CROATICA CHEMICA ACTA CCACAA 57 (5) 1065—1074 (1984)

YU ISSN 0011-1643
CCA-1511 UDC 541
Original Scientific Paper

The Molecular Electrostatic Potential as a Determinant of Receptor-Drug Recognition

Dušan Hadži, Milan Hodošček, Darko Kocjan, Tomaž Šolmajer, and Franc Avbelj

Lek — Pharmaceutical and Chemical Works and Boris Kidrič Institute of Chemistry, Ljubljana, Yugoslavia

Received March 23, 1984

The validity of the concept of the molecular electrostatic potential and its applicability in rationalizing drug-receptor interactions are discussed on hand of examples covering the qualitative aspects. The computational methods are briefly reviewed with respect to economy and quality of results.

INTRODUCTION

Molecular recognition is the prerequisite for the organization and functioning of biological systems. From the point of view of the theoretical chemist the treatment of the recognition is in essence reduced to the calculation of the energy of interaction between two molecules or between the relevant parts of macromolecules, *e. g.* the recognition site of an enzyme and the reacting part of the substrate. The entropy factor is usually not explicitely considered in this approach. For the understanding of the mechanism of recognition the analysis of the interaction energy into components such as the classical electrostatic, polarization, charge transfer, dispersion, *etc.* is important[1]. It is also helpful in the search for the molecular determinants of recognition that are needed in the design of molecules with predetermined properties. For the application of the full potential of the quantum methods developed in the study of intermolecular complexes some knowledge of the molecular structure is necessary. However, in one group of recognition phenomena the knowledge of the structure is lacking for one of the components of the complex. These are the receptors for hormones, neurotransmitters and drugs (we shall call them briefly ligands) involved in the regulation of various physiological processes. This fact throws the weight of the search of molecular determinants of recognition entirely onto the side of the small molecules with known structure that are recognized (or not) by the receptor. The task is not trivial since molecules having similar characteristics with respect to receptor recognition may be of quite dissimilar chemical structure. Although the established procedures for the theoretical investigations of the interactions between molecules with known structure cannot be directly applied to ligand-receptor interactions the results obtained on the former are used in developing model representations of the latter from which the search for the molecular determinants of recognition starts. After Scrocco and Tomasi had revived[2] the interest in the electrostatic

model of intermolecular interactions by demonstrating the parallels between the SCF energy and that arising from the unperturbed charges of interacting molecules the concept of the molecular electrostatic potential (MEP) is being increasingly used in biomolecular problems. For instance, it was applied in interpreting the reactivity of nucleic acids[2,3] and to the description of enzyme inhibitor interactions[4]. MEP appears promising[6] as a molecular determinant of ligand recognition by receptors and the MEP maps are particularly inviting for the visual search of similarities between chemically diverse ligands. There are, however, several issues about the use of MEP in dealing with the ligand--receptor interactions. They range from the general question of the validity of the concept to the heuritsic value of the topographic features in the MEP maps and to the quantitative aspects of MEP features what is most relevant to drug design. We shall deal with these issues in the first part of this paper on hand of some examples. The need to compute the MEP of series of rather large molecules raises problems of computational economy *vs.* quality of the results due to the approximations used and this is the second issue we shall deal with.

