Available online at www.sciencedirect.com
SCIENCE DIRECT
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 16 (2006) 1249–1254

The cytotoxicity of ortho alkyl substituted 4-X-phenols: A QSAR based on theoretical bond lengths and electron densities

R. J. Loader, N. Singh, P. J. O'Malley and P. L. A. Popelier*

School of Chemistry, Sackville Site, University of Manchester, Manchester M60 1QD, UK

Received 10 November 2005; revised 18 November 2005; accepted 19 November 2005
Available online 9 December 2005

Abstract—A new method called quantum topological molecular similarity (QTMS) was recently proposed [O'Brien, S. E.; Popelier, P. L. A. *J. Chem. Inf. Comp. Sci.* **2001**, *41*, 764] and has been shown to be successful in a variety of medicinal, ecological and physical organic QSAR/QSPRs. QTMS method uses electronic descriptors drawn from ab initio wavefunctions of geometry-optimized molecules. We investigated a remarkable and unusual set of ortho alkyl-substituted phenols [Selassie, C. D.; Verma, R. P.; Kapur, S.; Shusterman, A. J.; Hansch, C. *J. Chem. Soc., Perkin* **2002**, *II*, 1112], recently studied by the Hansch group. Our results do not support their proposal that a steric factor is important in the determination of the cytotoxicity of this set of substituted phenols. Thus, we conclude that the cytotoxicity of these sterically encumbered phenols is dependent primarily on electronic and radical effects, and that steric issues do not appear to be a critical distinguishing factor.

In recent years, many research groups have focused on the phenolic hydroxyl group due to its wide radius of activity.[1,2] On the one hand, it appears to act as an antioxidant or radical scavenger[3] whilst on the other hand, it demonstrates significant toxicity.[4,5] This dichotomy[6] in activity is believed to be associated with its hydrogen abstraction and subsequent formation of aryloxyl free radicals.

In a recent series of studies[7–10], the Hansch group examined the cytotoxicity of a set of simple and complex mono-substituted phenols towards a fast-growing murine leukaemia cell line. For electron-rich phenols a QSAR was obtained[9] comprising log *P* and the homolytic O–H bond dissociation enthalpy (BDE). The significant presence of the latter quantity implied a radical mechanism of toxicity. That study also reported that ortho mono-substituted phenols showed no significant dependence on either hydrophobic or steric parameters, since an excellent QSAR was obtained with a single BDE parameter.

In their most recent study[10], they investigated a set of poly-substituted electron-rich phenols, comprising analogues of the well-known and widely used antioxidants butylated hydroxyanisole (BHA) and butylated hydroxytoluene (BHT). Although regarded as safe food additives by the USA Food and Drug Administration, studies have shown that these molecules exhibit toxic (be it in very large doses) and carcinogenic properties in animals.[11] In an effort to gain insight into the mechanism of action of poly-substituted phenols like BHA and BHT, they studied the cytotoxicity towards the same leukaemia cell line (L1210). Remarkably, the best QSAR obtained (Eq. 6 in their paper) showed a considerable dependence on Taft's steric parameter E_{S-2} (for the larger of the two ortho substituents.) This strong dependence of cytotoxicity on E_{S-2} led them to examine the role of hydrophobicity. The resulting QSARs, all involving C log *P* and no steric parameters, gave rise to much poorer correlations. The authors concluded[12] that their ortho alkyl substituted phenol series 'deviates from most other phenols'. This conclusion was reinforced by three more QSARs of a series of 2,6-di-*tert*-butylphenols with substituents in the 4-position, linking the hydrogen abstraction reaction rate to their structure.

In this article, we re-examine the role of the steric effect by means of a new method developed in our laboratory, called quantum topological molecular similarity (QTMS). This method adopts the viewpoint of the quantum chemical electron density being the source of molecular properties. Originally it was designed as a

Keywords: Electron density; Atoms in molecules; Ab initio; Phenols; QSAR; Partial least square; Quantum chemical topology; QTMS.
* Corresponding author. Tel.: +44 161 3064511; fax: +44 161 3064559; e-mail: pla@manchester.ac.uk

0960-894X/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2005.11.079

practical implementation of the molecular similarity idea put forward by Carbó et al.,[13] in a similarity index depended as a 3D superposition of the electron densities of molecules *A* and *B*. We showed[14,15] that it suffices to compare the two densities by means of special points in space, as described by the topology of the electron density.[16,17] After successfully demonstrating this proof-of-principle, QSAR models[18] were constructed using partial least squares (PLS)[19] for the estimation of pK_a of carboxylic acids, anilines and phenols,[20] the prediction of σ_p, σ_m, σ_I and σ_p^0 parameters of mono-[14] and polysubstituted benzoic acids, phenylacetic acids and bicyclo carboxylic acids.[21] Most recent work[22] delivered a successful model for the prediction of σ^- Hammett constant of a set of para-substituted phenols and σ^+ of substituted toluenes and of bromophenethylamines. The action radius of QTMS did also extend into QSARs of medicinal[23–25] and ecological[21,26–28] nature. The QTMS approach inspired work in other groups (e.g., Ref. 29,30).

