COMPARATIVE STUDY OF THE ELECTROSTATIC POTENTIAL OBTAINED WITH VARIOUS ATOMIC BASIS SETS FOR THE CLUSTER MODEL OF THE TETRAHEDRAL SHEET [Si₂O₅H₂]∞.

Daniel Tunega, Ladislav Turi Nagy and Štefan Varga

Institute of Inorganic Chemistry, Slovak Academy of Sciences, Dúbravská 9, 842 36 Bratislava, Slovak Republik

Semiempirical (INDO/2) and ab initio (STO-3G, 6-31G and 6-31G* basis sets) quantum chemical calculations were done on the cluster model of the tetrahedral silicate sheet [Si₂O₅H₂]∞. The quality and reliability of electrostatic properties (charges from Mulliken population analysis and the electrostatic potential above the basal oxygen atom) from these calculations are compared. The shape of the semiempirical INDO/2 electrostatic potential for the tetrahedral silicate sheet is also compared with our previous results for a talc layer.

INTRODUCTION

Intermolecular interactions play an important role in various physical and chemical processes (adsorption, diffusion, solvatation, nucleation, etc.). Interaction potenials are often used as input data in statistical thermodynamic treatments and in evaluation of thermodynamic functions and dynamic properties of interacting systems [1]. Hence, for explantaion of the mentioned processes, it is important to have these interactions well described.

An intermolecular interaction energy is usually partitioned into the sum of several terms

$$E_{\text{int}} = E_c + E_{\text{pol}} + E_D + E_R + E_{\text{Ch}-T} \tag{1}$$

where the first two terms (E_c – coulombic and E_{pol} – polarization energy) are known to be of long-range, while E_D, E_R, $E_{\text{Ch}-T}$ terms represent short-range forces rapidly vanishing with increasing distance and correspond to dispersion, exchange repulsion, and charge-transfer energy, respectively.

The coulombic contribution is (due to its long-range character) an important and in some cases (e. g. highly ionic systems) also the dominant part of the intermolecular interaction energy (1). Thus, the importance of having a reliable and well described this contribution is obvious.

This part of the interaction energy can be evaluated with using an electrostatic potential $V(\rho(\mathbf{r}))$ that is generated around each interacting charge distribution with a correspondig charge density $\rho(\mathbf{r})$.

At some point $\mathbf{r}_i$ this potential is given as

$$V(\mathbf{r}_i) = \int |\mathbf{r}_i - \mathbf{r}|^{-1} \rho(\mathbf{r}) \, d^3\mathbf{r} \tag{2}$$

The main problem in solving of equation (2) is the determination of the charge distributions. Simple approximation to calculate the potential $V(\mathbf{r})$ is to assume point charges located at the centres of atoms,

hence reducing the integration to the sum of individual contributions of these point charges. Of course, such calculated electrostatic potentials strongly depend on the values of the pertinent point charges. Those can be determined in various ways, leading often to ambiguous results. Moreover, the description of the electrostatic potential near the atomic centers is wrong within the point charge approach [2]. Another way of the description of the electrostatic interaction is to express it via the multipole expansion. The latter leads to better results but calculations of this kind are more difficult and are often accompanied with some numerical problems [2].

In a quantum mechanical approach within the Born-Oppenheimer approximation the equation (2) can be rewritten into

$$V(\mathbf{r}_i) = \sum \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} - \int \frac{\rho(\mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|} d^3\mathbf{r} \tag{3}$$

where $V(\mathbf{r}_i)$ is the rigorously defined expectation value of the one electron operator $1/\mathbf{r}$. The first term represents electrostatic contributions from the nuclei of atoms located at positions R_A, and Z_A is the nuclear charge of the atom A. The second term corresponds to the electrostatic term generated by the electron charge distribution; the integration over the whole space is understood.