A. The validity of the electrostatic model has been demonstrated on numerous examples of complexes of small molecules, particularly those with hydrogen bonds where comparative calculations of the geometry and energy are feasible at the SCF level[2c,6]. Although for ligands with strongly polar groups, *e. g.* the catecholamines, the first order electrostatic interactions with the receptor are likely to be important it is not certain that other types of interactions may be neglected. In the absence of hard proofs, adequate evidence can be obtained, for instance, from quantitative structure-activity relationships (QSAR) using as parameters appropriately selected features of the MEP. The side chain amino group of biogenic amines is well known as a primary point of attachment to receptors and, on the other hand, the MEP of simple amines was thoroughly examined as a reactivity index in protonation and hydrogen bonding[7]. The obvious choice of the parameter to be correlated with the ligand--receptor dissociation constant was the value of the MEP at a determined point in the region of the nitrogen lone pair (V_N). We have calculated V_N from *ab--initio* wave functions of 11 amines representing the side chain of 35 β-adrenergic agonists[8] for which biological data were available. The statistical evaluation has demonstrated that V_N can explain more than 50 p. c. of the variance in affinity of these drugs for the β-receptor. The inclusion of the MEP of phenyl substituents in the regression analysis has improved the correlation[9] demonstrating the importance of the electrostatic forces for the ligand-receptor interaction, but obviously other contributions to the free energy of interaction are not negligible. This may be illustrated by the example of 2-phenyl-imino- and benzyl-imidazolidine type drugs that are important α-adrenergic agonists. Although they are likely to interact with the receptor primarily by electrostatic forces, the similarity of the dissociation constant (K_i) for the representatives shown in Figure 1 confronted with the character of the substituents intuitively suggests that the MEP could hardly determine the biological differences between them. Indeed, in the MEP's computed from MNDO wave functions for 23 representative molecules[10] no feature was found that would rationalize the trends in K_i. The MEP above the aromatic plane computed from the MNDO wave function for a test molecule (dopamine) in

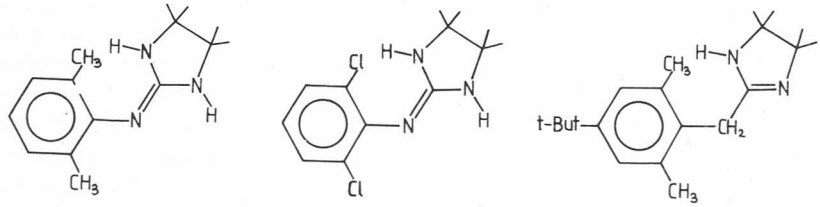

Figure 1. Chemical structure of representative phenylimidazoline type adrenergic drugs with their pK values I: 2 (2,6-dimethylphenyl-imino)-imidazolidine; II: clonidine, III: xylometazoline.

the point charge approximation was quite similar to the one computed in the same approximation from the *ab-initio* STO-3G function except for slightly smaller absolute values. Amongst other parameters tested in QSAR the best correlation of pK_i was found with the conformational entropy[10]. This is a very instructive case showing that the entropic factor may outweigh the differences in the enthalpic part of the free energy of association and be thus determining the differences between the members of a series of structurally related molecules.

Arguments on statistical basis require sufficiently large sets of data. However, suitable biological data (*e. g.* consistent K_i values) are scarce. Thus, in seeking good examples for the application of the MEP concept one is limited to rather qualitative considerations of the MEP concept for which molecules with pronounced differences in biological characteristics have to be chosen and it is desirable to produce additional evidence for the MEP as the determinant of recognition. Weinstein *et al.*[11] have examined a small series of 5-hydroxytryptamine congeners that are recognized by the serotonin receptor. These authors have introduced the electrostatic orientation vector[12] as the molecular determinant of recognition. The vector is obtained by connecting the MEP minima above the substituted indole ring. Its direction depends upon the position of the substituents. These drugs are supposed to become first attached to the receptor by the side chain like other biogenic amines and interact subsequently by stacking of the indole moiety with some aromatic amino acids of the receptor. This model was projected on hand of the known structures of indole complexes[13]. The interaction energy of the 5-hydroxytryptamine congeners with the receptor depends on the direction of their orientation vector relative to the electric field of the receptor presented by the model molecule, the imidazolium cation. Thus it was possible to explain the extreme differences in serotoninergic activity between 5- and 6-hydroxytryptamine and some other congeners in between. The analysis of the interaction energy with the receptor modelling imidazolium cation demonstrated the dominance of the electrostatic component[14].

