

29 June 2001

Chemical Physics Letters 341 (2001) 625-632



www.elsevier.nl/locate/cplett

Modeling physisorption with the ONIOM method: the case of NH₃ at the isolated hydroxyl group of the silica surface

I. Roggero, B. Civalleri, P. Ugliengo *

Dipartimento di Chimica IFM, Universitá di Torino, Via P. Giuria 7, 10125 Torino, Italy Received 2 March 2001; in final form 24 April 2001

Abstract

The ONIOM method has been used to model the isolated hydroxyl group at the silica surface using the octahydrosilasesquioxane cluster. Three different bi-layer clusters have been defined, using five levels of theory for the real system. For the model system, the B3-LYP/DZP level has been adopted for all calculations. Geometries, OH vibrational stretching frequency, NMR isotropic $\delta(^{1}H)$ and $\delta(^{29}Si)$ chemical shifts and the binding energies of NH₃ computed with the ONIOM approach have been compared with the full ab initio results. Whereas the ONIOM(B3-LYP/ DZP:HF/3-21G) level gives the best results, the cheapest ONIOM(B3-LYP/DZP:AM1) level can be adopted to model very large portions of siliceous materials. © 2001 Elsevier Science B.V. All rights reserved.

1. Introduction

The applications of computational chemistry to real world chemistry span from predicting the structure, spectra (either IR, NMR or UV–Vis) and reactivity of complicated molecules to understand the catalytic property of a given site in either a biological environment or in a crystalline nanoporous material as in silica-based zeolites. To serve as a predictive tool, however, the methods should be applicable to a large enough portion of the system, in such a way that the computed properties do reflect the features of the real system. For instance, one cannot restrict the study of Brönsted sites in H-exchanged zeolites by adopting too small clusters, because the structural and electrical

* Corresponding author. Fax: +39-011-670-7855.

memory imposed to the site by the infinite framework is simply not included in the model cluster.

This fact implies that only very approximate methods may be used to model such large systems to the detriment of accuracy. Indeed, accurate ab initio methods grow much faster than the number of particles, the growth being roughly proportional to the third power of the number of atomic basis functions used to solve the Schrödinger equation, at least within the DFT context.

However, it has soon been recognized that many important reactive chemical steps have a very local nature, i.e., they involve pairs of electrons in a limited number of bonds close to the reaction center. This opens the possibility for treating the local and reactive zone with a high level ab initio technique, whereas the atoms and bonds of the farther region in which the reactive zone is embedded may be treated at a (much) lower level.

For this purpose, a number of methods have been put forward in the literature with the

E-mail address: piero.ugliengo@unito.it (P. Ugliengo).

^{0009-2614/01/\$ -} see front matter © 2001 Elsevier Science B.V. All rights reserved. PII: S 0 0 0 9 - 2 6 1 4 (0 1) 0 0 5 3 0 - 9

corresponding computer code implementations. The interested reader may consider reference [1] for an overview of the modern proposals devoted to tackle large molecular systems with quantum mechanical methods.

Among them, the ONIOM method, originally proposed by Morokuma and co-workers [2-5], has assumed a prominent rôle in this respect, because it has been implemented in the very popular GAUSSIAN 98 computer code [6]. Whereas a number of ONIOM applications in the molecular realm have already been reported by the same Morokuma group (see [7] for a recent discussion), the applicability of the same methodology to study adsorption processes occurring at the surface of oxidic systems and inside the zeolite channels has been lacking. For the latter case, a procedure close to that of Morokuma has been independently proposed and successfully applied by Sauer and co-workers in their QM-Pot code [8], which nicely joins the quantum mechanical solution of the reactive zone with a molecular mechanics description of the infinite zeolite framework.

In the present note, the application of the ON-IOM method to model the adsorption of NH_3 at the isolated hydroxyl groups on a highly dehydrated silica surface is discussed. In a forthcoming paper the same technique will be discussed in the modeling of the structural, energetic and spectroscopic features of the Brönsted sites in H-Faujasite [9].

Since only methodological aspects of the ON-IOM method are studied in this note, no discussion with respect to either experimental data or absolute accuracy of the chosen levels of calculus is addressed. On the contrary, the purpose is to establish an acceptable level of theory within the ONIOM framework, which will allow to treat larger portions of siliceous materials, both free and in interaction with molecules.

