



Lone pairs vs. covalent bonds: conformational effects in bicyclo[3.3.1]nonane derivatives

Sergey A. Pisarev¹ · Dmitry A. Shulga² · Vladimir A. Palyulin^{1,2} · Nikolay S. Zefirov^{1,2}

Received: 1 September 2018 / Accepted: 15 November 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

Investigations on the relative energy of two least-strain conformers for bicyclo[3.3.1]nonane **1**, bicyclo[3.3.1]nonan-9-one **2**, and their heteroanalogues: 3,7-dimethyl-3,7-diazabicyclo[3.3.1]nonane **3**, 3,7-dimethyl-3,7-diazabicyclo[3.3.1]nonan-9-one **4** were performed using the calculations from the first principles (ab initio, DFT) as well as by semiempirical (NDDO, DFTB) and empirical (molecular mechanics, MM) techniques. For these quite simple structures, serious discrepancies in results of modeling between methods of different origins were revealed. Nonempirical calculations state that the “double chair” (CC) form is the most favorable for carbobicyclic structures **1** and **2**, while 3,7-dimethyl-3,7-diaza compounds **3** and **4** are in general more prone to adopt the “chair-boat” (CB) conformation. The classical rationalization of these quantum chemistry results leads to the hypothesis similar to one that underlies the Gillespie VSEPR concept, namely that the 3,7-repulsion of lone electron pairs is stronger than the corresponding interaction of hydrogen atoms of C–H bonds. The semiempirical NDDO calculations retain the qualitative correspondence of the results to those of the ab initio calculations, while the results of more recent DFTB approaches are closer to MM in their qualitative inconsistency with high-level ab initio methods. In particular, for **4** the relative energy of CC is severely underestimated, erroneously predicting the predominance of this form over CB. The origin of this failure could lie in the relatively coarse parameterization of common force fields when concerning the subtle interplay between different types of interatomic interactions and could be recovered, although only partially, by the proper choice of the charge scheme to use the atomic-centered charges in the explicit account for the non-valency interactions in the Coulombic form.

Keywords Bicyclo[3.3.1]nonanes · Bispindines · Non-bonding intramolecular interactions · Conformational analysis · VSEPR · Substituent effects · Quantum chemistry · Density functional theory · Force fields

Introduction

Chemical compounds generally exist as ensembles of certain different spatial molecular forms. Their content in the equilibrium mixture is dependent on their relative energy. Maxima of the population are attained for the potential

energy surface (PES) minima, i.e., dynamically stable structures called *conformers*. Conformers of lower energy, symmetry and rigidity are more populated and, therefore, more important for the description of *conformational behavior* of a molecule. Many methods of the conformer modeling are available, both nonempirical calculations “from the first principles” and parametric ones, that make use of transferable parameters fitted from the experimental data.

An instantaneous molecular structure could be characterized through its *conformation*. This semi-qualitative notion is recurrent in the sense that a structure is said to adopt a certain conformation on the basis of similar assignments made for its selected substructures, etc. Conformations of the simplest (sub)structures are derived directly from the values of selected dihedral (*torsional*) angles [1, 2]. These are some very basic and principal statements of the *conformational analysis*, an important part of the structural and physical organic chemistry [3, 4].

Nikolay S. Zefirov passed away in 2017

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11224-018-1240-z>) contains supplementary material, which is available to authorized users.

✉ Vladimir A. Palyulin
vap@qsar.chem.msu.ru

Extended author information available on the last page of the article.

The central problem of modern theoretical conformational studies is to find an affordable, yet reliable, technique of the potential energy evaluation for any reasonable instantaneous spatial molecular structure [5]. Since the state of a conformational equilibrium is controlled by relative energies of conformers, and these energy differences are mainly related to the differences in intramolecular non-valency interactions in the structures [6], the universal methodology of the accurate description of non-valency interatomic interactions is urgently required for greater reliability of results in conformational modeling [7].

Nowadays, most, if not all, properties of a compound are defined in the terms related to the dynamics of its molecules, so the methods of dynamic modeling are extremely useful for the characterization of the conformational behavior. These simulations mostly rely upon the classical molecular mechanics (MM), also known as the empirical force field (FF) theories [8, 9]. Quantum chemistry treats the interatomic interactions implicitly [10, 11], but classical methods of molecular modeling explicitly parameterize all interatomic interactions available in a certain force field.

The bicyclo[3.3.1]nonane system appeared to be an impressive touchstone for the conformational modeling methods. Deceptively simple, its derivatives represent the challenge for modeling due to the complexity of interactions affecting the relative energy of the conformers [12, 13]. Our present work is partially devoted to the investigation of close non-valency interaction of chemical bonds and non-bonding (*lone*) electron pairs, similar to that considered as the principal factor in forming the molecular structure geometry according to the VSEPR theory [14–16]. These interactions are crucial for the decent description of conformational behavior of 3,7-diheteroanalogues of bicyclo[3.3.1]nonane [17, 18]. The analysis is even more intriguing for 3,7-diazabicyclo[3.3.1]nonanes (*bispidines*) [19] due to the manifold structure-dependent biological activity observed for these heterocyclic derivatives [20].

Different computation methodologies target different energy quantities. The variety of computational techniques employed here required appropriate methods to analyze their predictive ability. Diagrams of conformational energy correlation exhibit qualitative features of the conformational behavior changes due to the introduction of substituents or heteroatoms in the reference structure. The related notion of “conformational effect of substitution”, also presented below, should not be confused with the formerly defined “conformational effect” that embraces the cases of “non-classical” behavior, where MM methods fail to predict the conformational properties consistently [21]. We found it more useful to connect the deviations in the conformational behavior of related compounds to the structural differences between their molecules within the same “model chemistry” framework.

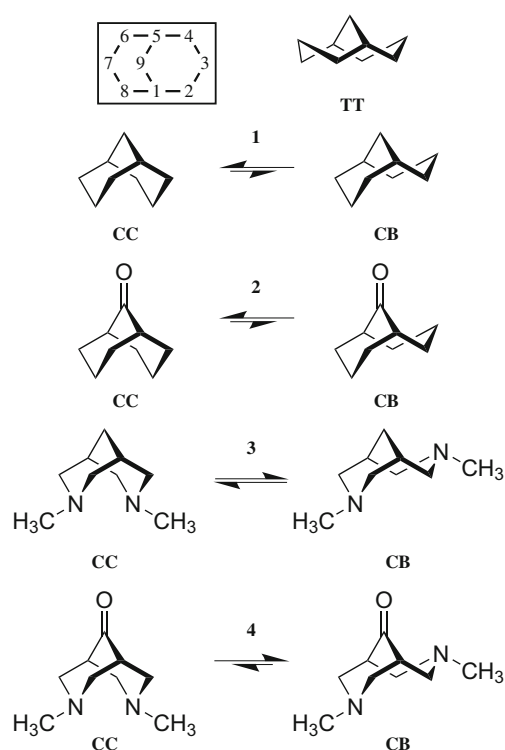


Fig. 1 The bicyclo[3.3.1]nonane system and conformational equilibria for 1–4

Conformations of bicyclo[3.3.1]nonane analogues

Three principal families of skeletal conformations are observed in saturated bicyclo[3.3.1]nonane derivatives: “chair-boat” (CB), “double chair” (CC), and “double twist” (TT). They all are named according to the conformations of six-membered rings comprising the corresponding bicyclic structure (Fig. 1). All TT conformers are left behind the scope of the present study due to their excessive strain. CB conformations are destabilized mainly due to the torsional eclipsing (Pitzer strain) in the “boat” ring, similar to that in the corresponding conformation of free six-membered cycles. It is the CC that is thought to be the most stable form of the bicyclic systems under the investigation due to the absence of eclipsing and the lack of the Pitzer strain [17].

Different types of 3,7-interactions in CC conformers (Fig. 2) could be observed and classified by means of the differential analysis of the electron density, obtained either computationally or from the X-ray diffraction experiments [12]. Their strength is directly affected by atoms and substituents in corresponding positions; several geometry parameters of the bicyclic system, however, would also affect it indirectly, e.g., through its characteristic interatomic distance $d_{3...7}$. The CC conformers are destabilized by repulsive interactions between *endo*-substituents in positions 3 and 7 of the bicyclic system (Fig. 2a). For

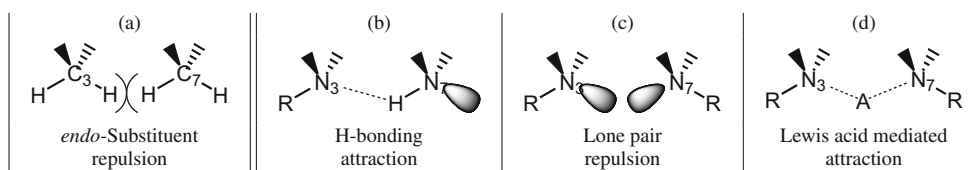


Fig. 2 Types of 3,7-interaction in **CC** conformations of bicyclo[3.3.1]nonanes **a** and their 3,7-diaza analogues **b–d**

3,7-diazabicyclo[3.3.1]nonanes (*bispidines* [19, 20]) with a hydrogen-substituted N atom, the **CC** conformation is, however, stabilized due to the formation of a bent N-H...N bridge as in Fig. 2b [22]. The attractive interaction between positions 3 and 7 increases the relative stability of the **CC** form. A non-hydrogen N substituent always adopts the equatorial position [23, 24], so that it is not sufficiently involved in the steric interactions within the ring system. The 3,7-interaction in **CC** conformations in 3,7-dialkyl-substituted bispidines in the absence of *endo*-substituents is described as “the lone-pair repulsion” (Fig. 2c); the strength of the repulsion is determined by the stereoelectronic state (*hybridization*[25]) of nitrogen atoms [26, 27]. The relative rigidity of the moiety in both conformations enables various effects, steric, electronic or combined, that could also affect the state of conformational equilibrium [28–30].

