Biological Communication via Molecular Surfaces

Timothy Clark[1], Kendall G. Byler[1] and Marcel J. de Groot[2]

[1]Computer-Chemie-Centrum, Universität Erlangen-Nürnberg, Nägelsbachstraße 25, 91052 Erlangen, Germany.

[2]Pfizer Ltd., Global Research and Development, Sandwich Laboratories, Sandwich, Kent CT13 9NJ, U.K.

E-Mail: * Tim.Clark@chemie.uni-erlangen.de

Abstract

The use and characteristics of local properties designed to describe intermolecular interactions projected onto molecular surfaces and based on semiempirical molecular orbital theory are described. After a discussion of the local properties themselves and their relationship to intermolecular interactions and chemical reactivity, two applications are described. The first, surface-integral models for physical properties, involve integrating a functional of the local properties over the molecular surface. In the second example, we discuss a possible approach to determining the potential specificity of biological interactions based on Shannon's theory of communication.

Introduction

The atomistic approximation (i.e., that molecules can be represented as an array of distinct atoms that are usually treated as points) is used almost universally for modelling, quantitative structure-activity (QSAR) and -property (QSPR) relationships, and chemoinformatics applications. The atomistic approximation is the basis of classical mechanical models of molecules (force fields) but it is often also derived from the results of quantum mechanical calculations. Wave functions or electron densities are "reduced" to atomistic descriptions by a variety of population analyses [1 – 6] or other techniques for partitioning the electron

density [7, 8] or by fitting atomic monopoles to the molecular electrostatic potential (MEP) [9 – 14] or any other local property around the molecule to atom-centred two-centre potentials. Although it has been argued strongly that "atoms in molecules" represent transferable and easily recognizable entities [8], a purist quantum mechanical view of molecules within the Born-Oppenheimer approximation is a cloud of electron density comprising the appropriate number of indistinguishable electrons in which a number of fixed point positive charges (the nuclei) exist. However, even techniques that nominally rely on surface descriptions of molecules often designate portions of the molecular surface to an underlying atom and assign them properties according to the appropriate element. Examples of such techniques include those which are purely classical such as MolFESD [15 – 17] and the PB-SA solvent techniques [18, 19] but also those based on quantum mechanics such as the SM_n solvation models [20] or COSMO-RS [21]. This should not be necessary if a wave function or electron density is available, but represents a simple approximation that allows the introduction of element-specific parameters that often improve model performance considerably. However, the need for different parameters for different elements means that the theory is incomplete. What is missing are the typical properties that were often defined early in the development of molecular orbital (MO) or density-functional theory (DFT) to describe the differences in the properties of elements, whole molecules, regions around them or even points in their vicinity. An incomplete summary of some such properties is given in Table 1.

Table 1. Some representative descriptive quantities relevant to molecular reactivity and intermolecular interactions. For the meanings of the various symbols, see the original references.

Property	Definition		Reference
Electronegativity, χ	Mulliken;	$\chi = \dfrac{I + A}{2}$	22,23
	Pauling;	$\chi_A - \chi_B \approx \sqrt{\left(E_{AB} + \dfrac{(E_{AA} + E_{BB})}{2} \right)}$	24,25
	Parr, Pearson;	$\chi = -\mu = \dfrac{\partial E}{\partial N}$	26,27
Hardness, η	Pearson;	$\eta = \dfrac{\varepsilon_{LUMO} - \varepsilon_{HOMO}}{2} \approx \dfrac{I - A}{2}$	27
	Parr;	$\eta = \dfrac{1}{2} \cdot \dfrac{\partial \mu}{\partial N} = \dfrac{\partial^2 E}{\partial N^2}$	26

Biological Communication via Molecular Surfaces

Property	Definition	Reference
Softness, σ	Parr; $\sigma = \dfrac{1}{\eta}$	26
Polarizabilty, α	$\alpha = \dfrac{P}{E}$	
MEP, V	$V(\mathbf{r}) = \sum_A \dfrac{Z_A}{|R_A - r|} - \int \dfrac{\rho(r')dr'}{|r' - r|}$	
Superdelocalizability, S_r	Fukui, $$S_r = (2-\nu)\sum_{j=1}^{m} \dfrac{C_r^{j^2}}{\alpha_h - \varepsilon_j}(-\beta) + \nu \sum_{j=m+1}^{n} \dfrac{C_r^{j^2}}{\varepsilon_j - \alpha_h}(-\beta)$$	28
Fukui function, f	Parr, $f(\mathbf{r}) = \left[\dfrac{\delta\mu}{\delta\eta(\mathbf{r})}\right]_N = \left[\dfrac{\partial\rho(\mathbf{r})}{\partial\mathbf{N}}\right]_v$	26
Average local ionization energy, IE_L	Murray, Politzer, $IE_L(\mathbf{r}) = \dfrac{\sum\limits_{i=1,HOMO} -\rho_i(\mathbf{r})\varepsilon_i}{\sum\limits_{i=1,HOMO} \rho_i(\mathbf{r})}$	29–33
Electron localization function, ELF	Becke, $ELF = \left[1 + \left(\dfrac{D_\sigma}{D_\sigma^0}\right)^2\right]^{-1}$	34
Local electron affinity, EA_L	Clark, $EA_L(\mathbf{r}) = \dfrac{\sum\limits_{i=LUMO,norbs} -\rho_i(\mathbf{r})\varepsilon_i}{\sum\limits_{i=LUMO,norbs} \rho_i(\mathbf{r})}$	35
Local polarizability, α_L	Clark, $\alpha_L(\mathbf{r}) = \dfrac{\sum\limits_{j=1}^{NAOs} \rho_j^1(\mathbf{r})q_j\bar{\alpha}_j}{\sum\limits_{j=1}^{NAOs} \rho_j^1(\mathbf{r})q_j}$	35
Local electronegativity, χ_L	Clark, $\chi_L(\mathbf{r}) = \dfrac{\left(IP_L(\mathbf{r}) + EA_L(\mathbf{r})\right)}{2}$	35
Local hardness, η_L	Clark, $\eta_L(\mathbf{r}) = \dfrac{\left(IP_L(\mathbf{r}) - EA_L(\mathbf{r})\right)}{2}$	35

