Journal of Computer-Aided Molecular Design, 6 (1992) 541–552
ESCOM

J-CAMD 179

Structure–activity relationship between the 3D distribution of the electrophilicity of sugar derivatives and their cytotoxic and antiviral properties

Alessandra Ricca[a,b], Jean M.J. Tronchet[a] and Jacques Weber[b,*]

[a]*Department of Pharmaceutical Chemistry and* [b]*Department of Physical Chemistry, University of Geneva, 30 quai E. Ansermet, 1211 Geneva 4, Switzerland*

Received 4 June 1992
Accepted 28 July 1992

Key words: Structure–property correlations; Reactivity index; Electrophilic reactivity; Michael acceptor sugars

SUMMARY

The cytotoxic activities of a series of sugar derivatives bearing electrophilic groups (1-cyanovinyl, 4-cyanochromen-2-yl and 3-nitrochromen-2-yl) have been correlated with their electrophilic properties. To this end, an electrophilic index was defined as an isovalue surface where the interaction energy with an incoming model nucelophile (H^-) was equal to a predefined value. This index, calculated from extended Hückel wave functions, allows one to quantify the electrophilic character of the substrates and to describe its spatial localization within the molecular volume (at Michael acceptor sites or on other parts of the molecules). Only sugars for which Michael acceptor reactivity was predicted were retained, and they were subdivided into two groups: those showing antiviral activity against a retrovirus and those devoid of such activity. Under these conditions, good correlations between cytotoxic activity and electrophilic reactivity – positive for the first group, negative for the second – were found. In addition, the ratio electrophilicity/sum of the absolute value of the dipole plus its projection along the principal axis of inertia, Z, of the molecule allows one to predict to which of these groups a sugar derivative belongs.

INTRODUCTION

We have previously described [1,2] both the synthesis and the cytotoxic and antiviral properties of a series of novel sugar electrophiles (Fig. 1), namely 1-cyanovinyl (**1**), 4-cyanochromen-2-yl (**3**), and 3-nitrochromen-2-yl (**4**) derivatives. In addition, some new species of nitroenose derivatives (**2**) were included in that study. These compounds, which a priori were representative of a homogeneous series of soft electrophiles, differed markedly in their cytotoxic and antiviral properties. This was strongly reminiscent of the situation encountered for derivatives of the ethacrinic acid

* To whom correspondence should be addressed.

type [3], where the structure of the group fixed onto the conjugated enonic electrophilic moiety exerted a considerable influence on biological activity. As we could not find a reasonable correlation between log P and biological activity, this ruled out an exclusive control of the differential toxicity of these compounds through their biodistribution properties. However, the observation that pretreatment of these compounds with soft nucleophiles such as L-cysteine abolished their biological activity (J.M.J. Tronchet and N. Dolatshahi, unpublished results) confirmed that their major mode of action was the blocking of vital nucleophiles. For these reasons, we decided to carry out a quantum chemical study of the electrophilic behavior of these molecules in an attempt to find a possible correlation with their cytotoxic and antiviral properties.

BIOLOGICAL METHODS

Cytotoxic activities were measured as previously described [1] for the compounds available at the start of this study (**1a**, **1b**, **1d$_1$**, **1d$_4$**, **1e**, **1h**, **1i**, **2b**, **2e**, **2h**, **4b**, **3g**) as well as for the geometrical isomers **1d$_2$ and 1d$_3$** prepared during this study to test our model (J.M.J. Tronchet and P. de Angelis, unpublished results). Antiviral activities against the oncogenic Polyoma virus have also been determined [1] for the same series of compounds.