Within the MO-SCF LCAO approximation eq. (3) becomes

$$V(\mathbf{r}_i) = \sum \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} - \tag{4}$$

$$- \sum \sum P_{\mu\nu} \int \frac{\chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r})}{|\mathbf{r}_i - \mathbf{r}|} d^3\mathbf{r}$$

where $\{\chi\}$ are the basis functions (atomic orbitals) and $P_{\mu\nu}$ is the matrix element of the first-order one-electron density matrix P and is obtained as

$$P_{\mu\nu} = 2 \sum_i c^*_{\mu i} c_{\mu i} \tag{5}$$

where the summation is over all orbitals i occupied in the electronic state considered and c_i are vectors of appropriate linear coefficients reffering to a chosen basis sets $\{\chi\}$.

It is evident from eq. (4) and (5) that the quality and reliability of the electrostatic potential $V(r_i)$ depends on the density matrix P and on the quality of basis set $\{\chi\}$. Comparison of the quality of electrostatic potentials obtained by many different quantum mechanical calculations for various sets of molecules has been a topic of several articles [3–14]. From these studies, the gross features of isopotential maps and shapes of potential hypersurface for each molecular system are well reproduced by any used quantum mechanical methods. Nevertheless, in details (positions and values of extrema) there are often great differences between the approaches. Some semiempirical methods (mainly CNDO and INDO) give considerable differences of calculated values contrary to the results from another semiempirical (e. g. MNDO or AM1) or ab initio methods [11–14]. This is caused by the fact that in the CNDO/INDO approximation atomic charge distributions tend to remain rather spherically averaged [14]. The MNDO/AM1 potentials have higher quality than CNDO/INDO ones and can be compared also in details with good ab initio potentials [13, 14]. It is important to remark that also in ab initio obtained potentials there are often considerable differences. Namely, results are strongly dependent on the quality of the used basis sets. It often happens that potentials obtained with minimal basis set and CNDO/INDO potentials have similar quality. Therefore, for obtaining reliable electrostatic potential by ab initio calculations it is unavoidable to take large basis sets with polarization and/or diffusion functions on all atoms [13, 14]. Similarly, values of the charges evaluated from ab initio calculations are also strongly depended on used basis sets. It is quite interesing, however, that charges from semiempirical calculations are often comparable with charges obtained using the modest basis sets in ab initio calculations.

In our previous studies [15, 16] we have presented the INDO/2 electrostatic potentials calculated above the surface of the isolated layer and above the surface of the crystal of the phyllosilicate mineral talc. However, as semiempirical results are often looked at with little confidence, we decided to perform a comparative study of both semiempirical and ab initio methods from the point of view of electrostatic potential calculations on a structural fragment of a silicate layer of a size manageable also by ab initio method. The quality of potentials obtained from INDO and Hartree-Fock ab initio calculations for a model silicate structure is compared in the present paper. The effective atomic charges from Mulliken population analysis are also compared and a correlation between their values and the values of the electrostatic potential minimum is shown. Finally, the INDO/2 potentials for the cluster of the isolated idealized talc layer [15] and for the tetrahedral silicate sheet are compared.

METHOD

The electrostatic properties (potenital and atomic charges) for the cluster modelling the tetrahedral silicate sheet $[Si_2O_5H_2]\infty$ were calculated at the following levels: 1) semiempirical INDO/2, 2) ab initio STO-3G minimal basis set, 3) ab initio 6-31G basis set, 4) ab initio 6-31G* basis set (polarisation functions on all atoms). While in semiempirical methods atomic orbitals are represented by Stater type functions (Slater type orbitals – STO), in ab inito method it si common that each atomic orbital is expanded as a sum over some set of primitive Gaussian functions (GF). The number of used Gaussian functions is usually included in the basis set description. Thus the STO-3G set is the minimal basis set with each atomic orbital (in this case STO) expanded by three GF. 6-31G basis set belongs to split-valence basis sets, where core atomic orbitals are expanded by one set of primitive GF containing six functions and valence orbitals are expanded by two sets of GF (the first set consists of 3 GF, the second set consists of 1 GF, respectively). As an ilustration, 6-31G basis set for silicon atom consists of the following atomic orbitals : a) core orbitals are 1s, 2s, 2p and each of them is described with six GF. b) split valence orbitals are 3s, 3p and each of them is described with two sets of GF (3,1). If, in addition, orbitals of higher angular momentum than required by the minimal basis set of a given atom are used (e. g. d orbitals for silicon atom) an asterix is added to description of basis set.

