

# JCTC Journal of Chemical Theory and Computation

## AM1/d Parameters for Magnesium in Metalloenzymes

Petra Imhof,<sup>†</sup> Frank Noé,<sup>†</sup> Stefan Fischer,<sup>‡</sup> and Jeremy C. Smith<sup>\*†</sup>

*Computational Molecular Biophysics and Computational Biochemistry, IWR University of Heidelberg, Im Neuenheimer Feld 368, 69120 Heidelberg, Germany*

Received March 10, 2006

**Abstract:** AM1/d parameters are derived for magnesium, optimized for modeling reactions in metalloenzymes. The parameters are optimized with a Monte Carlo procedure so as to reproduce the geometries and energies of a training set calculated with density functional theory. The training set consists of compounds with magnesium coordinated to the oxygen atom of typical biological ligands. Optimization of AM1 parameters without extension to *d* functions leaves serious errors. The new AM1/d parameters provide a clear improvement in accuracy compared to the standard semiempirical methods AM1 and MNDO/d and will be particularly useful for modeling reactions in large biological systems at low computational cost.

### 1. Introduction

Magnesium is the metal cofactor of numerous metalloenzymes. A popular modeling approach to understanding such reactions in enzymes is the combined quantum mechanical/molecular mechanical (QM/MM) ansatz, where the region of interest (usually the active site) is treated quantum mechanically and the remainder of the enzyme is described with an empirical force field.<sup>1–3</sup> Ab initio methods for the QM part are not only the most accurate but also the most computationally demanding and therefore used only in special cases. Alternatively, density functional (DFT) methods provide a more attractive balance of accuracy and computational cost than ab initio techniques and thus enjoy high popularity in the modeling of chemical reactions. However, although a single minimization step with DFT methods can be easily afforded, a complete optimization with thousands of such steps can become computationally costly. Especially when several of these minimizations are necessary, e.g., for the exploration of different reaction pathways, more economical methods are needed. Responding to this need, semiempirical methods provide a sufficiently accurate description of quantum regions in QM/MM setups of large systems for low computational cost.

Semiempirical methods derive their efficiency from explicit treatment of only valence electrons with a minimal basis set, the neglect of three- and four-center integrals, and the use of parametrized expressions for two-center integrals.<sup>4–8</sup> The parameters are usually obtained by a fit of properties (e.g., heats of formation) to a variety of very small compounds. Often these training sets are not representative of reactions in biological systems. However, the situation can be improved by the development of reaction-specific parameters, which are tuned to most accurately describe the specific biological systems under study, at the expense of losing generality.

The AM1 model is at present one of the most suitable semiempirical methods for studying reactions,<sup>8</sup> although it does have a tendency to predict bifurcated and too-weak hydrogen bonds.<sup>9</sup>

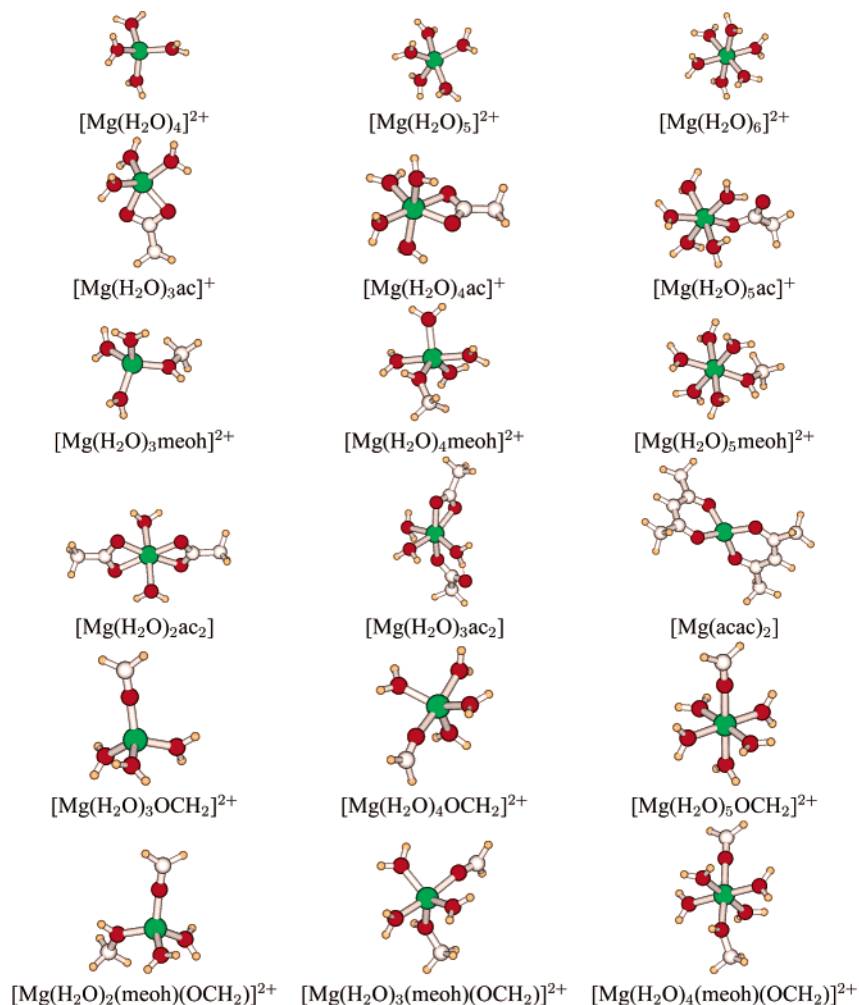
The standard AM1 parameters for magnesium have been developed for use in modeling the bacterial photosynthetic reaction center<sup>10</sup> and were fitted to reproduce mainly properties of divalent magnesium compounds. These parameters work quite well for most of the compounds listed in ref 10, including magnesium porphyrin, but yield wrong angles for the geometry of 6-fold coordinated magnesium (e.g.  $[\text{Mg}(\text{H}_2\text{O})_6]^{2+}$ ). The MNDO/d method<sup>11</sup> yields correct angle values but too long Mg–O bond lengths. Both methods use an *sp* basis for magnesium, and thus one cannot expect a proper description of hypervalent magnesium compounds.