The majority of applications of the MEP concept to pharmacological problems are limited to the demonstration of topological similarities between molecules

recognized by a certain type of receptor and differences between these and the molecules that are rejected by the receptor. For example, apomorphine may be taken as representative of catecholamine type molecules recognized by the dopamine receptor. Its rigidity is instrumental in resolving the problems of steric nature[15]. The molecular determinants of recognition deduced from the examination of a series of dopamine congeners are the aliphatic nitrogen with its negative MEP and the negative MEP region above the aromatic part which must be in proper spatial relations[15] as illustrated in Figure 2. Ergolene and some of its partial structures are also recognized by the dopamine receptor although they are different in chemical structure from the catecholamines. However, there is a definite topographical similarity between the MEP above the aromatic plane of both types of molecules. The MEP's in this study were obtained by the full *ab-initio* (STO-3G basis) procedure. To demonstrate the congruence of the MEP of catecholamine type ligands and those of the ergolene type the aliphatic nitrogen was taken as reference and the steric requirements of the dopamine receptor were properly observed. Figure 3 shows the congruence of the skeletons of apomorphine and ergolene, and the matching of the deepest negative region above the 11-OH group with that above the indole ring, respectively. The zero contours are also similar. In contrast, there is no acceptable congruence of the apomorphine and 2-azaergolene skeletons and the zero contours are almost perpendicular (Figure 4). This may explain the inactivity of the latter. The details of this work that include also partial ergolene structures will be published elsewhere[16].

The topographical congruence of the MEP of a prototype molecule with those of chemically dissimilar molecules as evidence that electrostatic interactions dominate the recognition can be corroborated by the exclusion of other possible determinants, *e. g.* reactivity indices based on HOMO and LUMO characteristics[16,17]. However, MEP based predictions of activity are the best and, from the practical point of view, the most useful evidence for the soundness of the MEP concept[5b,c].

B. The necessity of computing MEP's for a series of molecules in the size of 20 or more atoms requires economical procedures. However, this raises pro-

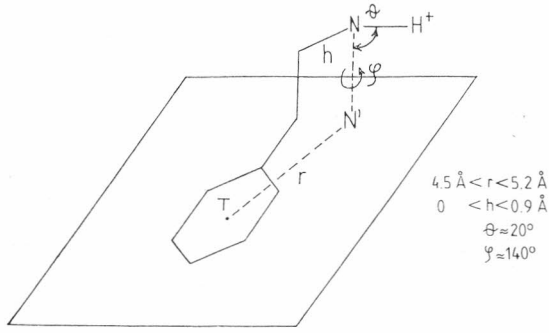

Figure 2. Topological requirements of the dopamine receptor for catecholamine type ligands.

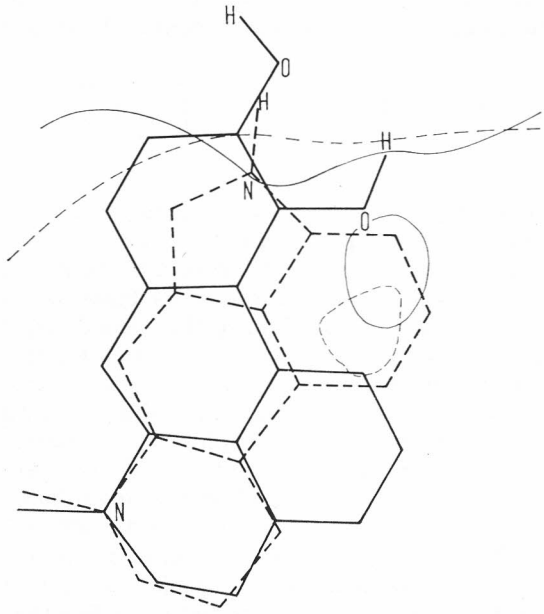

Figure 3. Superposition of apomorphine (full treat) and ergolene (dashed) skeletons and MEPs (deepest potential region and zero contour only, 1.6 Å above the aromatic plane.

