Deductive Molecular Mechanics of sp^3 Nitrogen Atom and Its Application to Analysis of a QM/MM Interface

A. L. TCHOUGRÉEFF, A. M. TOKMACHEV

Karpov Institute of Physical Chemistry, 10 Vorontsovo pole, Mosco, 105064, Russia

Center for Computational Chemistry, Keldysh Institute for Applied Mathematics of RAS, 4 Miusskaya pl., Moscow, 125047, Russia

ABSTRACT: The antisymmetrized product of strictly localized geminals (APSLG) wave function is used to construct a mechanistic picture for the sp^3 nitrogen atom. We previously showed [18] that the classical force fields for the sp^3 carbon atom can be sequentially derived from the APSLG-MINDO/3 method. Here we try to extend this approach. Special attention is given to the hybridization pattern for the nitrogen atom. Interpolation formulae for the inversion potential are obtained and the role of different physical contributions to the local geometry structure around nitrogen is elucidated. The results obtained are applied to construction of the junction in QM/MM hybrid procedures when a lone pair is assigned to the QM subsystem while three covalent bonds are considered to be a part of the MM subsystem. Changes in the hybridization due to QM/MM boundary are considered as a source of modification of the QM lone-pair characteristic as well as of that of the MM force fields.   © 2004 Wiley Periodicals, Inc. Int J Quantum Chem 100: 667–676, 2004

Key words: molecular mechanics; hybridization; inversion potential; QM/MM

Introduction

The mechanistic approach to the energy of molecules based on the combination of classical force fields is a useful tool for analysis of the potential energy surfaces (PESs) for large molecular systems. This approach is now called molecular mechanics (MM) and exists in many incarnations: CHARMM [1], OPLS [2], MM3 [3], and AMBER [4], to mention a few. These particular schemes differ in the number and structure of the force fields as well as in the sets of parameters used. It is important to stress that the purely empirical character of standard MM constructions makes these schemes computationally feasible but also precludes understanding and detailed analysis of physical interactions in the molecule. In practice, the ques-

Correspondence to: A. Tokmachev; e-mail: tokm@chat.ru

tion of whether to include some "weak" force fields like "off-diagonal" interbond interactions is solved on the "school-wise" basis [5]; that is, the solution depends more on the traditions of a particular scientific group than on the precise estimates of validity of that or another approximation.

The low computational cost of the MM treatment of molecules [especially as compared to high-quality quantum mechanical (QM) computations] has led to an idea [6] of constructing hybrid QM/MM (quantum mechanical/molecular mechanical) procedures, where a small part of the molecule is calculated by using precise QM method and the rest is described by the relatively cheap MM procedure. In past decades, progress in this field was enormous, which is shown in reviews [7–9]. At the same time the question about the interface between subsystems and the account of their interactions is not completely solved [9]. In practice, different recipes based on the intuitive models are used, and their success in description of that or another system is considered as an ultimate criterion on their value. In a series of articles [10–12], we proposed another way of thinking: The junction between subsystems should be obtained as a result of sequential derivation using explicit separation of electronic variables. This approach, based on the effective Hamiltonian technique, is quite general. In our opinion, the possibility of sequential separation of electronic variables is a prerequisite for multilevel treatment of a complex problem in which some parts of the system should be described by highly correlated wave functions, others are described by single determinant and the rest is taken into account by classical potentials.

The essential features of our approach [10–12] are (i) unique assignment of one-electron states (which are chosen to be hybrid orbitals, see below) to the subsystems and (ii) explicit use of an underlying wave function for the inert subsystem (i.e., that part of the molecule treated by the MM). The key feature of the required underlying QM procedure is its local character recovering the common concepts of chemical bonds and lone pairs. The antisymmetrized product of strictly localized geminals (APSLG) wave function is used here as an underlying function for the MM. Its semiempirical implementation, which we rely upon, has been described in Refs. [13, 14]. Here we give only a brief account of this method. The APSLG wave function is constructed as the antisymmetrized product of two-electron wave functions (geminals):

$$|\Phi\rangle = \prod_m g_m^+|0\rangle, \tag{1}$$

representing chemical bonds and lone pairs. Each geminal is expanded in the basis of hybrid orbitals (HOs). These orbitals are obtained by the orthogonal ($h^A \in SO(4)$) transformations within the sp-basis set of atomic orbitals (AOs) for each non-hydrogen atom A:

$$t_{m\sigma}^+ = \sum_{i \in \{s, x, y, z\}} h_{mi}^A a_{i\sigma}^+, \tag{2}$$

where σ is a spin projection variable. In the case of hydrogen the only HO coincides with the s-AO. Each HO is uniquely assigned to a chemical bond or a lone pair. The carrier space for a chemical bond is spanned by two HOs (four hybrid spin-orbitals) and the carrier space for a lone pair is formed by a single HO (two hybrid spin-orbitals). Each geminal is constructed as a linear combination of singlet two-electron configurations:

$$g_m^+ = u_m r_{m\alpha}^+ r_{m\beta}^+ + v_m l_{m\alpha}^+ l_{m\beta}^+ + w_m(r_{m\alpha}^+ l_{m\beta}^+ + l_{m\alpha}^+ r_{m\beta}^+), \tag{3}$$

where $r_{m\sigma}^+$ and $l_{m\sigma}^+$ are the electron creation operators on the right- and left-end HOs for the mth bond. The geminal amplitudes u_m and v_m correspond to the ionic contributions to the bond (both electrons are on the same end of the bond) while the amplitude w_m corresponds to the covalent (homeopolar, Heitler–London) contribution to it. The normalization condition due to orthogonality of basis functions reads:

$$u_m^2 + v_m^2 + 2w_m^2 = 1. \tag{4}$$

In the case of a geminal representing a lone pair, only one ionic contribution in Eq. (3) survives:

$$g_m^+ = r_{m\alpha}^+ r_{m\beta}^+. \tag{5}$$

The geminal amplitudes can be combined into the intrageminal elements of one- and two-electron density matrices [13, 14], entering the expression for the electronic energy (see below):

$$P_m^{tt'} = \langle 0|g_m t_{m\sigma}^+ t'_{m\sigma} g_m^+|0\rangle,$$

$$\Gamma_m^{tt'} = \langle 0|g_m t_{m\beta}^+ t'^+_{m\alpha} t'_{m\alpha} t_{m\beta} g_m^+|0\rangle,$$

$$P_m^{rr} = u_m^2 + w_m^2, \quad P_m^{ll} = v_m^2 + w_m^2,$$

$$P_m^{rl} = P_m^{lr} = (u_m + v_m)w_m, \quad \Gamma_m^{rr} = u_m^2,$$

$$\Gamma_m^{ll} = v_m^2, \quad \Gamma_m^{rl} = \Gamma_m^{lr} = w_m^2, \tag{6}$$

where t and t' are either r or l. In the APSLG approximation the total molecular energy is naturally subdivided on the bonding and nonbonding contributions like in the standard MM schemes.

The use of the APSLG wave function as an underlying one for the MM subsystem is based not only on the internal similarity between the APSLG and the MM energy expressions but also on the possibility of classical representation of the PES of an organic molecule on the basis of the APSLG approach. The key feature of the semiempirical implementation of the APSLG wave function, differing from others proposed in the field [15] and previously used to analyze additive systematics (e.g., in Ref. [16]), is the variational treatment of the strictly localized HOs which are determined on equal footing with other electronic structure parameters (ESPs), namely, the parameters characterizing the bonds (amplitudes of the geminal components), i.e., by minimizing the molecular total energy. Also, the important step here is the possibility to use the guess values of the ESPs and to construct the noniterative (MM-like) procedure based on the APSLG wave function [17]. The accepted way to treat the ESPs serves as a basis for classification of the approximate schemes [17]: FA or TA for the fixed or tuned geminal *amplitudes,* respectively, and FO or TO for the fixed or tuned hybrid *orbitals.*

Essential progress was achieved in Ref. [18] where the linear response approximation was used for analysis of interrelations between molecular geometry and hybridization for the sp^3 carbon atom. This construct may be qualified as a deductive molecular mechanics, because each of its components has a transparent counterpart in the underlying QM description and the approximations and simplifications used can be uniquely characterized and formulated. Analysis has shown that the APSLG approach can be a basis for introducing the transferable energy contributions, thus settling the QM foundation for the "force fields" known in the MM realm. Moreover, it allows us to find corrections to the transferable QM prototypes of the MM force fields, and in some cases to substantiate the presence of certain "weak" force fields. It is important to stress that the values of the elasticity constants obtained in the framework of deductive molecular mechanics and those accepted in the standard force fields are generally in a reasonable agreement, whereas those for the bending force fields are very close. At the same time the purpose of deductive molecular mechanics is not merely a substantiation of the standard MM force fields but also the detailed study of hybridization patterns in different situations, as is emphasized in Ref. [19].

In this article, we extend the treatment proposed in Ref. [18] to organic molecules containing sp^3 nitrogen. The next section gives a deductive MM description of the nitrogen-containing compounds with particular stress on the source of the pyramidalization potential. Then, the implications of our treatment for development of consistent hybrid QM/MM methods for analysis of PESs of chemical transformations of large molecular systems are considered.

Stereochemistry and Molecular Mechanics of sp^3 Nitrogen

For a century two fundamental facts shaped the area of stereochemistry: the terahedral carbon introduced by van t'Hoff and Le Bel and the pyramidal nitrogen. Despite that long history, no common viewpoint as to the source of the angular dependence of the energy has been developed within MM and stereochemistry itself. On the other hand, even simple quantum chemical methods reproduce the observed features. Also, in quite general terms, it is clear that the form of the coordination polyhedron is controlled by the relation between the bonding (two-center) energy, which favors population of excited states of an atom under consideration, and the excitation energies themselves, which tend to keep an atom in its ground state. Recently these notions again have been brought into the discussion in Ref. [20].

An alternative explanation is suggested by Gillespie [21]. According to this author, the angular dependence of energy appears due to Coulomb repulsion between electron pairs (valence shell electron pair repulsion, VSEPR). In the literature there exists quite a number of attempts to reconcile this qualitative and intellectually attractive picture with the results of the quantum chemical calculations, which are reviewed in Ref. [16]. These attempts, however, turned out to be discouraging. It has been found that the intraatomic energy terms responsible for the molecular shape cannot be identified with the interpair Coulomb interactions. At the

same time, the stereochemistry (form of the coordination polyhedron) and the hybridization pattern related to it can be fairly reproduced with no relation to any Coulomb repulsion at all, as in, say, the extended Hückel method [22]. This finding applies both to the carbon stereochemistry described previously and to the nitrogen one. An attempt to derive sequentially the stereochemistry rules with the forms of hybrids obtained semiquantitatively and a posteriori was undertaken in Ref. [23]. This simplified analysis based on the s-/p-ratio in HOs has led to qualitative description of stereochemistry with driving force differing from that claimed within the VSEPR approach. It was qualitatively shown in Ref. [23] that the known stereochemistry rules in fact rely on the difference between the atomic s- and p-levels. In what follows we analyze different contributions to the energy of the APSLG description of the model ammonia molecule and demonstrate that the same energy terms are responsible for determining the observed stereochemistry, and specifically are the source of the pyramidalization potential of the nitrogen atom.

