correcting the Q–J and J–J interactions at the MM level.[11] This is done by adding the following correction term to the QTCP corrected results:

$$E_{corr}^{ONIOM} = \sum_{i \in J\,with\,M1,\,j \in QM,\,i \neq j} \frac{f_{ij} Q'_i Q'_j}{4\pi\epsilon_0 r_{ij}} - \sum_{i \in J\,with\,HL,\,j \in QM,\,i \neq j} \frac{f_{ij} Q_i Q_j}{4\pi\epsilon_0 r_{ij}} \qquad (4)$$

The two sums run over the same atom pairs, but the first sum uses coordinates and charges of M1 atoms, whereas the second term uses instead coordinates and charges of the HL atoms. In fact, the total QM/MM energy with both the QTCP and ONIOM correction can simply be written as:

$$E_{QM/MM}^{EE,ONIOM} = E_{1+ptch2}^{QM,HL} + E_{12}^{MM,M1} - E_{1+ptch2}^{MM,HL} \qquad (5)$$

Here, the first term on the right-hand side appears already in Eqn. 2. The second term is the standard MM energy of full system with M1 coordinates and charges, whereas the last term is the MM energy of the QM system, with HL coordinates and charges, and including the point charge model of the MM system as a separate molecule (i.e. without applying exclusion rules for the QM/MM cross terms). By this simple approach, all errors introduced by the junctions are corrected, provided that the ESP charges give a proper description of the charge distribution in the QM system and the MM approximation is accurate enough to describe the difference between the HL and M1 atoms. We will call this the ONIOM correction in the following. We are not aware of any previous comparison of these approaches.

We will see below that the QTCP corrections sometimes become too large because the ESP charges on the HL atoms become strange, owing to overpolarisation by the point-charge model. Therefore, we also tested to obtain the ESP charges from a wavefunction that is calculated without the point-charge model (i.e. in vacuum). These charges will be called Q_0 and Q_0' in the following and the corresponding corrected results will be called QTCP$_0$ and ONIOM$_0$. Note that in these corrected energies, we still use the $E_{1+ptch2}^{QM,HL}$ term, so that the MM system still polarises the QM system in the energy; it is only when calculating the ESP charges for the QM system that the point-charge model is excluded. This requires an extra set of QM wavefunction calculations for each system.

Mechanical embedding

Another way to partly correct the use of HL atoms is to use mechanical embedding (ME),[17,13,11] although this approach is normally not introduced with this explicit aim. ME implies that the QM calculations are performed in vacuum (i.e. without any point charges), giving $E_1^{QM,HL}$. This energy is the QM-only energy of the isolated QM system and it will be called QM below. If it is combined with the two MM energy terms in Eqn. 2, we obtain a QM/MM energy that does not contain any electrostatic interactions between the QM and MM systems We will call this QM+vdW:

$$E_{QM/MM}^{QM+vdW} = E_1^{QM,HL} + E_{12,no1el}^{MM,M1} - E_{1,no1el}^{MM,HL}$$ (6)

The electrostatic interaction between the QM and MM systems can be introduced by calculating it at the MM level. This can be done by using the same two MM terms as in Eqn. 5 (i.e. without zeroing the charges of the QM system):

$$E_{QM/MM}^{ME} = E_1^{QM,HL} + E_{12}^{MM,M1} - E_1^{MM,HL}$$ (7)

This is the standard form of ME, e.g. used in the ONIOM approach.[19] The natural choice of charges for the QM system are the Q_0' and Q_0 charges for the second and third term in Eqn. 7, respectively, because the wavefunction used in the first (QM) term is obtained without any point-charge model. This approach will be called ME below. We have tested to use the Q_0' charges also for the $E_1^{MM,HL}$ term, which we call ME$_1$.

It should be noted that these two approaches by construction include the ONIOM correction in Eqn. 4 (i.e. the Q–J and J–J corrections for the HL atoms). To estimate the size of this correction, we have also considered a ME variant, in which this correction term is excluded:

$$E_{QM/MM}^{ME0} = E_{QM/MM}^{QM+vdW} + \sum_{i \in QM \cup MM \, with \, M1, \, j \in MM} \frac{f_{ij} Q'_i q_j}{4\pi \epsilon_0 r_{ij}}$$ (8)

8

i.e. where all electrostatic MM interactions between one atom in any of the two systems and one atom in the MM system with M1 positions and charges (i.e. Q–M, J–M, and M–M interactions; Eqn. 3) have been added to $E_{QM/MM}^{QM+vdW}$. This approach will be called ME_0 below.

A problem with the ME approach is that it completely ignores the polarisation of the QM system by the MM system. A simple way to partly fix this problem is to calculate the ESP charges with a wavefunction polarised by a point-charge model (i.e. to use the Q' and Q charges instead of Q_0' and Q_0 charges in Eqn. 7). We call such an approach ME'. Like $QTCP_0$ and $ONIOM_0$, it requires an extra set of wavefunction calculations. Moreover, the results will (slightly) depend on the charge-redistribution scheme used.

Unfortunately, such an approach is not fully consistent, because the cost of polarisation is not included in the energy. However, this cost can be included in a linear-response approximation by simply taking the average of the ME' and ME_0 energies:

$$E_{QM/MM}^{MEav} = \frac{E_{QM/MM}^{ME0} + E_{QM/MM}^{ME'}}{2} \tag{9}$$

which we will call ME_{av} in the following.

We have also calculated a sixth variant of ME, in which we instead take the average of the ME' and QM+vdW energies. This energy, which we call ME_{scal} simply scales down the electrostatic interaction energy by a factor of 2, which could be considered as a primitive model of the polarisation of the MM system. The rationale for this is that it has frequently been observed that electrostatic interaction energies are overestimated by QM/MM.[26,41,42]

A problem with all of the previous ME methods is that they are sensitive to the stability of the ESP charges used for the QM system. In particular, for a large QM system, any variation of the charges on the boundary of the QM system will make large contributions to the energy difference between various states, because they are close to the MM system. Such variation can have three causes. First, the charge redistribution that occurs in the centre of the QM system, i.e. in the actual chemical reaction, may induce changes further out (polarisation). Second, the charge-derivation scheme may be unstable so that small changes in the charge redistribution due to polarisation lead to large changes in the charges. Third, the wavefunction optimisation itself may be unstable so that the two calculations on the reactant and product states end up in different local minima with respect to a remote (typically not covalently linked) part of the system, which consequently acquires different charges. Whereas the first effect is clearly desirable (it makes the treatment of polarisation more self-consistent), the other two are artefacts.

A simple way to see if the desired effect is dominating is to eliminate all three effects and see if the results get worse. To this end, we tested a seventh ME method, in which the QM charges for the two studied states were forced to be identical (by averaging over the two reactants, except for the central core). Because equal charges do not contribute to the energy difference, this method in practice only includes QM/MM interactions between the central core and the MM system, whereas QM interactions are considered within the full QM system. Thus, all indirect effects of polarisation are ignored. We call this direct method ME_{dir}. All possible sizes of the central core were tested, but the best results were obtained with the smallest 46-atom QM system; consequently, all results in the tables are obtained with that selection.

All methods are summarised in Table 1. We will see that the various ME methods provide a convenient way to test the various correction terms used in this article.

Computational details

As a test case, we use a 446-atom model of the active site of [Ni,Fe] hydrogenase. It contains the central core, consisting of $[(CH_3COOH)(CH_3S)_2Ni(CH_3S)_2Fe(CO)(CN)_2(C_3N_2H_5)]^-$ as a 46-atom model of the [Ni,Fe] active site with four Cys ligands (Cys-72, 75, 543, and 546), as well as the second-sphere groups of His-79 and Glu-25 (Figure 3). To this system, we then have added 40 models of amino acids, according to their energy contribution in a QM/MM free-energy study.[26] Some of the added groups are covalently connected to the original model; these are shown in detail in Figure S1. Other groups are separated and more distant (up to 16 Å from the 46-atom QM system; cf. Figure 3).

In a second set of calculations, the same 40 groups were added, although in a different order: First, the original QM residues where capped with CH_3CONH- and $-CONHCH_3$ groups of the surrounding backbone. Next, this backbone was extended by one further CH_3CONH- and $-CONHCH_3$ group on each side. However, note that the full 446-atom system does not contain all such groups (because only the 40 groups with the largest QM/MM free energy components were included; cf. Figure S1). If the backbone groups are named after the residue containing the N atom, the following back-bone residues are included in the first system: Cys-72, Cys-75, His-79, Ala-80, Cys-543, Ile-544, Cys-546, and Gly-547. The second system included the backbone of Ala-71, Val-78, and Pro-542. After that, the remaining 29 groups were added one by one, according to their error in the EE calculations of set 1 with the BCS charge-redistribution scheme. These groups are not covalently connected to the other residues, except for Arg-70 and Gln-69 (the latter is added in two parts, one consisting of the backbone CH_3CONH- group and the other the inner part of the side chain. In addition, the carboxy-terminal His-549 is divided into two groups: the negatively charged backbone belongs to one group (called His-549), whereas the neutral side-chain is a ligand of the Mg site, which is added as a single group, consisting of Mg^{2+}, this imidazole group, three water molecules, the side-chain of Glu-53, and the backbone CO group of Leu-495.

We study the energy of the simple proton-transfer reaction shown in Figure 3. It involves the transfer of a proton from the S^γ atom of one of the bridging cysteine ligands (Cys-546 in *Desulfovibrio fructosovorans*) to the $N^{\epsilon2}$ atom of a second-sphere histidine ligand (His-79). We have calculated the energy difference between the form in which the proton resides on Cys-546 (called the HID state) and the form in which the proton resides on His-79 (called the HIP state). As a reference value, we use the QM energy for the full 446 model calculated in vacuum, 48.3 kJ/mol.[14] All methods give this energy for the largest QM system, because then no residues remain in the MM system. The aim of this investigation is to see how the QM/MM results converge towards this value as more and more residues are moved from the MM system to the QM system.

All calculations were performed with density-functional theory, using the Becke-1988 – Perdew-1986 functional[43,44] and the def2-SV(P) basis sets.[45] The calculations were sped up by expanding the Coulomb interactions in auxiliary basis sets, the resolution-of-identity approximation.[46,47] The Fe and Ni ions were assumed to be in the low-spin +II oxidation state, giving a closed-shell singlet state of the full system.[14,26] All calculations were performed on exactly the same two structures (one for the HID state and one for the HIP state; thus no QM/MM geometry optimisation was performed in this investigation), taken from QM/MM structures, obtained with the 46-atom QM system (i.e. only these 46 QM atoms had different positions in the HID and HIP states, whereas all the other 400 atoms had the same positions in the two states).[26]

The full 446-atom model consists of hydrogen-atom capped amino-acid fragments (Figures 3 and S1). Therefore, standard MM parameters could not be directly used for the MM system. Instead, we started from a full MM model of the whole protein, described by the Amber 1999 force field.[38,48] This system was truncated to the 446-atom model, filling all broken bonds with a hydrogen atom at a C–H distance of 1.101 Å. All internal parameters (bonds, angles, and dihedrals) that were not affected by the truncation were kept at the Amber

1999 force field. All angles and dihedrals involving the HL atoms were set to the corresponding parameters for the M1 atom, whereas the bonds involving the HL atom had an equilibrium distance of 1.101 Å and a force constant given by[27]

$$k_{Q1-HL} = k_{Q1-M1} \frac{1.101^2}{r_{Q1-M1}^{MM0\;2}} \qquad (10)$$

The Lennard-Jones parameters for the HL atoms were the same as for the Amber HC atom type (hydrogen bound to carbon). Finally, the MM charges (q_j charges above) were determined individually for each covalently connected fragment of the 446-atom model by ESP charges, calculated with the Merz–Kollman scheme,[39] as implemented in Turbomole.[28] For the large fragment involving the Ni and Fe ions, charges outside the smallest QM system were averaged between those of the HID and the HIP states. These charges were always used for the MM system, either as charges for the MM system in the mechanical-embedding calculations or as point charges in the QM calculations. They were always the same in all calculations (besides possible adaptations according to the charge-redistribution schemes).