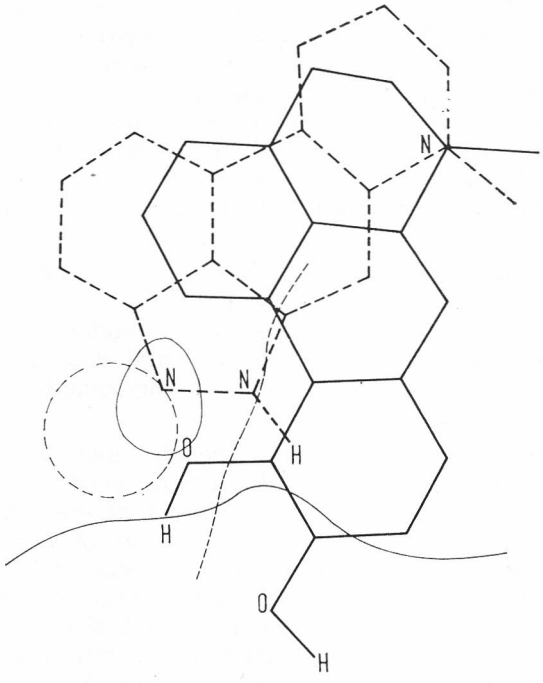

Figure 4. Superposition of apomorphine and 2-azaergolene skeletons and MEPs (as in Figure 3).

blems of the adequatness of the computed MEP's. To envisage the problems consider Eq. (1) defining the electrostatic potential V at point P:

$$V(\vec{P}) = \sum_{i=1}^{N} \frac{z_i}{|\vec{P} - \vec{r}_i|} - \int \frac{\varrho(\vec{r'})\, d\vec{r'}}{|\vec{P} - \vec{r'}|} \tag{1}$$

where $\varrho(\vec{r'})$ is the electron density, $\vec{r}_i$ are atomic coordinates and z_i atomic numbers. It is solved in two steps the first of which is the calculation of the density matrix from the molecular wave function. Assuming that a standard *ab-initio* method[18,19] with Gaussian functions as basis set was used to solve Roothan equations[20], we have to calculate in the second step the integrals over all Gaussian functions, *i. e.* atomic orbitals for each position of the probing charge (usually $+ e$) for which we need to know the electrostatic interaction energy with the molecule. For one point in space this is not much of a computer time consuming work, but for a decent graphic representation of the MEP of a twenty atom molecule about ten thousand points are nedded. This number can be reduced by some interpolation method[21] to about 300. Further computer time can be saved in the calculation of electrostatic attraction integrals over the atomic orbitals. These integrals when evaluated in the cartesian coordinate system split into three one dimensional integrals over each coordinate and in an additional integration which is related to the error function used in statistics. The error function integral cannot be solved analyticaly and therefore it is approximated by a series. Alternatively, it may be expressed by a quotient of two polynoms fitted to give accurate results. This latter method reduces the computer time to about five to ten times in comparison with the series approximation[22]. Since the basis functions are the same for all points in which the MEP has to be calculated one can put in tables many values for all the combinations of atomic orbitals. These tables are stored in the main computer memory during the calculation for all points. With these improvements a program was developed in our laboratory that allows computer time saving by a factor of twenty in comparison with Polyatom[19].

An often discussed point is the quality of the wave function. Within the *ab initio* scheme, the influence of the basis set upon the resulting MEP has been analyzed and the efficiency of pseudopotential calculation demonstrated[23]. The comparison for various basis sets shows no significant differences. The agreement for some basis sets with experimental data is still obtained by chance and an *ab initio* calculated MEP gives only qualitative results so that the minimal basis set is quite satisfactory[24,25].

The most controversial issue is the use of semiempirical methods. The calculation of ϱ/r integrals from zero differential overlap schemes results in differences both in the position and in the depth of the MEP minima in the region of lone electron pairs with respect to the *ab initio* calculated ones. A grave defect resulting from ZDO schemes is the sign of the MEP above the aromatic ring which is positive instead of negative. Improvements were obtained by calculating the exact nuclear attraction[26] to s orbitals in Eq. (1) or using empirical values and by deorthogonalizing the density matrix. Since all the attraction integrals in the summation of Eq. (1) have to be calculated in the latter procedure this requires more computer time. The results are comparable

to the non-empirically obtained ones for the MEP in the plane of substituted aromatics, but the MEP above the plane is still positive[27].