The cluster ($Si_6O_{18}H_{12}$) was chosen as the model of the infinite two-dimensional tetrahedral silicate sheet $[Si_2O_5H_2]\infty$. This sheet represents the basis structural framework of layer silicates, hence it is well suited also for a comparison with our previous calculations on the talc model. The size of the cluster (36 atoms) is a reasonable compromise between the reliability of the model and computational demands. It is represented by six ideal tetrahedra SiO_4^{-4} linked together over the bridging oxygen atoms, thus six membered ring with $\mathcal{C}_{6v}$ symmetry is formed (Fig. 1). The bridging oxygen atoms – O_b (also called basal oxygen atoms) lie in one plane and form the oxygen atom surface. On the opposite surface of the sheet there are hydroxyl groups (O_a-H) perpendicular to the surface. Dangling bonds on bridging oxygen atoms are saturated in the bond direction by hydrogen atoms (H_k) for keeping the charge neutrality of the cluster and for making the system to be closed shell. The bond length in tetrahedra d(Si-O) is 0.1619 nm and the
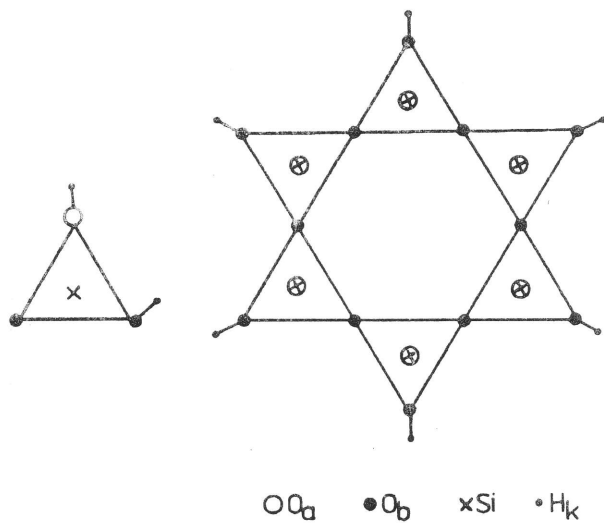
Fig. 1. 36 atomic cluster of the idealized tetrahedral silicate sheet (O_a a- apical oxygen atoms, O_b – basal (bridging) oxygen atoms, H_k – compensation hydrogen atoms).

bond length in the OH groups – d(OH) is 0.900 nm. No attempt to optimize the geometry of the cluster was made during any calculation.

The electrostatic potential for the cluster was calculated in the direction perpendicular to the surface of bridging oxygen atoms directly above one oxygen atom with the step of 0.01 nm. In the vicinity of the minimum of the potential curve the step was refined to 0.001 nm.

RESULTS AND DISCUSSION

In Fig. 2 the comparison of the electrostatic potential above the basal oxygen atom obtained with semiempirical INDO/2 method and *ab initio* method with mentioned basis sets is presented. It is evident that in gross features all the calculated potential curves are similar for the range of the calculations. Our potential curves can be compared with the potential above the basal oxygen atoms for the point charge model of the idealized talc layer presented in the work [17]. As expected, the potential from the point charge model shows large discrepancy mainly near the oxygen atom surface. For the point charge approximation no minimum on the potential curve is observed, thus the potential remains attractive along the whole distance from the oxygen atom. On the other hand, quantum mechanical calculations give a minimum on the potential curves (Fig. 1) and near to the surface the potentials become repulsive.