Our consideration is based on making reasonable approximations to the APSLG-MINDO/3 energy. The contributions to the energy beyond the APSLG approximation (electron delocalization between geminals and interbond dispersion interactions) are excluded from our analysis. At the same time we should mention that the semiempirical APSLG method gives perfect results for many properties of molecules, such as molecular geometries, heats of formation, and ionization potentials [13, 14, 24]. The neglected contributions can be taken into account by using formulae based on the perturbation theory applied to the geminal wave function. We are studying this problem now in more details. It is clear that the error induced by neglect of intergeminal excitations will vary with changes of the angles, and it will seriously affect the barriers for torsion rotations. This point is comprehensively discussed in Ref. [25]. An additional assumption made throughout the work is applicability of the FA picture proposed in Ref. [17]. It assumes that N—H and N—C bonds considered are not polarized. The validity of this approximation is confirmed by numerical calculations: Bonds N—H and N—C are only slightly polar [14].

As shown in Ref. [18], the form of the HOs within the FA class of the approximations leading to the MM-like description of the four-coordinated carbon atom is ultimately defined by the two-center resonance interactions. The SO(4) group structure

of the hybrid manifold significantly restricts the ability of the one-electron states residing on the atom to adjust themselves to the variations of geometry arrangement of the surrounding atoms (groups). For some geometries (hybridization incompatible) it is not possible to construct a system of HOs that would even approximately fit the direction of the internuclear axes. This rigidity, however, has nothing to do with the Coulomb repulsion of electrons populating the HOs. Moreover, in the case of the FA approximation the sum of the intraatomic Coulomb interaction terms for carbon atom does not depend on the hybridization at all. No significant bond–bond interaction occurs that can maintain the observed forms of the hybridization patterns (tetrahedra) and thus of the molecules themselves besides the topology of the hybridization manifold which assures the latter.

The situation with other organogenous atoms seems to differ significantly from this picture. In the case of nitrogen and oxygen atoms, even in the FA class of approximations the one-center energy is strongly hybridization dependent. But in contradiction with the VSEPR conjecture this happens because of the one-electron terms describing the core attraction of electrons in the lone pair and it is sensitive to the relative weights of the s- and p-AOs in the corresponding HO. The source of this sensitivity is, of course, the strong difference between the core attraction in the s- and p-subshells with large preference toward a purely s lone pair (in a line with qualitative reasoning of Ref. [23]). In the free atom (purely one-center energy) this immediately resulted in no hybridization for nitrogen and oxygen and in a 90° valence angles predicted by older theories [26] for water and ammonia with subsequent need to explain the observed form of these molecules with the valence angles only slightly smaller than the tetrehedral one and both exceeding 100°. It is curious, but the authors of the VSEPR model seem to overlook this simple mechanism of formation of nonequivalent HOs, well known for decades, and did not consider it as a starting point and incidentally the limiting case of the electron pair repulsion and started their theory from a scratch, inventing artificial interactions between the pairs of electrons in the valence shell.

If we reside for a while in the FA domain we have to admit that the only source of the observed stereochemistry can be found in the interplay between one-center hybridization-dependent terms and the resonance energy, which was clear to Coulson [26]. For nitrogen (in ammonia) one can easily

write the part of the electronic energy that depends on the hybridization of the nitrogen atom:

$$\left[U_s - U_p + \frac{1}{4}\left(3C_2 + 2C_3 + 4C_5\right)\right] s_4^2 + \frac{1}{4} C_3 s_4^4$$
$$- 4P^{rl} \sum_{m \neq 4} \left[\beta_{\sigma\sigma}^{NT_m} s_m + \beta_{\zeta\sigma}^{NT_m}(\vec{e}_{NT_m}, \vec{v}_m^N) \right], \quad (7)$$

where the subscript 4 stands for the lone pair; the quantities C_n are the linear combinations of the Slater–Condon parameters introduced in Ref. [14]; $(s_m, \vec{v}_m)$ is the 4-component vector of the mth HO in the sp-basis (s_m is the expansion coefficient at the s-function, $\vec{v}_m$ is a triple of the coefficients at the Cartesian p-functions) and the resonance integrals β are written in the diatomic coordinate frame (DCF) set by the nitrogen atom and the atom T_m constituting the mth bond characterized by the unit vector $\vec{e}_{NT_m}$ between N and T_m and the basis functions σ, ξ, v, and ζ. Inserting the characteristic values of the atomic parameters yields an instructive result: contributions depending on Coulomb integrals (on C_n) are rather small and can provide the total variation in energy less than 0.8 eV, whereas the difference of the core attractions results in a huge amount of about 10 eV. Thus the nontrivial equilibrium in such a system is only possible if the strong deforming potential exerted by the first term and tending to no hybridization is counterpoised by other contributions. Within the FA type of approximations the only counterpoise is the resonance energy.

It is instructive to parameterize the hybridization manifold. This problem was addressed in Refs. [18, 19], where the parameterization of the SO(4) group by Jacobi angles was used. This six-parametric group can be parameterized by the angles ω_{sx}^A, ω_{sy}^A, ω_{sz}^A, ω_{yz}^A, ω_{xz}^A, and ω_{xy}^A, corresponding to rotations in two-dimensional subspaces of the whole 4-dimensional space spanned by one s- and three p-functions. The first triple of (pseudorotation) angles ($\vec{\omega}_b^A = \{\omega_{sx}^A, \omega_{sy}^A, \omega_{sz}^A\}$) defines a hybridization tetrahedron—a shape uniquely defining relative weights of the s- and p-functions in the hybrids and by this also the interhybrid angles, while the second triple of (quasirotation) angles ($\vec{\omega}_l^A = \{\omega_{yz}^A, -\omega_{xz}^A, \omega_{xy}^A\}$) is responsible for rotation of the hybridization tetrahedron as a whole (The prefix "quasi" refers to the fact that no physical body rotates—only the hybridization tetrahedron defined right above). It is known that the SO(4) group is a direct product of two SO(3) subgroups. Using an intermediate pa-