For atoms in the QM system, new charges were calculated from the wavefunction for all atoms in each QM system and charge-redistribution scheme (Q_i charges above). They were Merz–Kollman ESP charges[39] obtained from a QM calculation of the entire QM system (not only on fragments as for the MM charges). In the case when the wavefunction was polarised by the point charges, the point charges were omitted in the ESP calculations, without reoptimising the wavefunction. These charges were used to describe the QM system in the QTCP and ONIOM energy-correction schemes and also in the mechanical-embedding calculations. To obtain charges for the M1 atoms (Q'_i charges above), the charges of the HL atoms were adapted so that the full system had the correct integer charge (so that they get a size typical for carbon atoms, rather than for hydrogen atoms).[27] For full amino acids, this adaptation is unambiguous, because each amino acid has an integer charge. However, in the present calculations, with QM charges calculated for fragments that may be connected, the fragments do not always have an integer charge (owing to charge transfer between the fragments; this charge transfer can be extensive, up to 0.5 e between ionic pairs). Therefore, the adaptation sometimes becomes ambiguous. We solved this problem by simply partitioning the ambiguous charge equally between all fragments with junctions. This partitioning can be based on either the QM (Q_i) or MM (q_j) charges on the M1 atoms, but this gave little difference in the final energies; the presented results are based on the QM M1 charges.

Finally, a third set of calculations was performed, based on a 12178-atom model of [Ni,Fe] hydrogenase, solvated in a spherical system (35 Å radius). The set-up of this system has been described before.[26,42] In these calculations, all 12178 atoms were included in the MM calculations, whereas 46 to 446 atoms were included in the QM system. The groups were moved between the QM and MM system in the same order as in the first set of calculations, described above and the QM systems were identical. Charges of the MM atoms were taken from the Amber-1999 force field,[38] except for the metal sites, for which the charges were taken from QM calculations.[26,42] The MM charges were the same for the HID and HIP states. Only the Z1 charge-redistribution scheme was tested.

Result and Discussion

Comparison of QM/MM methods

In this paper, we study how well QM/MM calculations reproduce a QM vacuum calculation for a 446-atom model of the active site of [Ni,Fe] hydrogenase. We study the energy of a simple proton transfer from the S^γ atom of the Ni ligand Cys-546 to the $N^{\varepsilon 2}$ atom

of the second-sphere group His-79. We have systematically studied this reaction with QM/MM methods, in which we increase the QM system from 46 to 446 atoms by moving 40 groups from the MM to the QM system. The smallest QM system contains the Ni^{2+} and Fe^{2+} ions and their ligands, as well as the proton acceptor and another second-sphere ligand that also shares a proton with a Cys ligand. All amino acids have been truncated in a standard way, i.e. Cys is modelled by CH_3S^-, His by imidazole, Asp and Glu by acetate, Lys by methylamine, and Arg by methylguanidine.

QM calculations with the full 446-atom model give a reaction energy of 48 kJ/mol in favour of the HIP state.[14] If the QM/MM calculations were perfect, they would always give this energy, irrespectively of the number of groups in the QM system. We have tested QM-only calculations, as well as 13 different variants of QM/MM with mechanical embedding (ME) or electrostatic embedding (EE), using 12 different charge-redistribution schemes (to avoid overpolarisation close to the covalent junctions between the QM and MM systems) and two different ways to correct errors introduced by the junctions. Finally, we have tested to add the 40 residues to the QM system in two different ways. In the first set of calculations, the 40 groups were added to the QM system in the order of the size of their contributions to the QM/MM free energy difference of this reaction.[26] These calculations showed that the largest errors come from junctions directly connected to the active site. Therefore, we constructed a second set of calculations, in which first the backbone of all residues involved in the active site were added to the QM system (including CH_3CO- and $-NHCH_3$ groups from the neighbouring residues), then the backbone of all neighbouring residues, and finally, the remaining 29 residues were added in the order of their QM/MM error with the BCS charge-redistribution scheme with the EE method (Figure S1 in the supplementary material shows details of the residues).

All these calculations give a large amount of data that are presented in Figures S2 and S3 in the supplementary material. Here, we will summarise the results and extract the most interesting conclusions.

First, we show in Figure 4 the error (compared to the QM calculation with 446 atoms, i.e. 48 kJ/mol) for the 40 (set 1) or 30 (set 2) individual calculations for all 13 methods, using the Z1 charge-redistribution scheme, when applicable. It can be seen that the error is smaller on average in set 2, but that the variation is large among the various calculations. Therefore, we will mainly discuss the results in statistical terms, using the mean signed error (MSE), the mean unsigned error (MUE), and maximum unsigned deviation among the 40 or 30 calculations for each method and charge-redistribution scheme.

We will start with discussing the 12 different charge-redistribution schemes tested (note that the QM, QM+vdW, ME_0, ME, ME_1, and ME_{dir} methods are not affected by the charge-redistribution schemes, because they do not use any point-charge model). In Figure 5, the mean unsigned errors (MUEs) for the two sets of calculations and two representative QM/MM methods, EE and ME_{av}. Many methods (in particular EE, $QTCP_0$, and $ONIOM_0$) give large errors with Z0, indicating that it is inappropriate to include M1 point charges in the QM calculations.

Moreover, for a small number of the BZ0, Z1, DZ1 (only set 2), BRCD, and BCS calculations, ME' and in particular QTCP and ONIOM give very large errors 100–430 kJ/mol; cf. Figure 4). The reason for this is that owing to the near-by point charges, the HID and HIP calculations end up in a different electronic states, giving different ESP charges for the QM atoms, leading to large QM–MM interaction energies, and especially QTCP and ONIOM energy corrections. Therefore, these combinations of methods and charge-redistribution schemes should be avoided. The ME_{av} and ME_{scal} methods are calculated as the average between ME' and either the ME_0 or QM+vdW energies. Therefore, these methods are also slightly affected by this problem for ME', but the effect is quite small, as can be seen in Figure 5. The EE results are not affected by this problem, because it never uses any point-charge description of the QM system. Likewise, the ME_0, $QTCP_0$, and $ONIOM_0$ methods are not

affected by this problem because the charge-model of the QM system is obtained without the point-charge model of the MM system (i.e. in vacuum).

From Figure 5, it can be seen that the ME_{av} method gives quite stable results, once the charge-redistribution schemes affected by the above problem are disregarded, with a MUE of ~14 kJ/mol for the first set and ~8 kJ/mol for set 2. It can also be seen that the variation is somewhat larger for set 1, indicating that the choice of charge-redistribution scheme matters mainly for junctions closest to the reaction. Most of the other QM/MM methods give similar variations among the various charge-redistribution schemes. The EE method is the prime exception to this rule, showing quite large variations among the charge-redistribution schemes. For set 1, it gives a clear improvement for all charge-neutralisation schemes (i.e. BZ0, DZ1, BRCD, and BCS give better results than Z1, RCD, and CS, respectively). The best results are obtained with the Z3 charge-redistribution scheme. Unfortunately, set 2 shows the opposite results, with the best results with Z1 and DZ2, and the worst results with Z3, BRCD, and BCS. This indicates either that the variation observed is only random, or that two different effects are present, viz. the effect of junctions, for which Z3, BRCD, and BCS give the best result, and other approximations in QM/MM, which the Z1 and DZ2 methods seem to treat better. Considering that most of the charge-redistribution schemes do not attempt any physical correction of the problems involved, but simply delete charges (which sometimes can be involved in important hydrogen bonds with the QM system, especially the backbone NH and CO groups) and that some of the schemes involve the questionable use of the charge on the M1 atoms, in the following we simply discuss the average over all 12 charge-redistribution schemes (for the methods affected by the occasional problem with QM charges, we also discuss the results when the affected methods are omitted from the averages).

The average performance of all methods for the two sets of calculations are summarised in Table 2. It can be seen that for set 1, the ME_{dir} method gives the lowest MUE (7 kJ/mol), which is a clear improvement from the QM-only approach (21 kJ/mol). It also gives the lowest maximum error (28 kJ/mol) and the fastest convergence to 10 and 20 kJ/mol errors. On the other hand, several other methods give a faster convergence to an error of less than 5 kJ/mol. The results do not change qualitatively if the four charge-redistribution schemes that give occasional spurious QM charges are omitted, but the values changes somewhat, and the variation among the various charge-redistribution schemes decreases, as can be seen from the standard deviations (σ_{MUE}, σ_{Max}, and σ_{MS} in Table 2).

For set 2, the results are somewhat more varying, because many methods give results of a similar quality. If all charge-redistribution schemes are considered, the $QTCP_0$ approach gives the lowest MUE (6 kJ/mol; but the ME_{dir}, $ONIOM_0$ and ME_{scal}, methods give similar results within 1 kJ/mol), whereas ME_{dir} gives lowest maximum error (16 kJ/mol and therefore is always converged to 20 kJ/mol). On the other hand, EE gives the fastest convergence to 5 kJ/mol and ME_{scal} gives the fastest convergence to 10 kJ/mol. If the four problematic charge-redistribution schemes are omitted from the average, ME_{scal} gives the lowest MUE, 5 kJ/mol.

Thus, we can conclude that the ME_{dir} approach gives outstanding results for set 1 and also among the best results for set 2. This is surprising, because it suggests that all efforts of including polarisation and correctly treating electrostatics around the junctions are meaningless unless a stable QM method is used; otherwise non-physical variations in the electrostatic description of the QM system (charges in the ME methods; charge density in the EE methods) will dominate the errors. To confirm this, we repeated the ME_{dir} calculations with all possible choices of the subsystem in which the charges were allowed to vary. Indeed, we found that the smallest QM system gave the best results, and more generally that all subsystems that gave a low MUE were small. However, taking this size-reduction to the extreme by using the same charges for all atoms (e.g. also the reacting atoms) gave larger errors (MUE = 22 and 19 kJ/mol for the two sets).

Energy contributions

Further understanding of the various methods and the components involved can be gained by a pair-wise comparison of the methods. The average magnitudes of these differences are shown in Table 3 for both sets. In Figure 6, this information has been combined with the mean signed and unsigned errors from Table 2 to provide a pictorial representation of how the various methods are related and how the effects behave (results are shown only for set 1).

The first column in Table 3 (vdW) shows the difference between the QM+vdW and QM methods, which estimates the effect of the van der Waals interactions (always involving a ONIOM-type correction for the HL atoms). It can be seen that the effect of the van der Waals interactions is small, 1–2 kJ/mol for the two sets, but it leads to a slight improvement in most quality estimates (Table 2).