The **CC** conformation is unambiguously stated to be optimal for hydrocarbon **1**, beginning with earlier MM studies [31, 32]. Recent high-level *ab initio* calculations of $\Delta\mathcal{E}_{conf}$ (2.76 kcal/mol from ZPE-corrected MP2(*full*)/6-311++G** [12] or 2.21 kcal/mol from ZPE + MP4SDTQ(*val*) + MP2(*core*)-corrected CBS MP2 // MP2(*full*) / aug-cc-pVDZ [13]) are in close agreement with the available experimental data: the gas phase electron diffraction studies assisted with the fitted force field gave $\Delta\mathcal{H}_{conf} \simeq \Delta\mathcal{E}_{conf} = 2.48$ kcal/mol; $\Delta\mathcal{S}_{conf} = 1.52$ cal/(mol·K) [33–35]. Due to the reduced conformational predominance of the **CC** form molecules of ketone **2** are conformationally disordered in the solid state at ambient temperatures [36–38]. Earlier MM studies [39] agreed in this part with more recent B3LYP and MP2 data [40] and with the results of microwave spectroscopy [41], which estimated $\Delta\mathcal{E}_{conf} \gtrsim 1.1$ kcal/mol for any conformer other than **CC**. The ^{13}C NMR spectroscopy consistently gives for the **CB** formation $\Delta\mathcal{G}^\circ = 0.95$ or 1.16 kcal/mol in liquid phase (without the lanthanide shift reagent or in its presence, respectively) [42, 43].

The conformer interconversion **CB** \rightleftharpoons **CC** in 3,7-diazabicyclo[3.3.1]nonanes proceeds with the planar inversion of the trigonal pyramidal nitrogen atom coupled with the torsional rotation of the corresponding N substituent (cf. Fig. 3). The barrier of the inversion is generally lower than that of the concerted conformational conversion of the six-membered ring (*semiversion* [44]) that occurs in cyclohexane [45]. This complexity of the mode coupling together with the sufficient anharmonism of the vibrations

involving amine atoms and large scale motions in general definitely requires a careful thermodynamic characterization. However we leave it out of scope of the present paper as a subject for further research. Nevertheless, we adjusted energy differences with the sufficient minimum of the corrections motivated by thermodynamics, e.g., ZPE, to estimate harmonic part of thermodynamic corrections.

Bispidines have generally lower transition energy for the conformational process than the respective carbocyclic derivatives [17, 19]. These compounds reportedly behave as “proton sponges” and ligands with high affinity to the transition metal ions due to their strong chelation ability accompanied with the alteration of the 3,7-interactions character (Fig. 2d) [46, 47]. For the diamine **3** [48] the MM, MNDO [49], and minimal basis HF model calculations [50] predict the **CC** form domination. Generally, it is **3** that constitutes the most complicated case among the compounds under the treatment for there is no clear conformational predominance for its derivatives (*vide infra*) [51, 52]. For **4** the results of MM and AM1 [53] calculations [54, 55] as well as the data from ^1H and ^{13}C NMR spectroscopy were initially interpreted in favor of the **CC** [56]; now on the basis of numerous structural studies, both X-ray diffraction [23] and NMR [52], the **CB** conformation is recognized to be commonly optimal for **4** and some its derivatives [24].

The aim of the present work is not only to find the reliable methods to obtain the consistent estimates of $\Delta\mathcal{E}_{conf}$

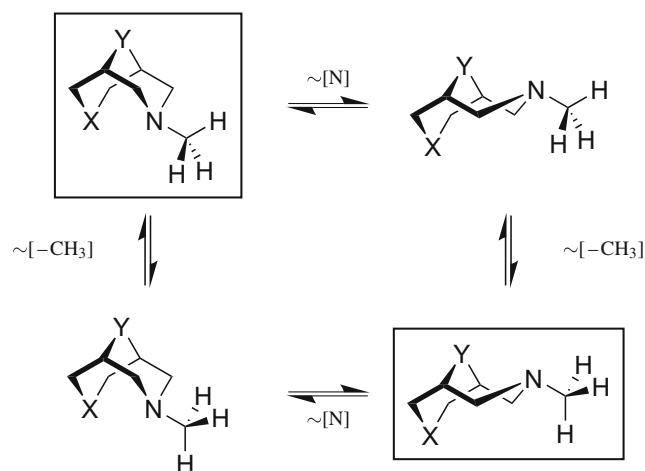


Fig. 3 Conformational motions coupling in N-methyl-3(7)-azabicyclo[3.3.1]nonanes

for all structures under investigation. The complicated picture of factors involved in the description of the relative stability of conformations makes us use molecules **1–4** as a benchmark for the wide variety of methods to find the appropriate techniques, both rigorous in underlying theory and affordable for the routine calculations on the practically important molecules.

Methods

While the search for conformers could be successfully completed by the known procedures of total energy minimization, after all what we need to characterize methods by themselves, is to quantify and analyze the similarity of their results using relative energy quantities that both describe the conformational behavior of compounds and could be uncontroversially defined for various methods. The *conformational energy* $\Delta\mathcal{E}_{conf}$ is defined hereafter as the energy of the **CB** relative to that of the **CC**: $\Delta\mathcal{E}_{conf} = E(\text{CB}) - E(\text{CC})$, so that values of $\Delta\mathcal{E}_{conf} > 0$ correspond to the **CC** conformer favored, while $\Delta\mathcal{E}_{conf} < 0$ indicates the favored **CB** form. This is exact for parametric methods. Thermodynamic calculations from the first principles for symmetric compounds use the non-zero threshold $\Delta\mathcal{E}_{conf}^0$ defined by the difference in axial symmetry of their conformers. Symmetric **CC** conformers are of C_{2v} symmetry with $\sigma = 2$, while their **CB** counterparts possess the planar symmetry (the C_s symmetry group) with the principal axis rotation number $\sigma = 1$. So $\Delta\mathcal{E}_{conf}^0 = k_B N_A \ln \frac{\sigma(\text{CC})}{\sigma(\text{CB})} \approx 0.38$ kcal/mol. Here k_B stands for the Boltzmann constant in the corresponding unit system and N_A for the Avogadro number [34].

The quantitative change in the conformational behavior between two analogous compounds with similar conformational behavior could be assessed using the concept of *conformational effect of substitution*. This notion differs substantially from the formerly defined *conformational effect*, concerning the qualitative difference between two types of the conformational behavior: the experimental one and the one predicted by a certain MM method, thus bound tightly to the classical FF theory [21]. According to the present approach, the difference in the values of the conformational energy is exclusively attributed to the difference in their structures.

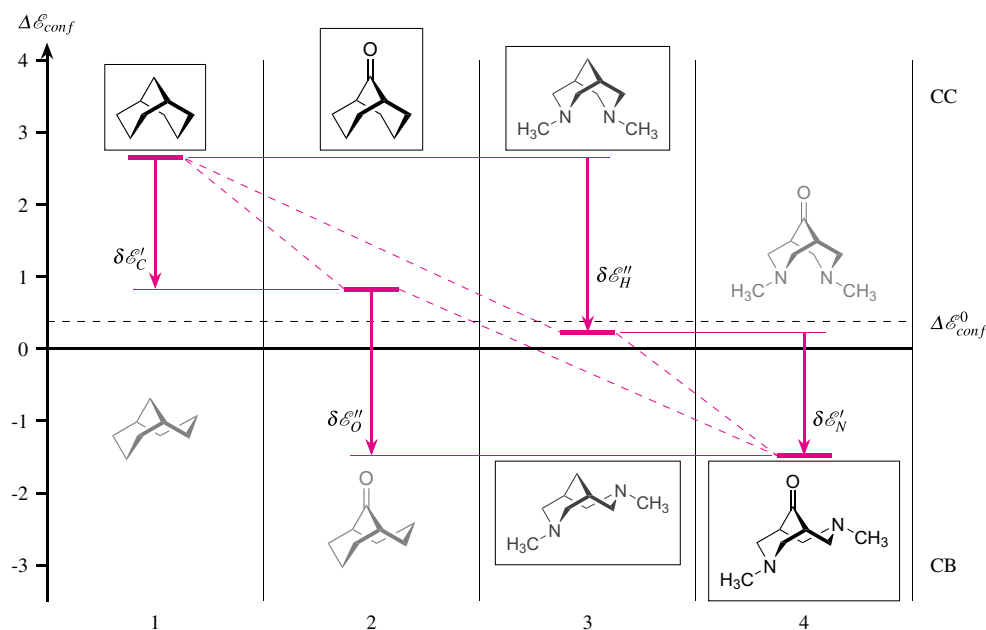
Four structures in our set are interrelated using two types of elementary structure changes. The first one is the carbonyl substitution at position 9 that generates **2** from **1** and **4** from **3**; the second one is the double heteroanalogous substitution at positions 3 and 7 that transforms **1** into **3** and **2** into **4**. The differences in $\Delta\mathcal{E}_{conf}$ corresponding to these changes are hereafter referred to by us as *conformational effects of substitution* of the corresponding type. These effects are denoted as $\delta\mathcal{E}'$ and $\delta\mathcal{E}''$ respectively. They could be estimated for any bound pair of structures as follows:

$$\left. \begin{aligned} \Delta\mathcal{E}_{conf}(\mathbf{2}) - \Delta\mathcal{E}_{conf}(\mathbf{1}) &= \delta\mathcal{E}'_C \\ \Delta\mathcal{E}_{conf}(\mathbf{4}) - \Delta\mathcal{E}_{conf}(\mathbf{3}) &= \delta\mathcal{E}'_N \end{aligned} \right\} \rightarrow \delta\mathcal{E}' \quad (1)$$

$$\left. \begin{aligned} \Delta\mathcal{E}_{conf}(\mathbf{3}) - \Delta\mathcal{E}_{conf}(\mathbf{1}) &= \delta\mathcal{E}''_H \\ \Delta\mathcal{E}_{conf}(\mathbf{4}) - \Delta\mathcal{E}_{conf}(\mathbf{2}) &= \delta\mathcal{E}''_O \end{aligned} \right\} \rightarrow \delta\mathcal{E}'' \quad (2)$$

with the cyclic redundancy relationship $\delta\mathcal{E}'_N + \delta\mathcal{E}''_H = \delta\mathcal{E}'_C + \delta\mathcal{E}''_O$. The simplest qualitative characteristic of the conformational behavior change is the sign of the effect, showing the trend in the conformation energy change induced by the substitution. The difference between

Fig. 4 The correlation diagram of conformational energy $\Delta\mathcal{E}_{conf}$ for **1–4**



estimates of the effect could be attributed to the change in the character of the intramolecular interactions between the compounds of different classes as predicted by the current theoretical scheme. The coincidence of the values for the same conformational effect for different pairs of structures exhibits the independence of its manifestation upon the structural difference relating these pairs.

For the purpose of qualitative analysis the above-mentioned quantities could be visualized using the *diagrams of conformational energy correlation*, as exemplified in Fig. 4, where the $\Delta\mathcal{E}_{conf}$ values of structures bound by a certain formal *elementary structure change* are connected. The $\Delta\mathcal{E}_{conf}^0$ threshold, where applicable, is also shown on the diagrams.

Symbols $\Delta\mathcal{E}^\eta(\mathbf{CB})$ and $\Delta\mathcal{E}^\eta(\mathbf{CC})$ stand for the energies of the formal reaction of *positional redistribution* of substituents and heteroatoms (Fig. 5) for the corresponding conformers. This reaction belongs to the hyperhomod-esmotic class with the reportedly improved accuracy for thermochemical calculations [57, 58]. The resulting values quantify the differences in the intramolecular strain between similar conformations of the compounds. In particular, $\Delta\mathcal{E}^\eta > 0$ reveals the cumulative stabilization of the leftmost structure due to both types of substitution in it, while $\Delta\mathcal{E}^\eta < 0$ denotes the corresponding destabilization [13]. For the present series of structures (with $X = X' = \text{N-CH}_3$ and $Y = \text{CO}$) $\Delta\mathcal{E}^\eta$ values quantify the interaction between double nitrogen and carbonyl substitutions for individual conformers of aminoketone **4**. The numerical consistency of $\delta\mathcal{E}'$, $\delta\mathcal{E}''$, and $\Delta\mathcal{E}^\eta$ values obtained by different methods should be considered as the general similarity estimate for the description of conformational behavior for those methods, independently on their origin.

Details of calculation

Calculations from the first principles were performed using the ORCA quantum chemistry package [59, 60]. We could not rely upon the double-zeta basis sets because of their deficiency demonstrated in correlated ab initio calculations [61], so the calculations, unless otherwise noted, are performed using the correlation-consistent triple-zeta polarized orbital basis set cc-pVTZ [62]. Hereafter we are focused on practical methods to calculate the electronic energy for the molecules with not less than

a dozen of heavy atoms; therefore, the “resolution of identity” (RI) approximations together with local-pair natural orbital (LPNO) approach [63, 64] are extensively used for calculations.

Semiempirical NDDO calculations were performed by MOPAC2016 software [65]. Most practically important methods based on the MNDO [49] formalism (AM1 [53], PM3 [66, 67], RM1 [68], PM6 [69], and PM7 [70]) are used. Calculations in the semiempirical density functional tight-binding (DFTB) formalism [71, 72] were done with DFTB+ [73, 74] software (version 17.0), together with corresponding Slater-Kostner parameterization for second-order (MIO) and third-order (3OB) binding [75, 76].

Force field calculations were performed by several packages. The sander program from AmberTools'16 suite [77] was used to calculate FF energy within the General Amber Force Field (GAFF [78, 79]) using several different charge schemes [80]: no charges, Gasteiger (G) [81, 82], DENR (D) [83], AM1-BCC (A) [84, 85], MMFF94 (M) [86]. The OpenBabel suite [87, 88] was used for molecule conversion and calculation of Gasteiger and MMFF94 charges; its in-house modification was used to implement DENR charges. The antechamber utility from AmberTools'16 was used to derive AM1-BCC charges.

Several other MM packages were also used, including version 7.1 of TINKER [89] for MM2 [90], version 8.0 of PC Model [91] for MMX, MM3 [92, 93], and MMFF94 [86]. The default values of the dielectric constant were used in the calculations, that is $\epsilon = 1.5$ for MM2, MMX, MM3, MMFF94, and $\epsilon = 1$ for GAFF.

Results and discussion

The structures of conformers were subjected to the unconstrained geometry optimization using several methods. For all **CB** conformers the C_s symmetry is found, while optimized **CC** structures are C_{2v} -symmetric. Values of conformational energies and effects, reported hereafter, are all in kcal/mol.

Structure optimization

The ab initio geometry optimization methods included HF (Fig. 6) and MP2 theory (Fig. 7). Of the latter, MP2 /

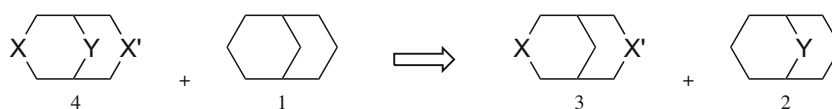


Fig. 5 The general form of the incomplete positional redistribution reaction in bicyclo[3.3.1]nonane 3,7,9-heteroanalogues

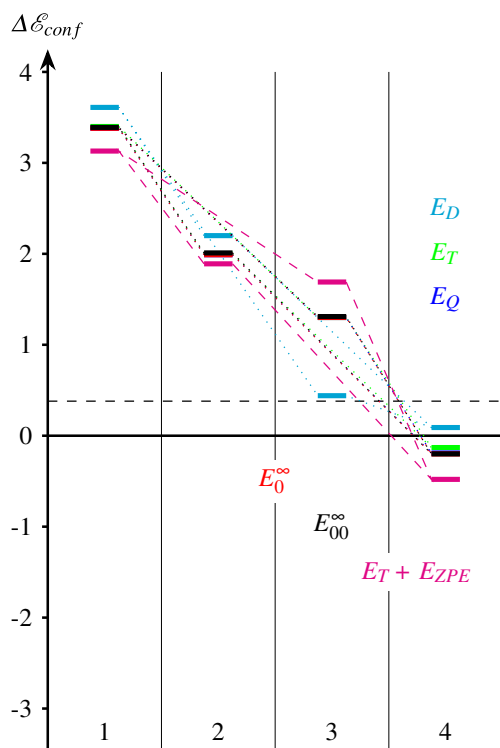


Fig. 6 The $\Delta\mathcal{E}_{conf}$ correlation in Hartree-Fock calculations, kcal/mol

cc-pVTZ is hereafter abbreviated as MP2T; E_2^T stands for the corresponding energy. MP2T is succeeded by MP2/aug-cc-pVTZ [94] (MP2A; E_2^A) and MP2/cc-pVQZ (MP2Q; E_2^Q) optimizations. The reference conformer geometries for the subsequent benchmark of the single-point energy calculations were obtained at MP2(full) level in core-polarized cc-pCVTZ basis set [95, 96] (MP2(2); $E_{2(2)}$), to account explicitly the most important core properties and effects. Single-point calculations of electronic energy in the present work are performed for the MP2(2)-optimized geometries, unless otherwise noted.