In metalloenzymes, 6-fold coordinated magnesium is quite common (ref 12 provides a survey of the Brookhaven Protein

\* Corresponding author phone: ++49 6221 8857; e-mail: [biocomputing@iwr.uni-heidelberg.de](mailto:biocomputing@iwr.uni-heidelberg.de).

<sup>†</sup> Computational Molecular Biophysics.

<sup>‡</sup> Computational Biochemistry.



**Figure 1.** Compounds used in the training set for the magnesium parametrization.

Data Bank<sup>13</sup> for X-ray and NMR structures of magnesium-bound proteins). To obtain a useful description of magnesium-containing active sites with different magnesium coordination spheres and magnesium-dependent reactions in metalloproteins at a semiempirical level, the present paper extends the AM1 parameters for magnesium to an *spd* basis in the AM1/d framework. The parameter set is derived specifically for oxygen-based ligands modeling magnesium coordination spheres that can be typically found in metalloproteins.

## 2. Methods

**2.1. The Training Set.** The AM1/d parameters for magnesium were derived by fitting properties of a set of magnesium compounds to a DFT training set consisting of model compounds for magnesium coordinated to the oxygen atom of typical biological ligands with coordination numbers 4, 5, and 6. These ligands are water, methanol (meoh), which models serine, threonine, and tyrosine amino acid side chains, acetate (ac) as a representative for aspartate and glutamate side chains, and formaldehyde (OCH<sub>2</sub>) modeling the coordination by a backbone carbonyl oxygen atom. The compounds used in the training sets are shown in Figure 1. The Cartesian coordinates of the DFT optimized structures used as a training set are given as Supporting Information.

The DFT data set was obtained by geometry optimization with the B3LYP functional<sup>15,16</sup> and a 6-31++G(d,p) basis set with subsequent single-point energy calculations using a 6-311++G(3df,2p) basis set. Normal-mode analysis on the optimized geometries was carried out to verify that a minimum energy structure has been obtained. All DFT calculations were performed using the Turbomole program package.<sup>17</sup> The B3LYP/6-311++G(3df,2p)//B3LYP/6-31++G(d,p) procedure is abbreviated as DFT in the remainder of this paper. This procedure followed here is similar to that described in ref 14 for the development of AM1/d parameters for phosphorus reaction-specific for nucleophilic attacks on biological phosphates.

In ref 14, *d*-orbitals are introduced only where necessary, e.g. on the phosphorus, while treating C, H and O atoms with standard AM1 parameters. We follow a similar approach, by extending the AM1 basis set to *d*-orbitals where necessary (here for magnesium) while keeping as much of the standard AM1 model as possible. Thus, the magnesium complexes are composed of the ligand molecules, which are treated with standard AM1, and the additional Mg<sup>2+</sup> ion, which is treated with the more extended AM1/d.

In a molecular orbital picture the basis functions of all atoms together form the molecular orbitals. Since mixed basis sets have to be used with care this would mean that, in a

**Table 1.** Optimized AM1/d and AM1 Parameters for Magnesium

parameter	AM1/d	AM1'
$U_{ss}/\text{eV}$	-16.63758	-12.83615
$U_{pp}/\text{eV}$	-11.97469	-9.51125
$U_{dd}/\text{eV}$	-10.90361	
$\beta_s/\text{eV}$	-3.60785	-1.26808
$\beta_p/\text{eV}$	-2.07794	-0.93230
$\beta_d/\text{eV}$	-3.30858	
$\zeta_s/\text{au}$	1.16850	1.57114
$\zeta_p/\text{au}$	1.07072	1.25833
$\zeta_d/\text{au}$	0.93469	
$\alpha/\text{\AA}^{-1}$	1.28263	1.80310
$a_1$ (dimensionless)	1.84869	1.99069
$b_1/\text{\AA}^{-2}$	4.22931	3.80477
$c_1/\text{\AA}$	0.66917	0.66033
$a_2$ (dimensionless)	0.03381	-0.00626
$b_2/\text{\AA}^{-2}$	3.57399	3.06817
$c_2/\text{\AA}$	2.33163	1.53666
$a_3$ (dimensionless)	0.02860	-0.00581
$b_3/\text{\AA}^{-2}$	2.27472	2.33455
$c_3/\text{\AA}$	2.89337	2.42691
$\rho^{\text{core}}/\text{au}$	0.94048	
$g_{sp}/\text{eV}$	7.48305	8.29115
$h_{sp}/\text{eV}$	0.67433	0.53547
$\bar{\zeta}_s/\text{au}$	1.61862	
$\bar{\zeta}_p/\text{au}$	1.48840	
$\bar{\zeta}_d/\text{au}$	1.07347	

semiempirical framework, all parameters used must be reoptimized. In trial calculations, further reoptimization of the AM1 parameters was performed (data not shown). Only changes of the parameters for oxygen, which is directly bound to magnesium and thus should be most affected, resulted in any significant influence on the energy and geometry data. However, no significant improvement was obtained, and thus, for simplicity, standard parameters were retained for all elements other than magnesium.