rameterization of the whole SO(4) manifold by a pair of quaternions in Ref. [19] allowed to determine the first-order correction to the HOs due to small quasi- and pseudorotations $\delta\vec{\omega}_l$ and $\delta\vec{\omega}_b$:

$$\delta^{(1)}s = -(\delta\vec{\omega}_b, \vec{v}), \qquad \delta^{(1)}\vec{v} = s\delta\vec{\omega}_b + \delta\vec{\omega}_l \times \vec{v}, \quad (8)$$

and corresponding results for the second-order variations. Here we use them to analyze the interplay of the molecular geometry and the hybridization manifold of the model ammonia molecule archetypic for all organic compounds containing an sp^3-hybridized nitrogen atom.

For symmetry reasons the resonance energy of the C_{3v} ammonia molecule is a function of only one pseudorotation angle ω_{sz} and of the pyramidalization angle δ, which equals to zero for the planar ammonia molecule. The resonance energy as a function of hybridization and geometry acquires the form (assuming equivalence of all bonds):

$$-4\sqrt{3}\, P^{rl}[\beta_{\sigma\sigma}\cos\omega_{sz} + \beta_{\zeta\sigma}\sin\delta\sin\omega_{sz} + \sqrt{2}\,\beta_{\zeta\sigma}\cos\delta]. \quad (9)$$

It is easy to see that the minimum of the above expression with respect to both its arguments is reached precisely for the planar configuration and the sp^2 hybridization ($\delta = 0$, $\omega_{sz} = 0$).

The hybridization-dependent part of the one-center energy of the nitrogen atom is

$$\left[U_s - U_p + \frac{1}{4}\left(3C_2 + 2C_3 + 4C_5\right)\right]\sin^2\omega_{sz}$$
$$+ \frac{1}{4} C_3 \sin^4\omega_{sz} \quad (10)$$

with obvious extrema: a minimum at $\omega_{sz} = \pi/2$ (no hybridization) and a maximum at $\omega_{sz} = 0$ (sp^2 hybridization). Thus, we arrive to a very simple (but internally consistent) picture for hybridization/stereochemistry of the nitrogen atom. Two contributions to the energy exist. One [Eq. (10)] tends to keep the interhybrid (and valence) angles at 90°, the other one [Eq. (9)] tends to place all substituents at the nitrogen atom to the plane. The observed pyramidal form is a result of the interplay between these two contributions. This results in a pyramidalization (inversion) potential in which no kind of interbond or bond-lone pair Coulomb interactions is directly involved and the Coulomb attraction of electrons to the core is the main factor.

Minimization of the sum of resonance Eq. (9) and one-center Eq. (10) energies on ω_{sz} allows to determine the optimal value of this pseudorotation angle as a function of the angle δ. This minimization is not possible analytically. So we try to find realistic estimates for the dependence $\omega_{sz}(\delta)$. First, we can determine the value of the pyramidalization angle δ optimal for a given value of pseudorotation angle ω_{sz}. It can be easily done by taking the derivative with respect to δ of the resonance energy, because the lone-pair contribution does not depend on δ explicitly. It should be noted that the relatively small specific correction to the core–core interaction adopted in semiempirical schemes is pyramidalization angle-dependent. Also, small polarization of the N—H bonds within the TA approximation framework modifies the energy expression. We neglect the effect of these contributions. It allows to obtain a simple relation:

$$\sin \omega_{sz} = \sqrt{2} \tan \delta, \qquad (11)$$

which holds for all equilibrium geometries—minima, maxima, and saddle points—but can be also considered as an interpolation formula for the intermediate values. The relation (11) also assumes that the HOs are directed along the bond vectors. It leads to appearance of certain pyramidalization potential of rather nontrivial form:

$$2\left[U_s - U_p + \frac{1}{4}\left(3C_2 + 2C_3 + 4C_5\right)\right]\tan^2\delta + C_3\tan^4\delta$$
$$+ -4\sqrt{6}\, P^{rl}[\beta_{\sigma\sigma}\sqrt{1/2 - \tan^2\delta} + \beta_{\zeta\sigma}\sin \delta \tan \delta$$
$$+ \beta_{\zeta\sigma}\cos \delta]. \quad (12)$$

The main source of this potential within the proposed picture is a purely quantum mechanical requirement of mutual orthogonality of the HOs centered on the nitrogen atom. Its physical nature may be characterized as the energy necessary for population of excited configurations of the nitrogen atom due to perturbation induced by the resonance interaction with surrounding bonded atoms. The admixture coefficients (weights) of the excited atomic configurations appear as functions of hybridization parameters, which can be even explicitly written according to Ref. [26]. Neither of these sources has anything to do with interpair Coulomb interaction.