An often used group of approximations are point charge methods. Total atomic charges obtained from the Mullikan population analysis[28] may be used in the evaluation of MEP by simple summation:

$$V(\vec{P}) = \sum_{i=1}^{N} \frac{q_i - z_i}{|\vec{P} - \vec{r}_i|} \tag{2}$$

where N is the number of atoms, q_i is the charge, z_i is the atomic number and $\vec{r}_i$ are the atomic coordinates. The values of the MEP at distance less than 1.5 Å from the nuclei are unreliable in the point charge approximation and Eq. (2) cannot be used for the calculation of in plane lone pair regions. The function $1/r$ has no repulsive character by itself and the atoms with opposite character of the one studied are too far for having any influence. However, the MEP over extended surfaces is quite comparable to the one obtained in the whole scheme. This is illustrated by Figure 5 with the MEP of 3-methylindole 1.6 Å above the aromatic plane.

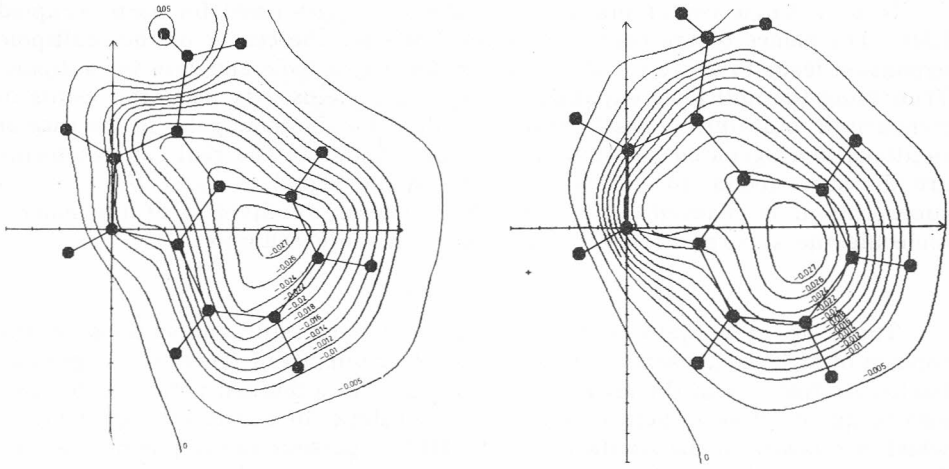

Figure 5. MEP maps (contours in a. u.) of 3-methylimidazole 1.6 Å above the aromatic plane computed from the *ab initio* STO-3G wave function. (a) — full integral evaluation, (b) — point charge method.

Tomasi and coworkers have made a further step in simplifying the evaluation of the MEP by which a previous evaluation of the wave function is avoided. The procedure is based on transferable group contributions to the MEP. The electrostatic potentials of characteristic electron pairs of σ and π bonded groups as well as of larger groups built from them do not seem to vary strongly from molecule to molecule and thus the sum of the model pair contributions is a reasonable approximation to the SCF calculated MEP[29]. In a further elaboration [30] of the concept of transferable group contributions the localized orbitals are modelled by point charges displaced from the electronic charge center of

the orbital. The direction of the displacement depends on the character (σ or π) of the orbital. For lone pairs the charge is placed at the orbital center. In aromatic systems σ and π bonds are treated separately[31]. Comparisons with the SCF calculated MEPs of simple molecules are promising to make the transferable group concept a useful and inexpensive tool for the investigation of reaction properties of large molecules, at least on a qualitative level. The results seem to be in certain respects even superior to those obtained by the simple summation over total atomic charges[30].

The final point concerns the MEP of large molecules for which even the nonempirical schemes are not practicable for calculating the density matrix. The only way to overcome this is to break down the system into smaller subunits for which the wave function and the MEP can be calculated. The free valencies appearing at the cuts are absorbed by fictitious hydrogens and the MEP of the whole system is obtained by the superposition of the subunits[32]. This leads obviously to a perturbed MEP of the macromolecule, but Pullman *et al.* have shown[33,34] that perturbations are generally small and confined to the vicinity of the junctions. The superposition method can be improved by using localized molecular orbitals and their expansion to multicentered multipoles. The *ab initio* SCF wave function is first calculated and is then localized with the Boys technique[35].