Next, we estimate the effect of the electrostatic interactions by taking the difference between the ME_0 and QM+vdW methods (the electrostatic interactions are calculated at the MM level, using charges of the QM system obtained in vacuum). From Table 3 (column ele), it can be seen that this gives a large effect, 26 kJ/mol in set 1 and 13 kJ/mol in set 2. It gives a small improvement in the MUEs, but a very strong improvement in MSE (to –2 kJ/mol). On the other hand, the maximum error increases, showing that the variation of the results becomes larger and in many cases, the results actually become worse. The direct electrostatic effect of the MM system on the reaction (i.e. the difference between the ME_{dir} and QM+vdw methods; column ele_{dir}) is significantly smaller (16 kJ/mol) for set 1 but equally large (13 kJ/mol) for set 2. This means that there are large indirect effects of charge redistribution whenever there are junctions close to the reaction. As discussed above, the direct electrostatic effect reduces the MUEs significantly.

The ME method differs from ME_0 only in that an ONIOM correction for the erroneous position and charge of HL of the Q–J and J–J interactions is attempted. From Table 3 (column juncorr), it can be seen that this amounts to a quite large correction in set 1 (12 kJ/mol), but a rather small correction in set 2 (where there are only a few junctions; 4 kJ/mol). From Table 2, it can be seen that in general, this correction leads to worse, rather than improved results. Thus, the point-charge model of the QM system does not seem to be accurate enough to make such corrections.

ME_1 is identical to the ME method, except that only the position of the HL atom is corrected, not the charges. From Table 3 (column q_{HL}), it can be seen that this makes a large difference. In fact, the results become very similar to those of the ME_0 method, indicating that the main effect of the ONIOM correction comes from the charge, rather than the position of the HL atoms.

In the ME' approach, the charges of the QM system are obtained from a wavefunction polarised by the MM system (the point charges). Therefore, they are polarised by the MM system. This approach gives results of a similar quality as to those of the ME_0 method (with unpolarised QM charges; Table 2), but the individual results (Table 3, column me') differ by 23 kJ/mol for set 1 and by 7 kJ/mol for set 2, showing that it makes a major difference whether the charges are calculated by a polarised or a vacuum wavefunction.

The ME' approach is inconsistent, because the cost of the polarisation of the QM system is not considered. This cost is included in the ME_{av} method, which is the average of ME' and ME_0. From Table 3, it can be seen that the effect of the consistent polarisation (column pol) is rather small: 8 kJ/mol for set 1 and 3 kJ/mol for set 2. Thus, the effect of polarisation is 3–4 times smaller than the electrostatic effect for this system. From Table 2, it can be seen that ME_{av} gives a consistent improvement in all quality measures, except MSE, compared to ME'.

Even better results are obtained if ME' is averaged with QM+vdW instead. This essentially means that the electrostatic interactions are scaled down by a factor of 2. This is not a physically consistent method, but it is instead based on the common observation that electrostatic interactions seem to be overestimated in QM/MM calculations.[41,49] A possible explanation for this overestimation is that electrostatic embedding actually is inconsistent (in

contrast to mechanical and polarised embedding): In EE, the QM system is polarisable, whereas the MM system is not polarisable (like mixing polarisable and non-polarisable force fields in MM calculations). This probably leads to an exaggeration of the (polarised) charges of the QM system, compared to those of the MM system. Several attempts have been made to cure this problem by using special van der Waals parameters for the QM system.[49,50,51,52] The effect of the scaling (column scal in Table 3) is appreciably larger than that of the polarisation, 19 kJ/mol for set 1 and 9 kJ/mol for set 2.

In the next column of Table 3 (elpol), we compare the results of the EE and QM+vdW methods. In EE, both the electrostatics and the polarisation are calculated at the QM level, without any attempt to correct errors arising from the junction atoms. This is the standard way to do QM/MM (with electrostatic embedding). As expected from the previous results, the difference between EE and QM+vdW is quite large, 15–17 kJ/mol. For the second set, this is quite close to the sum of the ele and pol effects, but for the first set, elpol is appreciably smaller. The EE method has the largest variation among the tested methods. Typically, it gives a slight improvement over QM+vdW.

EE and ME_{av} both include the same terms, although they are calculated at the QM level for EE and partly at the MM level for ME_{av}. Quite unexpectedly, the effect of this change (column qmeff) is quite large in both sets of calculations, 17 kJ/mol for set 1 and 6 kJ/mol for set 2. Apparently, there are instabilities in the calculations caused mainly by the junctions. The results in Table 2 quite clearly show that the gain of doing electrostatics and polarisation by QM is more than cancelled by the errors caused by the junctions, so that the ME_{av} method actually gives the better results.

Finally, we tried to improve the results of EE by using the QTCP and ONIOM corrections. From Table 2, it can be seen that both corrections are strongly affected by the occasional problems with the QM charges for some charge-redistribution schemes. Therefore, we first discuss the results obtained with QM charges based on a vacuum wavefunction ($QTCP_0$ and $ONIOM_0$).

The former correction (column qtcp in Table 3), is in general quite large, 18 kJ/mol for set 1 and 6 kJ/mol for set 2. It is close to the qmeff term, showing that this effect dominates the difference between EE and ME_{av} (note that by construction, the QTCP correction is avoided in all types of ME methods). For set 1, $QTCP_0$ gives results that are slightly worse than those of EE, whereas for set 2, $QTCP_0$ provides a clear improvement, giving the lowest MUE if all charge-redistribution schemes are considered. This indicates that the QTCP correction is advantageous, except when the junctions are too close to the centre of the studied reaction.

The ONIOM correction on the other hand is problematic. It gives a correction that is very similar to what is observed with mechanical embedding (compare columns juncorr and oniom in Table 3), 12 kJ/mol with set 1 and 5 kJ/mol with set 2. $ONIOM_0$ always gives worse results than $QTCP_0$ and it gives no or only a marginal improvement over the EE method. Again, we have to conclude that the QM charge model is not accurate enough to support a Q–J and J–J interaction correction for the HL atoms.

Finally, we compare QTCP with $QTCP_0$ and ONIOM with $ONIOM_0$ in the last two columns in Table 3, i.e. the effect of calculating the QM charges with a polarised or a vacuum wavefunction. It can be seen that in general, the effect is rather small, 10–14 kJ/mol with set 1 and 4 kJ/mol with set 2. However, for a few cases, very different results are obtained, as discussed above.

Residue contributions

Some further understanding can be gained by studying the contributions from each of the added groups. A proper QM/MM method should give the same results, no matter if a group is treated by QM or by MM. Therefore, the difference in QM/MM energy between two calculations that differ only in that one group is moved from the QM system to the MM

system can be considered as the QM/MM error caused by that group. Such group contributions, using the Z2 charge-redistribution scheme (if applicable), are listed in Table 4 for set 1 and Table 5 for set 2 (these tables also show in which order the residues are added in the two sets). It can be seen that for the electrostatic embedding methods and set 1, residues that give large errors are typically those directly connected with junctions to the active site (type A) or the first neighbours (type N), although the contributions vary quite extensively among the various methods. With two exceptions, these residues give errors of 6–32 kJ/mol with the $QTCP_0$ correction. For set 2, most of the junction problems are collected in the first two contributions (A and N) so that the remaining contributions are much lower. However, there are also problematic residues that are not covalently connected to the active site, e.g. Glu-S22 and Arg-428. For many residues, the energy corrections are large (e.g. up to 38 kJ/mol for QTCP) and they make the errors smaller or larger in a rather random manner. Again, the Oniom correction works poorly, giving maximum errors of 72 and 20 kJ/mol for the two sets (ignoring the A contribution for set 2), compared to 32 and 13 kJ/mol for QTCP.

For the mechanical embedding methods, the results are more unpredictable, with very large errors for some groups, e.g. Asp-63, Ile-544, and His-549. At first sight, this seems to indicate that interactions with these groups are poorly described by MM and thus that the mechanical embedding approach fails. However, the results with the ME_{dir} method are significantly better, with the maximum error reduced from ~100 kJ/mol to 20 kJ/mol for set 1 (~30 to 16 kJ/mol for set 2). This indicates that the problem is again related to instabilities in the description of the outer part of the QM system, leading to random errors when the interactions with the MM system are strong. In contrast, when only the direct interactions are considered, the MM description proves to be highly useful, and gradually gets more accurate as the distance from the reaction centre increases.

Calculations with the whole protein

It is conceivable that the calculations with the 446-atom model of [Ni,Fe] hydrogenase will overestimate the electrostatic interactions between the distant charged groups and the active site, although our previous results in vacuum and a continuum solvent with a dielectric constant of 4 differed by only 14 kJ/mol for the 446-atom model.[14] Therefore, we have also performed a set of calculations in which the protein and the surrounding solvent (in total 12178 atoms) were included in all calculations, i.e. a more typical QM/MM set-up. As before, we studied how the energy of the proton-transfer reaction changed as the QM system was systematically increased from 46 to 446 atoms in 40 steps. Unfortunately, it is no longer evident what energy to use as a reference, because all methods give different results also with the 446-atom QM system (because the MM system is no longer empty; note also that the calculation with the largest QM system now contains 49 junctions to the protein). In order not to bias the results towards any certain method, we decided to use the result obtained for each method with the 446-atom QM system as the reference (i.e. different references for all different methods). This reference value is 48 kJ/mol for QM and QM+vdW (i.e. without any electrostatics), 24–28 kJ/mol for the ONIOM and QTCP approaches, which we showed above give poor results, but 52–75 kJ/mol for the other methods (average 65 kJ/mol).

The calculations were performed for all 13 different QM/MM variants, but only for set 1 and only for the Z1 charge-redistribution scheme. The results are shown in Table 6. It can be seen that they are quite similar to those obtained with the 446-atom model. In particular, the results are not significantly improved by the inclusion of the surrounding. On the contrary, the MUE is worse for all methods, in particular for ME'. On the other hand, the MSE is closer to zero for five of the methods (ME, ME', ME_{av}, $QTCP_0$, and ONIOM) and the maximum errors are reduced for seven methods.

ME_{dir} still gives the lowest MUE among the 13 QM/MM methods (16 kJ/mol), but it only 3 kJ/mol lower than for ME_1. $QTCP_0$ gives the lowest maximum error (43 kJ/mol), but ME_{dir}

16

gives essentially the same result. ME_{av} gives the lowest MSE, whereas QM+vdW gives the lowest standard deviation, and EE the fastest convergence. Thus, the results with the 12178-atom model show that the conclusions based on the smaller model are not significantly changed.

Conclusions

In this paper, we study the accuracy of QM/MM calculations with respect to the size of the QM system. As a test case, we have used a simple proton transfer between a first-sphere Cys ligand and a second-sphere His group in a 446-atom model of [Ni,Fe] hydrogenase. As the reference, we use the QM results obtained for the whole model.[14] We study how the QM/MM results vary when the QM system is enlarged systematically. Many QM/MM variants are available.[17,13,22,23,11] We have tested QM-only and 13 different variants of QM/MM, including both standard mechanical and electrostatic embedding, as well as several new variants.