The starting point of our investigations was therefore to investigate whether we can define a set of local properties at or near the surfaces of molecules that allow us to formulate intermolecular forces and energies as the classical [36] combination of Coulomb, exchange repulsion, dispersion and donor-acceptor interactions.

LOCAL PROPERTIES

By far the most familiar local property is the MEP [37], which is often projected onto molecular surfaces and visualized using colour coding to represent the value of the MEP. However, although Coulomb interactions, which can be derived from the MEPs of two interacting molecules, are by far the strongest contributors to intermolecular interaction energies in the gas phase, they are strongly attenuated in polar solution and may even lead to a net destabilization if Coulomb attraction to the solvent is stronger than to the complexation partner. In situations such as this, interactions that are weaker in the gas phase become far more significant and may even dominate the total interaction energy. It is therefore necessary to include these weaker interactions in a complete model of intermolecular complexation, which is the basis of all molecular communications mechanisms. We can therefore consider appropriate surface properties that are related to the four types of interaction outlined above.

PAULI REPULSION = SHAPE

The Pauli repulsion between the molecule and a neighbour depends on the electron density. Therefore, using isodensity molecular surfaces [38] allows us to treat repulsion. Molecular shape analysis has been discussed in detail by Mezey [39] and virtual screening techniques based solely on the molecular shape are becoming well established [40, 41]. Currently, the Gay-Berne potential [42] is the best known of very few examples of repulsive/van der Waals potentials for anisotropic bodies. Describing the shapes of molecules using spherical-harmonic expansions [43 – 45] provides an analytical shape description that has, as far as we know, not yet been used to describe repulsion between molecules.

COULOMB INTERACTIONS

The MEP is well established as an observable molecular property that determines intermolecular Coulomb interactions. However, Coulomb interactions are usually calculated from atomic monopoles [1 – 14] or multipoles [46], distributed multipoles [47, 48], or from the electron density directly [37]. Electrostatic shell models in which charges are not centred on the atoms are common in materials modelling [49, 50]. The Coulomb interaction energy between two electrostatically anisotropic bodies is almost always calculated using a multi-centre approach such as atomic multipoles or distributed multipoles [46, 47]. Single-centre multipole expansions can be used [51], but may not converge as the order of the multipole expansion is increased. This approach is mathematically equivalent to our fitting of the molecular electrostatic potential at the surface of the molecule to a spherical-har-

monic expansion [45]. This approach promises to be very useful for cheminformatics applications, but may be less so for classical modelling applications such as molecular dynamics.

LOCAL POLARIZABILITY

The dispersion interactions of a molecule with neighbours are linked to the polarizability by the London equation [52 – 55]. We [56] have described a parameterized technique for calculating the molecular electronic polarizability tensor accurately using semiempirical MO-techniques and have proposed a partitioning scheme [57] similar to a population analysis that allows atomic (or even atomic orbital) polarizability tensors to be assigned. Note that any such partitioning scheme, like those used to assign net atomic charges [1 – 6] is arbitrary and that our scheme has been defined to give the molecular electronic polarizability as the sum of the atomic polarizability tensors, although this definition is also arbitrary. However, the "atomic orbital" polarizabilities can be used to define a local polarizability around the molecule that serves to indicate the anisotropy of the molecular polarizability. This is illustrated by the local polarizability of naphthalene projected onto an isodensity surface (calculated with AM1[58]) shown in Fig. 1. The π-face of the molecule is relatively more polarizable than the hydrogen atoms around the periphery and the 1-, 4-, 6-, and 9-hydrogens are less polarizable than their 2-, 3-, 7-, and 8-counterparts.
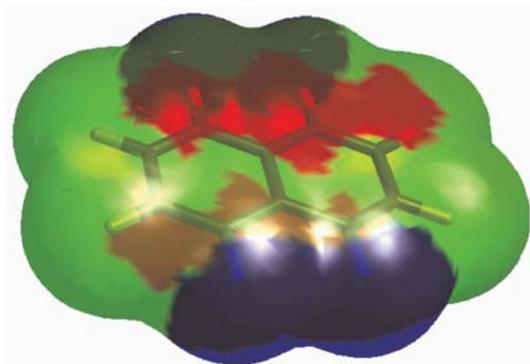

Figure 1. The AM1-calculated local polarizability of naphthalene projected onto an isodensity surface. The colour scheme ranges from red (most polarizable) to blue (least polarizable). The surfaces and the local electronegativity were calculated with ParaSurf'06 [59].

