CROATICA CHEMICA ACTA CCACAA **66** (1) 113–121 (1993)

ISSN 0011–1643
CCA–2122

Original Scientific Paper

Ab Initio Electronic Structure Calculations of Molecular Similarity: A Case Study of 4-Aminobutyric Acid and Its Agonists*

*Jerzy Cioslowski*** *and Eugene D. Fleischmann#*

Department of Chemistry and Supercomputer Computations Research Institute, Florida State University, Tallahassee, FL 32306–3006, USA

Received July 13, 1992

A new quantum mechanical measure of molecular similarity, based on the overlap between the first-order density matrices is applied in studies of the structure-activity relationships in the $GABA_A$ agonists: *trans*–4–aminocrotonic acid, 5–aminomethylisoxazol–3–ol, and 3–aminomethylisoxazol–5–ol. The geometries of these molecules are optimized at the HF/6–31G** level and their electronic structures compared to that of 4–aminobutyric acid. Factors affecting the GABAergic activity of these substances are discussed. The present study demonstrates that *ab initio* electronic structure calculations of molecular similarity are feasible for medium-sized molecules of biological interest.

INTRODUCTION

Interaction between the ligand molecule and the respective receptor is the molecular event which is responsible for the cellular response that ultimately gives rise to the physiological action of a drug. Were the details of the structure of a given receptor known, one could in principle predict the affinity of any substance that binds to that receptor. However, in most cases our knowledge is limited to the structures of several substances that induce a particular physiological action. There is usually one substance present in the living organism that is targeted by the cellular receptor. The other substances, usually exogenous, can exert either agonistic or antagonistic action depending on their similarity to the endogenous ligand. The agonists are substances that are

* This paper is dedicated to the brave citizens of the city of Dubrovnik that witnessed one of the most barbaric attacks by the Serbian hordes.

** Author to whom correspondence should be addressed.
# Presently at Cray Research Inc.

similar enough to the target molecule to cause the same physiological action. On the other hand, the antagonists possess lesser similarities that are sufficient to make them bind to the receptor, but not to trigger the response. Therefore, when present in sufficient concentration, antagonists' molecules are capable of blocking the physiological action of the endogenous ligand.

In light of these facts, one quickly realizes that the ability to assess the molecular similarity is of primary importance to a rational drug design. The structure-activity relationships (SAR) that are usually invoked in this context[1] require physical basis for the criteria used in the comparison. This requires either explicit knowledge (which is rarely available) or an assumption (which is prone to arbitrariness) about the structure of a particular receptor. Another possibility, which is explored in this paper, is to look for both the similar and dissimilar molecular fragments and, by correlating their presence (or absence) in a series of molecules with their agonistic or antagonistic properties, to identify the fragments responsible for affinity to a particular receptor.

While searching for the similar fragments one has to remember that they do not have to be necessarily composed of the same nuclei. Rather, they should have similar electron distributions that determine their »shape« perceived by the receptor. Taking this observation into account, the need for a practical and consistent method of assessing molecular similarity becomes obvious.

The recent progress in computational quantum chemistry has made it feasible to calculate the electronic structures of molecules of medium and large sizes. In principle, one may draw conclusions about the similarity of two molecules by comparing their wavefunctions, since they carry all the necessary information. Several definitions of the molecular similarity are possible.[2] Recently, Carbo *et al.*[2b] proposed a similarity index which is based on the overlap between the electron densities of two molecules, $\rho_A(\vec{r})$ and $\rho_A(\vec{r})$:

$$R_{AB} = \int \rho_A(\vec{r})\,\rho_B(\vec{r})\,d\vec{r} \Big/ \left[\int \rho_A^2(\vec{r})\,d\vec{r} \int \rho_B^2(\vec{r})\,d\vec{r}\right]^{1/2}. \tag{1}$$

