[g002]

# Telomerase inhibitory activity by stabilization of G-quartet: A QSAR approach using 2D autocorrelation descriptors.

Daimel Castillo-González [1, 2], Miguel Ángel Cabrera-Pérez [2], Maykel Pérez González [2,*], Alexander Durán-Martínez [1], Liane Saíz-Urra [2] and Marta Teijeira [3].

[1] Department of Chemistry. Central University of Las Villas. Santa Clara 54830, Villa Clara, Cuba

[2] Molecular Simulation & Drug Design Group. Centre of Chemical Bioactive. Central University of Las Villas. Santa Clara 54830, Villa Clara, Cuba.

[3] Department of Organic Chemistry, Vigo University, C.P. 36200 Vigo, Spain

[*]To whom correspondence should be addressed
Email: mpgonzalez76@yahoo.es
Phone: (53)-42-281473, (53)-42-281192
Fax: (53)-42-281130

**Keyword**: Telomerase inhibitors, G-quadruplex, molecular descriptors, QSAR.

**Abstract**

A QSAR approach on a dataset of 546 inhibitors of telomerase activity, by interaction with G-quadruplex DNA, was carried out using 2D autocorrelation descriptors. A linear discriminant analysis (LDA) was made on a training set of 437 compounds and it was assessed with 109 compounds belongs to the test set. The model good classified the 80.09% of the dataset and the percentage of good prediction was 78.89%. Only 5 compounds were not classified. The 2D autocorrelation descriptors are able to explain the factors that stabilize the G-quadruplex structure and consequently the inhibition of telomerase by this biological mechanism.

## Introduction

Telomeres are repetitive DNA sequences at the ends of linear chromosomes that protect the chromosome from recombination, end-to-end fusion and nuclease degradation.[1] In human cells, the telomeric DNA is typically composed of 5–15 kb of double-stranded pairs of tandem repeats of the guanine-rich sequence TTAGGG with a single-stranded 30-end overhang necessary to ensure complete chromosomal DNA replication. With each cell division, telomeres shorten by 50–200 bp because synthesis of the lagging strand of DNA is unable to replicate the 3'-end overhang. When the telomeres shorten to a critical length, "normal" cells stop growing and enter to state of senescence where end-to-end fusion and chromosomal instability leads to cell death. A cell can escape from this normal cycle and become immortal by stabilising (capping) the length of its telomeres.[2,3] This happens almost always under activation of the enzyme telomerase.[4]

Telomeres are believed to exist in different conformations together with several telomere-associated proteins, such as telomere repeat factors (TRF1, TRF2) and POT1.[5] The G-overhang is accessible for telomerase extension in the open state or inaccessible in a capped (or closed) conformation that involves the formation of a T-loop motif.[5] Although the T-loop structure has not been defined in detail, it may be created by the invasion of the G-overhang into the duplex region of the telomere.[6] Uncapping of the telomere ends leads to telomeric dysfunction characterized by end-to-end fusion, inappropriate recombination, anaphase bridges, and G-overhang degradation that may lead to either apoptosis or senescence.[7,8] Because of the repetition of guanines, the G-overhang is prone to formation of a four stranded G-quadruplex structure that has been shown to inhibit telomerase activity *in vitro.*[9,10]

The evidence that most cancer cells activate telomerase whereas normal cells are usually devoid of telomerase activity (with the exception of ongoing proliferating cells such as lymphocytes, basal keratinocytes, intestinal crypt cells, CD34 expressing peripheral blood stem cells, and germline cells)[11] has naturally lead to extensive investigations to detect this protein and its activity for a potential use in cancer diagnosis and prognosis, and to eventually monitor the tumor response to therapy. Finally, these data have greatly inspired the development of various strategies to target telomere and telomerase for cancer therapy.

Small molecules that stabilize G-quadruplexes are effective as telomerase inhibitors and several series of compounds have been identified. The ligands that stabilize G-quadruplex structures include cationic porphyrins,[12] perylenes,[13] amidoanthracene-9,10-diones,[14] 2,7-disubstituted amidofluorenones,[15] acridines,[16,17] ethidium derivatives,[18] disubstituted triazines,[19] fluoroquinoanthroxazines,[20] indoloquinolines,[21] dibenzophenanthrolines,[22] bisquinacridines,[23] pentacyclic acridinium,[24] telomestatin,[25] and the recently discovered bisquinolinium derivatives.[26,27] Due to the peculiar features of the quadruplex structure, as compared to classical double-stranded B-DNA, a selective recognition of telomeric G-quadruplex by small molecule ligands should be possible.[28,29] Some partial selectivity for G-quadruplex relative to duplex DNA was obtained with triazine[19] and with ethidium derivatives[30] and selectivity was significantly enhanced with the natural product telomestatin,[25] with a new series of 2,6-pyridin-dicarboxamide derivatives,[27] and with a porphyrin derivative.[29]