Properties used for the fitting reported here include geometries, Mg–O bond distances and O–Mg–O angles, and reaction energies for ligand exchange (see Tables B in the Supporting Information). The reaction energies included in the fits are listed in the results section (see Table 2 and Table C in the Supporting Information).

Although the aim of the fitting is to obtain parameters that reproduce DFT geometries and *relative* DFT energies (reaction and protonation energies), the absolute heat of formation of [Mg(acac)<sub>2</sub>] is included as a reference to keep the shift of the absolute energies moderate.

**2.2. The Error Function.** For AM1/d there are 25 adjustable parameters:  $U_{ss}$ ,  $U_{pp}$ , and  $U_{dd}$  for the one-electron integrals;  $\zeta_s$ ,  $\zeta_p$ ,  $\zeta_d$ ,  $\beta_s$ ,  $\beta_p$ , and  $\beta_d$  for the resonance integrals; and  $\alpha$ ,  $a$ ,  $b$ ,  $c$ , and  $\rho_{\text{core}}$  for the core–core interaction.<sup>6</sup> For one-center two-electron integrals only the parameters  $g_{sp}$  and  $h_{sp}$  are given explicitly in the implementation of the MNDO program which was employed here,<sup>18</sup> and the other one-center Coulomb integrals  $g_{ss}$ ,  $g_{pp}$ , and  $g_{dd}$  are calculated from orbital exponent parameters  $\bar{\zeta}_s$ ,  $\bar{\zeta}_p$ , and  $\bar{\zeta}_d$ .<sup>11,19,20</sup>

In the optimization procedure, the AM1/d parameter set  $\lambda = (U_{ss}, U_{pp}, \dots, \bar{\zeta}_d)$  was varied so as to minimize the deviation of geometries, reaction energies, and heats of formation with respect to the reference values. This deviation is measured by the following error function

$$\chi^2 = \sum_i \sum_a^{\text{compprop}} w_a [Y_{ia}^{\text{AM1/d}}(\lambda) - Y_{ia}^{\text{DFT}}]^2$$

where  $Y_{ia}^{\text{DFT}}$  is the DFT, and  $Y_{ia}^{\text{AM1/d}}$  is the AM1/d value for property  $a$  of compound  $i$ .  $w_a$  is the weighting factor used for each property: bond lengths, bond angles, reaction energies, and heats of formation.

As start parameters the standard  $sp$  MNDO/d parameters<sup>20</sup> and the standard AM1 core–core parameters<sup>10</sup> were taken. For the additional  $d$  specific start parameters were set:  $U_{dd} = U_{pp}$ ,  $\zeta_d = \zeta_p$ ,  $\beta_d = \beta_p$  and  $\bar{\zeta}_s = \zeta_s$ ,  $\bar{\zeta}_p = \zeta_p$ ,  $\bar{\zeta}_d = \zeta_d$ . The weighting factors used were as follows: absolute energies 0.1 (kcal/mol)<sup>-2</sup>, relative energies 1 (kcal/mol)<sup>-2</sup>, bond distances 100 Å<sup>-2</sup>, bond angles 10<sup>o-2</sup>.

In each iteration of the optimization procedure, the properties on the semiempirical level were computed for fully geometry-optimized structures using a prerelease version of the MNDO99 program.<sup>18</sup>

**2.3. Optimization.** The error function  $\chi^2$  was minimized using a Monte Carlo procedure. This was initialized with the starting parameters  $\lambda_0$ . At each step  $t + 1$ , a new parameter set  $\lambda_{t+1}$  was generated by randomly perturbing the previous parameter set  $\lambda_t$

$$\lambda_{j,t+1} := \lambda_{j,t} + s(r - 0.5)\sigma_j$$

where  $s$  is the step length,  $r \in [0, 1]$  is a random number, the index  $j$  runs over the parameters, and the standard deviations  $\sigma$  are identical to the initial parameter set  $\sigma = \lambda_0$ . A step and the new parameter set were accepted, if the new error function had a lower value than previously. Otherwise, it was rejected, and the old parameter set was kept. A step is also rejected, if one of the minimizations does not yield a true minimum (only positive vibrational frequencies).

The error function above was evaluated for each compound in each step, i.e. when the result for a compound produced terms whose sums were already larger than the old error value, the step was rejected immediately. The step length was changed adaptively. Upon an accepted step, the step length was multiplied by a factor of 1.5, otherwise it was divided by a factor of 2, while always remaining within a set of bounds, here:  $s \in [0.05, 0.3]$ .