We show that many groups provide sizeable (up to over 100 kJ/mol) contributions to the error in the QM/MM energies and that there is a large variation between the results obtained with different sizes of the QM system (cf. Figure 4). Thus, it is not enough to study only a few sizes of the QM system as in previous investigations if statistically significant results are to be obtained. There are at least three sources of these errors: The point-charge model of the MM group, the improper polarisation of the QM system (the MM system is not polarisable), and the approximations in treating the electrostatics around the junctions. The total QM/MM error is the sum of all these contributions for all groups in the MM system, most of which partly cancel; therefore, it is hard to reach any general conclusions, unless a large number of calculations are performed, to give proper statistics. This paper is based on 7140 separate QM/MM calculations, using 14 different methods, 12 different charge-redistribution schemes, and two different ways to systematically move 30 or 40 different groups from the MM to the QM system. This is appreciably more than in previous studies.[15,24,12,19,20,21,22,23] By systematically adding different components and corrections, we can estimate the size of the various terms and the accuracy of the various corrections. We have arrived at several interesting conclusions.

First, we show that various charge-redistribution schemes give similar results for most QM/MM methods. With junctions close to the active site, the Z0 scheme gives poor results, and for the ME', QTCP, and ONIOM methods, problems with the QM charges are observed for the BRCD and BCS schemes, as well as occasionally also for Z1, DZ1, and Z0. The standard EE method seems to be more sensitive to the charge-redistribution scheme than the other methods, but opposing results are obtained if junctions are close to the active site or not (set 1 and set 2). In general, there does not seem to be any consistent gain of using any charge-redistribution scheme, and we tend to recommend the Z1 scheme (i.e. include all point charges, besides that of M1, which is already included in the QM system as the HL atom), because it is the most simple method and there is no risk that any important interactions are omitted, as for the Z2 and Z3 schemes.

The best results in this investigation are obtained with the ME_{dir} method, which uses mechanical embedding but forces the charges used for the QM system to be identical for the reactant and product states except for a fairly small inner core, even when the QM system itself grows. It gives MUEs of 7 and 6 kJ/mol, MSEs of 1 and 4 kJ/mol, and maximum errors of 28 and 16 kJ/mol for the two sets of calculations, respectively. This is appreciably better than the QM-only calculations, by factors of 2–3 for both MUEs and maximum errors.

The second best method is the ME_{scal} method, which uses the average of the QM+vdW method and the ME' method, in which the charges of the QM system are obtained from a wavefunction polarised by MM point charges. Unfortunately, it is not physically consistent, but only motivated by the frequent observation that QM/MM methods with electrostatic

embedding seem to overestimate electrostatic interactions.[41,49] The ME' method itself, as well as the physically more consistent ME_{av} method, also give better results than QM-only, but with slightly worse MUEs and MSEs than ME_{scal}.

Among the methods with electrostatic embedding, standard EE and $QTCP_0$ give the best results, the latter especially if junctions are moved away from the active site. However, it is notable that this investigation indicates that mechanical embedding actually gives better results than electrostatic embedding, contrary to the common consensus that electrostatic embedding is a better approximation, because it includes the polarisation of the QM system by the MM system.[17,13] An advantage of mechanical embedding is that it avoids the junction problem for the electrostatic interactions between the QM and MM systems, because this term is calculated with M1 atoms and charges. Thus, it provides an alternative to the QTCP correction in Eqn. 3 and it avoids the risk of overpolarisation (and therefore no charge-redistribution is needed). Apparently, the errors introduced by letting the truncated QM system be polarised by a point-charge model are as severe as the omission of the polarisation of the QM system, and the errors can only be accurately corrected if the junctions are far from the active site (i.e. $QTCP_0$ in set 2). The reason why previous investigations have given better results with electrostatic embedding[15,19,22,23] is that the results depends on the tested system, i.e. the relative importance of polarisation and the junction errors, but also that we have in this investigation developed new variants of mechanical embedding (ME_{dir} and ME_{scal}) that give the best results.

Moreover, the results show that the ONIOM correction of the Q–J and J–J interactions within the QM system, using an ESP-charge model of the QM system, does not work properly. On the contrary, it consistently gives worse results than without this correction, both with mechanical embedding (i.e. ME compared to ME_0) and electrostatic embedding (ONIOM or $ONIOM_0$ compared to EE, QTCP, or $QTCP_0$). Apparently, the ESP-charge model of the QM system is too poor to allow for such a correction.

Finally, we show that the largest errors in QM/MM come from junctions between QM and MM close to the active site. In the second set of calculations, such junctions are avoided, and from Table 2, it can be seen that the results of all methods are strongly improved. Of course, in protein applications, junctions cannot be fully avoided.

Thus, our results lead to the following practical recommendations for QM/MM studies: First, junctions should always be moved into the backbone of the previous and next residue of each active-site group. In fact, the use of junctions, even far from the active site, should be minimised. Second, one should pay attention to the stability of the QM calculations with respect to the various reaction states calculated. Small perturbations may lead to non-physical differences in the charge distribution in the outer regions of the QM system, which although having small effect on the QM energies can have large effect on the QM/MM interaction energy because of the close proximity of the MM charges. This problem seems to be important for both mechanical and electrostatic embedding. When using mechanical embedding, the problem is amplified by instabilities in the charge-derivation scheme. On the other hand, we have devised a simple way of reducing this problem by using fixed charges outside a core system (the ME_{dir} method), whereas there seems to be no such quick fix to the electrostatic embedding approach.

It should be noted that the junction problem is not avoided by the ME_{dir} method. On the contrary, this method can be considered as the most basic QM/MM scheme, totally ignoring polarisation and treating junctions in a simple *ad hoc* manner. However, any approach to treat polarisation and junctions in a consistent way must be compared thoroughly to this basic approximation.

In our previous investigation of the same [Ni,Fe] hydrogenase system, we showed that it was very hard to obtain an accurate estimate of the proton-transfer energy in the protein by a QM-only approach.[14] For example, even if 40 groups from the surrounding protein were included (over 400 atoms), there was still a 60-kJ/mol difference if these groups were selected

as those closes to the active site or if they were selected from those giving largest energy contributions in a QM/MM calculation. In this paper, we show that the QM/MM approach also has convergence problems. However, once the junctions are moved away from the active-site residues the results become quite stable, with MUEs and maximum errors down to 5 and 16 kJ/mol, respectively. These results provide an estimate of the expected accuracy of the QM/MM approach. Thus, with wisely selected junctions, the QM/MM approach provides an appreciably faster convergence with respect to the size of the QM system than the QM-only approach.

Finally, we have also checked that the results do not change qualitatively if a 12178-atom model of the protein and surrounding solvent is included in the calculations. It should be noted that the reference results obtained with the largest QM system should not be considered as an accurate estimate of the true energy for this proton transfer in [Ni,Fe] hydrogenase. For such an estimate better methods (e.g. including dispersion) and basis sets should be used and dynamical effects need to be included. In a future publication, we will discuss how accurate estimates of protein reactions may be obtained by combining QM/MM calculations with large QM calculations.

Supporting Information Available

The chemical structure of groups directly connected to the active site, as well as the results of all methods and charge-redistribution schemes for set 1 and 2. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

1   Siegbahn, P. E. M.; Blomberg, M. R. A. *Annu. Rev. Phys. Chem.* **1999**, *50* 221-249.
2   Siegbahn, P. E. M.; Blomberg, M. R. A.   *Chem. Rev.* **2000**,*100* 421-438.
3   Ramos, M. J.; Fernades, P. A. *Acc. Chem. Res.* **2008**, *41*, 689-698.
4   Lonsdale, R.; Ranaghan, K.E.; Mulholland, A. J. *Chem. Comm.* **2010**, *46*, 2354-2372.
5   Siegbahn, P. E. M.; Borowski T., *Acc. Chem. Res.* **2006**, *39*, 729-738.
6   Himo F. *Theor. Chim. Acta* **2006**,*116*, 232-240.
7   Lin H.; Truhlar D. G. *Theor. Chem. Acc.*, **2007**, *117*, 185-199.
8   Senn H. M.; Thiel W. *Angew. Chem. Int. Ed.* **2009**, *48*. 1198-1229.
9   Hu H.; Yang W. *Annu. Rev. Phys. Chem.*, **2008**, *59*, 573-601.
10   Gao J. L.; Truhlar D. G., *Annu. Rev. Phys. Chem.* **2002**, *53*, 467-505.
11   Rod T. H.; Ryde U. *J. Chem. Theory Comput*. **2005**, *1*, 1240-1251.
12   Sumowski, C. V.; Ochsenfeld, C. *J. Phys. Chem. A* **2009**, *113*, 11734-11741.
13   Hu L.; Eliasson J.; Heimdal J.; Ryde U. *J. Phys. Chem. A* **2009**,*113*, 11793–11800.
14   Kaukonen M.; Söderhjelm P.; Heimdal J.; Ryde U. *J. Chem. Theory Comput.* **2008**, *4*, 985-1001.
15   Solt, I.; Kulhanek, P.; Simon, I.; Winfield, S.; Payne, M. C.; Csanyi, G.; Fuxreiter, M. *J. Phys Chem. B* **2009**, *113*, 5728-5735.
16   Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem. A.* **2000**, *104*, 1720-1735.
17   Amara, P; Field, M. J. *Theor. Chem. Acc.* **2003**, *109*, 43-52.

18   Rodriguez, A.; Oliva, C.; Gonzalez, M.; van der Kamp, M.; Mulholland, A. J. *J. Phys. Chem. B.* **2007**, *111*, 12909-12915.

19   Wang, B.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**. *6*. 359-369.

20   Lennartz, C.; Schäfer, A.; Terstegen, F.; Thiel, W. *J. Phys. Chem. B.* **2002**, *106*, 1758-1767.

21   Das, D.; Eurenius, K. P.; Billings, E. M.; Sherwood, P.; Chatfield, D. C.; Hodoscek, M.; Brooks, B. R. *J. Chem. Phys.* **2002**, *117*, 10534-10547.

22   Lin, H.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 3991-4004.

23   König, P. H.; Hoffman, M.; Trauenheim, T.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 9082-9095.

24   Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815-826.

25   Ryde U. *Protein Science* **1995**, *4*, 1124-1132.

26   Ryde U. *J. Comput.-Aided Mol Design* **1996**, *10*, 153-164.

27   Ryde U.; Olsson M. H. M. *Intern. J. Quant. Chem.* **2001**, *81*, 335-347.

28   *TURBOMOLE* V6.1 2009 a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH; available from http://www.turbomole.com.

29   Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, T. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. **AMBER 9**. University of California, San Francisco, 2006.

30   Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170-1179.

31   Antes, I.; Thiel, W. *J. Phys. Chem. A* **1999**, *103*, 9290-9295.

32   Zhang, Y.; Lee, T.-S.; Yang, W. *J. Chem. Phys.* **1999**, *110*, 46-54.

33   DiLabio, G. A.; Hurley, M. M.; Christiansen, P. A. *J. Chem. Phys.* **2002**, *116*, 9578-9584.

34   Jorgensen, W. L.; Maxwell, D. S.; irado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225-11236.

35   MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph- McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.

36   Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N.A.; Vincent, M. A.; Hillier, I. H. *Farady Discuss.* **1997**, *106*, 79-92.

37   Sherwood, P.; de Vries, A. H.; Guest, M. F. Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjovoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. *J. Mol. Struct. (Theochem)* **2003**, *632*, 1-28.