The use of R_{AB} is in practice hindered by the fact that calculations of the necessary four-center integrals are quite expensive. This expense precludes optimization of the mutual orientation of the molecules necessary for a unique definition of the molecular similarity. Also, the index R_{AB} was found to be very sensitive to the overlap between the respective nuclei in the molecules under comparison.[3] The index R_{AB} assumes the values between 0 and 1. This renders it useless for distinguishing between systems that possess small similar active sites with the remaining parts being dissimilar, and systems that possess a large dissimilar active site with the remaining parts similar to the target molecule. This is so, because in both comparisons the contributions to R_{AB} from the dissimilar parts yield a low value for the index, even if only the former systems are similar in the active site regions. It is the denominator in Eq. (1) that makes predictions of the affinity based on interpretation of R_{AB} quite difficult.

In a recent series of papers,[4] we have proposed a new measure of molecular similarity which is free from these difficulties. The similarity index, called the number of overlapping electrons (NOEL) is given by:

$$(A, B) = \iint \Gamma_A^*(\vec{x}, \vec{x}')\,\Gamma_B(\vec{x}, \vec{x}')\,d\vec{x}\,d\vec{x}', \tag{2}$$

where Γ_A and Γ_B are the first-order density matrices[5] of the molecules A and B, respectively. For the Hartree-Fock wavefunctions that dominate large-scale quantum-mechanical calculations, Eq. (2) reduces to

$$(A, B) = 2 \sum_{ij} | \langle \phi_{Ai} | \phi_{Bj} \rangle |^2 , \quad (3)$$

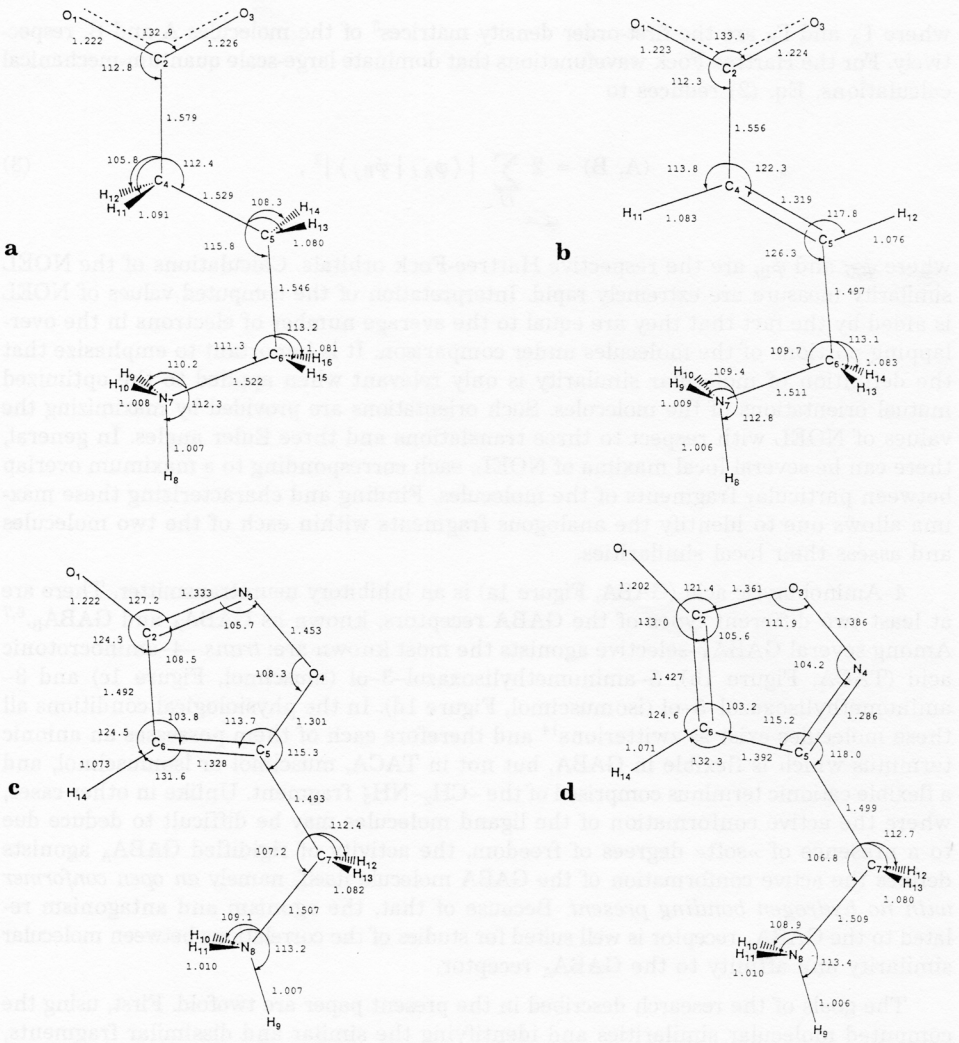
where ϕ_{Ai} and ϕ_{Bj} are the respective Hartree-Fock orbitals. Calculations of the NOEL similarity measure are extremely rapid. Interpretation of the computed values of NOEL is aided by the fact that they are equal to the average number of electrons in the overlapping portions of the molecules under comparison. It is important to emphasize that the definition of molecular similarity is only relevant when applied to the optimized mutual orientations of the molecules. Such orientations are provided by maximizing the values of NOEL with respect to three translations and three Euler angles. In general, there can be several local maxima of NOEL, each corresponding to a maximum overlap between particular fragments of the molecules. Finding and characterizing these maxima allows one to identify the analogous fragments within each of the two molecules and assess their local similarities.