2. Computational methods

The ONIOM2 approach [3–5] subdivides a molecule into two parts or layers, each of which is described at a different level of theory. The most important part of the whole molecule forms the innermost layer, the *model system*, which is

described at the highest level of theory. The model system usually includes the most chemically relevant portion of the entire molecule and the values of the computed properties become a function of the size and topology of the model system. The whole molecule is called the *real system*. The cut between the model system and the rest of the molecule is attained by adding a set of fictitious atoms, usually hydrogen atoms, along the pre-existing chemical bonds between the model system and the exterior layer. In the ONIOM2 approach, the energy E[HIGH:Low] of such a molecule, subdivided into two layers, is approximated as:

E[High: Low] = E(High, Model) + E(Low, Real)- E(Low, Model)

in which E(HIGH, Model), E(Low, Real) and E(Low, Model) are the total energies computed on the model system with the highest level of calculus and on both the real and model systems with the lowest level of calculus, respectively. The ONIOM2 formulation ensures, by definition, that if the low level of calculus is increased to reach the highest adopted level or if the model system is enlarged to become the real system, one gets:

E[High: High] = E(High, Real),

for $Model \rightarrow Real$

or

for Low \rightarrow High.

It is worth noting that, because of the definition of the model and real systems, neither direct charge transfer nor polarization is possible between the exterior layer and the model system. Some indirect effects on the model system are due to the equilibration of energy gradients acting on the link atoms, which force the atoms of the model system to follow those of the exterior layer and vice versa.

The ONIOM scheme, as implemented in GAUSSIAN 98 [6], allows to use any combination of Hamiltonians, by mixing, for instance, post-Har-tree–Fock methods for the model system with molecular mechanics description of the real system.

In the present Letter, the hybrid B3-LYP functional, which combines the exchange func-

tional proposed by Becke [10], with that proposed by Lee et al. [11], for the correlation, has been adopted as the high level method using the DZP basis set [12] (six primitives to represent the polarization function).

Five different low level Hamiltonians have been combined with the B3-LYP/DZP method, namely three semiempirical (MNDO, AM1 and PM3) and two based on the Hartree–Fock approximation with STO-3G and 3-21G Gaussian basis sets as implemented in GAUSSIAN 98 [6].

These moderate levels of theory allow to apply the ONIOM2 method to relatively large portion of siliceous materials, a fact which is an absolute need for modeling the molecule/surface interaction.

Geometry optimization for all types of structures was carried out using analytical gradient techniques, imposing a C_s symmetry for either the bare and interacting clusters.

Harmonic normal-mode frequencies have been computed adopting analytical second energy derivatives and solving the equations of nuclear motions by standard methods.

Isotropic NMR chemical shielding was evaluated by means of the GIAO method [13,14] as implemented in GAUSSIAN 98. The reference chosen for both ¹H and ²⁹Si nuclei was tetramethylsilane computed at the full B3-LYP/DZP level, for which $\sigma_{iso}(H) = 31.72$ and $\sigma_{iso}(Si) =$ 403.27 ppm, respectively. For the evaluation of the corresponding ONIOM2 screening constant values, the same formula derived for the energy expression has been used:

 $\sigma_{\rm iso}[B3-LYP/DZP:Low]$

$$= \sigma_{
m iso}({
m B3-LYP}/{
m DZP}, {
m Model}) \ + \sigma_{
m iso}({
m Low}, {
m Real}) - \sigma_{
m iso}({
m Low}, {
m Model}).$$

For the semiempirical methods, no implementation of the GIAO method has been encoded in GAUS-SIAN 98, so that the σ_{iso} [B3-LYP/ DZP: SEMIEMPIRICAL] coincides with the σ_{iso} (B3-LYP/ DZP,Model) value computed for the sole model system.

The calculation of the binding energies, BE, resulting from the interaction between the cage silica models and the NH_3 molecule, may be

worked out using the original formulae of the total energy of each subunit, resulting in:

$$BE[B3-LYP/DZP : Low](C_x/NH_3)$$

$$= E[B3-LYP/DZP : Low](C_x)$$

$$+ E[B3-LYP/DZP](NH_3)$$

$$- E[B3-LYP/DZP : Low](C_x/NH_3)$$

in which C_x indicates the silica cage with different settings of the model system (see Section 3 for a deeper discussion), namely Cs, Cm and Cb for the small, medium and big model systems.