The optimal conformer is **CC** for **1** and **2**, and **CB** for **4**, while **3** has no clear conformational predominance due to the alternation of the sign of the $\Delta\mathcal{E}_{conf}$ value while changing the level of calculations. These changes are comparable with $\Delta\mathcal{E}_{conf}^0$ by their magnitude. **CC** form is most stable and hence $\Delta\mathcal{E}_{conf}$ is most positive for **1**; conversely, the **CB** form is most stable for **4** with most negative $\Delta\mathcal{E}_{conf}$. The overall order of conformational energy obtained from high-level ab initio calculations could be summarized as the series of relative **CC** form stability:

$$\Delta\mathcal{E}_{conf} : 1 > 2 > 3 > 4 \quad (3)$$

(the **CC** predominance over **CB** is decreasing in this order from **1** to **4**) or in the equivalent generalized form for $\delta\mathcal{E}'$

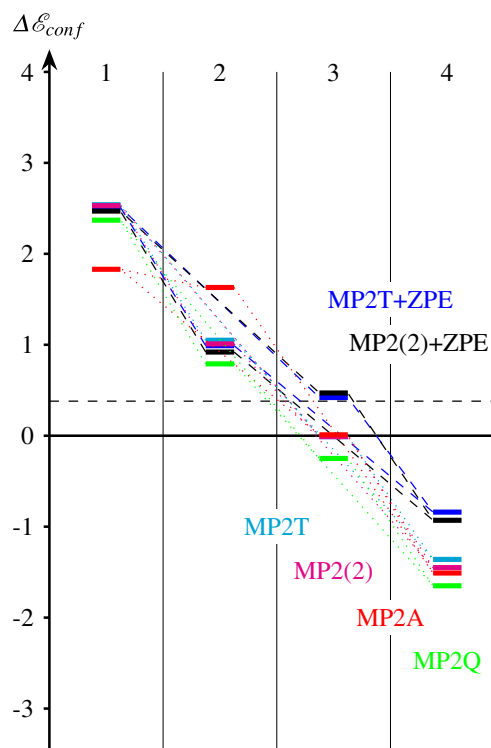


Fig. 7 The $\Delta\mathcal{E}_{conf}$ correlation from the MP2 optimization calculations, kcal/mol

and $\delta\mathcal{E}''$:

$$\delta\mathcal{E}'' < \delta\mathcal{E}' < 0 \quad (4)$$

expanding in the present case as follows:

$$\delta\mathcal{E}_H'' < \delta\mathcal{E}_C' < 0 \quad (5)$$

$$\delta\mathcal{E}_O'' < \delta\mathcal{E}_N' < 0 \quad (6)$$

$$\delta\mathcal{E}_O'' < \delta\mathcal{E}_C' < 0 \quad (7)$$

$$\delta\mathcal{E}_H'' < \delta\mathcal{E}_N' < 0 \quad (8)$$

As compared to the correlated ab initio methods, the HF calculations generally overestimate the relative stability of **CC** forms; the effect of the correlation energy account is to lower sufficiently all $\Delta\mathcal{E}_{conf}$ values.

These results are easily rationalized using the basic principles of the VSEPR theory [15], omitting the postulate that repelling electron pairs should belong to the same atom. Namely, the statement that lone electron pairs have greater steric volume than that corresponding to the covalent bonds allows the simple account for the following order of the 3,7-repulsion: **(1,2)** < **(3,4)**, i.e., the bispidine lone pair repulsion (Fig. 2c) is more energetic than that of the hydrogen atoms in C–H bond of carbocyclic compounds (Fig. 2a).

For all optimization methods, except MP2A and MP2Q, the Hessian calculations (finite differences for M06-2X, MP2T, and MP2(2), analytic elsewhere) were performed.

Results of the optimization were thus identified as PES minima in the “rigid rotor-harmonic oscillator” (RRHO) approximation, together with E_{ZPE} correction values supplied. The zero-point corrections contribute significantly to the stabilization of CC forms for compounds **3** and **4**, $\Delta\mathcal{E}_{conf}(ZPE) > 0$ for almost all cases, where E_{ZPE} is available. For **1** and **2** the $\Delta\mathcal{E}_{conf}(ZPE)$ values are generally negative and less in magnitude, so the effect of zero-point vibrations inclusion for carbobicycles is ambiguous.

Density functionals assessment

Representative, but not exhaustive, selection of density functional approximations for the structure optimization (Fig. 8) is made to span all important “rungs” of the “Jacob’s ladder”, representing the DFT approximations hierarchically [97]. Alongside with the local density approximation (LDA) in the SVWN5 form [98], a nonempirical GGA functional PBE96 [99], built upon the PW91 correlation [100], was used. Hybrid functionals include B3PW91 [101] with a fraction of 20% of non-local HF-type exchange (HFX) and PW91 correlation, B3LYP [102, 103] with 20% HFX and VWN5 local part in the LYP correlation [104], and PBE0 [105].

Additionally we employ a strongly non-local (i.e., with 54% of HFX) *meta*-GGA functional M06-2X [106], designed to overcome some of the most prominent handicaps of DFT, including the failure to properly reproduce the dispersion forces [107, 108]. For other, excluding the simplest LDA, the Becke-Johnson (BJ) corrections [109–111] were employed for the same purpose.

Recent criticism revealed several inherent flaws in the design methodology of density functionals [112], and this prevents us from further expanding the set of DFT approximations in favor of diving into analysis and creation of composite ab initio schemes (vide infra).

DFT methods commonly invert positions of **2** and **3** in the series (3) (cf. Fig. 8). Our DFT calculation results are similar in that they generally alter the ab initio conformational energy sequence. The zero-point correction account also sufficiently increases the relative stability of CB conformer. The introduction of BJ dispersion damping correction is not sufficient here since it does not alter the results somewhat significantly.

Second-order single point calculations

The complete basis set extrapolations for both HF ($E^{(n)} \rightarrow E_0^\infty$) and correlation ($\Delta E_c^{(n)} \rightarrow \Delta E_c^\infty$) energy components were obtained according to the CBS(3:4) scheme

$$E_0^\infty = E_0^{(n)} - A \cdot \exp[-\alpha n^\gamma] \quad (9)$$

$$\Delta E_c^\infty = \frac{n^\beta \cdot \Delta E_c^{(n)} - N^\beta \cdot \Delta E_c^{(N)}}{n^\beta - N^\beta} \quad (10)$$

Fitted values of $\alpha_{3:4} = 5.46$, $\gamma = \frac{1}{2}$ for Eq. 9 and $\beta_{3:4} = 3.05$ for Eq. 10 were employed for HF, MP2, LPNO-NCPF/1 and LPNO-CEPA/1 correlation using the cc-pVnZ basis sequence [113]. Additionally, the complementary auxiliary basis set (CABS) extrapolations were performed in

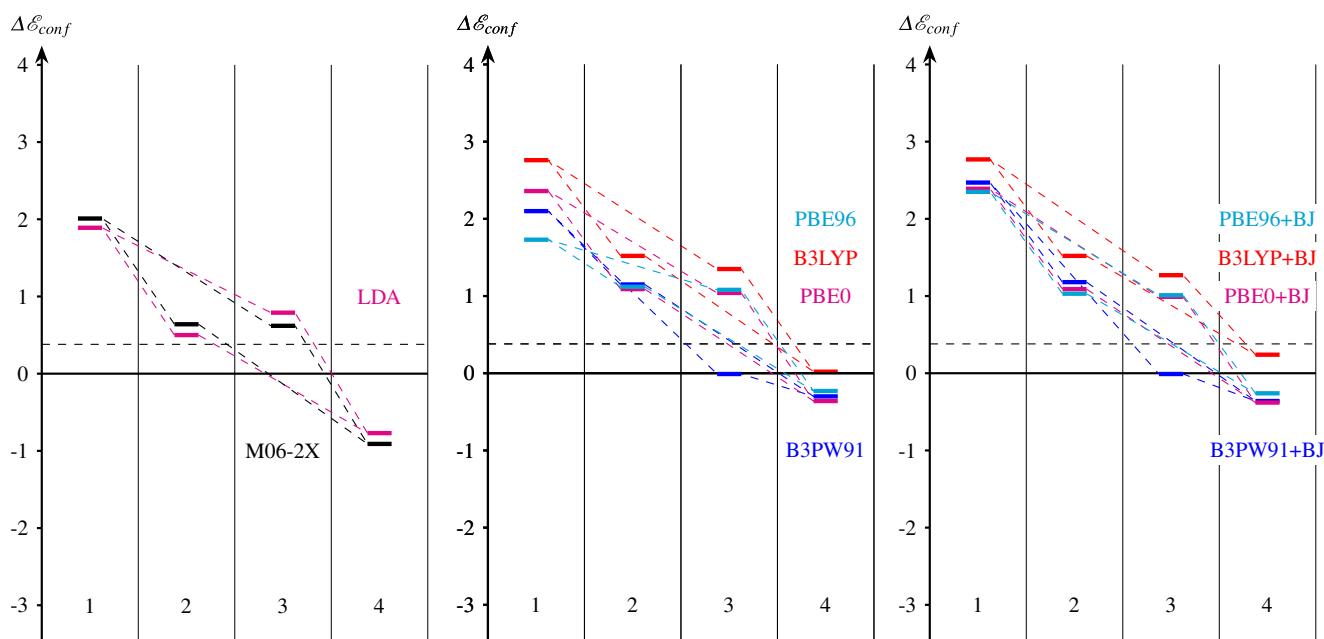


Fig. 8 The $\Delta\mathcal{E}_{conf}$ correlation from the DFT calculations, kcal/mol

the explicitly correlated MP2-F12 calculations using specialized cc-pVTZ-F12 and cc-pVQZ-F12 basis sets, abbreviated hereafter as CABS(T) and CABS(Q), respectively [114, 115]. Both CABS(T) and CABS(Q) results are expectedly almost indistinguishable and very close to those of MP2 CBS(3:4) (see Fig. 9). Of these calculations the method of choice is definitely the CABS(Q), most accurately approaching the complete basis set limits because of the increased size and flexibility of the corresponding basis set.