4–Aminobutyric acid (GABA, Figure 1a) is an inhibitory neurotransmitter. There are at least two different types of the GABA receptors, known as $GABA_A$ and $GABA_B$.[6,7] Among several $GABA_A$–selective agonists the most known are: *trans*–4–aminocrotonic acid (TACA, Figure 1b), 5–aminomethylisoxazol–3–ol (muscimol, Figure 1c) and 3–aminomethylisoxazol–5–ol (isomuscimol, Figure 1d). In the physiological conditions all these molecules exist as zwitterions[14] and therefore each of them possesses an anionic terminus which is flexible in GABA, but not in TACA, muscimol or isomuscimol, and a flexible cationic terminus comprised of the $-CH_2-NH_3^+$ fragment. Unlike in other cases, where the active conformation of the ligand molecules may be difficult to deduce due to a presence of »soft« degrees of freedom, the activity of rigidified $GABA_A$ agonists defines the active conformation of the GABA molecule itself, namely *an open conformer with no hydrogen bonding present.* Because of that, the agonism and antagonism related to the $GABA_A$ receptor is well suited for studies of the correlations between molecular similarity and affinity to the $GABA_A$ receptor.

The goals of the research described in the present paper are twofold. First, using the computed molecular similarities and identifying the similar and dissimilar fragments, we attempt to identify the structural and electronic features in the GABA agonists responsible for their affinity to the $GABA_A$ receptor. Second, using the GABA and its agonists as an example, we demonstrate how the NOEL similarity measure can be applied to systems of biological importance and how it may be used in a systematic way to determine the active sites of the biologically active molecules.

DETAILS OF COMPUTATIONS

The molecular geometries of the zwitterionic forms of GABA, TACA, muscimol and isomuscimol were fully optimized within C_S symmetry at the HF/6–31G** level.[8] The minimized energies, computed with the GAMESS package,[9] and the respective dipole moments are presented in Table I. The first-order density matrices were computed with the GAUSSIAN 88 suite of programs.[10]



Torsional angles							
a		**b**		**c**		**d**	
9–7–6–5	58.2	9–7–6–5	58.2	10–8–7–5	–58.1	10–8–7–5	–57.8
10–7–6–5	–58.2	10–7–6–5	–58.2	11–8–7–5	58.1	11–8–7–5	57.8
11–4–2–3	–123.1	13–7–6–5	118.2	12–7–5–4	62.0	12–7–5–4	62.5
12–4–2–3	123.1	14–7–6–5	–118.2	13–7–5–4	–62.0	13–7–5–4	–62.5
13–5–4–2	–57.2						
14–5–4–2	57.2						
15–6–5–4	118.0						
16–6–5–4	–118.0						

Figure 1. Optimized geometries of: a) GABA, b) TACA, c) muscimol, d) isomuscimol.