### 3. Results and Discussion

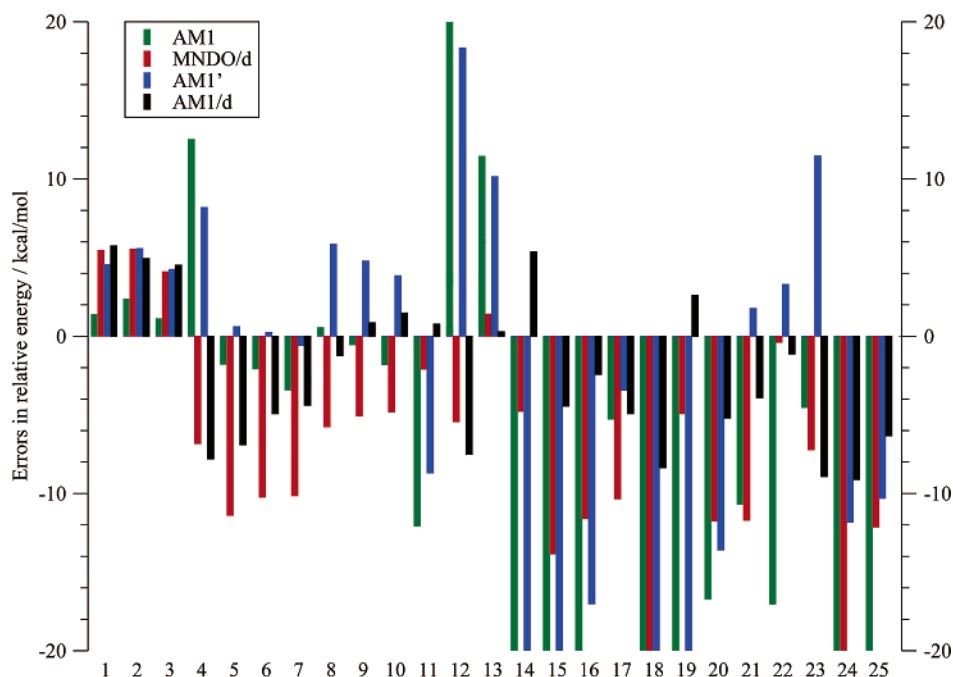
Table 2 shows Mg–O bond distances, O–Mg–O angles, and the reaction energies of ligand substitution at the central magnesium. The optimized parameters are listed in Table 1. AM1' denotes the adjusted  $sp$  parameters, and AM1/d has fitted parameters for a  $spd$  basis.

**Energies.** Figure 2 shows reaction energies for ligand exchange reactions at the magnesium center calculated at the different semiempirical levels (AM1/d, AM1', AM1, and MNDO/d) plotted versus the DFT reference. Table 2 lists the respective values.