Higher order correlation methods

The domain-based local-pair natural orbital coupled-cluster approach DLPNO-CCSD(T) [116] is used to improve the pair-correlation results by the perturbative approach to the triple-electron correlation due to its shallow size-scaling properties, even considering the perturbative account for the triples contributions. We used it both explicitly and in evaluation of terms for the multilevel additive schemes calculations. Additive schemes (11) are declared to improve the correlation energy account over a certain complete basis approximation M through the higher-level correction ΔE_k computed commonly in a lower basis set in the extrapolation sequence [117].

$$E = \underbrace{E_0^\infty + \Delta E_M^\infty}_{E_M^\infty} + \Delta E_k + E_{ZPE} + \dots \quad (11)$$

$$E_{\varpi} = \underbrace{E_0^\infty + \Delta E_M^\infty}_{E_M^\infty(n:n+1)} + \underbrace{[\Delta E_{SD(T)}^{(n-1)} - \Delta E_M^{(n-1)}]}_{\Delta E_k(n-1)} \quad (12)$$

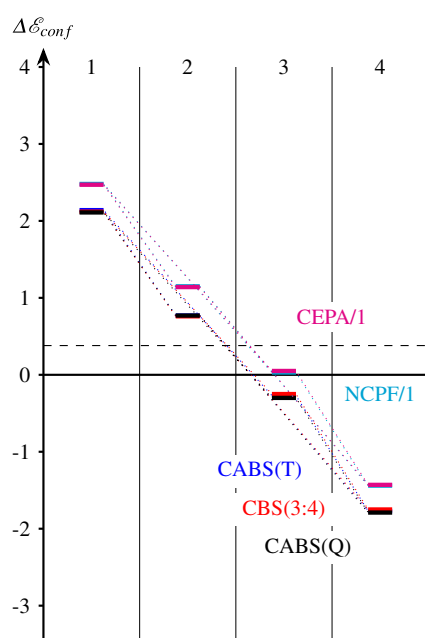


Fig. 9 The ΔE_{conf} correlation in extrapolated second-order correlated single point calculations, kcal/mol

The $EPn(M)$ schemes (12) were designed to make use of the $CBS(n:n+1)$ extrapolation for a certain lower-level correlated approximation M , and DLPNO-CCSD(T)/cc-pV $(n-1)Z$ for ΔE_k . We used EP3(LPNO-CEPA/1) scheme; LPNO-CEPA/1 was used as the intermediate correlation energy approximation M in accordance with the prior recommendations [118]. When using the scheme we found it possible to introduce several modifications into the original scheme based on our previous calculations. First, it has appeared that the realization of EP3(LPNO-CEPA/1) employed in ORCA software (version 3.0.3) directly approximate E_0^∞ with HF/cc-pVQZ results E_Q . We find it possible to use the CBS HF energy from three-point extrapolation (13), deduced from Eq. 9 with $\gamma = 1$. The result of such substitution is hereafter denoted as E_{\ddagger} . The corresponding quantity $E_{00}^\infty = (E_0^u + E_0^l)/2$, where the upper bound E_0^u is given by Eq. 13, and the lower is $E_0^l = E_0^u - (E_Q - E_0^u) = 2E_0^u - E_Q$ [119].

$$E_0^u = \frac{E_T^2 - E_D E_Q}{2E_T - E_D - E_Q} \quad (13)$$

This definition avoids the dependence upon any additional parameters, such as α in Eq. 9. The E_{00}^∞ values constructed from the results in cc-pVDZ (E_D), cc-pVTZ (E_T), and cc-pVQZ (E_Q), all available in the original EP3 scheme, are lower than any other corresponding HF limit estimate E_0^∞ , including CABS (see the Supporting Information).

Next, for the higher-level correlation part it became possible to substitute the coupled-cluster correction $\Delta E_k(D) = E_{SD(T)}^D - E_M^D$ from the cc-pVDZ basis set for the theoretically superior quantity $\Delta E_k(T) = E_{SD(T)}^T - E_M^T$ calculated in the cc-pVTZ basis. E_M^T is required for the E_M^∞ extrapolation within the original EP3(M) scheme, whereas $E_{SD(T)}^T$ is taken from our DLPNO-CCSD(T) calculations as described above. The resulted scheme is referred to hereafter as EP3' while E_{\ddagger} stands for the corresponding energy values.

To account for the chemical core, the total core-electron correction ΔE_K (14) and the generalized correction ΔE_{\varkappa} (15) are used. Here E_2^T stands for MP2 / cc-pVTZ energy of the MP2(2)-optimized geometry calculated in the procedure of CBS(3:4) extrapolation, and E_{ZPE} is calculated at the MP2(2) level.

$$\Delta E_K = E_{2(2)} - E_2^T \quad (14)$$

$$\Delta E_{\varkappa} = \Delta E_K + E_{ZPE} \quad (15)$$

These corrections are both applicable to the results of a certain additive valency electron scheme E_{ϖ} . For instance,

thus modified EP3 schemes are hereafter referred to as EP3⁺ (16) and EP3* (17).

$$E_{\varpi}^{+} = \underbrace{E_0^{\infty} + \Delta E_c^{\infty}}_{E_{\varpi}} + \Delta E_K \quad (16)$$

$$E_{\varpi}^{*} = E_{\varpi} + \underbrace{\Delta E_K + E_{ZPE}}_{\Delta E_{\infty}} \quad (17)$$

There is also an alternative way to estimate E_{tot} by E_{∇} from Eq. 18, that combines two different CBS schemes: E_{00}^{∞} from Eq. 13 and ΔE_c^{∞} from Eq. 10 using $\Delta E_{SD(T)}^D$, $\Delta E_{SD(T)}^T$ from DLPNO-CCSD(T) and fitted $\beta_{2:3} = 2.46$.

$$E_{\nabla} = E_{00}^{\infty} + \Delta E_{SD(T)}^{\infty}(2:3) \quad (18)$$

The overall accuracy of the schemes (16)–(17) based upon E_{\dagger} and E_{∇} is not the matter of investigation here. The attempts to account for the triple excitations using the coupled-cluster expansions result in the increase of the relative energy of the CC form of **3** (Fig. 10). This could sustain the assumption on the deficiency of the second-order methods for the purpose of the accurate description of strong 3,7-repulsive interactions in bispindines. The inequalities (5) and (7), that are obeyed for almost all the methods, witness the least susceptibility to the electronic effects of substitution in hydrocarbon **1**, possibly due to its least polarization and least polarizability comparing to those of **2** or **3**.

Semiempirical calculations

MNDO / NDDO calculations (Fig. 11) generally retain a nonempirical order of stability (3) with the only exception of the earlier AM1 method. However, the numerical

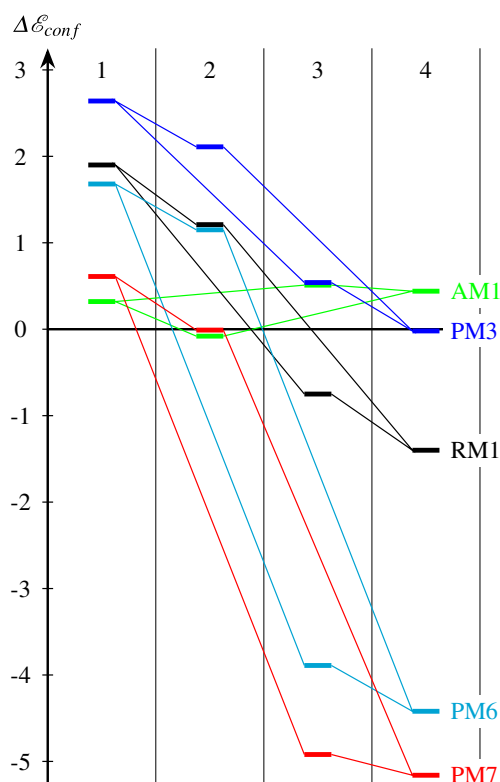
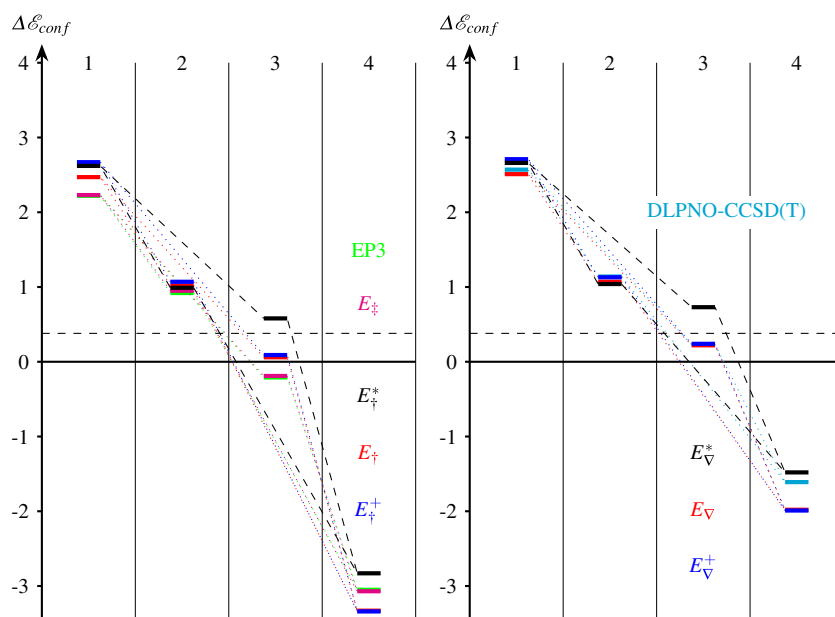


Fig. 11 The ΔE_{conf} correlation from the NDDO methods, kcal/mol

inconsistency of these results with those obtained from the first principles is significant. The best correspondence of the results to those of ab initio calculations is observed for RM1 method which is actually a deep reparameterization of AM1, obviously the worst one. The PM3 method is similar in results to more recent RM1, but further attempts

Fig. 10 The ΔE_{conf} correlation from the additive schemes based on EP3 (left) and DLPNO-CCSD(T) (right), kcal/mol



to improve semiempirical schemes lead to the significant degradation in the description quality. The values of both $\delta\mathcal{E}_H''$ and $\delta\mathcal{E}_O''$ in PM6 and PM7 are too negative, when compared to those calculated either from the first principles or using other semiempirical techniques.