38   Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

39   Besler B. H.; Merz K. M.; Kollman P. A. *J. Comput. Chem.* **1990**, *11*, 431-439.

40   Sigfridsson E.; Ryde U. *J. Comp. Chem.* **1998**, *19*, 377-395.

41   Källrot N.; NilssonK.; Rasmussen T.; Ryde U. *Intern. J. Quant. Chem.* **2005**, *102*, 520-541.

42   Söderhjelm P.; Ryde U. *J. Mol. Struct. Theochem* **2006**, *770,* 199-219.

43   Becke, A. D. *Phys. Rev. A* **1988**, *38,* 3098-3100.

44   Perdew J. P. *Phys. Rev. B* **1986**, *33,* 8822-8824.

45   Weigend F.; Ahlrichs R.; *Phys. Chem. Chem. Phys*. **2005,** 7, 3297-3305.

46   Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240,* 283-290.

47   Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997,** *97,* 119-124.

48   Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000,** *21,* 1049-1074.

49   Tu, Y.; Laaksonen, A. *J. Chem. Phys.* **1999**, *111,* 7519-7525.

50   Freindorf, M.; Gao, J. *J. Comp. Chem.* **1996**, 17, 386-395.

51   Murphy R. B.; Philipp D. M.; Friesner R. A. *J. Comput. Chem.* **2000**, *21,* 1442 – 1457.

52   Freindorf M.; Shao Y.; Furlani T. R.; Kong J., *J. Comput. Chem.* **2005**, *26,* 1270 – 1278.

Table 1. Description of the 14 methods discussed. The QM calculations can either be performed in vacuum (vac; mechanical embedding) or with a point-charge model of the surroundings (ptch; electrostatic embedding). Van der Waals interactions (vdW) can be included or not. The electrostatic interactions between the QM and MM systems (elstat) can be ignored, calculated by molecular mechanics (MM), or by a point-charge model in the QM calculations (QM). Errors introduced by the junction atom can be corrected by either the QTCP correction in Eqn. 3 or the Oniom correction in Eqn. 4. The charges of the QM atoms (Q_{QM}) can be obtained in a vacuum (vac) or with a wavefunction polarised by a point-charge model (pol). They can also be averaged, scaled down, or corrected. Finally, the charges of the junction atoms (Q_J) can be those of the HL or the M1 atoms.

Method	Eqn.	QM	vdW	elstat	QTCP corr	Oniom corr	Q_{QM}	Q_J
QM	Before 6	vac						
QM+vdW	6	vac	yes					
ME_0	8	vac	yes	MM			vac	M1
ME	7	vac	yes	MM		yes	vac	M1
ME_1	After 7	vac	yes	MM		yes	vac	HL
ME'	After 8	vac	yes	MM			pol	M1
ME_{av}	9	vac	yes	MM			(vac+pol)/2	M1
ME_{scal}	After 9	vac	yes	0.5MM			pol/2	0.5M1
ME_{dir}	7	vac	yes	MM			vac, only core	M1
EE	2	ptch	yes	QM			pol(QM)	HL
$QTCP_0$	After 5	ptch	yes	QM	yes		vac(corr)	M1
$ONIOM_0$	After 5	ptch	yes	QM	yes	yes	vac(corr)	M1
QTCP	3	ptch	yes	QM	yes		pol	M1
ONIOM	4,5	ptch	yes	QM	yes	yes	pol	M1

Table 2. The performance of the various methods. The Table list the mean signed and unsigned error (MSE and MUE), the maximum error (Max), the standard deviation of the MSE (Stdev), and the number of residues that can be added to the QM system before the error becomes larger than 5, 10, and 20 kJ/mol, respectively (n5, n10, and n20). If applicable, the presented results are the average over all 12 charge-redistribution schemes; n is the number of values in each average and σ_{MUE}, σ_{Max}, and σ_{MS} are the standard deviations of the MUE, Max, and MSE values in these averages. For some methods, one or four charge-redistribution schemes are omitted from these averages in the lower part of the table. These are Z0 for EE in set 1, and Z1, BRCD, BCS, and BZ0 (set 1) or DZ1 (set 2) for the ME', ME_{av}, ME_{scal}, QTCP, and ONIOM. The best value in each column is marked in bold face.

	n	Set 1										Set 2									
		MUE	Max	MSE	Stdev	n5	n10	n20	σ_{MUE}	σ_{Max}	σ_{MSE}	MUE	Max	MSE	Stdev	n5	n10	n20	σ_{MUE}	σ_{Max}	σ_{MSE}
QM	1	20.5	56.1	18.8	18.1	38	32	29				10.1	25.7	8.4	9.3	33	33	18			
QM+vdW	1	19.2	53.8	17.2	17.4	38	32	29				9.6	25.8	7.9	9.0	33	32	18			
ME_0	1	18.2	83.2	-1.5	25.2	**33**	33	31				9.3	32.5	-1.6	13.0	32	32	32			
ME	1	23.0	84.4	-7.9	30.7	**33**	33	32				9.1	36.5	-3.9	12.5	34	32	29			
ME_1	1	18.2	84.2	-2.2	25.5	**33**	33	32				9.4	32.5	-1.6	13.1	32	32	32			
ME_{dir}	1	**7.4**	**28.3**	**1.4**	**10.2**	38	**19**	**10**				5.9	**15.5**	-4.5	**6.1**	33	33	**0**			
ME'	12	18.0	64.6	-9.2	22.4	37	33.0	30.7	3.8	39.4	4.4	9.3	49.2	-1.4	14.3	38.0	24.5	17.7	5.4	44.6	5.3
ME_{av}	12	14.9	43.5	-5.4	18.7	**33**	33.0	31.0	1.7	18.0	2.2	8.8	35.0	-1.5	12.1	32.3	30.3	17.7	2.7	19.2	2.6
ME_{scal}	12	12.9	45.8	4.0	16.9	37	23.7	16.8	2.6	10.9	2.2	6.3	30.6	3.3	9.4	33.2	**20.1**	9.5	2.7	21.4	2.6
EE	12	16.8	57.4	7.2	21.4	37	27.7	19.2	8.8	31.4	7.6	9.0	32.5	-7.3	10.8	**28.4**	24.3	16.4	2.7	8.5	2.9
$QTCP_0$	12	20.6	58.3	-6.4	23.3	37	32.8	29.7	7.4	29.7	9.6	**5.6**	22.2	-1.3	8.0	32.7	22.9	9.1	0.9	4.7	0.7
$ONIOM_0$	12	23.2	81.2	-13.0	24.7	37	33.0	33.0	5.9	16.8	9.6	6.1	26.5	-5.0	7.7	34.9	26.0	17.8	0.5	3.8	0.7
QTCP	12	23.0	109.8	-6.8	35.0	37	33.0	30.4	16.0	132.3	8.7	15.1	109.2	2.4	28.4	30.8	24.8	17.8	19.5	149.6	16.4
ONIOM	12	32.1	154.2	-18.7	41.3	37	33.0	32.9	17.0	134.6	9.2	15.0	107.7	**-0.7**	27.1	34.0	25.3	18.4	17.5	134.5	15.4
EE	11	14.0	47.5	4.9	17.9	37	27.0	18.0	4.1	12.1	5.9										
ME'	8	16.0	49.5	-9.3	18.2	37	33.0	30.5	2.0	19.0	2.8	6.6	22.9	-3.5	8.5	38.0	23.0	13.9	0.3	2.5	0.1
ME_{av}	8	14.1	35.8	-5.4	17.0	**33**	33.0	31.0	1.0	4.7	1.4	7.5	25.5	-2.6	9.7	32.0	29.5	14.0	0.1	1.3	0.1
ME_{scal}	8	11.9	40.1	3.9	15.5	37	**21.4**	**14.0**	0.4	2.2	1.4	**5.0**	19.1	2.2	6.9	33.0	**17.5**	2.0	0.1	1.1	0.1
QTCP	8	15.4	46.2	-9.6	17.0	37	33.0	30.9	3.3	12.1	2.3	6.4	27.0	-4.3	8.4	30.0	23.0	14.1	0.8	4.6	1.0
ONIOM	8	23.9	94.6	-21.5	22.0	37	33.0	32.9	3.3	12.0	3.9	7.2	32.3	-7.1	8.6	34.0	23.6	14.9	0.9	4.1	0.9

Table 3. Effects of various components of the tested methods, presented as the difference in mean unsigned error between two methods. The components are: vdW = QM+vdW – QM; ele = ME_0 – QM+vdW; ele_{dir} = ME_{dir} – QM+vdW; juncorr = ME – ME_0; q_{HL} = ME_1 – ME; me' = ME' – ME; pol = ME_{av} – ME_0; scal = ME_{scal} – ME_0; elpol = EE – QM+vdW; qmeff = EE – ME_{av}; qtcp = $QTCP_0$ – EE; oniom = $ONIOM_0$ – $QTCP_0$; polc1 = QTCP – $QTCP_0$; polc2 = ONIOM – $ONIOM_0$. Av is the average over all charge-redistribution schemes, whereas in Av8, Z1, BRCD, BCS, and Z0 (set 1) or DZ1 (set 2) are omitted from the average. Note that the first four contributions, as well as the oniom contribution, are independent of the charge-redistribution scheme and therefore are only listed in the average rows.

	vdW	ele	ele_{dir}	juncorr	q_{HL}	me'	pol	scal	elpol	qmeff	qtcp	oniom	polc1	polc2
Set 1														
Z0						23.1	7.8	19.3	36.7	34.9	20.9		30.1	37.8
BZ0						20.2	7.4	17.6	14.2	16.3	26.5		8.9	15.4
Z1						24.7	10.7	17.9	11.3	29.9	29.8		12.6	14.9
DZ1						21.3	6.5	18.1	21.3	13.8	15.0		3.9	7.2
Z2						22.1	7.5	18.3	13.6	14.1	17.5		7.0	8.8
DZ2						22.3	7.4	18.8	11.8	21.0	23.4		7.2	8.7
NZ2						21.6	7.5	18.6	9.7	16.8	19.8		7.8	10.2
Z3						23.0	8.5	18.6	19.2	10.1	12.7		6.5	8.9
RCD						24.9	8.0	20.2	25.2	15.9	17.0		12.1	19.2
BRCD						27.1	11.8	21.8	16.2	18.1	19.4		56.6	61.8
CS						27.9	10.4	20.6	8.2	25.5	15.4		27.3	33.2
BCS						30.9	13.4	22.0	16.7	18.2	17.1		62.2	66.5
Av	1.6	26.1	16.2	11.6	11.8	24.1	8.9	19.3	17.0	19.6	19.5	11.8	20.2	24.4
Av8	1.6	26.1	16.2	11.6	11.8	22.9	7.9	18.8	15.4	16.7	18.4	11.8	10.1	13.9
Set 2														
Z0						7.2	3.2	9.0	16.6	5.0	4.9		4.7	4.8
Z1						10.1	5.0	10.4	12.2	3.9	1.4		6.7	9.6
DZ1						8.3	3.8	9.0	15.3	4.9	5.1		6.7	6.3
Z2						7.5	3.3	9.0	16.3	5.1	6.1		3.4	4.0
DZ2						7.2	3.2	9.0	13.2	2.9	2.9		2.4	2.8
NZ2						6.7	3.1	8.8	14.4	4.2	6.2		5.1	5.3
Z3						7.2	2.9	8.9	23.5	12.0	12.5		4.7	4.5
DZ3						7.1	3.1	8.9	16.3	4.7	5.8		2.2	2.5
CS						6.9	3.1	8.9	16.6	5.1	6.1		3.2	3.5
BCS						13.1	5.7	10.1	18.6	12.2	11.2		39.6	35.6
RCD						7.0	3.2	8.9	16.5	4.9	5.8		3.6	3.8
BRCD						18.8	8.7	12.3	17.7	14.7	10.9		68.4	64.6
Av	0.5	13.2	13.3	3.9	3.9	8.9	4.0	9.4	16.4	6.6	6.6	4.9	12.6	12.3
Av8	0.5	13.2	13.3	3.9	3.9	7.1	3.1	8.9	16.7	5.5	6.3	4.9	3.7	3.9

Table 4. Residue contributions (in kJ/mol) to the error in the QM/MM calculations for set 1, using the charge-redistribution scheme Z2. Dist is the shortest distance (in Å) of that residue to the smallest (46-atom) QM system. Type is the type of residue, either A – part of the active site, N – a neighbour to the active site, Ch – a charged residues, or O – another residue. The residue number (#) indicates the calculation in which this residue appears in the QM system for the first time. Likewise, the number of junctions (#J) is the number of junction atoms before this residue is moved to the QM system.