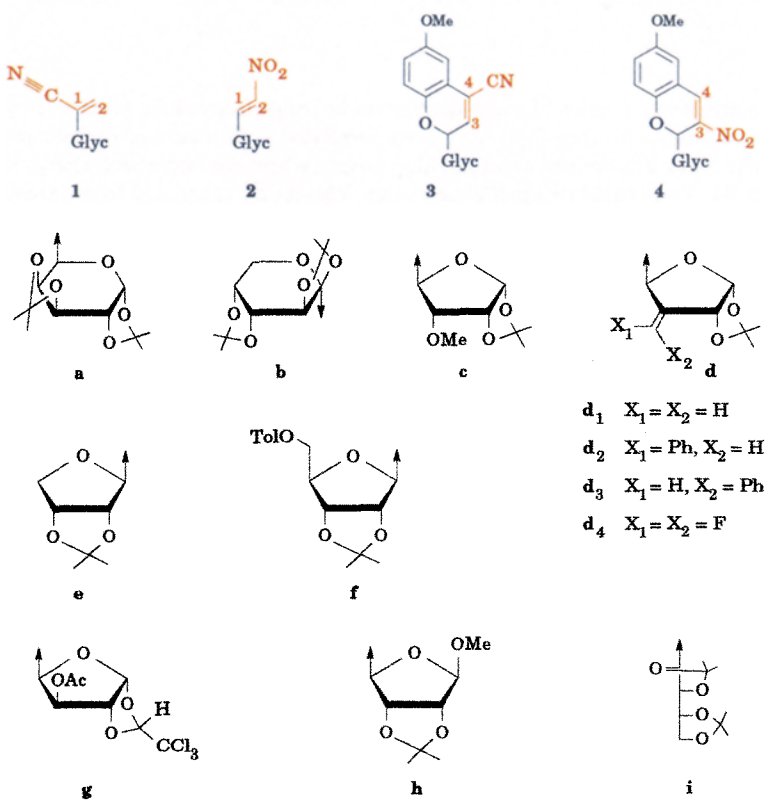

Fig. 1. Structures of the compounds studied (in red, the electrophilic part).

THEORETICAL MODEL

Our model is based on the evaluation of the interaction energy between a substrate, S, and a model reactant, R, exhibiting an electrophilic or a nucleophilic character. Calculated from extended Hückel (EH) wave functions, using a formalism we have recently developed [4,5], this energy has the advantage of being a local property which can play the role of a reactivity index. The interaction energy $E_{int}(\mathbf{r})$ between the substrate and the incoming reactant located in **r** is expressed as the sum of several components

$$E_{int}(\mathbf{r}) = E_{es}(\mathbf{r}) + E_{ct}(\mathbf{r}) + E_{ex}(\mathbf{r}) \tag{1}$$

with E_{es}, E_{ct}, E_{ex} being electrostatic, charge-transfer, and exchange-repulsion components, respectively. The regions where E_{int} is minimum are the sites of S most reactive to attack by R. To locate these zones, one may then display either the molecular surface of S color-coded according to E_{int} (Fig. 2A) or 3D solid models representative of isoenergy surfaces of E_{int} (Fig. 2B).

The reactant, R, is modeled as follows: a proton with a virtual 1s orbital for the electrophile and a H^- hydride ion with two electrons in the 1s orbital for the nucleophile.

Electrostatic component

In the case of electrophilic attack, E_{es} is equal to the molecular electrostatic potential (MEP) of substrate S [6]

$$E_{es}(\mathbf{r}) = \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{r}_A|} - \sum_\mu \sum_\nu P_{\mu\nu} \int \chi_\mu(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \chi_\nu(\mathbf{r}')\, d\mathbf{r}' \tag{2}$$

where the first term corresponds to nuclear repulsion, the summation running over all atoms A of

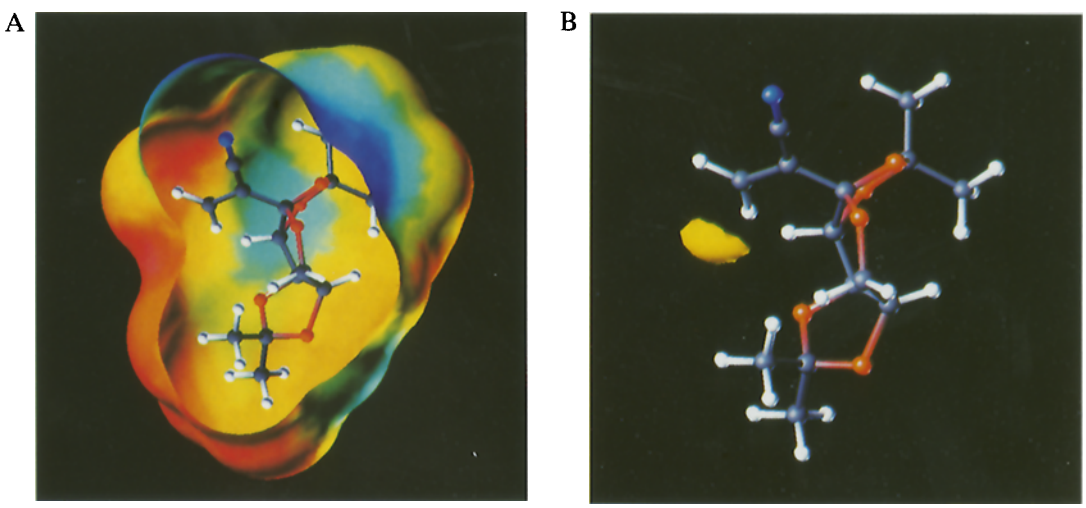

Fig. 2. Modeling of the interaction of compound **1a** with a nucleophile. A: molecular surface of **1a** colored according to E_{int} values (the red zones indicate the most negative sites). B: E_{int} isovalue surface (yellow lobe).