Next, a single center multipole expansion is performed for each occupied LMO. The choice of the centroids of the LMO for the center of the multipole expansion leads to a -2 (in atomic units) for a monopole and zero for a dipole. Truncating the multipole expansion at octopoles yields very accurate results in comparison to *ab-initio* ones[32]. Since the values for the monopoles in the case of fictitious hydrogenated subunits and of the subunit in the real supermolecule are the same (equal to -2) and for the dipole (equal to zero), the error in superposition is reduced. With a carefully chosen subdivision of the macromolecule the superposition error can be even more reduced.

CONCLUSION

There are now quite a few examples of successful applications of the concept of MEP in the search of molecular determinants of receptor recognition. Particular features in the MEP such as its value at a defined point can be used also in quantitative structure-activity relationships. In qualitative applications which are based on the similarity of the MEP a possible development is in the direction of removing subjectivity in the judgement of similarity by introducing some metrics. From the computational point of view the defects due to various approximations are by now well known so that the most efficient and yet sufficiently accurate method of computing MEP may be selected. Concerning the biomolecular problems to which the concept of MEP may be applied, it has to be borne in mind that the electrostatic component of interaction energy is not necessarily dominating in the recognition and, moreover, that recognition is only one link in the chain of events upon which the biological activity is depending. Therefore, a carefull judgement of the problem is advisable before embarking in lengthy computations.

Acknowledgement. — This work was supported by the Research Community of Slovenia and the LEK Works.

REFERENCES

1. K. Morokuma and K. Kitaura, in: *Molecular Interactions,* H. Ratajczak and W. J. Orville-Thomas (Eds.), John Wiley, Chichester, 1980, p. 21; P. Claverie, in: *Intermolecular Interactions: From Diatomics to Biopolymers* B. Pullman (Ed.), John Wiley, New York, 1978, p. 69.
2. (a) R. Bonaccorsi, L. Petrongolo, E. Scrocco, and J. Tomasi, *Theor. Chim. Acta* **20** (1971) 331;
   (b) E. Scrocco and J. Tomasi, *Adv. Quantum Chem.* II (1978) 115;
   (c) J. Tomasi, in: *Chemical Applications of Atomic and Molecular Electrostatic Potentials,* P. Politzer and D. G. Truhlar (Eds.), Plenum Press New York, 1981, p. 257.
3. A. Pullman and B. Pullman, *Q. Rev. Biophys.* **14** (1981) 289; P. M. Dean and L. P. G. Wakelin, *Phys. Trans. Roy. Soc. B* **287** (1979) 287.
4. P. K. Weiner, R. Langridge, J. M. Blaney, R. Schaefer, and P. A. Kollman, *Proc. Nat. Acad. Sci. U.S.A.* **79** (1982) 3754; D. M. Hayes and P. A. Kollman, *J. Amer. Chem. Soc.* **98** (1976) 3335.
5. C. Petrongolo, *Gazz. Chim. Ital.* **108** (1978) 445. J. J. Kaufman, P. C. Hariharan, H. E. Popkie, and C. Petrongolo, *Ann. N. Y. Acad. Sci.* **367** (1981) 452; H. Weinstein, R. Osman, and J. P. Green, in: *Computer Assisted Drug Design,* E. C. Olson and R. E. Christophersen (Eds.), *A. C. S. Symp. Ser* **112** (1979) 161; H. Weinstein, R. Osman, S. Topiol, and J. P. Green, *ibid.* p. 434; P. M. Dean, *Brit. J. Pharmacol.* **74** (1981) 39, 47.
6. P. A. Kollman, *J. Amer. Chem. Soc.* **99** (1977) 4875; M. D. Dolgushin and V. M. Pinchuh, *Theor. Chim. Acta* **45** (1977) 157.
7. R. Bonaccorsi, E. Scrocco, and J. Tomasi, *Theor. Chim. Acta* **43** (1976) 63; P. Kollman, J. McKlevey, A. Johansson, and S. Rothenberg, *J. Amer. Chem. Soc.* **97** (1975) 955.
8. T. Šolmajer, I. Lukovits, and D. Hadži, *J. Med. Chem.* **5** (1982) 1413.
9. T. Šolmajer, I. Lukovits, and D. Hadži, *Quant. Struct. Act. Relat.* **3** (1984) 51.
10. F. Avbelj and D. Hadži, *Mol. Pharmacol.,* in press.
11. H. Weinstein, D. Chou, S. Kang, C. L. Johnson, and J. P. Green, *Int. J. Quantum Chem. QBS* **3** (1976) 135.
12. H. Weinstein and R. Osman, *Int. J. Quantum Chem., QBS* **4** (1977) 253.
13. M. Shinitzky and E. Katchalski, in: *Molecular Associations in Biology* B. Pulman (Ed.), Academic Press, New York, 1968, p. 361. H. Sapper and W. Lohmann, *Mol. Pharmacol.* **12** (1976) 605.
14. H. Weinstein, R. Osman, W. D. Edwards, and J. P. Green, *Int. J. Quantum Chem., QBS* **5** (1978) 449.
15. D. Kocjan and D. Hadži, *J. Pharm. Pharmacol.* **35** (1983) 780.
16. D. Kocjan, M. Hodošček, and D. Hadži, to be submitted.
17. D. Kocjan, T. Šolmajer, M. Hodošček, and D. Hadži, *Int. J. Quantum Chem.* **23** (1983) 1121.
18. J. S. Binkley, R. A. Whiteside, R. Krishnan, R. Seeger, D. J. DeFrees, H. B. Schlegel, S. Topiol, L. R. Kahn, and J. Pople, GAUSSIAN 80, *QCPE, No. 406,* Indiana University, Bloomington, Ind.
19. D. B. Neumann, H. Basch, R. L. Komegay, L. C. Snyder, J. W. Moskowitz, C. Hornback, and S. P. Liebmann, Polyatom/2, *QCPE 199,* Chem. Dept., Indiana University, Bloomington, Indiana.
20. C. C. J. Roothaan, *Rev. Mod. Phys.* **23** (1951) 69.
21. J. S. Taylor, private communication.
22. L. J. Schad and G. O. Morrel, *J. Chem. Phys.* **54** (1971) 1965.
23. R. Osman and H. Weinstein, *Chem. Phys. Lett.* **49** (1977) 73.
24. W. D. Edwards and H. Weinstein, *Chem. Phys. Lett.* **56** (1978) 582.
25. J. Almlöf and A. Stögard, *Chem. Phys. Lett.* **29** (1974) 418.
26. C. Giessner-Prettre, and B. Pullman, *Theor. Chim. Acta* **25** (1972) 83.
27. C. Petrongolo, H. J.-T. Preston, and J. J. Kaufman, *Int. J. Quantum Chem.* **13** (1978) 459.
28. R. S. Mulliken, *J. Chem. Phys.* **23** (1965) 1833.
29. R. Bonaccorsi, E. Scrocco, and J. Tomasi, *J. Amer. Chem. Soc.* **98** (1976) 4049.