**Table 2.** Reaction Energies in kcal/mol for Magnesium Compounds<sup>a</sup>

no.	reaction	CN	charge	DFT	AM1/d	AM1'	AM1	MNDO/d
1	[Mg(H <sub>2</sub> O) <sub>6</sub> ] <sup>2+</sup> + meoh → [Mg(H <sub>2</sub> O) <sub>5</sub> meoh] <sup>2+</sup> + H <sub>2</sub> O	6 → 6	2 → 2	-3	3	1	-2	2
2	[Mg(H <sub>2</sub> O) <sub>4</sub> ] <sup>2+</sup> + meoh → [Mg(H <sub>2</sub> O) <sub>3</sub> meoh] <sup>2+</sup> + H <sub>2</sub> O	4 → 4	2 → 2	-6	-1	-1	-4	-1
3	[Mg(H <sub>2</sub> O) <sub>5</sub> ] <sup>2+</sup> + meoh → [Mg(H <sub>2</sub> O) <sub>4</sub> meoh] <sup>2+</sup> + H <sub>2</sub> O	5 → 5	2 → 2	-4	1	1	-3	0
4	[Mg(H <sub>2</sub> O) <sub>5</sub> ac] <sup>+</sup> → [Mg(H <sub>2</sub> O) <sub>4</sub> ac] <sup>+</sup> + H <sub>2</sub> O	6 → 6	1 → 1	15	7	23	27	8
5	[Mg(H <sub>2</sub> O) <sub>4</sub> ] <sup>2+</sup> + OCH <sub>2</sub> → [Mg(H <sub>2</sub> O) <sub>3</sub> OCH <sub>2</sub> ] <sup>2+</sup> + H <sub>2</sub> O	4 → 4	2 → 2	-5	-12	-4	-7	-16
6	[Mg(H <sub>2</sub> O) <sub>5</sub> ] <sup>2+</sup> + OCH <sub>2</sub> → [Mg(H <sub>2</sub> O) <sub>4</sub> OCH <sub>2</sub> ] <sup>2+</sup> + H <sub>2</sub> O	5 → 5	2 → 2	-3	-8	-3	-5	-13
7	[Mg(H <sub>2</sub> O) <sub>6</sub> ] <sup>2+</sup> + OCH <sub>2</sub> → [Mg(H <sub>2</sub> O) <sub>5</sub> OCH <sub>2</sub> ] <sup>2+</sup> + H <sub>2</sub> O	6 → 6	2 → 2	-1	-6	-2	-5	-11
8	[Mg(H <sub>2</sub> O) <sub>4</sub> ] <sup>2+</sup> + meoh + OCH <sub>2</sub> → [Mg(H <sub>2</sub> O) <sub>2</sub> (meoh)(OCH <sub>2</sub> )] <sup>2+</sup> + 2H <sub>2</sub> O	4 → 4	2 → 2	-11	-12	-5	-10	-16
9	[Mg(H <sub>2</sub> O) <sub>5</sub> ] <sup>2+</sup> + meoh + OCH <sub>2</sub> → [Mg(H <sub>2</sub> O) <sub>3</sub> (meoh)(OCH <sub>2</sub> )] <sup>2+</sup> + 2H <sub>2</sub> O	5 → 5	2 → 2	-7	-7	-3	-8	-13
10	[Mg(H <sub>2</sub> O) <sub>6</sub> ] <sup>2+</sup> + meoh + OCH <sub>2</sub> → [Mg(H <sub>2</sub> O) <sub>4</sub> (meoh)(OCH <sub>2</sub> )] <sup>2+</sup> + 2H <sub>2</sub> O	6 → 6	2 → 2	-4	-3	0	-6	-9
11	[Mg(H <sub>2</sub> O) <sub>4</sub> ] <sup>2+</sup> + 2H <sub>2</sub> O → [Mg(H <sub>2</sub> O) <sub>6</sub> ] <sup>2+</sup>	4 → 6	2 → 2	-57	-56	-65	-69	-59
12	[Mg(H <sub>2</sub> O) <sub>5</sub> ac] <sup>+</sup> → [Mg(H <sub>2</sub> O) <sub>3</sub> ac] <sup>+</sup> + 2H <sub>2</sub> O	6 → 5	1 → 1	31	24	49	55	26
13	[Mg(H <sub>2</sub> O) <sub>4</sub> ac] <sup>+</sup> → [Mg(H <sub>2</sub> O) <sub>3</sub> ac] <sup>+</sup> + H <sub>2</sub> O	6 → 5	1 → 1	17	17	27	28	18
14	[Mg(H <sub>2</sub> O) <sub>5</sub> ] <sup>2+</sup> + ac → [Mg(H <sub>2</sub> O) <sub>5</sub> ac] <sup>+</sup>	5 → 6	2 → 1	-231	-225	-256	-265	-235
15	[Mg(H <sub>2</sub> O) <sub>4</sub> ] <sup>2+</sup> + ac → [Mg(H <sub>2</sub> O) <sub>4</sub> ac] <sup>+</sup>	4 → 6	2 → 1	-246	-250	-268	-275	-260
16	[Mg(H <sub>2</sub> O) <sub>5</sub> ] <sup>2+</sup> + ac → [Mg(H <sub>2</sub> O) <sub>4</sub> ac] <sup>+</sup> + H <sub>2</sub> O	5 → 6	2 → 1	-216	-218	-233	-238	-228
17	[Mg(H <sub>2</sub> O) <sub>6</sub> ] <sup>2+</sup> + ac → [Mg(H <sub>2</sub> O) <sub>3</sub> ac] <sup>+</sup> + 3H <sub>2</sub> O	6 → 5	2 → 1	-173	-178	-176	-178	-183
18	[Mg(H <sub>2</sub> O) <sub>4</sub> ] <sup>2+</sup> + 2ac → [Mg(H <sub>2</sub> O) <sub>2</sub> ac <sub>2</sub> ] <sup>+</sup> + 2H <sub>2</sub> O	4 → 6	2 → 0	-354	-362	-374	-45	-379
19	[Mg(H <sub>2</sub> O) <sub>6</sub> ] <sup>2+</sup> + ac → [Mg(H <sub>2</sub> O) <sub>5</sub> ac] <sup>+</sup> + H <sub>2</sub> O	6 → 6	2 → 1	-204	-201	-225	-233	-209
20	[Mg(H <sub>2</sub> O) <sub>6</sub> ] <sup>2+</sup> + ac → [Mg(H <sub>2</sub> O) <sub>4</sub> ac] <sup>+</sup> + 2H <sub>2</sub> O	6 → 6	2 → 1	-189	-194	-203	-206	-201
21	[Mg(H <sub>2</sub> O) <sub>4</sub> ac] <sup>+</sup> + ac → [Mg(H <sub>2</sub> O) <sub>2</sub> ac <sub>2</sub> ] <sup>+</sup> + 2H <sub>2</sub> O	6 → 6	1 → 0	-108	-112	-106	230	-119
22	[Mg(H <sub>2</sub> O) <sub>4</sub> ac] <sup>+</sup> + ac → [Mg(H <sub>2</sub> O) <sub>3</sub> ac <sub>2</sub> ] <sup>+</sup> + H <sub>2</sub> O	6 → 6	1 → 0	-122	-124	-119	-140	-123
23	[Mg(H <sub>2</sub> O) <sub>5</sub> ac] <sup>+</sup> + ac → [Mg(H <sub>2</sub> O) <sub>3</sub> ac <sub>2</sub> ] <sup>+</sup> + 2H <sub>2</sub> O	6 → 6	1 → 0	-108	-117	-96	-112	-115
24	[Mg(H <sub>2</sub> O) <sub>6</sub> ] <sup>2+</sup> + 2ac → [Mg(H <sub>2</sub> O) <sub>2</sub> ac <sub>2</sub> ] <sup>+</sup> + 4H <sub>2</sub> O	6 → 6	2 → 0	-297	-306	-309	24	-320
25	[Mg(H <sub>2</sub> O) <sub>6</sub> ] <sup>2+</sup> + 2ac → [Mg(H <sub>2</sub> O) <sub>3</sub> ac <sub>2</sub> ] <sup>+</sup> + 3H <sub>2</sub> O	6 → 6	2 → 0	-312	-318	-322	-345	-324

<sup>a</sup> AM1/d values are calculated with fitted magnesium *spd* parameters, AM1' with fitted *sp* parameters. Column CN gives the change in coordination number, charge lists the change in charge of the magnesium complex.