Such descriptions of the polarizability are important if dispersion interactions such as, for instance, those that play a role in the π-stacking of aromatic rings. Using a dispersion term derived from the atomic polarizability tensors derived as described above [57] together with the London equation [52 – 55] and the Slater-Kirkwood approximation [60], we [61] were able to reproduce such interactions with MNDO[ref] semi-empirical MO-theory. Normally, neither semi-empirical MO-theory nor density-functional theory (DFT) is able to

reproduce dispersion, although several correction terms have been suggested for DFT [62 – 64]. The technique that we have described is atomistic, but can be formulated in terms of the local polarizability.
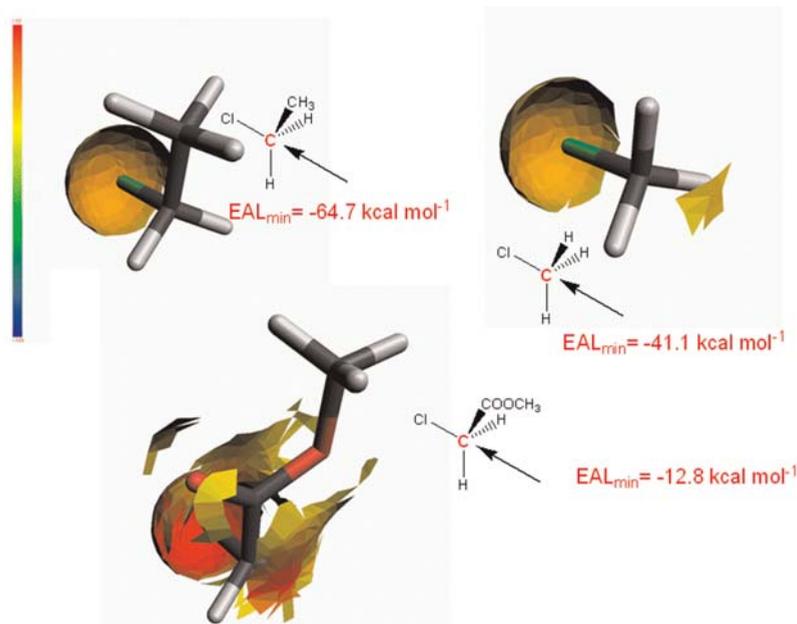
LOCAL IONIZATION ENERGY AND ELECTRON AFFINITY

Figure 2. The least positive areas of the local electron affinity at an isodensity surface (calculated with AM1 [58]) for three substrates for an S_N2 substitution reaction. The surfaces and the local electronegativity were calculated with ParaSurf'06 [59].

Electron donor-acceptor (Lewis acid-base) interactions are usually either ignored in classical modelling techniques or are implicit in more general interaction potentials. Interestingly, these interactions have traditionally played a dominant role in qualitative reaction theory [22 – 28]. Thus, the superdelocalizability [28] introduced by Fukui and the "Fukui function" introduced by Parr [27] follow similar concepts to the average local ionization energy [29 – 33] and the local electron affinity [35] in that they are based on perturbational molecular orbital theory. The Fukui function also uses the frontier orbital approximation and both the superdelocalizability and the local electron affinity rely on using virtual orbitals and therefore are limited in their current forms to minimal basis sets. Perhaps not surprisingly, the average local ionization energy and the local electron affinity describe not only the donor-acceptor component of intermolecular interactions, but also chemical reactivity. Figure 2, for instance, shows the areas of highest (least negative) local electron affinity for chloroethane, chloromethane and methyl chloroacetate. The activating influence of the ester substituent and the opposite effect of the methyl group in chloroethane relative

to S_N2-substitution at the carbon bearing the chlorine substituent are clearly visible. Similarly, the local ionization energy usefully indicates the reactivity of aromatic ring positions towards electrophilic aromatic substitution [35].

QSAR AND QSPR WITH DESCRIPTORS DERIVED FROM LOCAL PROPERTIES

The above discussion suggests that intermolecular interactions, which are the basis of biological communication and also determine many physical properties such as vapour pressure, boiling point, partition coefficients, solubility etc., are described well by the local properties. Therefore, these properties should be sufficient to describe intermolecular reactions, and thus for QSAR and QSPR applications. We have investigated two approaches to such models.
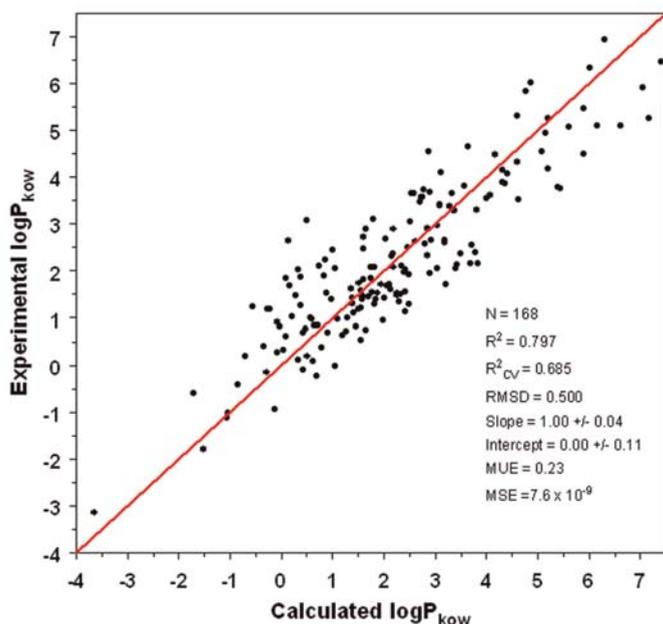

Figure 3. Experimental vs. calculated logP values for the SIM model described above.