30. R. Bonaccorsi, E. Scrocco, and J. Tomasi, *J. Amer. Chem. Soc.* **99** (1977) 4546.
31. A. Agresti, R. Bonaccorsi, and J. Tomasi, *Theor. Chim. Acta* **53** (1979) 215.
32. C. Etchebest, R. Lavery, and A. Pullman, *Theor. Chim. Acta, Berlin* **62** (1982) 17.
33. A. Pullman and H. Berthod, *Theor. Chim. Acta* **48** (1978) 269.
34. A. Pullman, in: *New Horizons in Biological Chemistry, Symposium in Honour of Prof. K. Yagi*, Japan Scientific Societies Press, 1981.
35. S. F. Boys, *Rev. Mod. Phys.* **32** (1960) 306.

SAŽETAK

Molekulski elektrostatski potencijal kao odlučujući faktor u određivanju prepoznavanja receptora i lijeka

Dušan Hadži, Milan Hodošček, Darko Kocjan, Tomaž Šolmajer i Franc Avbelj

Razmatrana je primjena koncepta molekulskoga elektrostatskog potencijala u određivanju i racionalizaciji interakcija između receptora i lijekova. Diskutirani su njegovi kvalitativni i kvantitativni aspekti. Dan je kratak opis odgovarajućih numeričkih postupaka s posebnim osvrtom na ekonomiju računa i kvalitet rezultata.