3. Molecular models

As formerly proposed by Sauer and Hill [15], a model of the class of cage-like hydridosilasesquioxanes carrying one hydroxyl group may be envisaged to represent the isolated hydroxyl group at the silica surface. The suitability of such clusters, and in particular of the H₇Si₈O₁₂(OH) molecule, has been proved by us, as far as the vibrational features [16], the isotropic NMR chemical shieldings [17] and the interaction with the NH₃ molecule [18] are concerned.

Fig. 1 shows the chosen model (hereafter indicated as 'C') both as free and in interaction with NH_3 (C/NH₃).

In order to model the feature of the hydroxyl group, the model systems defined within the ONIOM2 approach always include the OH group. For this purpose, three different model systems of increasing size, namely Cs, Cm and Cb, have been defined as shown in Fig. 1. The Cs, Cm and Cb models include a Si(OH), O₃Si(OH) and Si₃O₃Si(OH) portion of the entire cluster, respectively. In other words, the model system includes the first, second and third neighbors of the OH group. When the link atoms are added to the model system, the small model becomes the silanol molecule, the medium model is the orthosilicic acid whereas the big model is the first molecular tree-like cluster of formula $H_9O_3Si_4(OH)$.

For clusters interacting with NH_3 , the NH_3 molecule has always been included as part of the model system, as shown in Fig. 1.



Fig. 1. Balls and sticks representation of the siloxane cage, both alone (C) and interacting with NH_3 (C/NH₃). ONIOM2 layer models of the C cage interacting with NH_3 : small (Cs/NH₃), medium (Cm/NH₃) and big (Cb/NH₃). Free models not shown. Atoms belonging to the high level layer drawn as spheres.

4. Results and discussion

In the following sections the geometrical, vibrational, NMR features as well as the binding energies of both free and interacting clusters will be discussed. Because the present paper only deals with the internal consistency of the ONIOM method rather than making new predictions, the binding energies have not been corrected for the BSSE and no comparison with experimental data is addressed. For such details, the reader may refer to the recently published papers [16–18].

4.1. Features of the isolated silica cages

Table 1 shows the main geometrical features of the considered clusters, namely the Si–O bond length linking the hydroxyl group to the cage and the average value of the Si–O bond computed by averaging the Si–O values for bonds belonging to the cage itself.

Among the semiempirical methods adopted for the real system, MNDO behaves very satisfactorily, whereas PM3 and AM1 give too long values for the Si–O bond. This behavior is already shown

Si o bolid length of the terminal Si off group (A) and average ((Si of)) bolid length (A) for the free models						
Method	С	Cs	Cm	Cb		
B3-LYP/DZP	1.628 (1.641)					
B-LYP/DZP: MNDO	1.663 (1.631)	1.672 (1.636)	1.644 (1.634)	1.639 (1.644)		
B3-LYP/DZP: PM3	1.701 (1.659)	1.648 (1.659)	1.653 (1.658)	1.641 (1.654)		
B3-LYP/DZP: AM1	1.732 (1.709)	1.612 (1.706)	1.648 (1.702)	1.637 (1.691)		
B3-LYP/DZP:HF/STO-3G	1.648 (1.615)	1.608 (1.619)	1.627 (1.608)	1.628 (1.622)		
B3-LYP/DZP:HF/3-21G	1.629 (1.643)	1.614 (1.644)	1.631 (1.645)	1.628 (1.640)		

Table 1 Si–O bond length of the terminal Si–OH group (Å) and average ($\langle Si–O \rangle$) bond length (Å) for the free models

by the geometries computed with the pure semiempirical methods on the whole cluster C, as reported in Table 1. The [B3-LYP/DZP:HF/3-21G] level gives excellent agreement with the full B3-LYP/DZP level, already for the Cm model. For all considered methods, the Cs appears definitely too small.

The OH harmonic stretching frequency is shown in Table 2. As for the geometry, the [B3-LYP/DZP:MNDO] level gives excellent results already for the Cm medium model system. Lower performances are shown by both AM1 and PM3 levels. Besides the smallest Cs model system, the two ab inito HF low levels behave very satisfactory, giving quantitative agreement already with the medium Cm model system.