For every pair of molecules under comparison there are several optimal mutual orientations corresponding to local maxima of the NOEL similarity index. The optimization were carried out with the NOEL program, which is available from one of the authors (J.C.) upon request. Computation of the reported values of NOEL, which were converged to 10^{-10}, took about 2 hours of CPU time per single optimization on a VAX 3100 workstation.

TABLE I

*HF/6—31G** energies and dipole moments at the optimized geometries of GABA and its agonists*

Molecule	Energy / au	Dipole moment / D
GABA	–360.812971	19.481
TACA	–359.628978	19.705
muscimol	–413.421349	20.011
isomuscimol	–413.457521	17.344

RESULTS AND DISCUSSION

As already mentioned in the Introduction, in studies of structure-activity relationships a proper selection of a particular molecular geometry is often complicated by the presence of several »soft« degrees of freedom that result in relatively flat potential energy hypersurfaces. The fully optimized geometries, which describe molecules *in vacuo*, may differ significantly from the geometries of molecules in solution or in the receptor-ligand complex. For example, the GABA molecule has soft modes corresponding to rotations around four single carbon-carbon bonds. At the HF/STO–3G level of theory, the isolated zwitterion molecule of GABA is predicted[11] to form a pseudo six-membered ring composed of the atoms O_1, C_2, C_4, C_5, C_6 and H_{15} (Figure 1a). Existence of this hydrogen-bonded conformation is conceivable in the gas phase. However, such a structure bears little resemblance to the molecules of TACA, muscimol, or isomuscimol which have only the $-CH_2-NH_3^+$ fragment flexible and therefore cannot form cyclic structures, yet are potent $GABA_A$ agonists.[12,13] Lacking more detailed information on the geometries of the receptor-GABA (agonists) complex, we enforced C_S molecular symmetry in our calculations. Although to some degree arbitrary, this assumption is as reasonable and unbiased as possible under the circumstances. Within the C_S symmetry constraints all bond lengths, bond angles and dihedral angles were optimized. The optimized geometries are displayed in Figures 1a–1d.

TABLE II

Relative GABA receptor affinities of GABA and its agonists

Molecule	Affinity[a]
GABA	1.0
TACA	0.41
muscimol	5.5
isomuscimol	0.0011

[a] Relative to GABA.[11]

The affinities of GABA and its agonists to the $GABA_A$ receptor span a range of *ca.* 5000 (Table II). The affinities do not correlate with the degree of the molecular rigidity. This indicates that in this case the conformational factors play only a secondary role in determining the strength of binding to the receptor. In order to determine which molecular fragment is responsible for the GABAergic activity, one has to carefully correlate the experimental affinities with molecular similarities. As mentioned above, each of the molecules considered in this paper is composed of a cationic terminus, a connecting structure, and an anionic terminus. For the GABA/TACA and muscimol/isomuscimol pairs, optimization of NOEL result in the same final relative orientations regardless of the starting points. This means that in each of these two cases the geometrical similarity of the molecules under comparison makes simultaneous matching of the anionic and cationic termini possible. This is well illustrated in Figures 2a and 2f in which we display overlap of the respective molecular skeletons in their optimized orientations. This almost perfect overlap results in large values of NOEL (49.645 for GABA/TACA pair and 56.570 for the muscimol/isomuscimol pair, Tables III and IV) and makes it impossible to assess the similarities between the anionic and cationic termini alone.