In the detailed study of the potential curves form (Fig. 2) one can see evident differences mainly in the values of the minimum V_{min} and slight shifts in its position z_{min} respective to the used method. The positions and the values of the potential minima for all

calculations are summarized in Tab. I. The tendency to a slight shift of z_{min} away from the surface with the increasing quality of the used basis set can be observed.

The differences in values of V_{min} are much more distinct. The semiempirical INDO/2 method gives the highest value of the minimum, while the lowest value of the minimum is obtained with the 6-31G basis set. Adding polarization functions on all atoms (6-31G* basis set) shifts the potential curve between the INDO/2 and 6-31G curves (see Tab. I and Fig. 2). As expected, all *ab initio* potential curves lie below the INDO/2 curve. This is caused by approximations used in semiempirical INDO/2 method. This result is in agreement with results in comparative studies [3–14]. On the other hand, *ab initio* calculations with the minimal basis set (STO-3G) and without the polarisation functions (6-31G) result in too deep electrostatic minimum (Fig. 2). Adding the polarization functions leads to more reliable and accurate results,
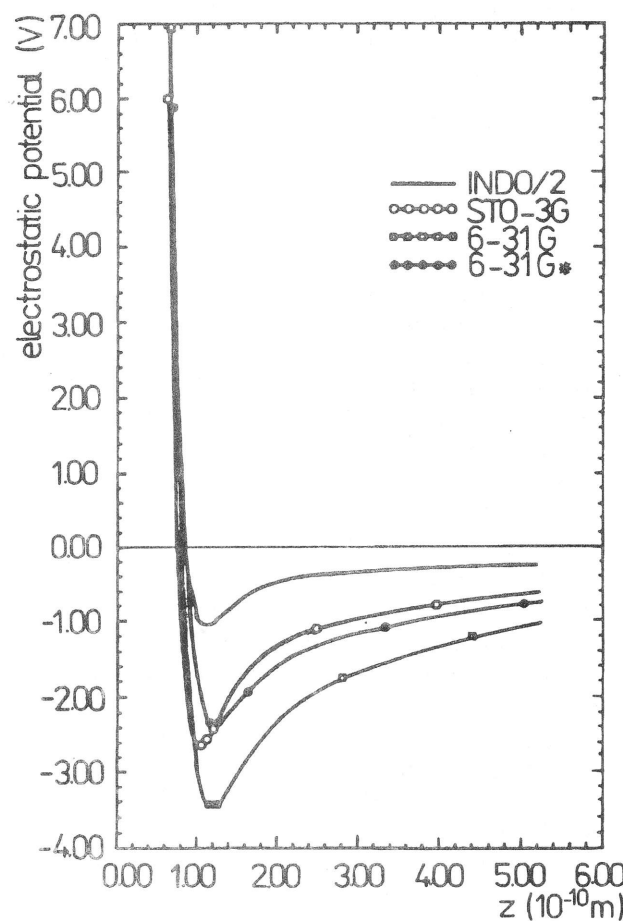

Fig. 2. The electrostatic potential curves above the basal oxygen atom of the cluster of the tetrahedral silicate sheet obtained with various basis sets.

Table I.

The positions and the values of the electrostatic potential minimum above the basal oxygen atom and the gross atomic charges in the cluster of the tetrahedral sheet.

Method	Z_{min}	V_{min}	gross atomic charges[+]		
	$[10^{-10}m]$	$[V]$	Si	O_b	O_a
INDO/2	1.11021	−1.06984	1.270	−0.610	−0.506
STO-3G	1.02661	−2.64089	1.409	−0.667	−0.643
6-31G	1.19594	−3.43399	2.275	−1.147	−1.076
6-31G*	1.22669	−2.34618	1.457	−0.729	−0.744

+ in the units of the electron charge

thus the 6-31G* electrostatic potential is believed to be our best approach.