#Residue	QM	QM+vdW	ME_0	ME	ME_1	ME_{dir}	ME'	ME_{av}	ME_{scal}	EE	$QTCP_0$	$Oniom_0$	QTCP	Oniom	Dist	Type	#J
1 Ile 544	3.6	4.8	-41.9	-35.3	-43.3	-13.5	-29.6	-35.7	-12.4	5.6	-38.2	-22.2	-32.0	-21.3	1.8	A	5
2 Arg 476	-30.1	-29.3	18.3	22.1	18.5	-6.3	11.3	14.8	-9.0	-2.1	20.1	24.0	9.0	17.1	2.2	Ch	6
3 Cys 546	-12.0	-12.1	15.2	3.3	16.1	3.8	-7.0	4.1	-9.5	-27.8	16.6	4.7	-8.2	-23.3	1.7	A	6
4 Asp 114	19.9	19.9	2.6	13.1	1.6	6.5	4.3	3.4	12.1	1.7	2.9	13.3	3.8	10.1	3.3	Ch	5
5 Cys 75	-4.5	-3.7	-15.8	-15.7	-16.4	-6.3	-9.3	-12.6	-6.5	4.3	-10.1	-10.0	-12.1	-6.7	1.6	A	5
6 Asp 541	22.5	22.4	16.2	21.1	16.0	13.2	-5.2	5.5	8.6	17.5	12.0	16.8	-2.7	0.1	4.8	Ch	4
7 Cys 72	3.3	3.4	-4.7	-13.8	-3.4	3.9	3.5	-0.6	3.4	-10.7	-6.7	-15.8	-0.2	-9.5	1.6	A	5
8 Gln 69	-5.6	-5.8	-10.0	5.1	-10.8	-1.9	1.3	-4.3	-2.3	12.1	-6.3	8.8	-0.6	14.0	3.2	O	6
9 Cys 543	-1.4	-1.2	8.9	-1.1	9.8	6.4	11.8	10.4	5.3	-17.8	5.1	-4.9	11.8	1.4	1.6	A	8
10 Glu S22	12.1	12.2	22.7	18.1	23.4	4.4	10.7	16.7	11.4	15.6	20.8	16.2	17.6	11.4	5.6	Ch	8
11 His 481	-18.3	-17.0	-11.6	-16.1	-10.9	-8.8	-11.3	-11.5	-14.1	-7.7	-16.2	-20.6	-6.9	-10.7	2	O	8
12 His 79	-3.7	-3.9	-13.6	-4.6	-15.6	-20.2	-16.0	-14.8	-10.0	-11.1	-20.7	-11.7	-17.5	-2.2	1.6	A	8
13 Wat	-7.5	-8.0	0.5	0.5	0.2	-1.7	-2.3	-0.9	-5.2	-4.2	-7.5	-7.5	-4.7	-3.1	2.7	O	8
14 Pro 542	0.0	-0.1	0.7	25.8	-0.5	5.9	3.0	1.8	1.5	-6.8	-1.6	23.6	6.1	44.2	3.5	N	8
15 Mg 558	-9.2	-9.3	6.5	27.4	4.9	-12.7	-19.1	-6.3	-14.2	2.1	3.6	24.6	-1.6	-4.9	4.2	Ch	6
16 Arg 428	-7.7	-7.7	4.6	10.2	4.3	-0.3	-25.4	-10.4	-16.5	-15.8	-15.3	-9.7	-23.9	-20.0	7.8	Ch	7
17 Arg 70	0.8	0.8	-12.5	-46.2	-11.1	-2.2	15.3	1.4	8.1	-7.1	-3.9	-37.5	-0.9	-38.0	7.3	Ch	7
18 His 115	-4.5	-4.6	-39.0	-34.7	-40.2	1.8	5.3	-16.8	0.4	-2.3	-2.5	1.8	-2.3	7.2	6.4	Ch	8
19 His 538	-9.9	-9.9	-3.1	-47.7	1.2	-3.4	-6.3	-4.7	-8.1	-1.2	-4.1	-48.7	-6.9	-71.9	8.8	Ch	8
20 Ala 80	1.6	1.7	12.1	31.0	11.6	13.8	3.7	7.9	2.7	-12.9	11.8	30.7	6.0	42.8	2.9	A	8
21 Arg 103	-0.3	-0.3	6.0	16.4	4.9	-6.1	-6.7	-0.4	-3.5	0.2	-4.0	6.4	2.3	20.2	12	Ch	8
22 Glu S46	12.7	12.7	0.2	2.7	0.0	7.4	13.3	6.8	13.0	4.1	11.1	13.5	3.6	4.5	8.7	Ch	8
23 Glu S75	8.3	8.3	1.3	-6.2	1.9	3.4	2.3	1.8	5.3	0.8	-0.8	-8.3	1.2	-2.1	7	Ch	8
24 Asp 88	-10.4	-10.4	-10.3	-15.4	-9.3	-4.6	-6.0	-8.2	-8.2	2.1	5.1	0.0	4.8	3.8	13.9	Ch	8
25 Arg 85	6.1	6.1	-1.5	7.6	-2.5	0.1	-2.0	-1.7	2.1	-6.5	-13.9	-4.8	-12.4	-7.9	9.6	Ch	8
26 Val 78	3.6	3.8	7.9	-1.9	9.0	-3.6	8.3	8.1	6.0	27.2	8.8	-1.0	10.4	-4.6	4.4	N	8
27 Ala 71	2.8	4.1	-1.6	11.8	-1.8	2.8	-2.4	-2.0	0.8	0.0	-4.4	9.0	-2.4	16.9	4.2	N	7
28 Arg 23	-2.9	-2.9	-10.4	-7.7	-10.5	1.8	-3.4	-6.9	-3.1	-0.3	1.2	4.0	-2.7	-0.2	11.9	Ch	3
29 Asp 126	3.8	3.8	4.3	3.7	4.4	-0.6	1.9	3.1	2.9	0.2	-8.9	-9.5	-0.8	-0.9	14	Ch	3
30 Asp 63	-6.7	-6.7	99.3	101.8	99.8	-1.6	-6.3	46.5	-6.5	6.5	18.0	20.5	1.0	-0.3	13	Ch	3
31 Gln 69'	4.0	4.0	13.9	13.6	14.8	-2.7	-0.6	6.7	1.7	-0.1	-0.9	-1.2	-5.3	1.9	3.1	O	3
32 His 549	-2.8	-2.8	-66.2	-88.9	-62.4	-1.5	8.0	-29.1	2.6	0.3	-1.3	-24.0	5.9	-11.2	7.4	Ch	2
33 Glu 334	-7.8	-7.8	-34.5	-30.9	-34.7	-4.5	3.1	-15.7	-2.3	-1.2	2.9	6.5	1.7	4.8	9.8	Ch	1
34 Gly 547	-11.3	-10.8	14.0	31.9	9.4	3.5	16.6	15.3	2.9	-7.2	12.1	30.0	18.3	30.3	1.7	A	1
35 Thr S18	0.2	0.2	-1.4	-1.3	-1.4	-2.1	-0.3	-0.9	-0.1	-0.6	-0.6	-0.5	-0.6	-0.6	3.4	O	0
36 Arg 388	-3.9	-3.9	2.1	2.1	2.1	0.2	-2.6	-0.2	-3.2	-0.2	-0.2	-0.3	-0.2	-0.2	6.4	Ch	0
37 Arg 59	1.6	1.6	1.5	1.4	1.5	-2.6	-2.5	-0.5	-0.5	-1.8	-1.8	-1.8	-1.8	-1.8	13.8	Ch	0
38 Asp 89	-0.3	-0.3	-0.8	-0.8	-0.8	4.2	5.5	2.4	2.6	2.7	2.6	2.7	2.7	2.6	13.8	Ch	0
39 Asp 60	2.3	2.3	4.6	4.6	4.6	6.6	3.6	4.1	2.9	3.1	3.1	3.1	3.1	3.0	16.3	Ch	0
40 Glu 431	3.1	3.1	-2.4	-2.4	-2.4	-0.7	-1.2	-1.8	0.9	-0.4	-0.3	-0.4	-0.4	-0.3	8.8	Ch	0

Table 5. Residue contributions (in kJ/mol) to the error in the QM/MM calculations for the second set of calculations using the Z2 charge-redistribution scheme, if applicable. Dist is the shortest distance (in Å) of that residue to the smallest (46-atom) QM system. The first two systems involve the residues of the original QM system (Cys-72, Cys-75, His-79, Ala-80, Cys-543, Ile-544, Cys-546, and Gly-547; A) and the neighbouring residues (Ala-71, Val-78, and Pro-542; N). The residue number (#) indicates the calculation in which this residue appears in the QM system for the first time. Likewise, the number of junctions (#J) is the number of junction atoms before this residue is moved to the QM system.