It is worth noting that, when the cluster C is treated with the low level only (either semiempir-

ical or ab initio), the OH stretching values span a much larger region in comparison with the ON-IOM data.

A more stringent test is addressed in Table 3, where the isotropic NMR chemical shifts of the proton belonging to the OH group and those of the topmost silicon atom of the cage have been collected. It is worth noting that the Si chemical shift is a sensitive test of the chemical perturbations due to the definition of the model system. As explained in Section 2, data computed with the semiempirical level of calculus are less representative, because in these cases, only the values computed at B3-LYP/DZP for the model system are available. For all considered cases, results for the Cs cluster are erratic, with the $\delta(^{1}H)$ showing in some cases the wrong sign whereas the $\delta(^{29}Si)$ is severely underestimated. The situation improves

Table 2 Harmonic OH frequency $\omega_{\rm b}(\rm cm^{-1})$ for the free models

Method	С	Cs	Cm	Cb	
B3-LYP/DZP	3925				
B3-LYP/DZP: MNDO	4151	3917	3929	3926	
B3-LYP/DZP: PM3	3924	3926	3937	3928	
B3-LYP/DZP: AM1	3601	3972	3949	3934	
B3-LYP/DZP:HF/STO-3G	4299	4030	3924	3928	
B3-LYP/DZP:HF/3-21G	4040	3947	3925	3930	

Table 3

Chemical shifts $\delta({}^{1}H)$ and $[-\delta({}^{29}Si)]$ (ppm) for the free model

Method	С	Cs	Cm	Cb
B3-LYP/DZP	1.9 [105.9]			
B3-LYP/DZP: MNDO		-0.2 [30.7]	1.2 [121.7]	1.3 [97.7]
B3-LYP/DZP: PM3		-0.5 [25.2]	1.3 [113.8]	1.5 [98.9]
B3-LYP/DZP: AM1		0.3 [37.7]	1.3 [124.7]	1.4 [99.7]
B3-LYP/DZP:HF/STO-3G	-2.0 [3.0]	2.4 [-13.3]	1.0 [97.7]	1.9 [96.7]
B3-LYP/DZP:HF/3-21G	3.0 [91.7]	2.4 [127.7]	2.0 [117.7]	1.8 [101.1]

Data relative to B3-LYP/DZP shield constants for tetrametilsilane: $\delta(^{1}H) = 31.7 - \sigma(^{1}H), \ \delta(^{29}Si) = 403.3 - \sigma(^{29}Si).$

slightly with the Cm cluster, in which the $\delta(^{29}\text{Si})$ values systematically overshoot the reference B3-LYP/DZP value. For the largest Cb model system, the $\delta(^{1}\text{H})$ becomes very close to the reference ab inito value, whereas the value of the $\delta(^{29}\text{Si})$ is still underestimated. Calculations using the lower level only (namely HF/STO-3G and HF/3-21G) give values of the chemical shifts much worse than those computed adopting the ONIOM recipe.

This trend could be interpreted in terms of the electronic effects acting on the SiOH moiety. As stated above, in the ONIOM method neither charge transfer nor polarization is possible between the two layers. Thus, in order to recover the correct electronic environment, the model system must be enlarged up to including at least the third neighbors of the examined nucleus, that is, the Cm for the H nucleus and the Cb for the Si nucleus, in which the model system is, respectively, orthosilicic acid and trisiloxysilanol molecule.

As for the OH frequency, data at [B3-LYP/ DZP:HF/3-21G] are in excellent agreement with the reference B3-LYP/DZP ones.

4.2. Features of the cages interacting with NH₃

Table 4 shows the intermolecular $N \cdots H$ distance (see Fig. 1 for details), as well as the binding energy of the NH₃ with respect to the separated components. As expected from the accumulated experience with the semiempirical methods in dealing with intermolecular interactions [19], both MNDO and PM3 overestimate the $N \cdots H$ distance by almost 0.1 Å with respect to the reference B3-LYP/DZP value, whereas AM1 slightly underestimates the $N \cdots H$ distance by about 0.05 Å. When the HF level is adopted for the real system, the agreement with the reference value is excellent, irrespective of the basis set, at least when data for the smallest Cs cluster are excluded.