The second derivatives matrix $\nabla_{\vec{\omega}}^2 E$ which is necessary for construction of linear response relations between the geometry variations and the ESPs can

be readily calculated for the sp^3 nitrogen atom in the vicinity of the equilibrium values of the quasi- and pseudorotation angles. It can be shown that the second-order correction to the resonance energy is an average of the 6×6 matrix over vector parts of the intermediate quaternions $\mathbf{p}$ and $\mathbf{q}$ (for details, see Ref. [19]). It has the form:

$$\delta^{(2)}E_{\text{res}} = 2\sqrt{3}\, P^{rl}\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{B}^+ & \mathcal{A} \end{pmatrix}, \quad \mathcal{A} = a\mathcal{I},$$

$$a = \beta_{\sigma\sigma}\cos \omega_{sz} + \beta_{\zeta\sigma}(\sqrt{2}\cos \delta + \sin \delta \sin \omega_{sz}),$$

$$\mathcal{B} = -\begin{pmatrix} b & d & 0 \\ -d & b & 0 \\ 0 & 0 & c \end{pmatrix},$$

$$b = \beta_{\sigma\sigma}\cos \omega_{sz} - \beta_{\zeta\sigma}\sin \delta \sin \omega_{sz},$$

$$c = \beta_{\sigma\sigma}\cos \omega_{sz} - \beta_{\zeta\sigma}(\sqrt{2}\cos \delta - \sin \delta \sin \omega_{sz}),$$

$$d = \beta_{\sigma\sigma}\sin \omega_{sz} + \beta_{\zeta\sigma}\sin \delta \cos \omega_{sz}. \qquad (13)$$

By going to the basis of the small variations of the pseudo- and quasirotation angles $\delta\vec{\omega}_b$ and $\delta\vec{\omega}_l$, we immediately find a pair of eigenvalues of this matrix that correspond (very naturally) to variations of the ω_{sz} and ω_{xy} angles:

$$2\sqrt{3}\, P^{rl}(a \pm c). \qquad (14)$$

The remaining 4×4 matrix in the basis of the quasi- and pseudorotation angle variation acquires the form:

$$2\sqrt{3}\, P^{rl}\begin{pmatrix} a-b & 0 & 0 & d \\ 0 & a-b & -d & 0 \\ 0 & -d & a+b & 0 \\ d & 0 & 0 & a+b \end{pmatrix} \qquad (15)$$

with the eigenvalues evidently pairwise-degenerate:

$$a \pm \sqrt{b^2 + d^2}. \qquad (16)$$

The eigenvectors corresponding to these values can be easily evaluated. For the planar case we get $d = 0$, so then the second-order correction remains diagonal in the basis of pseudo- and quasirotations.

It is interesting to find how the equilibrium value of pseudorotation angle ω_{sz} is reproduced in the linear response approximation:

$$\omega_{sz} \simeq \omega_{sz}^0 - \frac{\partial E / \partial \omega_{sz}(\omega_{sz}^0)}{\partial^2 E / \partial w_{sz}^2(\omega_{sz}^0)}. \qquad (17)$$

The relevant first and second derivatives have the form:

$$\partial E / \partial \omega_{sz} = \left[U_s - U_p + \frac{1}{4}\,(3C_2 + 2C_3 + 4C_5) \right]$$
$$\times \sin 2\omega_{sz} + C_3\sin^3\omega_{sz}\cos\omega_{sz}$$
$$+ 4\sqrt{3}\,P^{rl}(\beta_{\sigma\sigma}\sin\omega_{sz} - \beta_{\zeta\sigma}\sin\delta\cos\omega_{sz});$$

$$\partial^2 E / \partial \omega_{sz}^2 = 2\left[U_s - U_p + \frac{1}{4}\,(3C_2 + 2C_3 + 4C_5) \right]$$
$$\times \cos 2\omega_{sz} - C_3\sin^4\omega_{sz} + 3C_3\sin^2\omega_{sz}\cos^2\omega_{sz}$$
$$+ 4\sqrt{3}\,P^{rl}(\beta_{\sigma\sigma}\cos\omega_{sz} + \beta_{\zeta\sigma}\sin\delta\sin\omega_{sz}). \quad (18)$$

It is clearly seen that the estimate Eq. (17) is δ-dependent.

We give here the estimates for experimental geometry of ammonia with interbond angle equal to 106.7° (it corresponds to $\delta \approx 22.1°$). The QM APSLG-MINDO/3 calculation gives the hybrids corresponding to $\omega_{sz} \approx 49.7°$, which must be taken as an exact value for the purposes of this article because it is a value coming from the underlying QM method. All other estimates are only approximations to this value based on the assumptions discussed above. Linear response estimate Eq. (17) depends on the initial approximation to the angle ω_{sz}. If we take $\omega_{sz}^0 = 0$ (it corresponds to sp² hybridization), then the following estimate appears for dependence of ω_{sz} on geometry (angle δ): $\omega_{sz} \approx 3.42 \sin\delta$. Thus the obtained value ($\omega_{sz} \approx 73.8°$) for the equilibrium geometry is too large. If we take $\omega_{sz}^0 = \pi/2$ (no hybridization) then Eq. (17) gives $\omega_{sz} \approx 56.3°$, which is close to the APSLG one. Use of relation (11) for defining ω_{sz}^0 leads to a corrected value, $\omega_{sz} \approx 55.2°$. So the linear response estimate gives quite reliable HOs in the last two cases.

Nitrogen Frontier Atoms in Hybrid QM/MM Methods

One purpose of this article is to elucidate the precise form of the junction between the QM and MM regions in the hybrid QM/MM methods. The latter is more or less covered by the polarization effects when it goes about the junctions across space, i.e., when no chemical (covalent) bond binds the regions to be treated by the MM and QM, respectively. An opposite situation arises when there are chemical bonds between the QM and MM regions—as in the case of enzymatic reactions or even reactions of large organic molecules where only a small fraction of bonds breaking or forming is considered with use of QM. In these cases, it is now accepted that the interregion boundary must not cut bonds in halves, but is assumed either to pass "through nuclei," leaving some of the bonds completely in the MM region and keeping others completely in the QM region, or is replaced by a kind of "grey zone" for which neither fully MM nor fully QM treatment applies. Theoretical substantiation of a "through nuclei" choice of the interregion border is given in Ref. [11]. However, in either approach there remains a problem of how to select the HOs residing on the frontier atoms, which do not belong to the QM region completely, and how the participation of an atom in the QM subsystem affects its MM parameters. The list of recipes existing in the field is quite impressive and includes "capping" hydrogen atoms, "link" atoms, and others. In all these cases the question arises how to set the parameters of these atoms (particularly if advanced *ab initio* schemes are employed to treat the QM part). It is also clear that the HOs included in the QM region but residing on the frontier atoms must adjust at least their orientation to the variation of the positions of the MM region atoms attached to the frontier atom. This obvious fact did not escape the attention of workers in the field (see Ref. [28] and references therein). However, the solutions of the problem proposed so far largely reduce to ad hoc recipes of adjusting the orientation of the QM-residing HOs to the geometry variations in the MM part. These recipes are introduced without much explanation or substantiation. An obvious point, completely missed by this approach, is the modification of the parameters of the MM part of the system due to variations of the ESPs on the frontier atom caused by its participation in the QM subsystem.