The first uses the statistical descriptors based on local properties at the molecular surface first introduced by Murray and Politzer [65, 66] and later extended by us to other local properties [67]. These descriptors are derived by first calculating a triangulated molecular surface such as an isodensity or solvent-excluded surface. The calculated values of the local properties, in the case of the descriptors introduced by Murray and Politzer the MEP, are then used to calculate statistical descriptors such as the variance, maximum, minimum, mean value, range etc. These values serve as descriptors that are completely independent of the 2D-structure of the molecule, by which we mean that they do not contain information

such as atom counts, numbers of aromatic rings, numbers of hydrogen-bond donors or acceptors etc. The descriptors can then be used in combination with an interpolation technique such as multiple regression or artificial neural nets to construct a classical QSPR-model.

The second type of QSPR model is known as a surface-integral model (SIM) and has been used in connection with the MolFESD technique [15 – 17]. We [68] have presented SIMs for the free energies of solvation in water, *n*-octanol, and chloroform and for the enthalpy of solvation in water. Strictly speaking, solvation energies are not local properties, but the concept of a hydration free-energy density (HFED) was introduced by Scheraga [69] and has proved useful. The target property (in the following example $logP_{kow}$) is calculated as the integral of a functional of the local properties over the entire isodensity surface of the molecule. The functional is determined by regression using potential functional expressions based on one or more of the local properties, as outlined in reference [68]. In order to demonstrate the generality of the concept of SIMs, we have trained a SIM-model based on the gas phase AM1 wave-function for the water/*n*-octanol partition coefficient, $logP_{kow}$. The statistical characteristics of the resulting 8-term model and a plot of the calculated vs. experimental $logP_{kow}$ values are shown in Fig. 3.

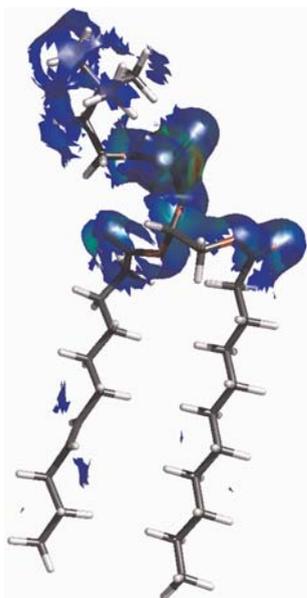**Figure 4.** Local $logP_{kow}$ values projected onto an isodensity surface of a phospholipid. The calculations are based on the AM1 [58] wave function and the surface and SIM values were calculated with ParaSurf'06 [59].

One of the attractive features of SIMs is that the values of the functional themselves represent a local property that can be visualized in order to help understand the system. Figure 4 shows the areas of the surface of a phospholipid that have the largest positive contribution to $logP_{kow}$ (i.e. the hydrophilic regions). The distribution corresponds exactly to our qualitative ideas of the hydrophilic/hydrophobic regions of the phospholipid.

SURFACE INFORMATION

Clearly, biologically active molecules carry information. Instinctively, we expect, for instance, an octapeptide to "carry more information" than, say, cyclohexane. This leads to the expectation that we should be able to quantify the information content of molecules. This idea is not new. For instance Kuz'min *et al*. [70] have discussed molecular information fields. However, if we turn to Shannon's classical work on information theory [71], we can define analogies and differences between signal transfer in communications systems and in biology.

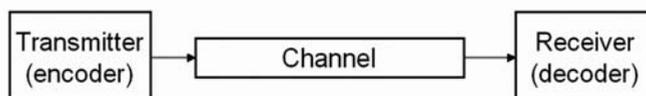Figure 5 shows Shannon's original scheme of a communications system.

Figure 5. Schematic diagram of a communications system consisting of transmitter, channel and receiver (after reference [71]).

Shannon was mostly concerned with the capacity of the channel and with the influence of noise. Biological communication can be described by a modified scheme, as shown in Fig. 6.
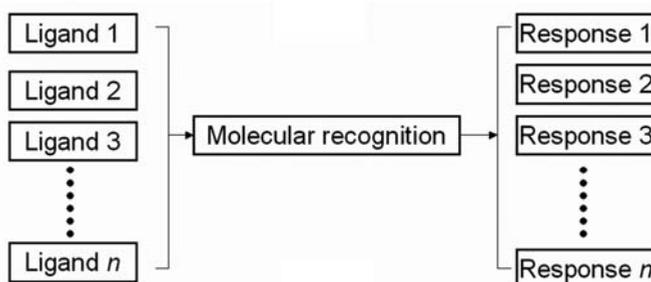

Figure 6. Schematic diagram of biological communication. For our purposes, the capacity of the channel ("Molecular Recognition") can be regarded as infinite. The dominant question is whether a given legend carries enough information to elicit one and only one response.

We can assume that the process of information transfer (molecular recognition) has enough capacity to satisfy the needs of the system. The pertinent question then becomes whether a given ligand carries enough information to be able to distinguish between its own response and all others. Note that "carries enough information" in this context means exactly the reverse of the concept of information content defined by Shannon. We are actually interested in the ligand carrying no information at all (i. e. being able to interact with only one receptor), rather than in Shannon's information content, which we might define as the

"degree of ambiguity" in the context of Fig. 6. Thus, the most selective ligands should have the lowest Shannon entropy. Following Shannon [71], we can define the "amount of information" using the Shannon entropy, H:

where P_i is the probability of finding symbol i in the message.