The binding energies closely match the behavior described by the $N \cdots H$ distance: with all semiempirical methods, a too low binding energy is computed, particularly for PM3, which is the worst among all of them. On the other hand, AM1 recovers almost 90% of the binding energy for the largest Cb cluster. The binding energy as a function of the size of the model cluster oscillates in all cases, with the minimum value for the Cm cluster. As already found for the other quantities, the [B3-LYP/DZP:HF/3-21G] level for the Cb cluster gives a binding energy in excellent agreement with the reference B3-LYP/DZP datum.

Values computed using the low level only for the whole C cluster show an erratic behavior, particularly for all semiempirical methods, for the reasons already mentioned. The pure HF/3-21G level gives a too short $N \cdots H$ distance corresponding to an overestimated value of the binding energy.

Table 5 shows the separate contributions to the ONIOM binding energy for the B3-LYP/ DZP:HF/3-21G case: the BE₁(Low, Real) term is obviously almost constant for the three considered layers, whereas both BE₂(HIGH, Model) and BE₃(Low, Model) values change as a function of the Model size and topology. The correction term BE_C is, however, remarkably constant along the adopted Model regions, and imparts the same numerical stability to the final value of the ON-IOM binding energy.

Table 6 shows the values of the OH frequency shift caused by the interaction with the NH_3 molecule. The OH frequency in the complex is shifted to low wave numbers compared to unperturbed values, a well-known fact for hydrogen bonded complexes, in which the larger the shift,

Table 4

Intermolecular $N \cdots H$ distance (Å) and binding energies (BE) (kJ/mol) for the models in interaction with NH_3

Method	C/NH ₃	Cs/NH ₃	Cm/NH ₃	Cb/NH ₃
B3-LYP/DZP	1.75 (52.2)			
B3-LYP/DZP: MNDO	4.42 (1.3)	1.83 (39.1)	1.85 (34.6)	1.80 (42.0)
B3-LYP/DZP: PM3	1.87 (-6.1)	1.91 (30.4)	1.87 (31.6)	1.82 (42.8)
B3-LYP/DZP: AM1	2.59 (6.2)	1.71 (46.4)	1.70 (36.8)	1.70 (46.3)
B3-LYP/DZP:HF/STO-3G	1.83 (34.0)	1.66 (59.9)	1.76 (51.5)	1.73 (54.9)
B3-LYP/DZP:HF/3-21G	1.69 (70.0)	1.71 (55.9)	1.74 (51.5)	1.74 (52.0)

Simponents of the ONIOW binding energy BE[High . Low] (kJ/mol) for the B5-L1F/DZF . III/5-210 level				
BE components	Cs/NH ₃	Cm/NH ₃	Cb/NH_3	
BE ₁ (Low, Real)	65.5	64.7	64.5	
BE ₂ (High, Model)	44.5	31.2	47.7	
$BE_3(Low, Model)$	54.2	44.3	60.2	
$BE_C = BE_2 - BE_3$	-9.7	-13.1	-12.5	
$BE[HIGH : Low] = BE_1 + BE_C$	55.9	51.5	52.0	

Table 5 Components of the ONIOM binding energy BE[HIGH : Low] (kJ/mol) for the B3-LYP/DZP : HF/3-21G level

Table 6

Harmonic OH frequency shifts, $\Delta \omega_h = \omega_h (C_x / \text{NH}_3) - \omega_h (C_x) (\text{cm}^{-1})$ for the models in interaction with NH₃

Method	C/NH ₃	Cs/NH ₃	Cm/NH ₃	Cb/NH ₃
B3-LYP/DZP	-679			
B3-LYP/DZP: MNDO	-5	-488	-458	-564
B3-LYP/DZP: PM3	-189	-370	-403	-543
B3-LYP/DZP: AM1	-32	-620	-476	-595
B3-LYP/DZP:HF/STO-3G	-99	-937	-661	-751
B3-LYP/DZP:HF/3-21G	-765	-793	-702	-701

the larger the value of the corresponding binding energy. This fact is well accounted for by the ONIOM results, as shown in Fig. 2, in which a remarkable correlation exists between BE and $-\Delta \omega_h(OH)$, for all cases and within a given