**Figure 2.** Errors of semiempirical reaction energies compared to DFT. AM1/d values are calculated with fitted magnesium *spd* parameters, AM1' with fitted *sp* parameters. The plot for energy errors is cropped at  $\pm 20$  kcal/mol, and some AM1 values exceed this range. The respective reactions are listed in Table 2.

The ligand exchange reactions can be separated into four classes: 1. reactions with change in coordination number (reactions 11–13), 2. reactions with change of the charge of the magnesium compound (19–25), 3. both of the above (14–18), and 4. neither of the above (1–10). As anticipated, the largest reaction energies are found for those reactions in

which the charge of the magnesium complex is decreased, i.e. opposite charges are brought together from an infinite distance. These reaction energies are less in solution<sup>12,21</sup> or in a protein environment than in vacuo. One effect which leads to a reduction of the reaction energies is charge screening by the solvent. In addition, in a protein environ-

**Table 3.** Performance of the Semiempirical Methods AM1/D, AM1', AM1, and MNDO/d for the Magnesium Complexes in Figure 1<sup>a</sup>

property (number of comparisons)	AM1/d	AM1'	AM1	MNDO/d
relative energies (25): mean abs. error/kcal/mol	5	10	14	9
relative energies: max abs. error/kcal/mol	9	25	39	26
bond lengths (93): mean abs. error/Å	0.02	0.15	0.07	0.07
bond lengths: max abs. error/Å	0.07	0.21	0.18	0.23
angles (199): mean abs. error/degree	4	11	15	4
angles: max abs. error/degree	24	40	93	31

<sup>a</sup> AM1/d values are calculated with fitted magnesium *spd* parameters, AM1' with fitted *sp* parameters.

ment reaction energies would not be calculated as the difference between infinitely separated reactants and products but rather would include electrostatic interaction of the reacting partners in both reactant and product states. For this type of reaction, the semiempirical methods show the largest difference from the DFT reference. MNDO/d is closer than standard AM1 but still deviates 10–20 kcal/mol from the DFT reaction energies. Fitting of the *sp* parameters brings the AM1' values close to those of MNDO/d and is even better in two cases. With inclusion of *d* parameters, however, the reaction energy errors are significantly reduced, to at most 9 kcal/mol.

Changes in the coordination number with conservation of the charge are reproduced better by MNDO/d than standard AM1. Optimized *sp* parameters (AM1') do not significantly improve the results. An extension to *d* functions is clearly necessary for a proper energetic description of reactions with hypervalent magnesium compounds.

For those reactions in which neither the coordination number nor the charge of the magnesium complex change, all semiempirical methods perform quite well.

The average absolute error of all reactions evaluated is 5 kcal/mol for AM1/d and is significantly lower than those for AM1', AM1, and MNDO/d, see Table 3. The larger and thus more flexible *spd* basis clearly provides a more balanced description of the different types of ligand exchange reaction.

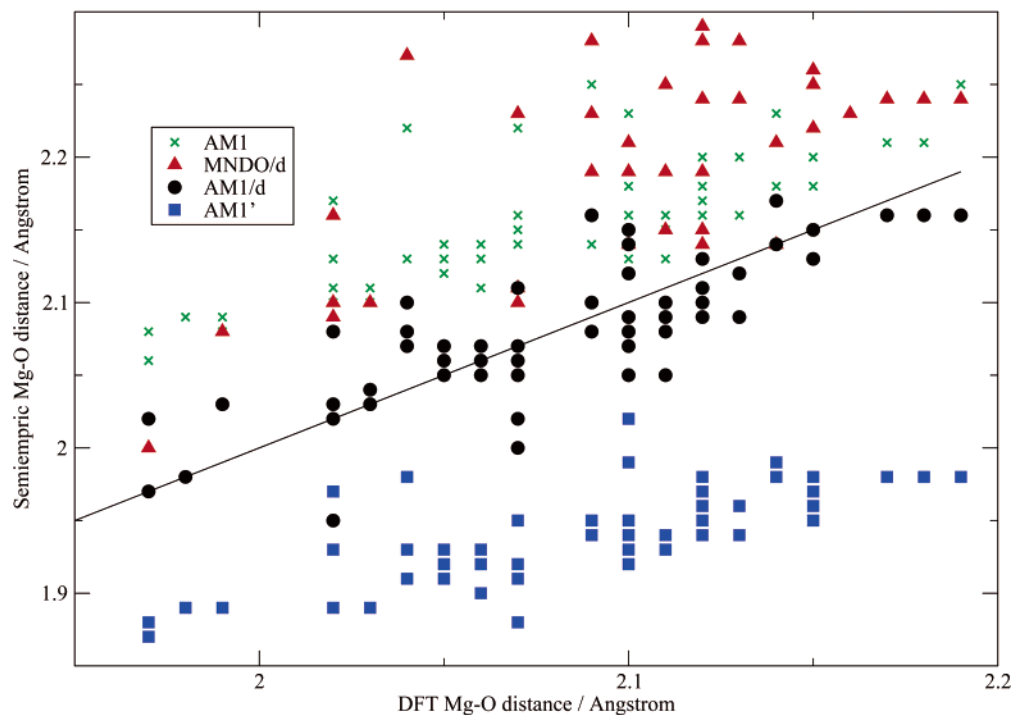
**Geometries.** Mg–O distances in magnesium compounds calculated with AM1/d deviate by at most 0.07 Å, i.e. 3%, from the DFT values to larger and smaller distances, the mean absolute error being 0.02 Å. With both MNDO/d and standard AM1 the distances are too long, by 0.07 Å on average (see Table 3). AM1' with fitted *sp* parameters strongly underestimates the Mg–O bond lengths, which are uniformly shifted by –0.15 Å relative to the AM1/d bond lengths. The improvement in bond distances by AM1/d is due to the fitting procedure, in which specific parameters have been derived for a class of compounds in which the magnesium atom is directly bound only to oxygen atoms. Standard parameters derived for more general applicability must simultaneously represent other bond types such as Mg–C, Mg–H, or Mg–X (X = halogen), which is a more difficult task. In MNDO/d a partial tuning is achieved by interaction specific core parameters  $\alpha$  for Mg–H, Mg–C, and Mg–S.<sup>20</sup>