When the NOEL similarity measure is maximized starting with the mutual orientation corresponding to a maximal overlap of the cationic termini, the optimized values of NOEL are almost constant in the series comprised of the GABA/muscimol, GABA/isomuscimol, TACA/muscimol, and TACA/isomuscimol pairs (Table III). For each of these pairs, the mismatch between the respective anionic termini is so great that the small differences in values of NOEL essentially reflect the small differences in the degree of similarity of the cationic termini. The values of 37.172–37.453 correspond roughly to the 36 electrons in the $C–CH–C–CH_2–NH_3^+$ present in both GABA and its agonists together with some residual overlap. Should GABAergic activity depend on the details of the electronic structure of the cationic terminus, one could expect little variation in it. Therefore, one can rule out the cationic terminus as the active site of GABA and its agonists.

TABLE III

Values of NOEL among GABA and its agonists (matching of the cationic termini)

	GABA	TACA	muscimol	isomuscimol
GABA	56.000			
TACA	49.645	54.000		
muscimol	37.417	37.453	60.000	
isomuscimol	37.408	37.172	56.570	60.000

TABLE IV

Values of NOEL among GABA and its agonists (matching of the anionic termini)

	GABA	TACA	muscimol	isomuscimol
GABA	56.000			
TACA	49.645	54.000		
muscimol	36.267	35.729	60.000	
isomuscimol	35.657	34.671	56.570	60.000

On the other hand, the optimized values of NOEL corresponding to overlap of the anionic termini exhibit quite a different pattern (Table IV). In this case, the substantial mismatch in the cationic termini (Figures 2b–2e) in the GABA/muscimol, GABA/isomuscimol, TACA/muscimol and TACA/isomuscimol pairs means that the corresponding values of NOEL measure the degrees of similarity of the anionic termini alone. For both the GABA and TACA molecules, the similarity to the muscimol molecule is higher than to the isomuscimol one. This is reflected by the corresponding NOEL which are higher by 0.6–1.0 for the pairs involving muscimol. Since the molecular geometries of muscimol and isomuscimol are quite similar, this trend indicates *electronic* mismatch between isomuscimol and the other molecules. The same conclusion was reached previously[14] from analysis of the Mulliken atomic charges. One should point out, however,