The adjustment of the QM-residing HOs must follow the variational principle for energy when the geometry (there is nothing else) of the MM part changes. Even in such a setting, the energy parameters at the QM-residing HOS vary following the variations of the geometry changes in the MM part, because the latter may result also in variation of the form of the HOs (i.e., of the relative weights of the s- and p-contributions to the HOs). Our previous approach [11, 12] describes only the energy adjust-

ment with relatively unclear prospects of including the variation of the form of the HOs in the QM region. As opposed to this, we consider here the modification of the form of the HOs and the modification of the QM and MM parameters corresponding to the frontier atom. More precisely, we consider a special case when the frontier atom serving as the QM/MM junction is the sp^3 nitrogen atom supplying its lone pair to the QM subsystem. Such a setting seems to be quite natural, given that the basicity or the nucleophilicity functions of the nitrogen atom are both due to interaction of its lone pair with acceptor orbitals. This interaction is naturally to be treated by some kind of QM technique while leaving the rest of the nitrogen neighbors in the MM region.

The one- and two-electron density matrix elements for the QM-residing HO are evaluated by the corresponding (QM) procedure, thus invoking the TA type of description for this HO. In the case of involving the lone pair into QM subsystem the corresponding density matrix elements will vary from their invariant values ($P_4^{rr} \neq 1$, $\Gamma_4^{rr} \neq 1$). The density matrix elements assigned to HOs of the MM region are fixed at their invariant values according to the FA setting. The nonvanishing intersystem two-electron densities reduce to the products of the corresponding one-electron density matrix elements, and the interaction is described by the reduced Coulomb integrals [29]. It is a consequence of neglect of one-electron transfers between subsystems. With these assumptions, we get the corrected hybridization-dependent one-center energy for the frontier nitrogen atom:

$$[(1 + 2\delta P_4^{rr})(U_s - U_p + C_5 + C_3/2)$$
$$+ (3/4 + \delta\Gamma_4^{rr})C_2]\sin^2\omega_{sz}$$
$$+ (1/4 + \delta\Gamma_4^{rr} - \delta P_4^{rr})C_3\sin^4\omega_{sz}. \quad (19)$$

In the case of $\delta P_4^{rr} = \delta\Gamma_4^{rr} = 0$, this expression reduces to Eq. (10). In practice, variation of the one- and two-electron densities on the lone pair leads to modification of the nitrogen pyramid (δP_4^{rr}, $\delta\Gamma_4^{rr} < 0$). Numerical estimate shows that the correction to piramidalizing momentum is

$$[-45.020\delta P_4^{rr} + 8.424\delta\Gamma_4^{rr}]\frac{eV}{rad}. \quad (20)$$

The derivatives of the energy correction with respect to the angles $\vec{\omega}_b$, $\vec{\omega}_l$ yield additional quasi-

and pseudotorques ($\vec{K}_4'$ and $\vec{N}_4'$, respectively) acting upon the hybridization tetrahedron:

$$\vec{K}_4' = 0,$$

$$\vec{N}_4' = -2s_4^N\vec{v}_4^N[\delta P_4^{rr}\{2U_s - 2U_p + C_3 + 2C_5 - 2C_3(s_4^N)^2\}$$
$$+ \delta\Gamma_4^{rr}\{C_2 + 2(s_4^N)^2\}]. \quad (21)$$

We see that the quasitorque induced by the small variations of the one-center ESPs is vanishing, resulting in no quasirotation of the hybridization tetrahedron. At the same time the pseudotorque appears due to involvement of the frontier atom in the density redistribution within the QM part of the combined system. This contribution to the QM-induced pseudotorque is itself collinear to the QM-residing HO ($m = 4$).

In the QM part of the system variation of the bond orders can take place as well. In variance with the pure APSLG picture [13, 14], accepted in the present work as the underlying QM method for the MM part of the system, the atoms in the QM part of the combined system may have off-diagonal elements of the one-electron density matrix between arbitrary one-electron states ascribed to the QM subsystem. The latter are obviously the bond orders for the QM part of the system. The corresponding contribution to the energy reads:

$$E'_{res} = -2\sum_A\sum_\mu P_{r_4\mu}\beta_{r_4\mu}^{NA}, \quad (22)$$

where $P_{r_4\mu}$ are the elements of the one-electron density matrix (spin-bond orders) between the r_4th HO for lone pair assigned to the QM subsystem and the μth AO in the QM system on an arbitrary atom A within the latter. The resonance integrals between the lone pair HO residing on the frontier atom and the AOs on any atom in the QM region are functions of six independent angles ($\vec{\omega}_b$ and $\vec{\omega}_l$) defining the hybridization pattern for the nitrogen atom. They take the following form (in the corresponding DCF):

$$\beta_{r_4\sigma}^{NA} = \beta_{\sigma\sigma}^{NA}s_4^N + \beta_{\zeta\sigma}^{NA}v_{4\zeta}^N,$$

$$\beta_{r_4\zeta}^{NA} = \beta_{\sigma\zeta}^{NA}s_4^N + \beta_{\zeta\zeta}^{NA}v_{4\zeta}^N,$$

$$\beta_{r_4\xi}^{NA} = \beta_{\pi\pi}^{NA}v_{4\xi}^{NA}, \quad \beta_{r_4v}^{NA} = \beta_{\pi\pi}^{NA}v_{4v}^N. \quad (23)$$