The DFTB (“density functional tight binding”) results (Fig. 12) look far more erroneous than those of NDDO, because of their qualitative fail in the treatment of nitrogen heteroanalog, as from (4), numerically similar to that of MM force fields (vide infra).

Molecular mechanics

Molecular mechanics in its default implementation (Fig. 13) predicted the optimality of CC conformers for all compounds 1–4 in line with the earlier results [31, 32, 48, 54] in severe contrast to the ab initio trend expressed in (3). The satisfactory result for $\delta\mathcal{E}'$ in MM2 [90] and MMX description of 1 is more likely due to the inclusion of corresponding values in the parameterization set. Most of the $\delta\mathcal{E}''$ values are sufficiently positive in MM calculations. The deviations in $\Delta\mathcal{E}_{conf}$ predictions are assumed to be caused by the improper explicit parameterization of the important interactions, treated implicitly by high-level methods with decent accuracy.

Earlier force fields of MM were parameterized mostly using experimental data, which were relatively scarce to allow for reliable derivation of parameters for a diverse chemical space of structures, which are of interest to molecular modeling community. Contemporary classical force fields are likely parameterized against results of ab initio and DFT calculations as a reference, but different levels of QC theory being used to derive different types of MM contributions. Since for the studied structures conformational preference apparently changes with the

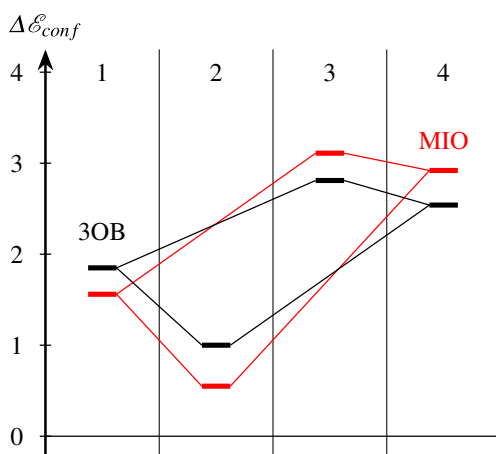


Fig. 12 The $\Delta\mathcal{E}_{conf}$ correlation from the DFTB calculations, kcal/mol

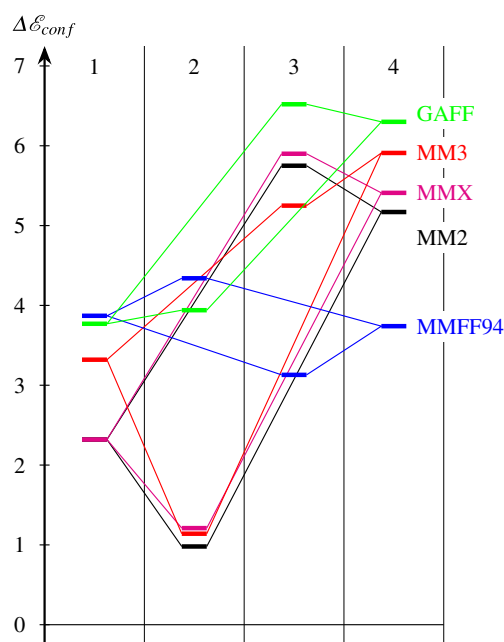


Fig. 13 The $\Delta\mathcal{E}_{conf}$ correlation from MM calculations with default electrostatics, kcal/mol

increase of polarity of their substructures, we expected that the electrostatic part of non-valency interactions should play the crucial role. A complication is that there is no unique way to derive electrostatic interaction parameters: earlier MM methods relied upon the bond dipole approximation, more suitable for non-polar to slightly polar structures, whereas more contemporary classical force fields rely generally on atomic charge approximation.

It seems that years passed do not improve the situation with force fields that should have been augmented with terms to correct the inconsistency in the description of the interactions critical for the conformational behavior of small molecules containing about a dozen of non-hydrogen organogenic atoms (i.e., C, N, O). To describe non-valency interaction three different contributions are basically involved in classical force fields: electrostatic interaction, van der Waals interactions (*dispersion*) and, implicitly, the torsional term E_τ . Historically the latter has been parameterized to minimize the residual discrepancy between the force field and reference QC description when all other MM terms have been parameterized. Thus, it implicitly includes some portion of non-valency interaction not covered by classical terms. Electrostatics dominates at longer and plays a significant role at shorter separations, while dispersion contribute more sufficiently at the distances close in magnitude to the sum of van der Waals radii of atoms.

The impact of the explicit electrostatics account

The results of General AMBER [120] Force Field (GAFF [78, 79]) shown in Fig. 13 are obtained without specifying any atomic charges. To check the influence of using different charge schemes on and to estimate the contribution of electrostatic interactions for the conformational preference, GAFF has been used with charges produced by substantially different charge schemes (Fig. 14).

First, it should be noted that while using zero charges (no electrostatic interaction approximation), the qualitatively incorrect conformational preference is predicted with the **CC** form almost completely dominating ($\Delta\mathcal{E}_{conf} = 3.5 \div 7.0$ kcal/mol). Second, the effect of charge schemes is absent for compound **1** and almost negligible for **2**, as could be expected due to the lack of space separating polar parts in them. On the contrary, the effect of charges is significant for **3** and **4** where the interplay between electrostatic attraction and repulsion had to be anticipated. Third, the increase of the degree of polarization in charge schemes in the following order: none (default), Gasteiger (G), [81, 82]

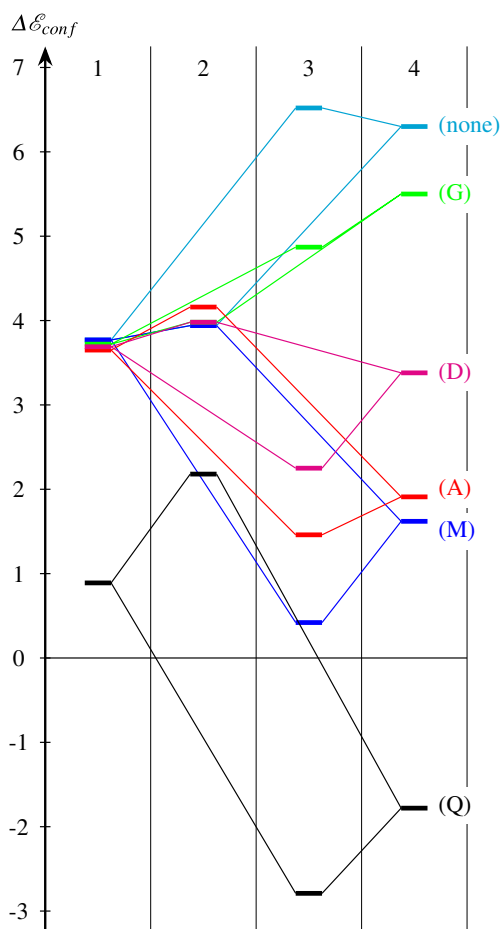


Fig. 14 The $\Delta\mathcal{E}_{conf}$ correlation from GAFF calculations with different charge schemes, kcal/mol

DENR (D) [83], AM1-BCC (A) [84, 85], MMFF (M) [86] — results in qualitative improvement of the description, tending to favor **CB** forms for the more polar structures **3** and **4**. Moreover the purely electrostatic contribution $E_q + E_q^{(1-4)}$ to the conformational preference is -5.8 , -4.8 , -4.0 , -1.7 kcal/mol for **3** and -4.6 , -4.4 , -2.9 , -0.8 kcal/mol for **4** using MMFF, AM1-BCC, DENR and Gasteiger charges, respectively. Thus, the classical electrostatic interactions correctly describe the decrease in preference of **CC** form in going from non-polar **1** and **2** to polar **3** and **4**. Also, it should be noted that the first three schemes were parameterized to reproduce the molecular electrostatic potential (MEP) of the molecules calculated at HF/6-31G(d) level and are more by default compatible with GAFF force field employing the same parameterization for its electrostatics terms. Fourth, the careful examination of relative contributions come from different MM terms to the overall values reveals that a significant handicap for the preference of **CC** form surprisingly comes from E_τ ($3.5 \div 5.3$ kcal/mol) (see the Supporting Information for the details). Thus, electrostatic atomic charge based interactions tend to overcome the initial preference for **CC** but are not sufficient at the current parameters values in GAFF. The use of Mulliken charge values obtained for the optimized MM structures at the HF/6-31G[d] level (Q) improves sufficiently the $\delta\mathcal{E}''$ reproduction, while the $\delta\mathcal{E}'$ estimates (**1**) are still evaluated to the erroneously positive values.