Supposing the 6-31G* potential is the most accurate of our results, we will considered it for the purpose of this comparative study as the reference. Notice that the differences

$$|V_{min}^{INDO/2} - V_{min}^{6-31G^*}|$$

and

$$|V_{min}^{6-31G} - V_{min}^{6-31G^*}|$$

are almost equal. From this point of view the INDO/2 and 6-31G results are of comparable quality.

The similar comparison for atomic charges is even less favourable for *ab initio* results (see. Tab. I). There are dramatic differences in the obtained atomic charges depending on used basis set. For example for Si the difference in the atomic charge between 6-31G and 6-31G* is about 0.8 electrons. Incorporation of the polarization and diffuse functions causes spreading of the charge distribution farther out from the nuclei. For this reason not only the value of the minimum is changed but also the effective atomic charges are of lower value than those obtained with standard basis sets.

Comparison of the INDO/2 atomic charges with 6-31G* atomic charges (again considered as the reference state) evidently shows that the INDO/2 values are closer to the reference state than the 6-31G values. Therefore, in this case the discrepancy in the values of atomic charges from *ab initio* calculations with basis sets without polarization or diffuse functions is greater than for semiempirical calculations. The large values of the 6-31G atomic charges might indicate almost an ionic character of bonds in terahedral silicate

sheet, but it is well known that in silicate structures the bonds are covalent with a polar character. The 6-31G* values of the atomic charges evidently confirm this polar covalent character of bonds in the silicate sheet.

The obvious consequence of these results is that for obtaining more reliable and accurate electrostatic properties of silicates the inclusion of polarisation and/or diffuse functions into the *ab initio* basis sets of all atoms turns out to be unavoidable. Our results have shown, that semiempirically obtained values of electrostatic properties are almost on the same level of reliability as the *ab initio* results with basis sets without polarisation functions; for atomic charges the semiempirical results are even better. Moreover, semiempirical results were obtained more effectively and are less time consuming than the *ab initio* ones.

The semiempirical INDO/2 electrostatic potential above the cluster of the silicate sheet was also compared with the potential calculated in our previous paper [15] for the cluster of talc with idealized structure (details of the talc model are described in [15]). Both potentials are shown in Fig. 3. One can see that near the cluster surfaces both curves are almost identical up to the distance of about 0.25 nm. This is evident also from Tab. II, where the positions and the values of the potential minima for both systems are presented. Up to the above mentioned distance the electrostatic potential is fully governed by the charge distribution of the surface oxygen atoms. In the medium and large distance from the oxygen atoms surfaces the difference between the potential curves is increased, hence the influence of the octahedral sheet and the second tetrahedral sheet in the talc layer is felt. As can be seen from Fig. 3 the electrostatic po-
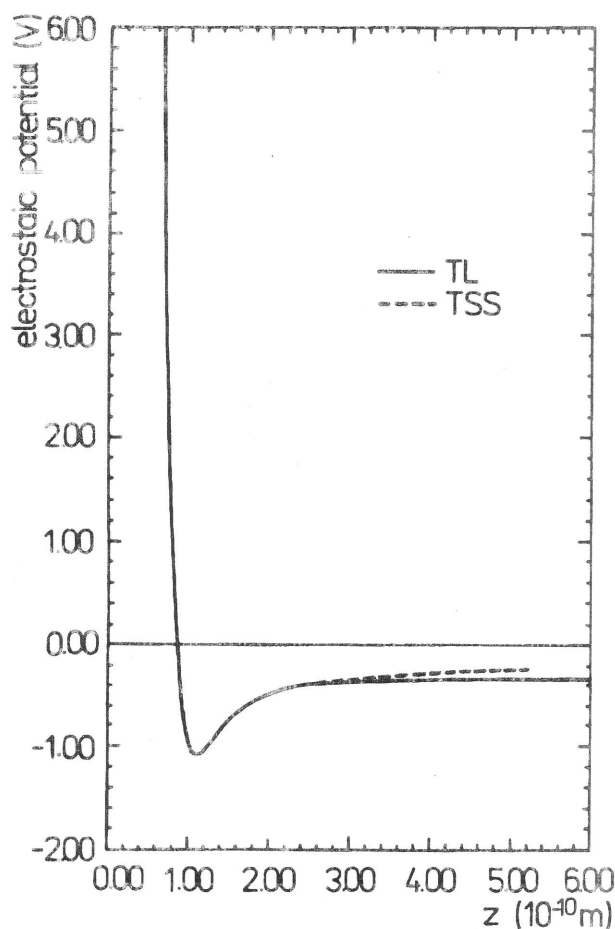
Fig. 3. The INDO/2 electrostatic potentials above the basal oxygen atoms of the cluster of the tetrahedral silicate sheet (TSS) and above the basal oxygen atom of the cluster of the talc layer (TL).