Taking into account the definitions of the components of the vector part of the HO in the DCF:

$$v_{4\zeta}^N = (\vec{v}_4^N, \vec{e}_{NA}),$$

$$v_{4\xi}^N = (\vec{v}_4^N, \vec{e}_{NA}^\xi),$$

$$v_{4v}^N = (\vec{v}_4^N, \vec{e}_{NA}^v), \tag{24}$$

and expressions (8) for the variations of the HO coefficients with respect to pseudo- and quasirotation angles ($\vec{\omega}_b$ and $\vec{\omega}_l$), we get explicit expressions for the resonance contribution to the pseudo- and quasitorque at the frontier atom:

$$\vec{N}'_{\text{res}} = 2 \sum_A \{(P_{r_4\sigma}\beta_{\sigma\sigma}^{NA} + P_{r_4\zeta}\beta_{\sigma\zeta}^{NA})\vec{v}_4^N$$

$$- (P_{r_4\sigma}\beta_{\zeta\sigma}^{NA} + P_{r_4\zeta}\beta_{\zeta\zeta}^{NA})s_4^N\vec{e}_{NA} - \beta_{\pi\pi}^{NA}s_4^N(P_{r_4\xi}\vec{e}_{NA}^\xi + P_{r_4v}\vec{e}_{NA}^v)\},$$

$$\vec{K}'_{\text{res}} = 2 \sum_A \{(P_{r_4\sigma}\beta_{\zeta\sigma}^{NA} + P_{r_4\zeta}\beta_{\zeta\zeta}^{NA})\vec{e}_{NA}$$

$$+ \beta_{\pi\pi}^{NA}(P_{r_4\xi}\vec{e}_{NA}^\xi + P_{r_4v}\vec{e}_{NA}^v)\} \times \vec{v}_4^N, \tag{25}$$

where $\vec{e}_{NA}^\xi$, $\vec{e}_{NA}^v$, and $\vec{e}_{NA} = \vec{e}_{NA}^\zeta$ are the orts of the DCF defined by the NA pair of atoms.

The total pseudo- and quasitorques then become

$$\vec{N}' = \vec{N}'_4 + \vec{N}'_{\text{res}}, \qquad \vec{K}' = \vec{K}'_{\text{res}}, \tag{26}$$

and appear due to quantum behavior of electrons in the QM region. Finally, in the linear response approximation, after being treated by the $(\nabla^2_{\vec{\omega}}E)^{-1}$ matrix Eq. (18), they produce the pseudo- and quasirotations of the hybridization tetrahedron on the frontier atom N. The corrections to the pseudo- and quasirotation angles of the hybridization tetrahedron result both in a new form and the orientation of the latter. It is thus inconsistent with the positions of the atoms bonded to the frontier atom from the MM side of the system. The multiplying the angular corrections by the $(\nabla_{\vec{\omega}}\nabla_{\vec{\varphi}_m}E)^\dagger$ matrix results in a torque acting upon the T_m atom of the mth bond incident to the frontier atom N. Also the additional one- and two-electron densities on the frontier atom give additional forces acting upon its MM neighbors. They can be easily obtained of the variations of the quasi- and pseudorotation angles are multiplied by the mixed second derivatives matrix $(\nabla_{\vec{\omega}}\nabla_{r_{NT}m}E)^\dagger$. These forces are directed along the $\vec{e}_{NT_m}$ vectors. This comprises the effect (forces and torques) exerted by the QM subsystem upon the atoms attached to the frontier atom on the side of the MM system due to changes of hybridization of the frontier atom.

On the other hand, any deformation in the MM system results in variation of the pseudo- and quasirotation angles. The shifts of the positions of the MM neighbors of frontier atom result in quasi- and pseudotorques acting upon the hybridization tetrahedron. In its turn, this produces variations of both one-center parameters corresponding to the QM-residing HO and of the resonance parameters for the QM-residing HO and all other orbitals in the QM region. The variation of the one-center matrix elements of the Hamiltonian corresponding to the QM HO is

$$\delta U_4^r = 2(U_s - U_p)s_4^N\delta^{(1)}s_4^N = 2(U_s - U_p)s_4^N(\delta\vec{\omega}_b, \vec{v}_4^N),$$

$$\delta(r_4r_4 \mid r_4r_4)^N = 2C_2s_4^N\delta^{(1)}s_4^N + 4C_3(s_4^N)^3\delta^{(1)}s_4^N$$

$$= 2[C_2s_4^N + 2C_3(s_4^N)^3](\delta\vec{\omega}_b, \vec{v}_4^N). \tag{27}$$

The numerical estimate for the modification of parameters by Eq. (27) is

$$\delta U_4^r \simeq -9.53\delta\omega_{sz} \text{ eV},$$

$$\delta(r_4r_4 \mid r_4r_4)^N \simeq -1.22\delta\omega_{sz} \text{ eV}. \tag{28}$$