The increase in dielectric constant ϵ for the MM calculation would lead to further increase in calculated **CC** predominance due to more rapid cancellation of the electrostatic part of 3,7-repulsion (Coulombic $E_q \sim \epsilon^{-1}r^{-1}$), resulting in overall decrease of **CC** total energy.

The MMFF, AM1-BCC and DENR perform sufficiently better than the Gasteiger scheme; this is exactly the order of the decreasing of the polarization degree in the charge schemes. Nevertheless, the high absolute values of both $\Delta\mathcal{E}^\eta$ values show the implicit flaws of the present parameterizations, possibly, concerning the structures containing the amine nitrogen atom. The more special and thorough analysis of the classic force fields failures for the 3,7-diheterobicyclo[3.3.1]nonane derivatives is left behind the scope of the present paper and reserved to the forthcoming publications.

Conclusions

The relative energies of two least-strain conformers were evaluated using a variety of practically important computational chemistry methods for four closely related structures: bicyclo[3.3.1]nonane **1**, bicyclo[3.3.1]nonan-9-one **2**, 3,7-dimethyl-3,7-diazabicyclo[3.3.1]nonane **3**, and

3,7-dimethyl-3,7-diazabicyclo[3.3.1]nonan-9-one **4**. These compounds are known to be governed in their conformational properties by strong oriented non-valency interactions. Diagrams of conformational energy correlation were utilized to treat qualitatively the ability of different methods to model the conformational effects of substitution properly.

While ab initio calculations on the **CB** energy relative to that of **CC** result in the order $1 > 2 > 3 > 4$, DFT methods exhibit an altered sequence $1 > 3 \gtrsim 2 > 4$. To avoid the implicit basis set dependence of correlated ab initio results we relied mostly upon the complete basis set extrapolation together with the MP2 dynamic correlation (Fig. 7). The most advanced MP2-F12 CABS(Q) calculations result in 2.11, 0.77, -0.30 , and -1.79 kcal/mol (for $\Delta\mathcal{E}_{conf}$ of **1**, **2**, **3**, and **4**, respectively), attributing the slight preference to the **CB** conformation for **3**.

Due to the significance of the lone-pair interaction in conformational behavior of compounds **3** and **4** an account of triple excitations is found to be valuable for the consistency of its description. While the DLPNO framework makes the CCSD(T) calculations sufficiently more feasible, our highest level theory, combining the CBS of single-reference pair correlation function CEPA/1 with DLPNO-CCSD(T), ZPE and core corrections, gives the $\Delta\mathcal{E}_{conf}$ values 2.62, 0.99, 0.58, and -2.83 kcal/mol, while the ZPE and core-corrected CBS-extrapolated CCSD(T) calculations estimated these energies at 2.66, 1.04, 0.73, and -1.48 kcal/mol (both series are given here in the order of structures indicated above), changing the optimal conformer for **3** to **CC**. So the reliable methods of choice in the evaluation of conformational energies for the bicyclo[3.3.1]nonane derivatives and heteroanalogues are the complete basis techniques with the electron correlation account up to the double excitations or higher. While MP2-based techniques are known to account properly all doubles within the reasonable time scale [121], DLPNO-CCSD(T) methods go farther in electron correlation, but require further, more extensive evaluation. These systematic ab initio approaches are easily extrapolated to the full basis limit and look currently more preferable than DFT methods. An open question here is that of relative contributions of triples and higher excitations into the conformation energy.

There are serious discrepancies in the results between calculations from first principles and parametric schemes. Of the latter, the NDDO semiempirical methods generally retain the order of **CB** vs. **CC** relative stability, obtained from the nonempirical calculations, while DFTB and MM inevitably overestimate the stability of the **CC** conformer in all cases. These deviations could be attributed to the intrinsic deficiencies in the description of strong intramolecular non-valency interactions within the present framework of these methods. The introduction of explicit electrostatic account using charge schemes leads to the partial improvement of

the description, although there was no complete qualitative consistency observed.

The accuracy of general force fields is now seemingly not sufficient to describe the intricate balance of interaction types for such a hard case of strong non-valency interactions represented by **1-4**. Generally, this agrees well with the estimated accuracy (one standard deviation of error, σ) of the most widespread force fields to be *ca.* $1 - 2$ kcal/mol, leading to a conclusion that for about 5% of energies tested the error could be more than $2 - 4$ kcal/mol (*ca.* 2σ). The account of the electrostatic interactions even in its simplest Coulombic form of atomic point charges would sufficiently improve the description. Also the impact of the proper choice of the charge scheme to investigate the conformational preference of the molecules should be stressed. This also makes even more promising the development of atomic multipole electrostatics models [122].

The non-electrostatic MM contributions could also be fitted directly to reproduce QC results with high accuracy, but in that scenario, the generality of parameters would be sacrificed for the accuracy. The question is still open to which extent a fine-tuning of MM force constants could be done so that to reach the increase in accuracy of conformational preference prediction without losing the generality of parameters obtained. We believe that there is still a plenty of possibilities for improvement of general parameters including electrostatic, van der Waals and torsional ones, stemming from very old studies, and therefore being the obvious target for refitting.

Funding information This work was financially supported by the Russian Foundation for Basic Research (project 18-03-01065).

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Eliel EL, Wilen SH, Mander LN (1994) Stereochemistry of organic compounds. Wiley, New York. ISBN 978-0-471-01670-0
2. Eliel EL, Wilen SH, Doyle MP (2001) Basic organic stereochemistry. Wiley-VCH, New York. ISBN 978-0-471-37499-2
3. Glass RS (ed) (1989) Conformational analysis of medium-sized Heterocycles. Wiley-VCH, New York. ISBN 3-527-26789-1
4. Juaristi E (ed) (1995) Conformational behavior of six-membered rings. Analysis, dynamics and stereoelectronic effects. VCH Publishers, New York. ISBN 9780471186052
5. Case DA (1989) Theoretical and Computational Methods of Conformational Analysis in Heterocyclic Rings, chap 1, pp 1–34 in [3]