S, with nuclear charge, Z_A, located in $\mathbf{r}_A$. The second term describes the electronic attraction, $P_{\mu\nu}$ being the extended Hückel first-order density matrix element correponding to atomic orbitals χ_μ and χ_ν. As the evaluation of the integrals $\int \chi_\mu(\mathbf{r}') \frac{1}{|\mathbf{r}-\mathbf{r}'|} \chi_\nu(\mathbf{r}')\, d\mathbf{r}'$ over Slater-type orbitals is tedious, the NDDO approximation, based on the neglect of all three-center integrals, is used [7]. Therefore, the second term of Eq. 2 becomes [4]:

$$\sum_\mu \sum_\nu P_{\mu\nu} \int \chi_\mu(\mathbf{r}') \frac{1}{|\mathbf{r}-\mathbf{r}'|} \chi_\nu(\mathbf{r}')\, d\mathbf{r}' = \sum_A \sum_{\mu\varepsilon A} \sum_{\nu\varepsilon A} P_{\mu_A \nu_A} \int \chi_{\mu_A}(\mathbf{r}') \frac{1}{|\mathbf{r}-\mathbf{r}'|} \chi_{\mu_A}(\mathbf{r}')\, d\mathbf{r}' \tag{3}$$

In the case of nucleophilic attack, we assume that the electrostatic interaction between S and H^- reduces to that between S and a negative point charge, which is correct for rather large S–R distances where the so-called penetration integrals vanish. The electrostatic component is then taken as $-E_{es}$.

Charge-transfer component

The charge-transfer or 'orbital interaction energy' component is calculated as the difference between the total energies of the supermolecule and those of the separate fragments. Indeed, as EH total energies represent the sum of covalent energies within chemical bonds, the S–R charge-transfer energy may be approximated as [5]:

$$E_{ct}(\mathbf{r}) = E^{tot}(S-R,\mathbf{r}) - E^{tot}(S) - E^{tot}(R) \tag{4}$$

where $E^{tot}(X)$ represents the EH total energy of system X calculated as

$$E^{tot}(X) = \sum_i n_i \varepsilon_i \tag{5}$$

n_i and ε_i being occupation number and energy, respectively, of the ith molecular orbital (MO) of X.

Exchange-repulsion component

For electrophilic attack there is no exchange component as the reactant has no electrons. In the case of nucleophilic attack, we use a parametrized potential of a Buckingham type [8]

$$E_{ex}(\mathbf{r}) = C_1 \sum_{A\varepsilon S} k_{AH^-} \exp\left[-C_2 \frac{|\mathbf{r}-\mathbf{r}_A|}{R_A + R_{H^-}} \right] \tag{6}$$

where k_{AH^-} is an energy parameter dependent on atoms A and H^-, and R_A and R_{H^-} are the van der Waals radii of A and H^-, respectively.

Using the values of k_{AH^-}, C_1 and C_2 suggested by Hobza and Zahradnik [8], Eq. 6 becomes

$$E_{ex}(\mathbf{r}) = 828'000 \sum_{A\varepsilon S} k_{AH^-} \exp - 13.587 \left[\frac{\mathbf{r}-\mathbf{r}_A}{R_A + R_{H^-}} \right] \tag{7}$$

Computational details

All the EH calculations were performed using the single-zeta-Slater-type atomic orbitals of Clementi and Roetti [9]. For all hydrogen atoms and for the reactant, a 1s exponent of 1.0 was used. For the evaluation of E_{es}, self-consistent charge and configuration (SCCC) calculations were performed with quadratic dependence of the valence-state ionization energies (VSIEs) of all atoms. The calculation of E_{ct} was carried out without the SCCC procedure and with the VSIE H_{RR} of reactant systematically chosen at $\varepsilon_{HOMO}+0.2$ (eV) for electrophilic attack and $\varepsilon_{LUMO}-0.5$ (eV) for nucleophilic attack.

RESULTS

As the electrophilic sugars studied belong to the class of Michael acceptors, a theoretical study of their reactivity towards a nucleophile was performed. It was then possible to locate the most reactive site(s) of the sugar by representing isoenergy surfaces corresponding to low values of E_{int}. In order to make significant comparisons of the sizes and positions of these lobes, we chose for every compound:

$$E_{contour} = E_{int}(\text{minimum}) + 8.4\,\text{kJmol}^{-1} \tag{8}$$

With this representation, the information is more accessible and more accurate than with a classical color-coding of the molecular surface.