As shown in Figure 4 O–Mg–O bond angles range from about 60° (at the bidentate acetate ligand) to linear (180°). The standard AM1 values strongly deviate from the angles calculated with DFT (mean absolute error: 15°, see Table

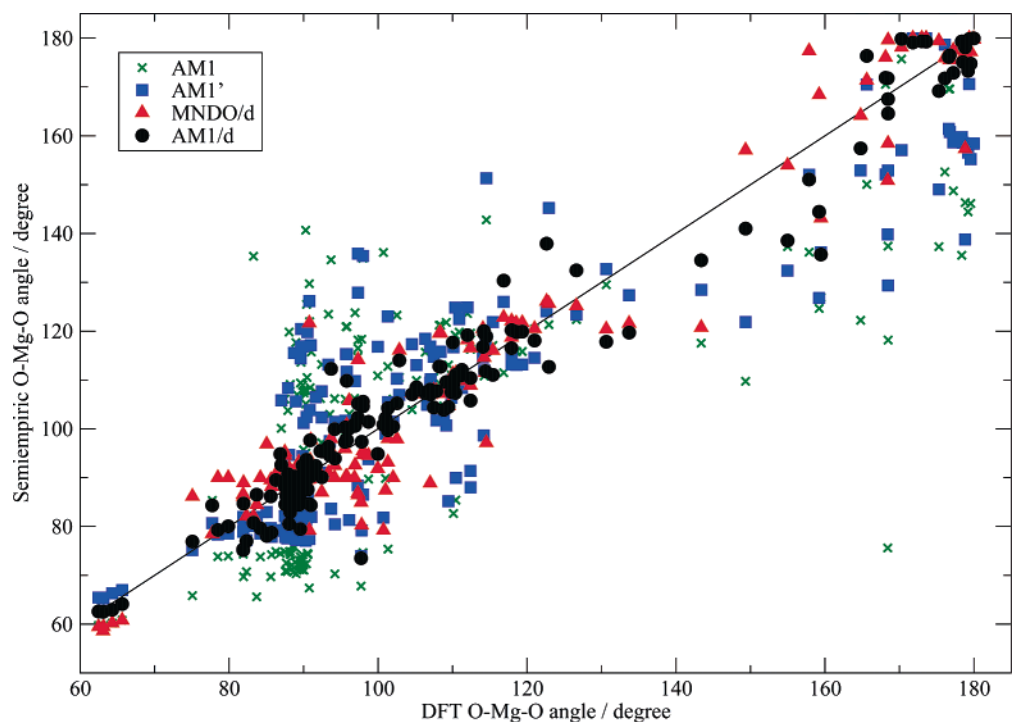
3) and cannot be improved significantly by fitting the *sp* parameters. MNDO/d shows an average error in bond angles of only 4°, the same as is achieved with fitted *spd* AM1 parameters. The maximum AM1/d error for O–Mg–O angles is 24° compared to 31° calculated for MNDO/d. For both methods, the largest angle errors can be attributed to errors in the treatment of intramolecular hydrogen bonds, rather than inaccuracies in the magnesium parameters: a too weak O<sub>a</sub>–H···O<sub>b</sub> interaction leads to a too large O<sub>a</sub>–Mg–O<sub>b</sub> angle. This effect is particularly pronounced for those complexes including acetate. The structure of [Mg(H<sub>2</sub>O)<sub>3</sub>-ac<sub>2</sub>] is the worst case in this regard: the hydrogen atoms point in different directions compared to the DFT optimized structure, leading to bifurcated hydrogen bonds, to which AM1 is known to be prone.<sup>9</sup> Interestingly, the bite angle of the acetate ligand (ca. 119°, which is not included in the training set properties) is also reproduced best for all complexes with AM1/d. However, the improvement on standard AM1 is marginal. This may be attributed to the fact that the use of standard parameters for first row elements (H, C, O) leads to well-reproduced O–C–O angle values. However, this agreement shows that the presented optimized magnesium parameters indeed work in concert with these standard parameters and lead to an overall improvement. As an additional test, we combined the AM1/d parameters for phosphorus from ref 14 with our AM1/d parameters for magnesium (and standard parameters for H, C, and O) and evaluated the reaction of pentaquomagnesium dimethyl phosphate with water to pentaquomagnesium methyl phosphate plus methanol (for structures see Figure 1 in the Supporting Information). The hydrolysis of dimethyl phosphate has been taken into account in the parametrization for phosphorus in ref 14. The Mg–O and P–O distances agree on average within 0.02 Å with the DFT-optimized distances (the maximal error is 0.07 Å in 20 distances), and the O–Mg–O and O–P–O angles differ on average 7° (42 angles). The largest geometric differences from DFT optimized values for the magnesium phosphates is 26° for one O–Mg–O angle. The AM1/d calculated reaction energy of –4 kcal/mol agrees well with the DFT value of 1 kcal/mol. This shows that a combination of specific AM1/d parameters for phosphorus and magnesium can be combined together and with standard (C, H, O) AM1 parameters to give sufficiently reliable results.