6. Eliel EL, Allinger NL, Angyal SJ, Morrison GA (1965) Conformational analysis. Interscience Publishers, New York
7. Meitei OR, Heßelmann A (2017) *J Comp Chem* 38:2500
8. Ermer O (1981) *Aspekte von Kraftfeldrechnungen* (Wolfgang Baur Verlag München)
9. Burkert U, Allinger NL (1982) ACS Monograph 177 (American Chemical Society, Washington, DC)
10. Bader RFW, Johnson S, Tang T-H, Popelier PLA (1996) *J Phys Chem* 100:15398
11. Chesnut DB (1644) *J Phys Chem A* 104(1):2000
12. Bushmarinov IS, Fedyanin IV, Lyssenko KA, Lapteva VL, Pisarev SA, Palyulin VA, Zefirov NS, Antipin MY (2738) *J Phys Chem A* 115(1):2011
13. Pisarev SA, Palyulin VA, Zefirov NS (2013) *Doklady Chemistry* 450:131
14. Gillespie RJ, Nyholm RS (1957) *Q Rev Chem Soc* 11:339
15. Gillespie RJ, Robinson EA (2005) *Chem Soc Rev* 34:396
16. Gillespie RJ, Hargittai I (2012) *The VSEPR model of molecular geometry*. Dover Publications, New York. ISBN 9780486486154
17. Zefirov NS, Palyulin VA (1991) *Top Stereochem* 20:171
18. Breuning M, Paasche A, Steiner M, Dilsky S, Gessner VH, Strohmman C, Engels B (2011) *J Mol Struct* 1005:178
19. Jeyaraman R, Avila S (1981) *Chem Rev* 81:149
20. Tomassoli I, Gündisch D (2016) *Curr Top Med Chem* 16:1314
21. Zefirov NS (1977) *Tetrahedron* 33:3192
22. Brukwicki T (1998) *J Mol Struct* 446:69
23. Potekhin KA, Levina OI, Struchkov YT, Svetlanova AM, Idrisova RS, Palyulin VA, Zefirov NS (1991) *Mendeleev Commun* 1:87
24. Palyulin VA, Grek OM, Emets SV, Potekhin KA, Lysov AE, Zefirov NS (2000) *Doklady Chemistry* 370:4
25. Alabugin IV, Bresch S, dos Passos Gomes G (2015) *J Phys Org Chem* 28:147
26. Levina OI, Potekhin KA, Kurkutova EN, Struchkov YT, Baskin II, Palyulin VA, Zefirov NS (1985) *Dokl Akad Nauk SSSR* 281:1367
27. McCabe PH, Milne NJ, Sim GA (1985) *J Chem Soc Chem Commun.* p 625
28. Eliel EL (1970) *Acc Chem Res* 3:1
29. Hoffmann R (1971) *Acc Chem Res* 4:1
30. Alabugin IV (2016) *Stereoelectronic effects: a bridge between structure and reactivity*. John Wiley & Sons, New York. ISBN 978-1-118-90634-7
31. Peters JA, Baas MA, van de Graaf B Jr, van der Toorn JM, van Bekkum H (1978) *Tetrahedron* 34:3313
32. Ishiyama J-I, Senda Y (1996) *Research Reports, Fukushima National College of Technology* 32:37
33. Osina EL, Mastryukov VS, Vilkov LV, Belikova NA (1976) *J Chem Soc Chem Commun.* pp 12–13
34. Mastryukov VS, Osina EL, Dorofeeva OV, Popik MV, Vilkov LV, Belikova NA (1979) *J Mol Struct* 52:211
35. Mastryukov VS, Popik MV, Dorofeeva OV, Golubinskii AV, Vilkov LV, Belikova NA, Allinger NL (1981) *J Am Chem Soc* 103:1333
36. Sim GA (1990) *Acta Cryst B* 46:676
37. White MA, Perrott A (1991) *Solid State Chem* 90:87
38. Mora AJ, Fitch AN (1999) *Z Kristallogr* 214:480
39. Jaime C, Osawa E, Takeuchi Y, Camps P (1983) *J Org Chem* 48:4514
40. Choo J, Kim S, Joo H, Kwon Y (2002) *J Mol Struct (THEOCHEM)* 619:113
41. Li Y-S, Li S (1989) *J Mol Struct* 213:155
42. Raber DJ, Janks CM, Johnston MD Jr, Raber NK (1980) *Tetrahedron Lett* 21:677
43. Grilli S, Lunazzi L, Mazzanti A (2000) *J Org Chem* 65:3563
44. Cremer D, Szabó KJ (1995) *Ab initio studies of six-membered rings: Present status and future development*, chap 3, pp 59–135 in [4]
45. Brouwer AM, Krijnen B (1995) *J Org Chem* 60:32
46. Toom L, Kütt A., Kaljurand I, Leito I, Ottosson H, Grennberg H, Gogoll A (2006) *J Org Chem* 71:7155
47. Comba P, Kerscher M, Schiek W (2007) In: Karlin KD (ed) *Progress in Inorganic Chemistry*, John Wiley & Sons, Inc., vol 55, pp 613–704
48. Douglass JE, Ratliff TB (1968) *J Org Chem* 33:355
49. Dewar MJS, Thiel W (1977) *J Am Chem Soc* 99:4899
50. Livant P, Roberts KA, Eggers MD, Worley SD (1981) *Tetrahedron* 37:1853
51. Galasso V, Goto K, Miyahara Y, Kovač B, Klasinc L (2002) *Chem Phys* 277:229
52. Zefirov NS, Palyulin VA, Yefimov GA, Subbotin OA, Levina OI, Potekhin KA, Struchkov YT (1991) *Dokl Akad Nauk SSSR* 320:1392
53. Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) *J Am Chem Soc* 107:3902
54. Palafox MA (1992) *J Mol Struct (THEOCHEM)* 262:21
55. Palafox MA, Boggs JE (1993) *J Mol Struct (THEOCHEM)* 285:33
56. Arias MS, Galvez E, del Castillo JC, Vaquero JJ, Chicharro J (1987) *J Mol Struct* 156:239
57. Wheeler SE, Houk KN, Schleyer PVR, Allen WD (2009) *J Am Chem Soc* 131:2547
58. Wheeler SE (2012) *WIREs: Comp Mol Sci* 2:204
59. ORCA. An ab initio, DFT and semiempirical SCF-MO package, <http://orcaforum.cec.mpg.de/> (2014–2017)
60. Neese F (2012) *WIREs: Comp Mol Sci* 2:73
61. He Y, Cremer D (2000) *Theor Chem Acc* 105:110
62. Dunning TH Jr (1989) *J Chem Phys* 90:1007
63. Wennmohs F, Neese F (2008) *Chem Phys* 343:217
64. Riplinger C, Sandhoefer B, Hansen A, Neese F (2013) *J Chem Phys* 139:134101
65. Stewart JJP (2016) MOPAC2016. A semiempirical SCF-MO package, <http://www.openmopac.net/>
66. Stewart JJP (1989a) *J Comp Chem* 10:209
67. Stewart JJP (1989b) *J Comp Chem* 10:221
68. Rocha GB, Freire RO, Simas AM, Stewart JJP (2006) *J Comp Chem* 27:1101
69. Stewart JJP (2007) *J Mol Model* 13:1173
70. Stewart JJP (2013) *J Mol Model* 19:1
71. Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G (1998) *Phys Rev B* 58:7260
72. Wahiduzzaman M, Oliveira AF, Philipsen P, Zhechkov L, van Lenthe E, Witek HA, Heine T (2013) *J Chem Theory Comput* 9:4006
73. (2017). DFTB+ package, <http://dftbplus.org>
74. Aradi B, Hourahine B, Frauenheim T (2007) *J Phys Chem A* 111:5678
75. Gaus M, Cui Q, Elstner M (2011) *J Chem Theory Comput* 7:931
76. Gaus M, Goez A, Elstner M (2013) *J Chem Theory Comput* 9:338
77. Case DA, Betz RM, Cerutti DS, Cheatham TE III, Darden TA, Duke RE, Giese TJ, Gohlke H, Goetz AW, Homeyer N et al (2016) AMBER 2016, <http://www.ambermd.org>
78. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) *J Comp Chem* 25:1157
79. Wang J, Wang W, Kollman PA, Case DA (2006) *J Mol Graph Model* 25:247
80. Oliferenko AA, Pisarev SA, Palyulin VA, Zefirov NS (2006) In: *Adv. Quantum Chem.* Elsevier, vol 51, pp 139–156
81. Gasteiger J, Marsili M (1978) *Tetrahedron Lett* 19:3181
82. Gasteiger J, Marsili M (1980) *Tetrahedron* 36:3219

83. Shulga DA, Oliferenko AA, Pisarev SA, Palyulin VA, Zefirov NS (2008) SAR QSAR Env Res 19:153
84. Jakalian A, Bush BL, Jack DB, Bayly CI (2000) J Comp Chem 21:132
85. Jakalian A, Jack DB, Bayly CI (2002) J Comp Chem 23:1623
86. Halgren TA (1996) J Comp Chem 17:490
87. OpenBabel. The open source chemistry toolbox, <http://www.openbabel.org/> (2011a)
88. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) J Cheminform 3:33
89. TINKER. Software tools for molecular design, <http://dasher.wustl.edu/tinker/> (2015)
90. Allinger NL (1977) J Am Chem Soc 99:8127
91. PC Model. Molecular modeling software for personal workstations, <http://www.serenasoft.com/> (2011b)
92. Lii J-H, Allinger NL (1989) J Am Chem Soc 111:8566
93. Allinger NL, Yuh YH, Lii J-H (1989) J Am Chem Soc 111:8551
94. Kendall RA, Dunning TH Jr, Harrison RJ (1992) J Chem Phys 96:6796
95. Woon DE, Dunning TH Jr (1995) J Chem Phys 103:4572
96. Peterson KA, Dunning TH (2002) J Chem Phys 117:10548
97. Perdew JP, Ruzsinszky A, Tao J, Staroverov VN, Scuseria GE, Csonka GI (2011) J Chem Phys 123(06):2005
98. Vosko SH, Wilk L, Nusair M (1980) Can J Phys 58:1200
99. Perdew JP, Burke K, Ernzerhof M (1996) Phys Rev Lett 77:3865
100. Perdew JP, Wang Y (1992) Phys Rev B 45:13244
101. Becke AD (1993a) J Chem Phys 98:1372
102. Becke AD (1993b) J Chem Phys 98:5648
103. Hertwig RH, Koch W (1997) Chem Phys Lett 268:345
104. Lee C, Yang W, Parr RG (1988) Phys Rev B 37:785
105. Adamo C, Barone V (1999) J Chem Phys 110:6158
106. Zhao Y, Truhlar DG (2008a) Theor Chem Acc 120:215
107. Zhao Y, Truhlar DG (2008b) Acc Chem Res 41:157
108. Cohen AJ, Mori-Sánchez P, Yang W (2012) Chem Rev 112:289
109. Becke AD, Johnson ER (2005a) J Chem Phys 122:154104
110. Becke AD, Johnson ER (2005b) J Chem Phys 123:154101
111. Becke AD, Johnson ER (2006) J Chem Phys 124:014104
112. Medvedev MG, Bushmarinov IS, Sun J, Perdew JP, Lyssenko KA (2017) J Sci 355:49
113. Neese F, Hansen A, Wennmohs F, Grimme S (2009) Acc Chem Res 42:641
114. Neese F, Valeev EF (2011) J Chem Theory Comput 7:33
115. Liakos DG, Izsák R, Valeev EF, Neese F (2013) Mol Phys 111:2653
116. Riplinger C, Neese F (2013) J Chem Phys 138:034106
117. Tajti A, Szalay PG, Császár AG, Kállay M, Gauss J, Valeev EF, Flowers BA, Vázquez J, Stanton JF (1999) J Chem Phys 121(1):2004
118. Liakos DG, Neese F (2012) J Phys Chem A 116:4801
119. Schmidt MW, Ruedenberg K (1979) J Chem Phys 71:3951
120. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) J Comp Chem 26:1668
121. Cremer D (2011) WIREs: Comp Mol Sci 1:509
122. Shulga DA, Titov OI, Pisarev SA, Palyulin VA (2018) SAR QSAR Env Res 29:21

Affiliations

Sergey A. Pisarev¹ · Dmitry A. Shulga² · Vladimir A. Palyulin^{1,2} · Nikolay S. Zefirov^{1,2}

Sergey A. Pisarev
 serge@qsar.chem.msu.ru

¹ Institute of Physiologically Active Compounds, Russian Academy of Sciences, 1 Severny proezd, Chernogolovka, Moscow Region, 142432 Russia

² Department of Chemistry, Lomonosov Moscow State University, 1/3 Leninskie Gory, Moscow, 119991 Russia