Translated into a continuous surface described by the four local properties V, I_L, E_L,

$$H = -\sum P_i \log(P_i)$$

and α_L, which we assume to be orthogonal to each other (as is approximately the case [35]), we can write the Shannon entropy as the numerical integration of the triangulated surface:

$$H = -\sum_{i=1}^{k} \left[p(V_i)\log_2 p(V_i) + p(I_{L,i})\log_2 p(I_{L,i}) + p(E_{L,i})\log_2 p(E_{L,i}) + p(\alpha_{L,i})\log_2 p(\alpha_{L,i}) \right] \cdot A_i$$

where k is the number of triangles on the surface, A_i the area of triangle i, and V_i, $I_{L,i}$, $E_{L,i}$ and $\alpha_{L,i}$ are the average values of V, I_L, E_L, and α_L, respectively, for triangle i. the probability of finding a given value of a local property x is defined as $p(x)$.

Thus, we can calculate a molecular Shannon entropy by numerical integration over the molecular surface analogously to a SIM-model. In this case, however, the Shannon entropy is a true local property [71]. The only question that remains is that of the probability distribution appropriate for calculating the Shannon entropy. Here, there are two possibilities. If we consider the ligand in isolation, we can use the distribution of the local properties on its own surface to define the probability function.

This results in what we term the "internal" Shannon entropy. Alternatively, we can consider an ensemble of ligands all competing to convey their messages, in which case we need a probability function for the complete set of ligands (termed the "external" Shannon entropy). We have approximated this probability by calculating the local properties at the surfaces of all the ligands contained in the PDBBIND dataset [72, 73]. In order to eliminate size effects, we define a surface-entropy density $\rho_{Shannon}$ as:

$$\rho_{Shannon} = \frac{H}{A}$$

where H is the molecular Shannon entropy and A the total surface area of the molecule.

Figure 7 shows the distribution of the calculated information densities for the PDBBIND ligands.
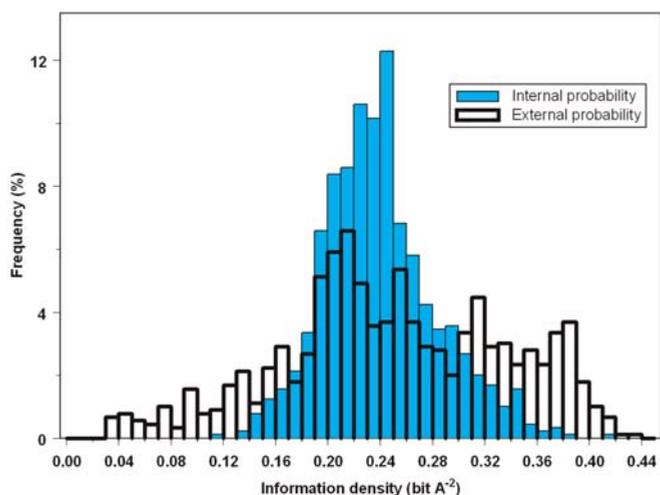
Figure 7. Calculated distribution of molecular information densities ("internal" and "external", as defined in the text) calculated based on the AM1 [58] wave-functions for the ligands of the PDBBIND database [72, 73]. The Shannon entropies were calculated with ParaSurf'06 [59].
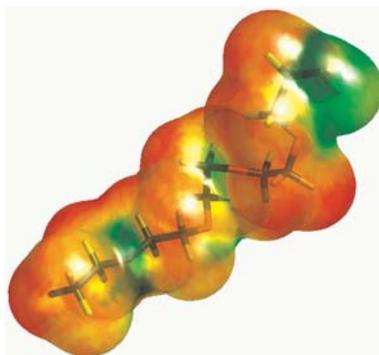

Figure 8. The "external" Shannon entropy projected onto an isodensity surface for a fragment of polyethylene glycol, PEG. The surface area is 345 Å^2, the internal and external Shannon entropies 97.5 and 134.8 bits, respectively, and the internal and external surface-entropy densities 0.280 and 0.389 bits Å^{-2}, respectively. The surface and the Shannon entropies were calculated with ParaSurf'06 [59].

The "internal" information content shows a narrower distribution than the "external". This is an effect of the charge of the ligands, which shifts the V_L values strongly in the case of the "external" Shannon entropy and therefore broadens the distribution. Thus, as might be expected, the "internal" Shannon entropy removes the effect of charge on the ligand, whereas the "external" equivalent provides a more global view of the ligands.

Two examples serve to indicate the possible meaning of the molecular Shannon entropy, although this quantity needs to be investigated more thoroughly before we can reach firm conclusions. Figure 8 shows the Shannon entropy projected onto the surface of a model segment of polyethylene glycol, which is often found bound non-specifically to proteins in crystal structures.

This can be compared with the model FPN tripeptide, which is also a neutral molecule but could be expected to bind more specifically to proteins. Its Shannon-entropy surface (on the same colour scale as Fig. 8) is shown in Fig. 9.
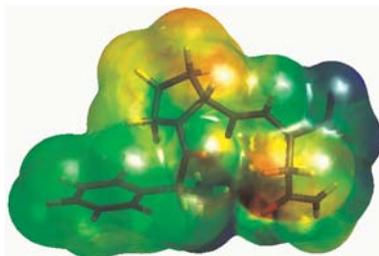

Figure 9. The "external" Shannon entropy projected onto an isodensity surface for a fragment of a model FPN tripeptide. The surface area is 380 $Å^2$, the internal and external Shannon entropies 95.7 and 106.7 bits, respectively, and the internal and external surface-entropy densities 0.250 and 0.280 bits $Å^{-2}$, respectively. The surface and the Shannon entropies were calculated with ParaSurf'06 [59].