Regioselectivity

Before trying to find any structure–activity relationship, it was important to establish the predicted regioselectivity of the reaction, i.e., the relative propensity of each substrate to behave as a Michael acceptor or to undergo side-reactions (acid-base, substitutions). First of all, the geometries of the compounds were obtained from a conformational analysis performed with the semiempirical packages AMPAC [10] or SCAMP [11] and using the AM1 Hamiltonian. The torsion of the electrophilic group was obtained by successive rotations from 0° to 350° in 10° increments and the geometry of the molecule optimized for each value of the dihedral angle. The most stable structure was then submitted to an extended Hückel calculation. The use of two different semiempirical models in this study, namely AM1 and EH, is easily justified on the following grounds: it is well known that the AM1 and PM3 models are in general adequate for predicting the geometry of organic compounds such as the carbohydrates investigated here [12], which is obviously not the case for EH. However, the EH formalism may be used with advantage as a first approach to calculate intermolecular interaction energies, using the methodology reported here, as its clear physical meaning allows one to evaluate the charge–transfer component in a simple way [5]. The calculation of E_{int} would therefore be more complicated in all other quantum chemical models, were it of a semiempirical nature, as it requires second-order perturbation theory [13]. In our opinion, there is no contradiction in using two different models to calculate two different molecular properties.

A clear distinction appears between the sugars bearing the 1-cyanovinyl group (**1**) and those bearing the 2-nitrovinyl group (**2**). In the first case, the attack takes place mainly at carbon 2 of the 1-cyanovinyl group. This selectivity is in agreement with a Michael-type reaction. In the se-