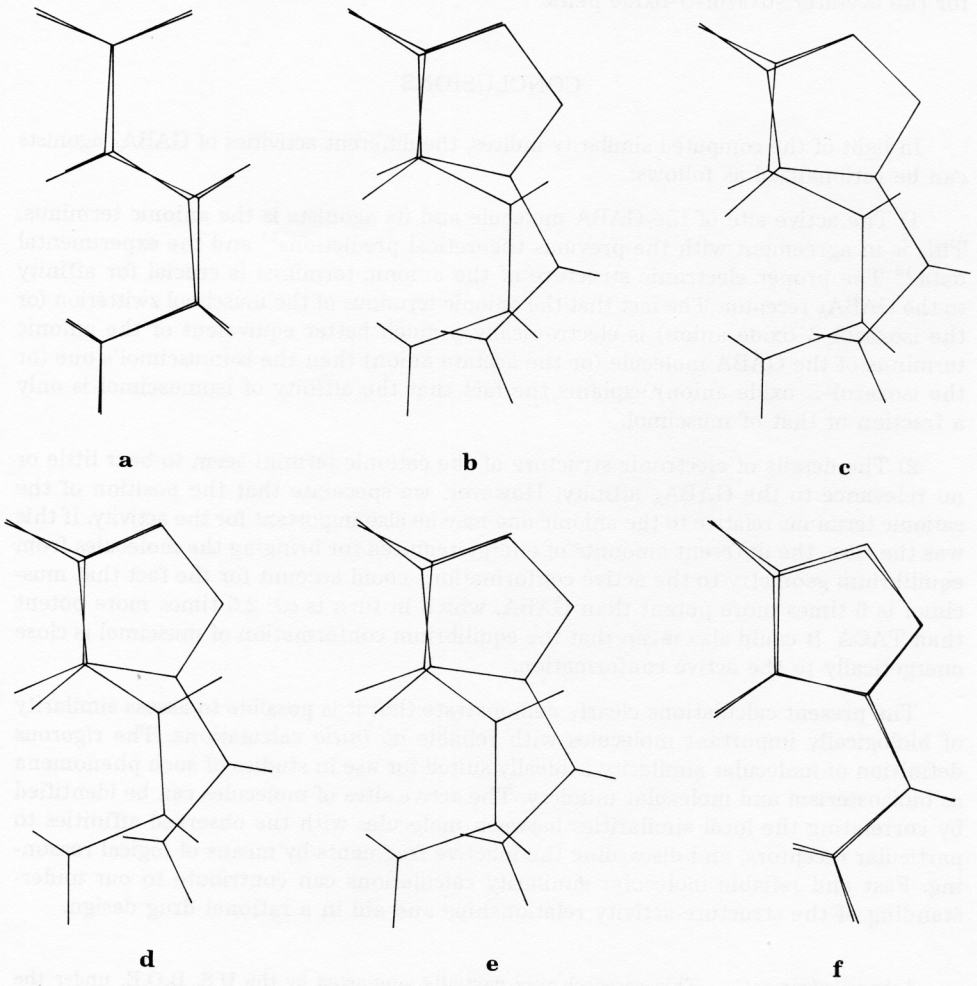


Figure 2. a) GABA-TACA, b) GABA-muscimol, c) GABA-isomuscimol, d) TACA-muscimol, e) TACA-isomuscimol, f) muscimol-isomuscimol.

that the present analysis is more rigorous since it uses a well defined definition of molecular similarity. The Mulliken atomic charges has been demonstrated repeatedly to bear little or no relation to the actual electron distribution.[15]

To elaborate further the above observations and to prove their independence from the assumptions on the molecular geometries, we computed the similarity indexes for the truncated analogs of GABA, muscimol and isomuscimol, namely the acetate, isoxazol–3–oxide and isoxazol–5–oxide anions. These molecules, which possess only the anionic termini, are much less flexible than the corresponding molecules of GABA and its agonists. Using geometries optimized without any constraints, we found the same electronic mismatch as in the case of GABA agonists. The mismatch is indicated by the values of NOEL equal to 28.480 for the acetate/isoxazol–3–oxide and only 27.786 for the acetate/isoxazol–5–oxide pairs.

CONCLUSIONS

In light of the computed similarity indices, the different activities of $GABA_A$ agonists can be rationalized as follows:

1) The active site of the GABA molecule and its agonists is the anionic terminus. This is in agreement with the previous theoretical predictions[14] and the experimental data.[16] The proper electronic structure of the anionic terminus is crucial for affinity to the $GABA_A$ receptor. The fact that the anionic terminus of the muscimol zwitterion (or the isoxazol–3–oxide anion) is electronically a much better equivalent of the anionic terminus of the GABA molecule (or the acetate anion) than the isomuscimol's one (or the isoxazol–5–oxide anion) explains the fact that the affinity of isomuscimol is only a fraction of that of muscimol.

2) The details of electronic structure of the cationic termini seem to bear little or no relevance to the $GABA_A$ affinity. However, we speculate that the position of the cationic terminus relative to the anionic one may be also important for the activity. If this was the case, the different amounts of energy required for bringing the molecules from equilibrium geometry to the active conformations could account for the fact that muscimol is 6 times more potent than GABA, which in turn is *ca.* 2.5 times more potent than TACA. It could also mean that the equilibrium conformation of muscimol is close energetically to the active conformation.

The present calculations clearly demonstrate that it is possible to assess similarity of biologically important molecules with reliable *ab initio* calculations. The rigorous definition of molecular similarity is ideally suited for use in studies of such phenomena as bioisosterism and molecular mimicry. The active sites of molecules can be identified by correlating the local similarities between molecules with the observed affinities to particular receptors, and discarding the inactive fragments by means of logical reasoning. Fast and reliable molecular similarity calculations can contribute to our understanding of the structure-activity relationships and aid in a rational drug design.

Acknowledgement. — This research was partially supported by the U.S. D.O.E. under the contract DE–FC05–85ER250000 and by the Florida State University through time granted on its VAX and ETA10–G digital computers. One of the authors (J.C.) also acknowledges support by the Camille and Henry Dreyfus Foundation New Faculty Award Program.