The tripeptide has very similar, but slightly lower total "internal" Shannon entropy and surface-entropy density than the PEG fragment, but the difference between the two becomes apparent when we consider the "external" Shannon entropy and surface-entropy density. The tripeptide derives its specificity (blue and green areas) from the zwitterionic end groups and from the phenyl group, whereas the praline ring is considered to be very unspecific.

SUMMARY AND CONCLUSIONS

Modelling and simulation using molecular surfaces is certainly technically more difficult than the common atomistic techniques. However, molecular surfaces provide new opportunities to view molecules in a different light and to reconsider whether our continued preference for atomistic techniques is not influenced by the fact that we, as chemists, have learnt to view molecules atomistically. It remains to be seen whether techniques such as surface-integral models or surface Shannon entropy provide significant advantages in treating biological systems and, above all, in understanding biological communication. However, after 30 years of atomistic modelling it is probably time to consider alternatives.

ACKNOWLEDGMENT

REFERENCES

[1]   Coulson, C.A., Longuet-Higgins, H.C. (1947) The electronic structure of conjugated systems. 1. General theory. *Proc. R. Soc.* **A191:**39 – 60.

[2]   Mulliken, R.S. (1955) Electronic population analysis on LCAO-MO molecular wave functions. 11. Overlap populations, bond orders, and covalent bond energies. *J. Chem. Phys.* **23:**1833 – 1840.

[3]   Löwdin, P.-O. (1950) The nonorthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.* **18:**365 – 375.

[4]   Weinhold, F., Carpenter, J.E. (1988) *The Structure of Small Molecules and Ions.* Plenum, New York.

[5]   Hall, D., Williams, D.E. (1975) Effect of coulombic interactions on the calculated crystal structures of benzene at atmospheric and 25 kbar pressure. *Acta Crystallogr. A* **31:**56 – 58.

[6]   Li, J., Williams, B., Cramer C.J., Thruhlar, D.G. (1999) A class lV charge model for molecular excited states. *J. Chem. Phys.* **110:**724 – 733.

[7]   Streitwieser, Jr. A., Williams, Jr. J.E., Alexandratos, S., McKelvey,J.M. (1976) Ab initio SCF-MO calculations of methyllithium and related systems. Absence of covalent character in the carbon-lithium bonds. *J. Am. Chem. Soc.* **98:**4778 – 4784; Klein, J., Kost, D., Schriver, G.W., Streitwieser, Jr. A. (1982) Ab initio calculations of dilithiopropenes. *Proc. Natl Acad. Sci. U S A* **79:**3922 – 3926.

[8]   Bader, R.F.W. (1990) *Atoms in Molecules: A Quantum Theory* .Oxford University Press, Oxford.

[9]   Chirlian, L.E., Francl, M.M. (1987) Atomic charges derived from electrostatic potentials: a detailed study. *J. Comput. Chem.* **8:**894 – 905.

[10]   Breneman C.M., Wiberg, K.B. (1990) Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **11:**361 – 373.

[11]   Besler, B.H., Merz, K.M., Kollman, P.A. (1990) Atomic charges derived from semiempirical methods. *J. Comput. Chem.* **11:**431 – 439.

[12]   Ferenczy, G.G., Reynolds C.A., Richards, W.G. (1990) Semiempirical AM1 electrostatic potentials and AM1 electrostatic-potential derived charges: a comparison with ab initio values. *J. Comput. Chem.***11:**159 – 169.

[13]   Orozco, M., Luque, F.J. (1990) On the use of AM1 and MNDO wave functions to compute accurate electrostatic potentials. *J. Comp. Chem.* **11:**909 – 923.

[14] Beck, B., Glen, R.C., Clark, T. (1997) VESPA: a new, fast approach to electrostatic potential-derived atomic charges from semiempirical methods. *J. Comput. Chem.* **18:**744 – 756.

[15] Pixner, P., Heiden, W., Merx, H., Möller, A., Moeckel, G., Brickmann, J. (1994) Empirical method for the quantification and localization of molecular hydrophobicity. *J. Chem. Inf. Comput. Sci.* **34:**1309 – 1319.

[16] Jäger, R., Schmidt, F., Schilling, B., Brickmann, J. (2000) Localization and quantification of hydrophobicity: the molecular free energy density (MolFESD) concept and its application to sweetness recognition. *J. Comput.-Aided Mol. Des.* **14:**631 – 646.

[17] Jäger, R., Kast, S.M., Brickmann, J. (2003) Parametrization strategy for the Mol-FESD concept: Quantitative surface representation of local hydrophobicity. *J. Chem. Inf. Comput. Sci.* **43:**237 – 247.

[18] Massova, I., Kollmann, P.A. (2000) Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Persp. Drug Discov. Des.* **18:**113 – 135.

[19] Kollman, P.A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D.A., Cheatham II, T.E. (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **33:**889 – 897.

[20] Kelly, C.P., Cramer, C.J., Truhlar, D.G. (2005) SM6: A density functional theory continuum solvation model for calculating aqueous free energies of neutrals, ions and solute-water clusters. *J. Chem.Theory Comput.* **1:**1133 – 1152 and references therein.

[21] Klamt, A. (2005) *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design.* Elsevier, Amsterdam.

[22] Mulliken, R.S. (1934) New electroaffinity scale; together with data on valence states and on valence ionization potentials and electron affinities. *J. Chem. Phys.* **2:**782 – 793.