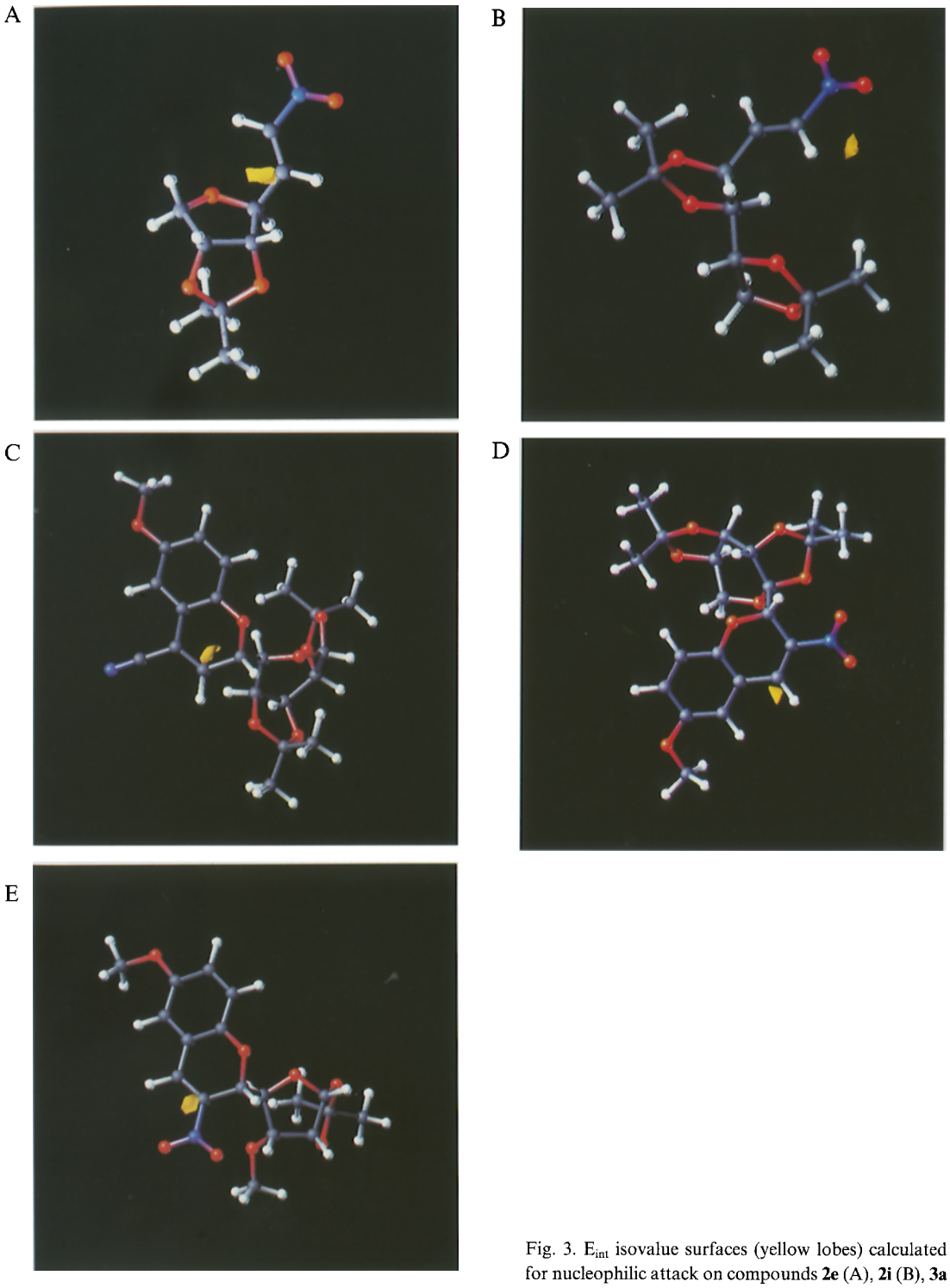

Fig. 3. E_{int} isovalue surfaces (yellow lobes) calculated for nucleophilic attack on compounds **2e** (A), **2i** (B), **3a** (C), **4b** (D) and **4c** (E).

cond case, the favorable site is found either at carbon 1 of the nitrovinyl group (Fig. 3A), which is consistent with a Michael-type reaction, or at carbon 2 (Fig. 3B), depending on the structure of the sugar moiety. This difference in regioselectivity can be explained by the fact that the nitro group is a better leaving group than the cyano group, which allows an addition–elimination reaction. Experimentally, sugars of type **2** were indeed shown to lose their nitro group easily [14].

The same calculations performed on a limited number of cyano- (**3**) and nitrochromenes (**4**) show the same general tendencies as those noted for **1** and **2** respectively, with compound **3** acting as a pure Michael acceptor (Fig. 3C) and compound **4** acting as a Michael acceptor (Fig. 3D) or preferring nucleophilic attack on the carbon bearing the nitro group (Fig. 3E). Compound **1g** is preferentially attacked at the carbon atom bearing the trichloromethyl group, whereas the *p*-nitro derivative of $\mathbf{1d_2}$ constitutes the best example of an acid–base reaction, implying the removal of the H-C4 proton. The Michael reactions are also predicted to be generally highly stereoselective (cf. Fig. 2, 3A, 3C, 3D). Such a stereoselectivity is commonly observed in nucleophilic addition on the carbonyl group of aldehydo-sugars bearing these types of sugar moieties (see, for example, [15]).

Structure–activity relationships

We found that the cytotoxic activities expressed as $-\log$ (minimum cytotoxic concentration (M)) correlated reasonably well with the surface of the electrophilic lobe at the Michael acceptor site. The surface of the lobe is represented as the extension of space having the same value of $E_{contour}$ as defined in Eq. 8. It is conveniently evaluated by the number of its constituent triangles. Actually, the size of the lobe is roughly proportional to the aperture angle of the cone defining the accessibility of the reactive site to a nucleophile; a lobe with a large size reflects that the site is accessible from a wide range of directions.

For antiviral activity we used an all-or-none variable, all compounds showing antiviral activity at half the cytotoxic concentration being considered active.

TABLE 1
CYTOTOXIC, ANTIVIRAL ACTIVITIES [−LOG (MINIMUM ACTIVE MOLAR CONCENTRATION)] AND MOLECULAR DESCRIPTORS OF SOME SUGAR ELECTROPHILES OF TYPES **1** AND **3**

Compounds	Cytotoxic activity	Antiviral activity[a]	Surface of the lobe	$\lvert\mu\rvert + \mu_z$
1b	3.449	4.148 (PI)	24	2.082
$\mathbf{1d_1}$	4.316	5.319 (TI)	72	4.112
$\mathbf{1d_2}$	4.452	4.753 (48h D)	60	3.760
$\mathbf{1d_4}$	4.988	5.29 (24h D)[b]	100	4.826
1e	3.892	4.194 (6h D)	48	3.371
1i	4.703	5.004 (PI)	40	3.064
1a	4.148	NA	40	5.256
$\mathbf{1d_3}$	4.753	NA	36	4.686
1h	4.654	NA	32	3.689
3g	5.284	NA	28	3.490

[a] TI total inhibition; PI partial inhibition; D delay in the manifestation of viral effect; NA no activity.
[b] Measured on retrovirus SV40.