REFERENCES

1. W. G. Richards, *Quantum Pharmacology*, Butterworths, London, 1977.
2a. For a recent review see: *Concepts and Applications of Molecular Similarity*, M. A. Johnson and G. Maggiora, Eds., John Wiley, New York, 1990.
2b. R. Carbo, L. Leyda and M. Arnau, *Intern. J. Quant. Chem.* **17** (1980) 1185
3. P. E. Bowen-Jenkins and W. G. Richjards, *Intern. J. Quant. Chem.* **30** (1986) 763.
4. J. Cioslowski and E. D. Fleischmann, *J. Amer. Chem. Soc.* **113** (1991) 64; J. Cioslowski, *J. Amer. Chem. Soc.* **113** (1991) 6756.
5. E. R. Davidson, *Reduced Density Matrices in Quantum Chemistry* Academic Press, New York, 1976.
6. S. J. Enna, *The GABA Receptors*, Humana Press, Clifton, NJ, 1983.
7. P. Krogsgaard-Larsen, L. Nielsen, L. Falch, and F. S. Jørgensen, in: *Neurotox '88, Molecular Basis of Drug and Pesticide Action*, G. G. Lunt, Ed., Excerpta Medica, New York, 1988, p. 145.
8. P. C. Hariharan and J. A. Pople, *Theor. Chim Acta* **28** (1973) 213.
9. Original version by: M. Dupuis, D. Spangler and J. J. Wendoloski, NRCC Program QG01, 1980, ETA10–G conversion by D. Moncrieff and E. D. Fleischmann.
10. *GAUSSIAN 88*, M. J. Frisch, M. Head-Gordon, H. B. Schlegel, K. Raghavachari, J. S. Binkley, C. Gonzales, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, L. M. Fluder, S. Topiol, and J. A. Pople, GAUSSIAN, Inc., Pittsburgh PA.
11. T. Boulander, D. P. Vercauteren, F. Durant, and J. M. Andre, *J. Theor. Biol.* **127** (1987) 479 and the references cited therein.
12. P. Krogsgaard-Larsen and Roldskov-Christiansen, *Eur. J. Med. Chem. Chim. Ther.* **14** (1979) 157 and the references cited therein.
13. G. A. R. Johnston, D. R. Curtis, P. M. Beart, C. J. A. Game, R. M. McCulloch, and B. Twitchin, *J. Neurochem.* **24** (1975) 157 and the references cited therein.
14. T. Boulander, D. P. Vercauteren, F. Durant, and J. M. Andre, *Int. J. Quant. Chem.* **S15** (1988) 149.
15. J. Cioslowski, P. J. Hay, and J. P. Ritchie, *J. Phys. Chem.* **94** (1990) 148.
16. P. R. Schofield, M. G. Darlison, N. Fujita, D. R. Burt, F. A. Stephenson, H. Rodriguez, L. M. Rhee, J. Ramachandran, V. Reale, T. A. Glenscore, P. H. Seeburg, and E. A. Barnard, *Nature* **328** (1987) 221.

SAŽETAK

Ab initio računi sličnosti molekula: studij slučaja 4–aminomaslačne kiseline i njenih agonista

Jerzy Cioslowski i Eugene D. Fleischmann

Nova kvantno-kemijska mjera sličnosti molekula, zasnovana na prekrivanju među matricama gustoće prvog reda, primijenjena je u proučavanju odnosa strukture i aktivnosti slijedećih $GABA_A$ agonista: *trans*–4–aminolimunske kiseline, 5–aminometilizooksazol–3–ol, i 3–aminometilizooksazol–5–ol. Geometrija tih molekula optimizirana je na HF/6–31G** razini, a njihova elektronska struktura usporedena je s onom 4–aminomaslačne kiseline. Razmotreni su činitelji koji utječu na GABA-orgičnu aktivnost ovih tvari. Studija pokazuje da su *ab initio* računi molekulske sličnosti dostupni za biološki zanimljive molekule srednje veličine.