#	Residue	QM	QM+vdW	ME_0	ME	ME_1	ME_{dir}	ME'	ME_{av}	Me_{scal}	EE	$QTCP_0$	$Oniom_0$	QTCP	Oniom	Dist	#J
8	A	-24.1	-21.6	-9.6	-1.1	-10.5	-12.8	-17.6	-13.6	-19.6	-52.2	4.7	17.8	-29.8	-14.8	1.6	5
11	N	0.5	1.8	-6.9	1.9	-6.7	-4.8	-3.1	-5.0	-0.7	13.0	10.4	13.0	-3.5	9.5	3.5	4
12	Arg 428	-13.5	-13.5	-18.5	-20.4	-18.6	-6.1	-18.4	-18.4	-15.9	-9.9	-29.8	-21.0	-12.8	-14.0	7.8	3
13	Hid 481	-18.4	-17.1	-5.4	-4.6	-5.4	-8.9	-10.4	-7.9	-13.8	-8.8	-13.5	-12.9	-9.3	-9.0	2	3
14	Arg 85	5.0	5.0	-6.5	-7.8	-6.5	-1.0	-4.6	-5.5	0.2	-11.9	-11.7	-11.0	-4.4	-4.8	9.6	3
15	Glu S22	13.3	13.3	17.9	20.1	17.9	5.5	4.4	11.2	8.9	7.4	13.8	12.6	9.0	7.8	5.6	3
16	Asp 541	14.8	14.8	3.6	12.5	3.7	8.8	11.1	7.4	13.0	0.0	-2.1	0.1	9.0	20.0	4.8	3
17	Wat	-9.8	-10.3	-6.6	-6.7	-6.6	-4.0	-5.4	-6.0	-7.8	-3.9	9.9	17.8	-3.9	-3.8	2.7	2
18	Glu S46	8.4	8.4	8.1	8.7	8.0	3.1	7.6	7.9	8.0	3.0	2.5	2.5	-0.1	-0.8	8.7	2
19	Hip 115	-12.7	-12.7	-2.8	-4.3	-2.6	-6.4	-5.1	-3.9	-8.9	-1.2	1.7	2.3	0.9	1.5	6.4	2
20	Gln 69	-5.2	-5.4	-0.2	2.7	-0.6	-1.4	0.6	0.2	-2.4	14.6	7.9	6.1	-2.6	-0.3	3.2	2
21	Gln 69'	-0.3	-0.3	-2.4	-3.8	-2.3	-0.4	-2.2	-2.3	-1.3	0.2	-13.3	-10.5	-1.8	-3.5	3.1	4
22	Arg 70	4.7	4.7	2.4	-1.7	2.7	0.3	2.0	2.2	3.4	0.7	-1.7	-3.1	1.5	-2.6	7.3	1
23	Hip 538	-7.4	-7.4	-1.8	-1.8	-1.8	-0.9	-3.7	-2.8	-5.6	-1.5	-0.8	-4.8	-1.6	-1.5	8.8	0
24	Asp 60	1.2	1.2	8.5	8.5	8.5	5.5	7.4	8.0	4.3	6.2	6.2	6.2	6.2	6.2	16.3	0
25	Asp 89	0.3	0.3	1.2	1.2	1.2	4.7	6.9	4.1	3.6	3.6	3.7	3.6	3.7	3.6	13.8	0
26	Asp 88	-1.4	-1.4	3.0	3.0	3.0	4.3	-2.8	0.1	-2.1	0.3	0.2	0.2	0.2	0.3	13.9	0
27	Arg 59	0.5	0.5	-2.0	-2.0	-2.0	-3.7	-2.9	-2.4	-1.2	-0.9	-0.9	-0.9	-0.9	-1.0	13.8	0
28	Mg 558	9.5	9.4	4.5	-4.8	4.6	6.0	1.8	3.2	5.6	1.7	1.9	1.7	1.7	-6.4	4.2	0
29	Glu S75	0.5	0.5	28.0	30.6	27.9	-4.4	-1.3	13.3	-0.4	3.8	3.6	-5.8	3.5	1.5	7	1
30	Asp 114	0.7	0.7	-14.1	-13.2	-14.0	-12.7	-0.6	-7.4	0.0	1.3	2.0	4.5	1.2	0.4	3.3	1
31	Asp 63	0.5	0.5	2.5	2.0	2.5	5.6	2.6	2.5	1.5	0.9	1.0	1.9	1.1	3.2	13	1
32	Arg 103	-0.3	-0.3	0.6	1.9	0.6	-6.1	-1.1	-0.2	-0.7	-1.6	-1.7	-2.2	-1.8	-3.3	12	1
33	Glu 334	-4.5	-4.5	-21.4	-22.5	-21.4	-1.1	7.6	-6.9	1.6	-1.7	-1.4	-0.1	-1.6	-1.6	9.8	1
34	Arg 476	-7.6	-6.8	0.3	-3.4	0.3	16.2	-5.9	-2.8	-6.3	-2.4	-4.3	-3.6	-2.2	-1.2	2.2	1
35	Hid 549	-3.9	-3.9	-0.4	9.3	-0.4	-2.6	1.4	0.5	-1.2	0.7	0.3	-3.5	0.8	10.1	7.4	1
36	Thr S18	-0.8	-0.9	-1.1	-1.1	-1.1	-3.1	-0.3	-0.7	-0.6	-0.7	-0.5	9.1	-0.7	-0.8	3.4	0
37	Arg 23	-2.7	-2.7	-2.0	-2.1	-2.0	2.0	0.8	-0.6	-1.0	-0.2	-0.2	-0.2	-0.2	-0.2	11.9	0
38	Glu 431	0.1	0.1	6.7	6.7	6.7	-3.7	4.2	5.5	2.1	0.3	0.3	0.3	0.3	0.3	8.8	0
39	Arg 388	-0.1	-0.1	-1.2	-1.3	-1.2	4.0	-5.0	-3.1	-2.6	-0.5	-0.5	-0.5	-0.5	-0.5	6.4	0
40	Asp 126	4.5	4.5	-2.4	-2.4	-2.4	6.0	-1.0	-1.7	1.8	0.0	0.0	0.0	0.0	0.0	14	0

Table 6. The performance of the various methods for the QM/MM calculations including a 12178-atom model of the protein and the surrounding water. The quality measures are the same as in Table 2, but calculations were performed only with the Z1 charge-redistribution scheme (if applicable) and only for set 1. Reference is the reference value for the proton-transfer reaction energy (in kJ/mol) for the calculation with the 446-atom QM system.

	MUE	Max	MSE	Stdev	n5	n10	n20	Reference
QM	20.5	56.1	18.8	18.1	38	32	29	48.3
QM+vdW	19.2	53.7	17.2	**17.4**	38	32	29	47.9
ME_0	19.2	52.8	8.9	21.6	39	33	33	67.7
ME	23.2	83.5	-6.1	28.6	39	39	35	52.2
ME_1	19.6	54.6	9.3	22.0	39	33	31	66.0
ME_{dir}	**16.1**	43.1	3.1	19.1	38	34	33	67.5
ME'	43.8	83.1	-7.9	50.1	**34**	34	32	72.4
ME_{av}	27.7	58.4	**0.5**	31.8	**34**	33	32	70.1
ME_{scal}	26.2	63.8	4.7	30.6	36	32	32	60.1
EE	26.4	68.3	25.7	19.6	39	**26**	**25**	57.4
$QTCP_0$	21.0	**42.7**	2.2	25.0	39	33	32	74.5
$ONIOM_0$	27.3	73.0	-12.8	31.2	39	39	33	59.0
QTCP	29.5	82.8	-9.3	35.6	39	39	32	28.3
ONIOM	39.2	103.3	-8.5	48.9	36	34	34	24.4