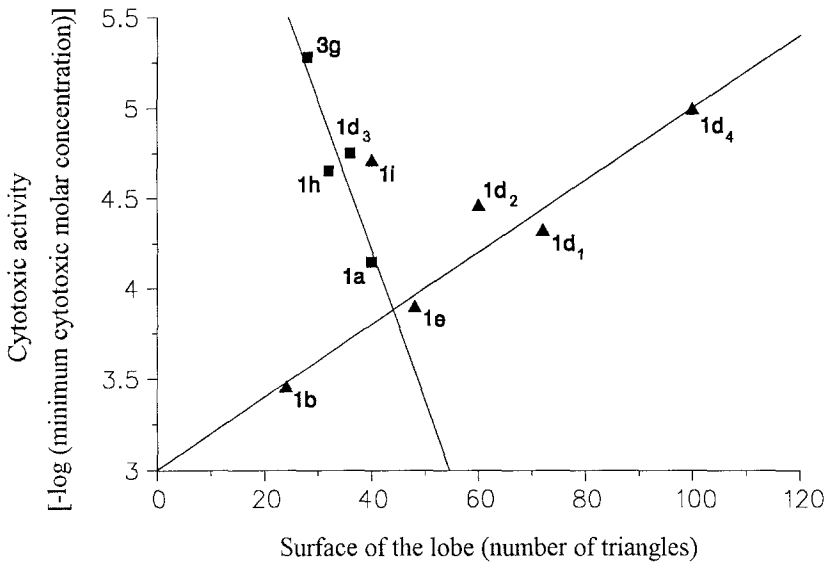

Fig. 4. Surface of the E_{int} isovalue lobe calculated for nucleophilic attack versus the cytotoxic activity for antiviral (▲) and non-antiviral (■) sugar electrophiles of type **1** and **3**.

Sugars of type 1 and 3

The data relative to these compounds are collected in Table 1. The compounds are aligned along two lines according to whether they showed antiviral activity or not (Fig. 4). The single exception is compound **1i**, the only linear sugar of this series, which exhibited antiviral activity but did not align with the other active compounds. Sugar **1i** was not taken into account in the regression analysis. The same QSAR model, 9, is obtained for both series:

$$(\text{Cytotoxic activity}) = b\,(\text{Surface of the lobe}) + a \qquad (9)$$

The following statistical parameters are obtained for compounds showing antiviral activity:

$a = 3.0043\ (\pm 0.6234);\ b = 0.0200\ (\pm 0.0095);\ r = 0.9683;\ s = 0.1679;\ F = 45.0738.$

For compounds devoid of antiviral activity the statistical parameters are shown below:

$a = 7.5224\ (\pm 3.7410);\ b = -0.0827\ (\pm 0.1093);\ r = 0.9175;\ s = 0.2267;\ F = 10.6481.$

For a given cytotoxic activity, however, these relationships are unable to predict whether a compound has antiviral activity or not. It was then necessary to find another descriptor able to make this difference. After trial and error, we found that a plot of the surface of the lobe versus the sum of the absolute value of the dipole moment plus its projection along the principal axis of inertia Z of the molecule can discriminate between the antiviral sugars and those devoid of antivi-

ral activity. Relationship 10 is obtained for both series:

$$(|\mu| + \mu_z) = b\,(\text{Surface of the lobe}) + a \tag{10}$$

The statistical parameters for antiviral compounds are the following:

$a = 1.5564\ (\pm 0.6304);\ b = 0.0345\ (\pm 0.01);\ r = 0.9784;\ s = 0.2171;\ F = 89.5939.$

For non-antiviral compounds the statistical parameters are shown below:

$a = -1.0705\ (\pm 3.8594);\ b = 0.1574\ (\pm 0.1127);\ r = 0.9735;\ s = 0.2339;\ F = 36.2084.$

Members of each subgroup aligned perfectly along their own line and the correlations were better than in the previous case (Fig. 5). In particular, the properties of compound **1i** were well predicted by these relationships.

One may wonder why the $|\mu| + \mu_z$ descriptor leads to such a clear discrimination between antiviral and inactive compounds. While the answer to this question is not entirely clear yet, one may suggest the following interpretation. There is no doubt that the surface of the isoenergy lobes of E_{int} is a nondirectional property, i.e., it reflects only, in an approximate way, the aperture angle of the cone defining the accessibility of the reactive site but not the direction of the cone axis. Therefore, as the surface of the lobe and several other nondirectional descriptors used in this study did not correlate with antiviral activity, we have inferred that directionality should be introduced