[23] Mulliken, R.S. (1955) Electronic population analysis on LCAO molecular wave functions l. *J. Chem. Phys.* **23:**1833 – 1840.

[24] Pauling, L. (1932) Nature of the chemical bond. lV. The energy of single bonds and the relative electronegativity of atoms. *J. Am. Chem. Soc.* **54:**3570 – 3582.

[25] Pauling, L. (1960) *The Nature of the Chemical Bond.* Cornell University Press, Ithaca, New York.

[26] Parr, R.G., Yang, W. (1989) *Density Functional Theory in Atoms and Molecules.* Oxford University Press, New York.

[27]   Pearson, R.G. (1991) Density functional theory: Electronegativity and hardness. *Chemtracts – Inorg. Chem.* **3:**317 – 333.

[28]   Fukui, K., Yonezawa, T., Nagata, C. (1957) Interrelations of quantum-mechanical quantities concerning chemical reactivity of conjugated molecules. *J. Chem. Phys.* **26:**831 – 841.

[29]   Sjoberg, P., Murray, J.S., Brinck, T., Politzer, P.A. (1990) Average local ionization energies on the molecular surfaces of aromatic systems as guides to chemical reactivity. *Can. J. Chem.* **68:**1440 – 1443.

[30]   Politzer, P.A., Murray, J.S., Grice, M.E., Brinck, T., Ranganathan,S. (1991) Radial behavior of the average local ionization energy of atoms. *J. Chem. Phys.* **95:**6699 – 6704.

[31]   Politzer, P.A., J.S. Murray, J.S., Concha, M.C. (2002) The complementary roles of molecular surface electrostatic potentials and average local ionization energies with respect to electrophilic processes. *Int. J. Quant. Chem.* **88:**19 – 27, and references therein.

[32]   Hussein, W., Walker, C.G., Peralta-Inga, Z., Murray, J.S. (2001) Computed electro-static potentials and average molecular ionization energies on the molecular surfaces of some tetracyclines. *Int. J. Quant. Chem.* **82:**160 – 169.

[33]   Murray, J.S., Abu-Awwad, F., Politzer, P.A. (2000) Characterization of aromatic hydrocarbons by means of average local ionization energies on their molecular surfaces. *J. Mol. Struct. (Theochem.)* **501 – 502:**241 – 250.

[34]   Becke, A.D., Edgecombe, K.E. (1990) A simple measure of electron localization in atomic and molecular systems. *J. Chem. Phys.* **92:**5397 – 5403.

[35]   Ehresmann, B., Martin, B., Horn, A.H.C., Clark, T. (2003) Local molecular proper-ties and their use in predicting reactivity. *J. Mol. Model.* **9:**342 – 347.

[36]   Stone, A.J. (1997) *The Theory of Intermolecular Interactions.* Oxford University Press, Oxford.

[37]   Politzer, P., Murray, J.S. (1991) Molecular electrostatic potentials and chemical reactivity. In: *Rev. Comput. Chem.* (Lipkowitz, K., Boyd, R.B. Eds). Vol. 2, pp. 273 – 312. VCH, New York.

[38]   Loew, L.M., MacArthur, W.R. (1977) A molecular orbital study of monomeric metaphosphate. Density surfaces of frontier orbitals as a tool in assessing reactivity. *J. Am. Chem. Soc*. **99:**1019 – 1025.

[39]   Mezey, P.G. (1993) *Shape in Chemistry.* VCH, New York.

[40]   Grant, J.A., Gallardo, M.A., Pickup, B. (1996) A fast method of molecular shape comparison: a simple application of a Gaussian description of molecular shape. *J. Comput. Chem.* **17:**1653 – 1666.

[41]   Rush, T.S., Grant, J.A., Mosyak, L., Nicholls, A. (2005) A shape-based 3D scaffold hopping method and its application to a bacterial protein-protein interaction. J. Med. Chem. **48:**1489 – 1495.

[42]   Gay, J.G., Berne, B.J. (1981) Modification of the overlap potential to mimic a linear site-site potential. *J. Chem. Phys.* **74:**3316 – 3319.

[43]   Duncan, B.S., Olson, A.J. (1995) *Approximation and Characterization of Molecular Surfaces*. Scripps Institute, San Diego, California.

[44]   Ritchie, D.W., Kemp, G.J.L. (1999) Fast computation, rotation and comparison of low resolution spherical harmonic molecular surfaces. *J. Comput. Chem.* **20:**383 – 395.

[45]   Lin, J.-H., Clark, T. (2005) An analytical, variable resolution, complete description of static molecules and their intermolecular binding properties. *J. Chem. Inf. Model.* **45:**1010 – 1016.

[46]   Horn, A.H.C., Lin, J.-H., Clark, T. (2005) Multipole electrostatic model for MNDO-like techniques with minimal valence spd-basis sets. *Theor. Chem. Acc.* **114:**159-168. Erratum available online *(DOI:* 10.1007/s00214 – 006 – 0167 – 4).

[47]   Stone, A.J. (1981) Distributed multipole analysis, or how to describe a molecular charge distribution. *Chem. Phys. Lett.* **83:**233 – 239.

[48]   Stone, A.J., Alderton, M. (1985) Distributed multipole analysis methods and applications. *Mol. Phys.* **56:**1047 – 1064.

[49]   Fincham, D., Mitchell, P.J. (1993) Shell model simulations by adiabatic dynamics. *J. Phys. Condens. Matter* **5:**1031 – 1038.