achieved only with using more complete basis sets with polarization and/or diffuse functions. On the other hand, using the well-parametrized semiempirical methods for these silicate systems can often lead to more effective and reliable calculations of the mentioned properties than using the minimal basis sets in *ab initio* calculations. Moreover, for such large systems as the silicates are, the use of large basis sets is usually computationally hard to manage.

The comparison of the electrostatic potentials above the basal oxygen atom for the silicate sheet and for the talc layer shows the dominant influence of the electron charge distribution of the oxygen atom on the potential near the surfaces of both clusters. The influence of the octahedral sheet and the second tetrahedral sheet in the talc layer on the potential becomes more evident in the medium and large distance from the layer surface. In these distances the electrostatic potential becomes the dominant part of the whole interaction potential.

Finally, we would like to remark, that for the correct description of the whole interacting potential in the region close to the surface the short-range terms $E_{D'}$, E_R (eq. 1) should be also considered.

tential for talc layer vanishes more slowly than for the tetrahedral sheet.

CONCLUSIONS

In the presented paper the electrostatic properties (Mulliken atomic charges, electrostatic potential) of the cluster of the tetrahedral sheet obtained with semiempirical INDO/2 method and *ab initio* method with several basis sets are compared. The gross features of the electrostatic potential above the surface of the basal oxygen atoms are preserved for all used methods and basis sets. The main differences are in the value of the potential minimum and in the values of the atomic charges. The discrepancies for the INDO/2 and 6-31G electrostatic potential minimum are of a comparable magnitude. For the atomic charges the INDO/2 results are even better than the 6-31G ones. Consequently, on the level of SCF *ab initio* calculations more reliable and accurate results of the electrostatic properties of silicate systems can be

References

[1.] Viegné-Maeder, F. and Auroux, A., J. Phys. Chem. *94*, 316 (1990)
[2.] Larson, E. C., Li M. and Larson, G. C., Int. J. Quant. Chem.: Quant. Chem. Symposium *26*, 181 (1992)
[3.] Giessner-Prettre, C. and Pullman, A., Theor. Chim. Acta *25*, 83 (1972)
[4.] Giessner-Prettre, C. and Pullman, A., Theor. Chim. Acta *33*, 91 (1974)
[5.] Giessner-Prettre, C. and Pullman, A., Theor. Chim. Acta *37*, 335 (1975)

[6.] Osman, R. and Weinstein, H., Chem. Phys. Letters *49*, 69 (1977)

[7.] Edwards, W. D. and Weinstein, H., Chem. Phys. Letters *56*, 582 (1978)

[8.] Petrongolo, C., Preston, H. J. T. and Kaufman, J. J., Int. J. Quant. Chem. *13*, 457 (1978)

[9.] Ramos, M J. and Webster, B., J. Chem. Soc. Faraday Trans. II *79*, 1389 (1983)

[10.] Webster, B., Hiczer, M., Ramos, M. J. and Carmichael, J., J. Chem. Soc. Faraday Trans. II *81*, 1761 (1985)