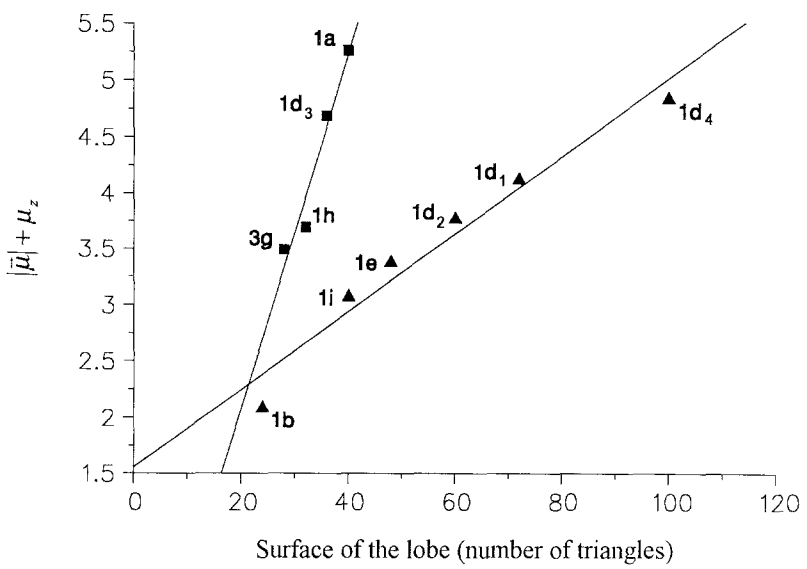

Fig. 5. Relationships between the surface of the E_{int} isovalue lobe and the sum of the dipole moment plus its projection along Z (Debye) for antiviral (▲) and non-antiviral (■) sugar electrophiles of types **1** and **3**.

somehow in a new descriptor, in particular through the dipole moment. Indeed, though the $|\boldsymbol{\mu}| + \mu_z$ property is a scalar, it reflects the trend of a compound to exhibit polarity oriented along the principal axis of inertia Z of the molecule; a compound exhibiting a large $|\boldsymbol{\mu}| + \mu_z$ obviously has a large dipole moment not necessary aligned along the Z axis. The situation is less clear, however, for compounds with small $|\boldsymbol{\mu}| + \mu_z$ values, as these may arise either from a small dipole moment or from a larger one with a negative μ_z component.

When examining Fig. 5, one notices that, compound **1b** excepted, the $|\boldsymbol{\mu}| + \mu_z$ descriptor has a value larger than 3 for all sugars, which indicates that the polarity of these compounds is rather large. However, when looking separately at their $|\boldsymbol{\mu}|$ and μ_z calculated values, no clear trend emerges as to their μ_z components; some of them are characterized by small μ_z values whereas others have large μ_z values, independently of their belonging to any antiviral or non-antiviral family. This shows that the Z axis does not play a particular role in characterizing these species, but rather one should consider the whole $|\boldsymbol{\mu}| + \mu_z$ descriptor. Nevertheless, Fig. 5 clearly shows that for given values of $|\boldsymbol{\mu}| + \mu_z$, the compounds that exhibit antiviral activity are those displaying large lobes of E_{int}, i.e., a good accessibility to the electrophilic site. This suggests that whereas active compounds are probably not all characterized by a dipole moment oriented in the same direction, site accessibility must be large enough to involve antiviral activity.

*Sugars of types **2** and **4***

The relevant data are collected in Table 2. Compounds of these two series are devoid of antiviral activity. Those compounds reacting as Michael acceptors exhibit a good correlation with a negative slope (Fig. 6), as already noted for the non-antiviral members of the two first series. QSAR model 9 applies with the following statistical parameters:

$a = 5.2702\ (\pm 0.8640)$; $b = -0.0227\ (\pm 0.0185)$; $r = 0.9654$; $s = 0.2011$; $F = 27.3984$.

CONCLUSION

Even if this study suffers from a number of limitations, the most severe ones being that only one conformer (the most stable) of each compound was submitted to computation and that the series of tested compounds was not very large, several interesting conclusions can be drawn. The biological activity of these compounds is clearly associated with their electrophilicity, but no linear relation is found before subdividing these compounds into two subgroups. Compounds showing anti-

TABLE 2
CYTOTOXIC ACTIVITY [−LOG (MINIMUM ACTIVE MOLAR CONCENTRATION)] AND SURFACE OF THE ELECTROPHILIC LOBE OF SOME SUGARS OF TYPES **2** AND **4**

Compounds	Cytotoxic activity	Surface of the lobe
2b	5.081	16
2e	4.635	24
2h	4.088	44
4b	3.639	76