Figure 1. Illustration of the partitioning between the QM and MM systems and the naming of the atoms.
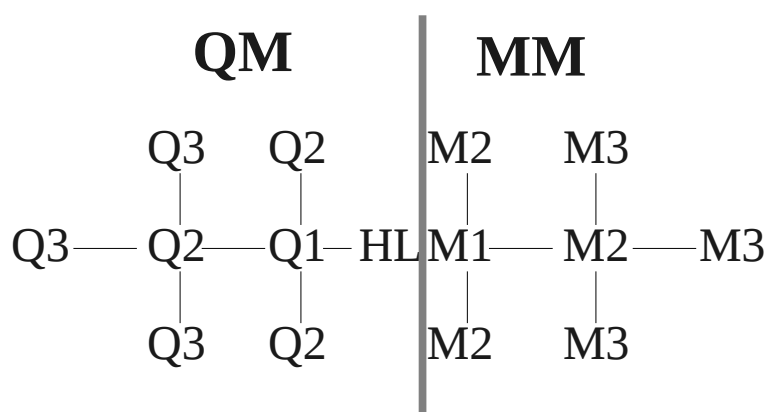
Figure 2. Illustration of the excluded-atom problem, using ethanol as an example, with the HOH moiety as the QM system and the rest as the MM system. The HO and H23 atoms are separated by four bonds and therefore their interaction is included fully in both the MM and QM calculations. However, the O and C2 atoms are separated by only two bonds and their interaction is therefore excluded in the MM calculations, whereas it is included in the QM calculations.
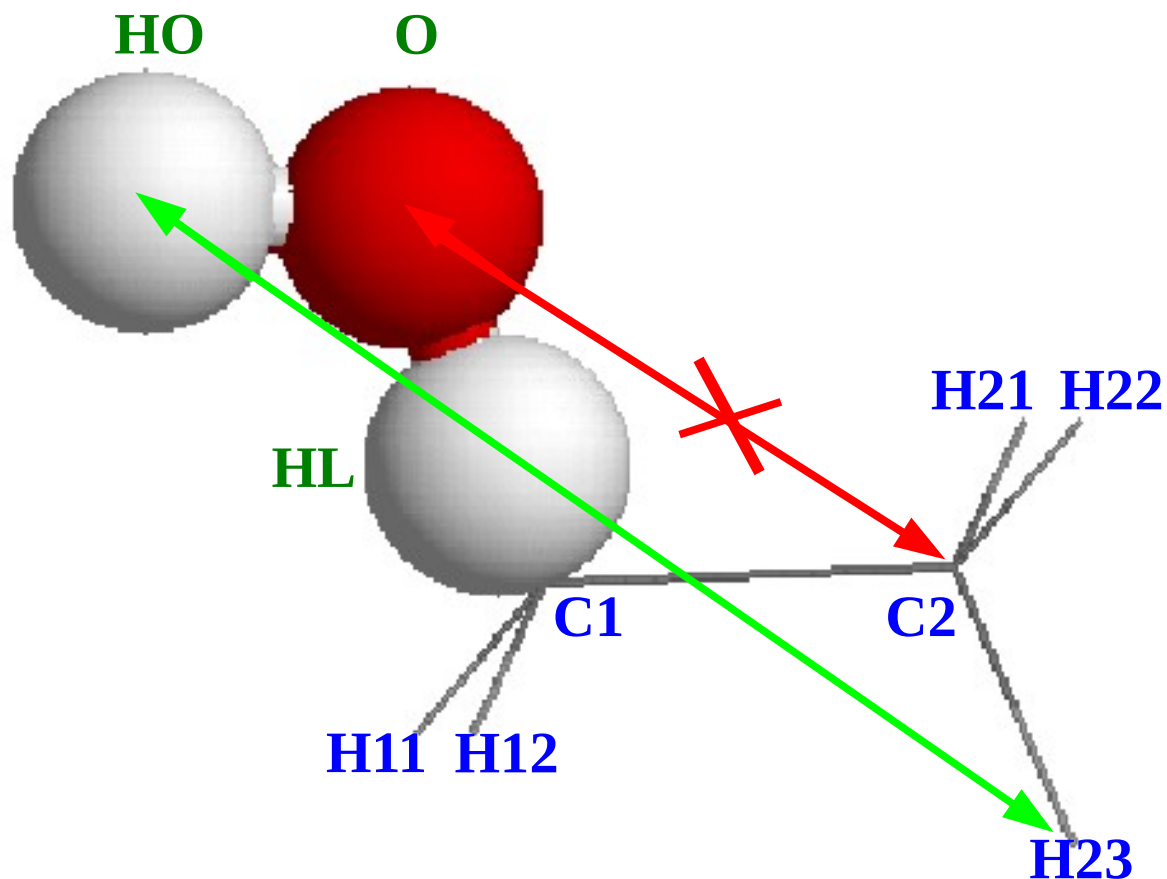
Figure 3. Atoms included in the 446-atom test system. The Mg ligand Gln540 is mainly hidden behind His79. Note that there are four water molecules in the calculation: one at the arrow, one just to left of the label, and two just to the right of the label. All are Mg ligands, except the upper to the right of the label. The smallest (46-atom) quantum system is shown in balls and sticks and with blue bold-face labels.
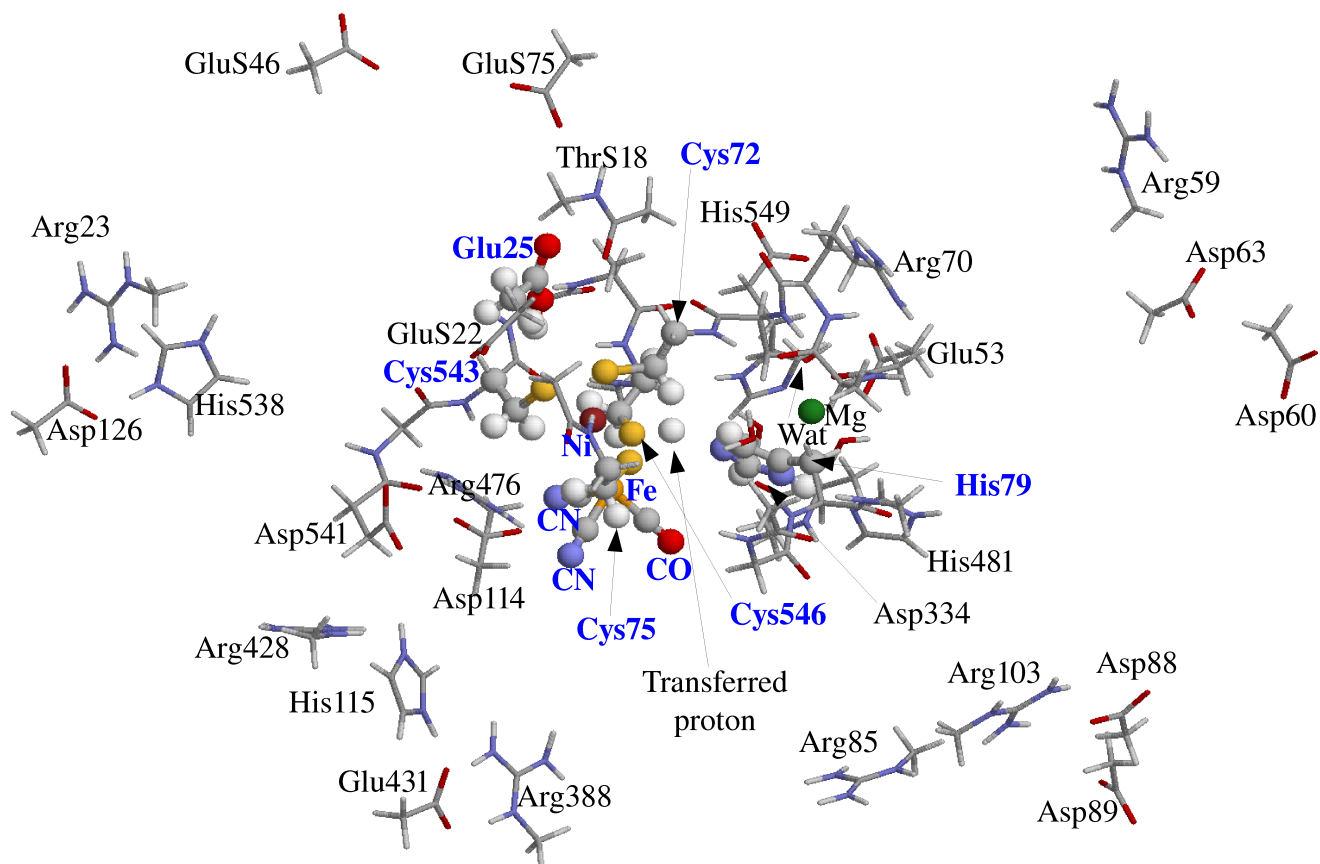
Figure 4. Errors for the 14 methods tested for the two sets (a, b, c, set 1; d, e, f set 2) of calculations (using the Z1 charge-redistribution scheme when applicable; results for the other charge-redistribution schemes are shown in Figures S2 and S3 in the supplementary material). Note that the QM+vdW method almost entirely coincidence with QM and ME1 with ME.
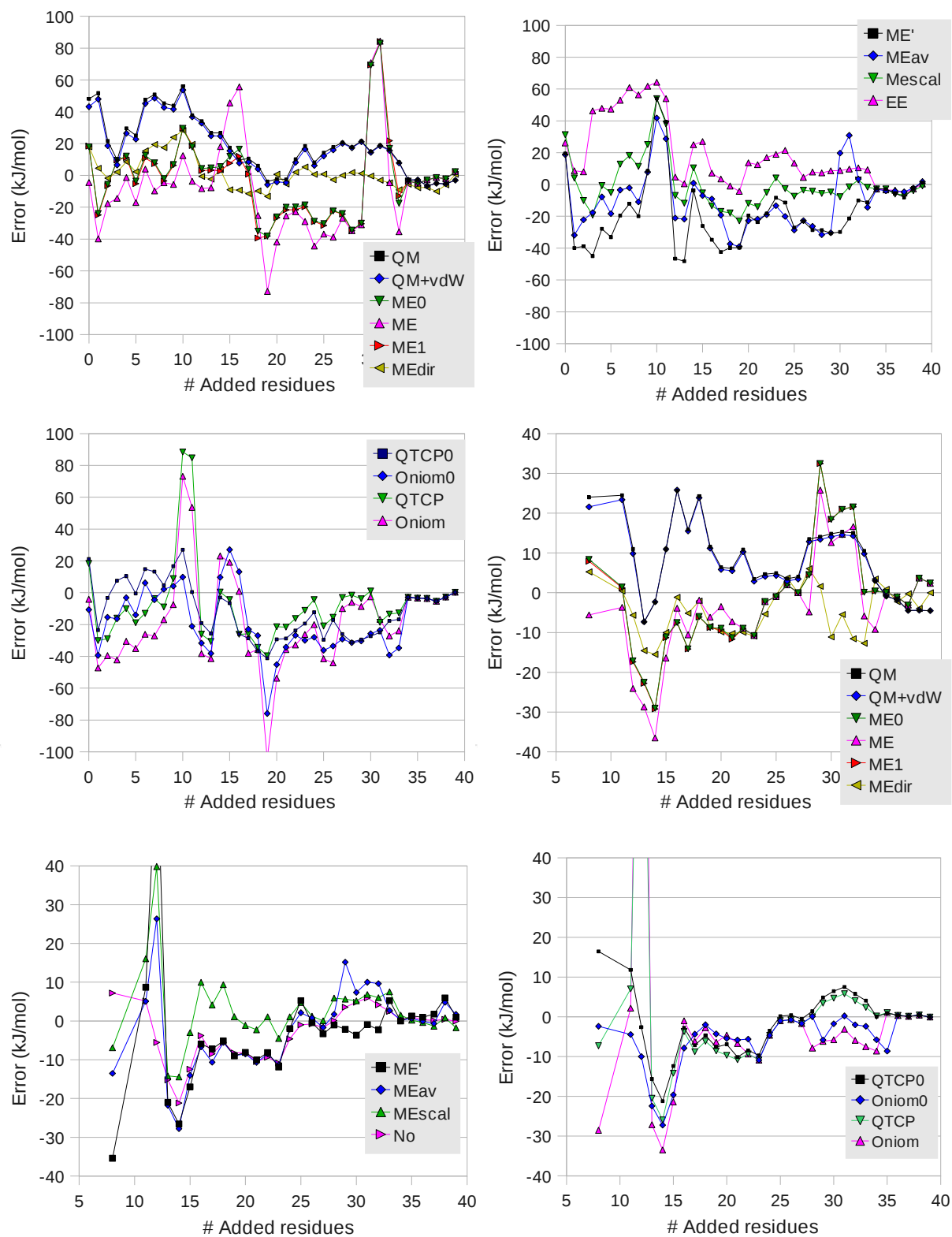
Figure 5. The mean unsigned error (MUE) of two methods, EE and ME_{av}, as a function of the charge-redistribution scheme for the two sets of calculations.
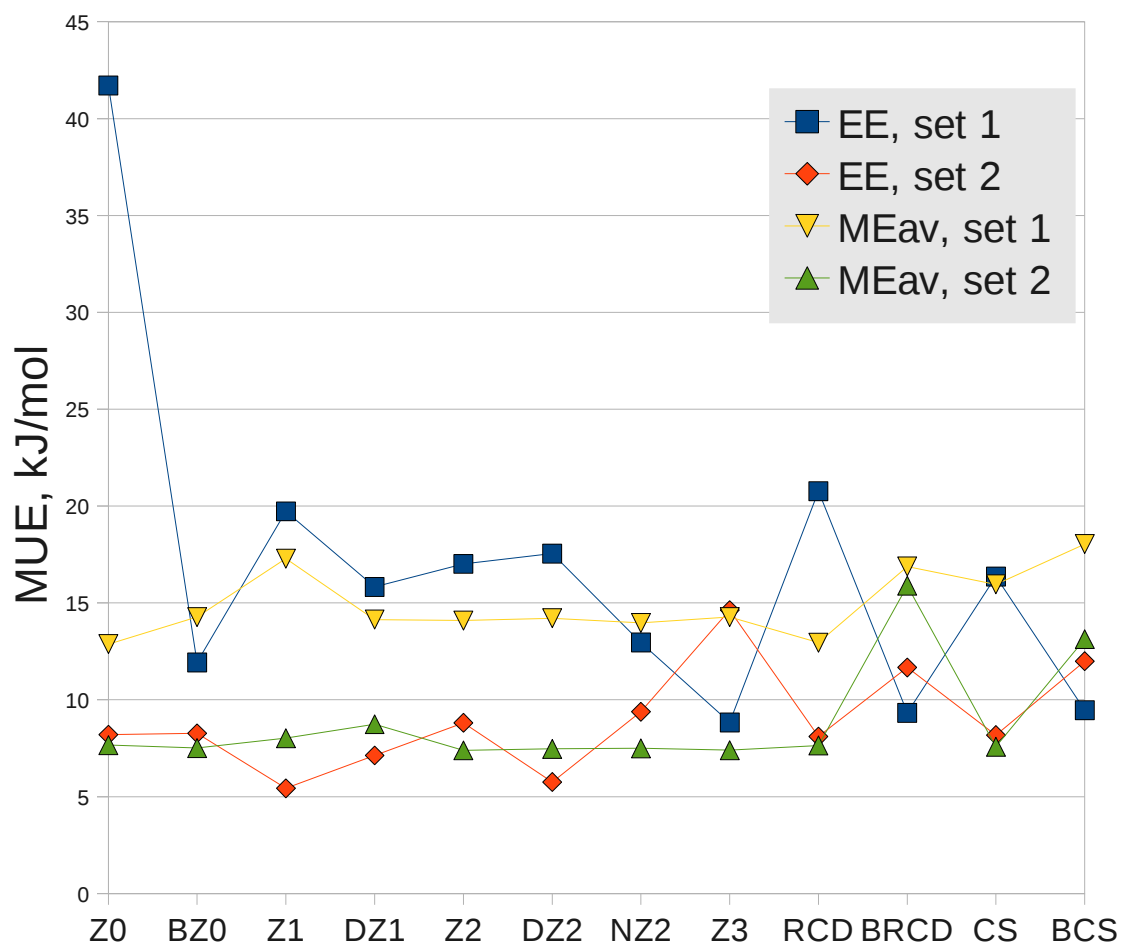
Figure 6. Position of each method in the space described by the mean unsigned error (MUE) and the mean signed error (MSE) for set 1, with the value at each arrow giving the size of the effect as described in Table 2. All values are average energies in kJ/mol over the various charge-redistribution schemes. For all methods where charge redistributions occur, the results shown are Av8, i.e. omitting Z1, BRCD, BCS, and Z0.