[11.] Oroczo, M. and Luque, F. J., Chem. Phys. Letters *160*, 305 (1989)

[12.] Luque, F. J. and Oroczo, M., Chem. Phys. Letters *169*, 269 (1990)

[13.] Luque, F. J., Illas F. and Oroczo, M., J. Comput. Chem. *11*, 416 (1990)

[14.] Alkorta, I., Villar, H. O. and Arteca, G. A., J. Comput. Chem. *14*, 530 (1993)

[15.] Tunega, D., Turi Nagy, L. and Varga, Š., Chem. Papers *47*, 273 (1993)

[16.] Tunega, D. and Turi Nagy, L., Chem. Phys. Miner., accepted (1993)

[17.] Bleam W.: Clays & Clay Miner. *38*, 522 (1990)

Submitted in English by the authors

márne náboje a priebeh elektrostatického potenciálu – získané *ab initio* výpočtami za použitia rôznych báz a tiež semiempirickou INDO/2 metódou. Všetky použité metódy reprodukujú hlavné črty elektrostatického potenciálu nad vrstvu bazálnych kyslíkových atómov. Hlavné rozdiely sa našli v hĺbke minima potenciálu a v hodnotách atomárnych nábojov. Spoľahlivé hodnoty nábojov v priebehu elektrostatického potenciálu pre silikátové systémy poskytujú na SCF úrovni *ab initio* výpočty v rozsiahlych bázach, zahŕňajúcich aj polarizačné funkcie. Odchýlky od hodnôt, poskytovaných takýmito výpočtami, sú v prípade semiempirickej INDO/2 metódy v absolútnej hodnote zhruba rovnaké, ako pre výsledky, získané v báze 6-31G, pričom znamienka odchýliek sú opačné.

Zaujímavé je zistenie, že dobre parametrizované semiempirické metódy môžu pre silikátové systémy poskytnúť zhruba rovnako dobré výsledky pre elektrostatické vlastnosti, ako poskytujú *ab initio* výpočty v malých bázach, pričom semiempirické výpočty sú väčšinou rýchlejšie. V situáciách, kde pre riešenie konkrétneho problému treba častokrát použiť veľké klastrové modely, neumožňujúce výpočet *ab initio* v dostatočne veľkej báze, je toto zistenie dôležité. Rovnako dôležité je zistenie, že hodnoty elektrostatického potenciálu budú semiempirickými metódami nadhodnotené približne rovnako, ako by boli výpočtom v malej báze podhodnotené.

POROVNÁVACIE ŠTÚDIUM ELEKTROSTATICKÉHO POTENCIÁLU, VYPOČÍTANÉHO POUŽITÍM RÔZNYCH BÁZ PRE KLASTROVÝ MODEL TETRAEDRICKEJ VRSTVY [Si$_2$O$_5$H$_2$]∞

DANIEL TUNEGA, LADISLAV TURI NAGY, ŠTEFAN VARGA

Ústav anorganickej chémie SAV, Dúbravská cesta 9, 842 36 Bratislava, Slovenská republika

Pre klastrový model tetraedrickej vrstvvy [Si$_2$O$_5$H$_2$]∞ sa porovnali elektrostatické vlastnosti – Mullikenove ato-

Obr. 1. 36 atómový klaster ideálnej tetraedrickej siete (O$_a$ – vrcholové atómy kyslíka, O$_b$ – bazálne atómy kyslíka, H$_k$ – kompenzačné atómy vodíka)

Obr. 2. Elektrostatické potenciálové krivky nad bazálnym atómom kyslíka z klastra tetraedrickej silikátovej siete získané s rôznymi bázami.

Obr. 3. Priebeh elektrostatických potenciálov nad bazálnym atómom kyslíka z klastra tetraedrickej silikátovej siete (TSS) a nad bazalánym atómom kyslíka z klastra izolovanej vstvy mastenca (TL) získané semiempirickou INDO/2 metódou.