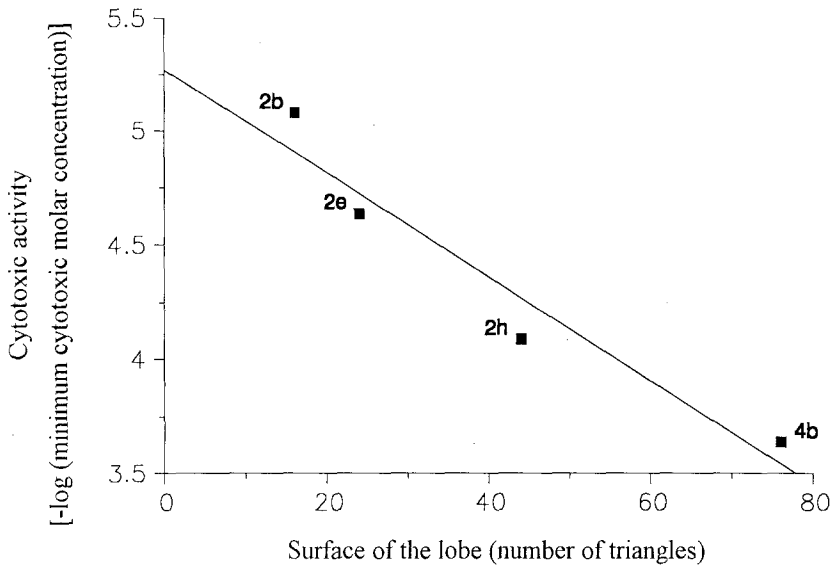

Fig. 6. Surface of the E_{int} isovalue lobe versus the cytotoxic activity of sugar electrophiles of types **2** and **4** devoid of antiviral activity.

viral activity constitute a homogeneous cluster, the cytotoxicity of which is positively related to their electrophilicity. A second homogeneous group corresponds to compounds devoid of antiviral activity, the cytotoxicity of which is negatively related to their electrophilicity. To interpret this behavior, one may suggest that the former group undergoes specific bonding with some vital nucleophile, which confers it with general (somewhat selective) toxicity. The latter group could be involved in aspecific bonding.

However, the biochemical relevance of the molecular descriptor used to assign each compound to one of the two groups, namely the absolute value of the dipole moment plus its projection along the principal axis of inertia, cannot be fully substantiated at this time and further investigations in this direction are required.

ACKNOWLEDGEMENTS

This work was supported by the Swiss National Science Foundation, grants No 20–31259.91 and 20–29856.90. The authors are grateful to Dr T. Clark for providing a copy of the SCAMP package.

REFERENCES

1 Tronchet, J.M.J., Zerelli, S., Dolatshahi, N. and Tuerler, H., Chem. Pharm. Bull., 36 (1988) 3722.
2 Tronchet, J.M.J., Pallie, K.D., Graf-Poncet, J., Tronchet, J.F., Werner, G.H. and Zerial, A., Eur. J. Med. Chem., 21 (1986) 111.
3 Cragoe, E.J., Jr, In Cragoe, Jr, E.J., (Ed.), Diuretics. Chemistry, Pharmacology and Medicine, John Wiley & Sons, New York, 1983, pp. 201–266.

4 Weber, J., Flükiger, P., Morgantini, P.Y., Schaad, O., Goursot, A. and Daul, C., J. Comput.-Aided Mol. Design, 2 (1988) 235.
5 Daul, C., Goursot, A., Morgantini, P.Y. and Weber, J., Int. J. Quantum Chem., 38 (1990) 623.
6 Scrocco, E. and Tomasi, J., Top. Curr. Chem., 42 (1973) 95.
7 Pople, J., Santry, D.P. and Segal, G., J. Chem. Phys., 43 (1965) S129.
8 Hobza, P. and Zahradnik, R., Weak Intermolecular Interactions in Chemistry and Biology, Academia Prague, 1980, p. 104.
9 Clementi, E. and Roetti, C., Atom. Data Nucl. Data Tables, 14 (1974) 177.
10 Dewar, M.J.S. and Stewart, J.J.P., Q.C.P.E. Bull., 6 (1986) 24, QCPE 527, AMPAC.
11 Clark, T., SCAMP version 4.2, private communication.
12 Stewart, J.J.P., J. Comput.-Aided Mol. Design, 4 (1990) 1.
13 Moriishi, H., Kikuchi, O., Suzuki, K. and Klopman, G., Theoret. Chim. Acta, 64 (1984) 319.
14 Tronchet, J.M.J., Bonenfant, A.P., Pallie, K.D. and Habashi, F., Helv. Chim. Acta, 62 (1979) 1622.
15 Tronchet, J.M.J., Pallie, K.D. and Barbalat-Rey, F., J. Carbohydr. Chem., 4 (1985) 29.