It should be stressed that QTMS (in its current stage of development) offers a reliable alternative to electronic parameters only.[22] In other words, we have accumulated substantial evidence that the two other types of parameters, hydrophobic and steric, are *not* modelled by QTMS. Of course, the QTMS descriptors can be combined with externally provided steric and/or lipophilicity parameters, but QTMS does not generate them independently. The important point is that because of this clear discriminatory capacity, QTMS can determine, based on quantum chemical electron densities, whether a given activity is due to steric effects or electronic effects. Here, we use QTMS to re-examine the remarkable and unusual set of ortho alkyl substituted phenols investigated by the Hansch group.[10] We wish to determine if the cytotoxic activity is captured by a QSAR model built from electronic effects or instead from steric effects, as claimed before by Selassie et al.[10]

The full details of the QTMS can be found in O'Brien et al.[15] but here we reiterate salient features. QTMS consists of three stages. First, geometry optimisation is performed on the dataset to obtain bond lengths and wavefunctions at a number of levels of theory. The lowest level is the semi-empirical AM1 method,[31] the next level is HF/6-31G(d),[32] and the third is B3LYP/6-31+G(d,p).[32] For convenience these levels are referred to as I, II and III. The ab initio calculations are performed with the GAUSSIAN03 suite of programs.[33]

Second, the so-called bond critical points (BCPs) are localised by a local version of the program MORPHY.[34] In short, BCPs are points where the gradient of the electron density vanishes. One BCP is found for every bond in the common skeleton, shown in Figure 1. Four topological descriptors (ρ, $\nabla^2\rho$, ε, and K(r)) are evaluated at each BCP, and together with the computed bond lengths (r_e), these five descriptors comprise the variable inputs. The topological descriptors are defined and extensively discussed elsewhere.[17] Loosely speaking, the first three descriptors can be interpreted as a measure of bond order, covalency and π character, respectively. We note here that optimised bond lengths on their own (not in conjunction with topological descriptors) can also provide a successful model, as discussed below (Table 2). Because of the lack of core electron densities in the AM1 method, level I does not yield BCPs, and a pure bond length model is the only type of model it generates.

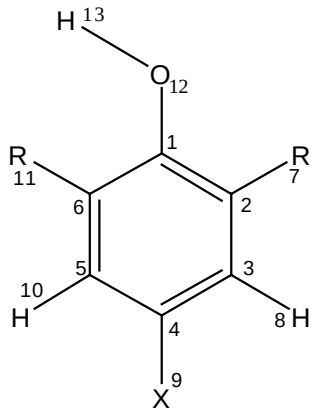

Figure 1. Labelling scheme for the series of 2-alkyl and 2,6-dialkyl-4-X-phenols.

Third and finally, a chemometric analysis employs PLS to construct the models from the descriptor set (''X-variables'') and experimentally obtained activity data (''Y-variables''). As a supervised method PLS combines linear least-squares with principal component analysis and constructs linear combinations of the X-variables, called latent variables (LV). The quality of QTMS-generated models is assessed by the correlation coefficient, r^2 and the cross-validated correlation coefficient, q^2. The 'leave one out' (LOO) q^2 coefficient is now thought to be a less trustworthy measure of internal predictivity and may contribute to optimistic q^2 values quoted in QSAR work.[35] Instead, the QTMS methodology uses the 'leave one seventh out' q^2 coefficient, which is more reliable and is the default used by the program SIMCA-P[36] software performing the PLS analysis. Further validation of the constructed models is measured by the randomisation validation statistics r^2(int) and q^2(int). These measures safeguard against correlations determined by chance by estimating the probability that a good fit can be obtained after random reorganisation of the dependent variables. That is, the activities become deliberately associated with the wrong descriptors, which should give rise to a deterioration of models fitted to permuted data. Each r^2 and q^2, generated for each 'scrambled' dataset, is then plotted against the absolute value of the correlation coefficient between the original set of activities and its permutation. Lines are drawn through the r^2 and q^2 values and the intercepts are examined. A model is deemed valid if $r^2(\text{int}) < 0.4$ and $q^2(\text{int}) < 0.05$. The randomisation test must be performed at least 10 times to ensure that a good model is not merely the result of pure chance. Together these four statistical measures provide a robust framework to assess the quality of the models constructed. The significance of the independent variables to the model