Fig. 2. Correlation between the redshift, $-\Delta \omega_h(OH)$, and the binding energy, BE, for the considered clusters. Arrows with the 'b' label mark the results computed for the largest model system, Cb/NH₃, for each low level of theory. The grayed region marks the points closest to the reference ab initio data.

method. Data in Table 6 and close inspection of Fig. 2 show that the AM1 level for the real system is the best among the semiempirical methods, the worst being the PM3 level, whereas the [B3-LYP/ DZP:HF/3-21G] gives values of the frequency shift in remarkable agreement with the reference value, already with the Cm cluster. As expected, results derive from a compensation between the accuracy due to the combination of high and low levels of theory and the size of the adopted model systems. In this respect, as already found for the other observables, the OH values computed adopting the lower level only on the whole C cluster show an erratic behavior, the semiempirical results showing the largest variations.

5. Conclusions

The ONIOM2 approach has been applied to study the properties of a cluster model mimicking the isolated hydroxyl group of the silica surface, both free and when interacting via H-bond with NH₃. Among the three semiempirical levels (MNDO, AM1 and PM3) adopted for the treatment of the real system, the AM1 method is the level of choice. Results obtained at [B3-LYP/ DZP: AM1] level can be used for semiquantitative prediction on larger portion of siliceous materials, providing that a large enough number of chemical relevant atoms are included in the model system.

If a more refined approach is needed, the [B3-LYP/DZP:HF/3-21G] level of treatment can provide data in excellent agreement with the results computed at the full B3-LYP/DZP level, at a fraction of computational cost. In that respect, the [B3-LYP/DZP:HF/3-21G] calculation of the NMR shielding constants is about seven times faster than the full B3-LYP/DZP calculation and uses one third of the needed disk space. The speedup is quite sizeable and encourages the adoption of the ONIOM approach for the treatment of very large systems which are, at present, too demanding for a full ab initio treatment.

Acknowledgements

The present work is part of a project coordinated by A. Zecchina and co-financed by the Italian MURST (Cofin98, Area 03). Computer facilities from CINECA are kindly acknowledged.

References

 G. Frenking, Quantum chemical methods for large molecules, J. Comput. Chem. 21 (2000) (special issue).

- [2] F. Maseras, K. Morokuma, J. Comput. Chem. 16 (1995) 1170.
- [3] M. Svensson, S. Humbel, R.D.J. Froese, T. Matsubara, S.N. Sieber, K. Morokuma, J. Phys. Chem. 100 (1996) 19357.
- [4] S. Humbel, S. Sieber, K. Morokuma, J. Chem. Phys. 105 (1996) 1959.
- [5] S. Dapprich, I. Komáromi, K.S. Byun, K. Morokuma, M.J. Frisch, J. Mol. Struct. (Theochem) 461–462 (1999) 1.
- [6] M. J. Frisch, GAUSSIAN 98, Revision A.7, Gaussian Inc., Pittsburgh PA, 1998.
- [7] T. Vreven, K. Morokuma, J. Comput. Chem. 21 (2000) 1419.
- [8] J. Sauer, M. Sierka, J. Comput. Chem. 21 (2000) 1470.
- [9] I. Roggero, B. Civalleri, P. Ugliengo, in preparation.
- [10] A.D. Becke, J. Chem. Phys. 98 (1993) 5648.
- [11] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B 37 (1988) 785.
- [12] T.H. Dunning Jr., J. Chem. Phys. 90 (1989) 1007.
- [13] J.R. Cheeseman, G.W. Trucks, T.A. Keith, M.J. Frisch, J. Chem. Phys. 104 (1996) 5497.
- [14] P.B. Karadakov, K. Morokuma, Chem. Phys. Lett. 317 (2000) 589.
- [15] J. Sauer, J.-R. Hill, Chem. Phys. Lett. 218 (1994) 333.
- [16] B. Civalleri, E. Garrone, P. Ugliengo, Chem. Phys. Lett. 294 (1998) 103.
- [17] B. Civalleri, E. Garrone, P. Ugliengo, Chem. Phys. Lett. 299 (1999) 443.
- [18] B. Civalleri, E. Garrone, P. Ugliengo, Langmuir 15 (1999) 5829.
- [19] J.J.P. Stewart, in: P. von Ragué Schleyer (Eds.), Encyclopedia of Computational Chemistry, vol. 2, Wiley, New York, 1998, p. 1283.