The modification of the QM resonance integrals to which the HO at hand is involved is somewhat more complex. It nevertheless uses the same (DCF) representation of the resonance integrals as previously:

$$\delta\beta_{r_4\sigma}^{NA} = \beta_{\sigma\sigma}^{NA}\delta^{(1)}s_4^N + \beta_{\zeta\sigma}^{NA}\delta^{(1)}v_{4\zeta}^N,$$

$$\delta\beta_{r_4\zeta}^{NA} = \beta_{\sigma\zeta}^{NA}\delta^{(1)}s_4^N + \delta_{\zeta\zeta}^{NA}\delta^{(1)}v_{4\zeta}^N,$$

$$\delta\beta_{r_4\xi}^{NA} = \beta_{\pi\pi}^{NA}\delta^{(1)}v_{4\xi}^N, \qquad \delta\beta_{r_4v}^{NA} = \beta_{\pi\pi}^{NA}\delta v_{4v}^N. \tag{29}$$

If the above variations are taken into account in the calculations on QM part of the complex system, the effect of the MM system on the parameters of the effective Hamiltonian for the QM part turns out to be taken into account in the first order. It should be stressed that changes in the hybridization of the frontier atom due to participation of one orbital in the QM subsystem are not taken into account in the existing QM/MM schemes. This effect is not very large, and the first-order correction for taking it into account seems to be adequate.

Conclusion

In this article, we extended the QM substantiation of MM of aliphatic hydrocarbons proposed in Ref. [18] to compounds containing aliphatic nitrogen. We analyzed the energy expression for ammines and revealed sources of pyramidality of the nitrogen atom. That allowed us to construct approximate (interpolation) formulae for the pyramidalization potential for the sp^3 nitrogen atom starting with the APSLG description of electronic structure of ammonia molecule and related (locally) C_{3v} tertial ammines. From another point of view it leads to the special treatment of frontier atoms responsible for the junction between the QM and MM parts of the system under the hybrid QM/MM treatment. The effect of deformation in the MM-treated part upon the parameters of the effective QM Hamiltonian and the reverse effect of the QM driven variation of the elements of the one- and two-electron density matrices upon the MM-treated part of the combined system have been considered, and explicit formulae and numerical estimates of these effects have been presented.

References

1. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. J Comput Chem 1983, 4, 187.

2. Jorgensen, W. L.; Tirado-Rives, J. J Am Chem Soc 1988, 110, 1657.

3. Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. J Am Chem Soc 1989, 111, 8551.

4. Cornell, W. D.; Ciepak, P.; Bayly, C. I.; Gould, I. R.; Merz, K.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. J Am Chem Soc 1995, 117, 5179.

5. Meyer, A. Y. In Theoretical Models of Chemical Bonding, Part 1; Maksić, Z. B., Ed.; Springer: Heidelberg, 1989.

6. Warshel, A.; Levitt, M. J Mol Biol 1976, 103, 227.

7. Gao, J. In Reviews in Computational Chemistry; Boyd, D. B.; Lipkowitz, K. B., Eds.; VCH Publishers: New York, 1995.

8. Sherwood, P. In Modern Methods and Algorithms of Quantum Chemistry, Proceedings, 2nd ed.; Grotendorst, J., Ed.; NIC Series 3; John von Neumann Institute for Computing: Jülich, 2000.

9. Tchougréeff, A. L.; Tokmachev, A. M. In Advanced Topics in Theoretical Chemical Physics; Maruani, J.; Lefebvre, R.; Brändas, E., Eds.; Kluwer: Dordrecht, 2003.

10. Tchougréeff, A. L. Phys Chem Chem Phys 1999, 1, 1051.

11. Tokmachev, A. M.; Tchougréeff, A. L.; Misurkin, I. A. J Mol Struct (Theochem) 2000, 506, 17.

12. Tokmachev, A. M.; Tchougréeff, A. L. Int J Quantum Chem 2001, 84, 39.

13. Tokmachev, A. M.; Tchougréeff, A. L. J Comp Chem 2001, 22, 752.

14. Tokmachev, A. M.; Tchougréeff, A. L. J Phys Chem A 2003, 107, 358.

15. Surján, P. R. In Theoretical Models of Chemical Bonding, Part 2; Maksić, Z. B., Ed.; Springer: Heidelberg, 1989.

16. Daudey, J.-P.; Malrieu, J.-P.; Rojas, O. In Localization and Delocalization in Quantum Chemistry, Vol. 1, Atoms and Molecules in the Ground State; Chalvet, O.; Daudel, R.; Diner, S.; Malrieu, J.-P., Eds.; D. Reidel: Dordrecht, 1975.

17. Tokmachev, A. M.; Tchougréeff, A. L. Int J Quantum Chem 2002, 88, 403.

18. Tchougréeff, A. L.; Tokmachev, A. M. Int J Quantum Chem 2004, 96, 175.

19. Tchougréeff, A. L. J Mol Struct (Theochem) 2003, 630, 243.

20. Nicolaides, C. A.; Komninos, Y. Int J Quantum Chem 1999, 71, 25.

21. Gillespie, R. J.; Hargittai, I. The VSEPR Model of Molecular Geometry; Allyn & Bacon: Boston, 1991.

22. Hoffmann, R. J Chem Phys 1963, 39, 1397.

23. Mayer, I. Struct Chem 1997, 8, 309.

24. Tokmachev, A. M.; Tchougréeff, A. L.; Misurkin, I. A. Int J Quantum Chem 2001, 85, 109.

25. Surján, P. R.; Mayer, I.; Kertész, M. J Chem Phys 1982, 77, 2454.

26. Coulson, C. A. Valence; Oxford University Press: Oxford, 1961.

27. Elliot, J. P.; Dawber, P. G. Symmetry in Physics; MacMillan Press: London, 1979.

28. Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. J Phys Chem A 1998, 102, 4714.

29. McWeeny, R. Methods of Molecular Quantum Mechanics, 2nd ed.; Academic Press: London, 1992.