## 4. Conclusions

The present paper presents the results of the development of AM1/d parameters for magnesium. These parameters



**Figure 3.** Semiempirical vs DFT Mg–O bond distances. AM1/d values are calculated with fitted magnesium *spd* parameters, AM1' with fitted *sp* parameters.



**Figure 4.** Semiempirical vs DFT O–Mg–O angles. AM1/d values are calculated with fitted magnesium *spd* parameters, AM1' with fitted *sp* parameters.

provide a significantly improved description of biologically important magnesium complex geometries and reaction energies on a semiempirical level relative to standard semiempirical methods. Attempts to fit AM1 parameters for an *sp* basis are of limited success, showing, that for a proper semiempirical description of hypervalent compounds, the extension of the basis to *d* orbitals is necessary.

For the systems investigated in this work MNDO/d turns out to be superior to standard AM1. The quality of the

specifically parameterized AM1/d results, however, is clearly superior to that of the standard methods. This shows that the effort of developing reaction- or system-specific parameters is worthwhile when high accuracy is desired, rather than covering a large variety of compounds.

Remaining deviations from the DFT values can be traced back to the underestimation of hydrogen-bond strengths on the AM1 level. The compounds used in the magnesium training set cover a variety of possible coordination spheres

for biological magnesium and may thus be used in applications to a broad range of magnesium-containing proteins. They also work well for magnesium phosphates, when combined with the phosphorus parameters reported in ref 14. However, to cover all possible magnesium-coordination partners in proteins, the parametrization has to be extended to include Mg–N bonds such as magnesium–histidine complexes. This is subject of ongoing work.

Particularly when used in combined QM/MM calculations the new AM1/d parameters reported here furnish a method for modeling magnesium-containing biological systems with reasonable accuracy at low computational cost.

**Acknowledgment.** We thank Walter Thiel for support with a prerelease version of the MNDO program. This work was funded by the Deutsche Forschungsgemeinschaft as part of the SFB 623.

**Supporting Information Available:** Cartesian coordinates of the B3LYP/6-31++G(d,p) optimized structures of the training set and Mg–O distances and O–Mg–O angles derived therefrom as well as reaction energies and total energies. This material is available free of charge via the Internet at <http://pubs.acs.org>.

### References

- (1) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505.
- (2) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425–443.
- (3) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523–2544.
- (4) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.
- (5) Thiel, W.; Voityuk, A. A. *Adv. Chem. Phys.* **1996**, *93*, 703–757.
- (6) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (7) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (8) Bredow, T.; Jug, K. *Theor. Chim. Acta* **2005**, *113*, 1–14.
- (9) Dannenberg, J. J.; Evleth, E. M. *Int. J. Quantum Chem.* **1992**, *44*, 869–885.
- (10) Hutter, M. C.; Reimers, J. R.; Hush, N. S. *J. Phys. Chem. B* **1998**, *102*, 8080–8090.
- (11) Thiel, W.; Voityuk, A. A. *Int. J. Quantum Chem.* **1992**, *44*, 807–829.
- (12) Dudev, T.; Cowan, J. A.; Lim, C. *J. Am. Chem. Soc.* **1999**, *121*, 7665–7673.
- (13) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
- (14) Lopez, X.; York, D. M. *Theor. Chim. Acta* **2003**, *109*, 149–159.
- (15) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040–1046.
- (16) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (17) Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H.; Kolmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- (18) Thiel, W. *MNDO Version 6.1*; Max-Planck-Institut fuer Kohlenforschung, Mülheim a.d. Ruhr, Germany, 2004.
- (19) Thiel, W.; Voityuk, A. A. *Theor. Chim. Acta* **1992**, *81*, 391–404.
- (20) Thiel, W.; Voityuk, A. A. *J. Phys. Chem.* **1996**, *100*, 616–626.
- (21) Mayaan, E.; Range, K.; York, D. M. *J. Biol. Inorg. Chem.* **2004**, *9*, 807–817.

CT600092C