[50]   Lindan, P.J.D., Gillan, M.J. (1993) Shell-model molecular dynamics simulation of superionic conduction in calcium fluoride.*J. Phys. Condens. Matter* **5:**1019 – 1030.

[51]   Mezei, M., Campbell, E.S. (1977) Efficient multipole expansion: choice of order and density partitioning techniques. *Theoret. Chim. Acta* **43:**227 – 237.

[52]   Eisenschitz, R., London, F. (1930) The relation between the van der Waals forces and the homeopolar valence forces. *Z. Physik.* **60:**491 – 527.

[53]   London, F. (1930) Theory and systematics of molecular forces. Z. Physik. **63:**245 – 279.

[54]   London, F. (1927) The number of dispersion electrons in the mechanics of undulation. Z. Physik. Chem. (B). **15:**322 – 326.

[55]   London, F. (1937) The general theory of molecular forces.*Trans. Faraday Soc*. **33:**8 – 26.

[56] Schürer, G., Gedeck, P., Gottschalk, M., Clark, T. (1999) Accurate parameterized variational calculations of the molecular electronic polarizability by NDDO-based methods. *Int. J. Quant. Chem.* **75:**17 – 31.

[57] Martin, B., Gedeck, P., Clark, T. (2000) Additive NDDO-based atomic polarizability model. *Int. J. Quant. Chem.* **77:**473 – 497.

[58] Dewar, M.J.S., Zoebisch, E.G., Healy, E.F., Stewart, J.J.P. (1985) Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **107:**3902 – 3909; AM1. - Holder, A.J.(1998) *Encyclopedia of Computational Chemistry.* (Schleyer, P. v. R., Allinger, N.L., Clark, T., Gasteiger, J., Kollman, P.A., Schaefer, H.F. III, Schreiner, P.R. Eds). Vol. 1, pp.8 – 11. Wiley, Chichester.

[59]   *ParaSurf'06* (2006) Cepos InSilico Ltd, Ryde, U.K.

[60] Slater, J.C., Kirkwood, J.G. (1931) The van der Waals forces in gases. *Phys. Rev.* **37:**682 – 697.

[61] Martin, B., Clark, T. (2006) Dispersion treatment for NDDO-based semiempirical MO-techniques. *Int. J. Quant. Chem.* **106:**1208 – 1216.

[62] Dewar, M.J.S., Thiel, W. (1977) Ground states of molecules. 38. The MNDO method. Approximations and parameters. *J. Am. Chem. Soc.* **99:**4899 – 4907; Dewar, M.J.S., Thiel, W. (1977) Ground states of molecules. 39. MNDO results for molecules containing hydrogen, carbon, nitrogen, and oxygen. *J. Am. Chem. Soc.* **99:**4907 – 4917; MNDO. Thiel, W. (1988) *Encyclopedia of Computational Chemistry.* (Schleyer, P.v.R., Allinger, N.L., Clark, T., Gasteiger, J., Kollman, P.A., Schaefer, H.F. lll, Schreiner, P.R. Eds). Vol. 3, pp.1599 – 1604. Wiley, Chichester.

[63] Sponer, J., Leszczynski, J., Hobza, P. (1996) Base stacking in cytosine dimer. A comparison of correlated ab initio calculations with three empirical potential models and density functional calculations. *J. Comput. Chem.* **17:**841 – 850.

[64] Grimme, S. (2004) On the importance of telectron-correlation effects for the π-π interactions in cyclophanes. *Chemistry-A European Journal.* **10:**3423 – 3429; Hyla-Kryspin, I., Haufe, G., Grimme, S. (2004) Weak hydrogen bridges: A systematic theoretical study on the nature of C-H...F-C interactions. *Chemistry-A European Journal* **10:**3411 – 3422.

[65] Murray, J.S., Politzer, P.A. (1998) Statistical analysis of the molecular surface electrostatic potential: an approach to describing noncovalent interactions in condensed phases. *J. Mol. Struct. (Theochem.)* **425:**107 – 114.

[66] Murray, J.S., Lane, P., Brinck, T., Paulsen, K., Grice, M.E., Politzer, P.A. (1993) Relationships of critical constants and boiling points to computed molecular surface properties. *J. Phys. Chem.* **97:**9369 – 9373.

[67]   Ehresmann, B., de Groot, M.J., Alex, A., Clark, T. (2004) New molecular descriptors based on local properties at the molecular surface and a boiling- point model derived from them. *J. Chem. Inf. Comp. Sci.* **44:**658 – 668.

[68]   Ehresmann, B., De Groot, M., Clark, T. (2005) Surface-integral QSPR-models: Local energy properties. *J. Chem. Inf. Model.* **45:**1053 – 1060.

[69]   No, K.T., Kim, S.G., Cho, K.-H., Scheraga, H.A. (1999) Description of hydration free energy density as a function of molecular physical properties. *Biophys. Chem.* **78:**127 – 145.

[70]   Kuz'min, V.E., Ognichenko, L.N., Artemenko, A.G. (2001) Modeling of the informational field of molecules. *J. Mol. Model.***7:**278 – 285.

[71]   Shannon, C.E., Weaver, W. (1963) *The Mathematical Theory of Communication.* University of Illinois Press, Urbana.

[72]   Wang, R., Fang, X., Lu, Y., Wang, S. (2004) The PDBbind Database: Collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J. Med. Chem.* **47:**2977 – 2980.

[73]   Wang, R., Fang, X., Lu, Y., Yang, C.-Y., Wang, S. (2005) The PDBbind Database: Methodologies and updates. *J. Med. Chem.***48:**4111 – 4119.