The search for new drugs against cancer plays a central role in the research programs of pharmaceutical companies and many governmental organizations due to the impact of this disease. Computational models that are able to predict the biological activity of compounds by its structural properties are powerful tools to design highly active molecules. In this sense, quantitative structure–activity relationships (QSAR) studies have been successfully applied for modeling biological activities of natural and synthetic chemicals.[31] Graph-theoretical and topological methods are included in the most QSAR studies. Among these methods, 2D spatial autocorrelations has been successfully used for modeling log *P* values,[32] biological activities,[33,34] for pharmaceutical research[35] and toxicological research. [36]

## Material and Methods

### *2D- autocorrelation descriptors*

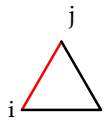
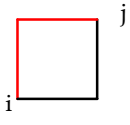
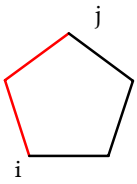
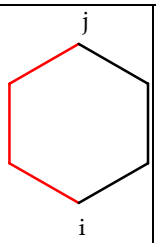
In this research were used the 2D autocorrelation descriptors to developed the QSAR function. The mathematical details of the method have been largely reported, [37-42] thus we will outline only the fundamental remarks.

As is well-know the binding of a substrate to its receptor dependents on the shape of the substrate and a variety of effects such as the molecular electrostatic potential, polarizability, hydrophobicity and lipophilicity.[43] Therefore, in a QSAR study the strategy for encoding molecular information must in some way, either explicitly or implicitly, account for these physicochemical effects.[44-48] Furthermore, usually data sets include molecules of different size with different numbers of atoms, so the structural encoding structures must allow comparison of such molecules.[49-52]

Autocorrelation vectors have several useful properties. First, a substantial reduction in data can be achieved by limiting the topological distance, *l*. Second, the autocorrelation coefficients are independent of the original atom numberings, so they are canonical. Finally, the length of the correlation vector is independent of the size of the molecule.[53]

For the autocorrelation vectors, H-depleted molecular structure is represented as a graph *G* and physico-chemical properties of atoms as real values assigned to the vertices of *G* (Table 1).

**Table 1.** Representation of different molecular graphs G and topological distances or spatial lags $d_{ij}$

| Molecular Graphs G |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| $d_{ij}$ | 1 | 1 | 2 | 2 | 3 | 4 |

These descriptors can be obtained by summing up the products of certain properties of two atoms located at given topological distances or spatial lag in *G*.

Three spatial autocorrelation vectors were employed for modeling the inhibitory activity.

Moran's index [54]

$$I(p_k, l) = \frac{N\sum_{ij}\delta_{ij}(p_{ki}-\overline{p_k})(p_{kj}-\overline{p_k})}{2L\sum_i(p_{ki}-\overline{p_k})}$$

Geary's index [54]

$$c(p_k, l) = \frac{(N-1)\sum_{ij}\delta_{ij}(p_{ki}-\overline{p_k})(p_{kj}-\overline{p_k})}{4L\sum_i(p_{ki}-\overline{p_k})}$$

Broto-Moreau's autocorrelation coefficient [33]

$$A(P_k, l) = \sum_i \delta_{ij} p_{ki} p_{kj}$$

where $I(P_k, l)$, $c(p_k, d)$ and $A(p_k, l)$ are Moran's index, Geary's coefficient and Broto–Moreau's autocorrelation coefficient at spatial lag $l$ respectively; $p_{ki}$ and $p_{kj}$ are the values of property $k$ of atom $i$ and $j$ respectively; $P_K$ is the average value of property $k$ and $\delta(l, d_{ij})$ is a Dirac-delta function defined as:

$$\delta(l, d_{ij}) = \begin{cases} 1 \text{ if } d_{ij} = 1 \\ 0 \text{ if } dij \neq 1 \end{cases}$$

where $d_{ij}$ is the topological distance or spatial lag between atoms $i$ and $j$. Spatial autocorrelation measures the level of interdependence between properties, and the nature and strength of that interdependence. It may be classified as either positive or negative. In a positive case, all similar values appear together, while a negative spatial autocorrelation has dissimilar values appearing in close association.[54,55] In a molecule, Moran's and Geary's spatial autocorrelation analysis tests whether the value of an atomic property at

one atom in the molecular structure is independent of the values of the property at the neighboring atoms. If dependence exists, the property exhibits spatial autocorrelation.[33,34] The calculation of the 2D-autocorrelations descriptors was carried out by means of the software package DRAGON version 5.4.[56] We used atomic masses, atomic van der Waals volumes, atomic Sanderson electronegativities and atomic polarizabilities as weighting properties. Autocorrelation vectors were calculated at spatial lags *l* ranging from 1 up to 8. The total number of computed descriptors was 96. Descriptors with constant or near to constants values were discarded.

***Data set and Computational strategies***

Linear discriminant analysis (LDA) has been chosen the statistical technique in most of the QSAR studies carried out using molecular descriptors.[57-60] In the present work, a similar expression for the QSAR is derived where TI is the acronym of Telomerase Inhibition. This variable reaches the values TI = 1 for actives compounds or TI = -1 for the non-active ones. Deciding whether a compound may be classified as an inhibitor is based on the information extracted from the literature.[14-20,22,27,61-79]

Forward stepwise was fixed as the strategy for variable selection.[80] To develop the QSAR for active/non-active compounds discrimination, we use the 2D autocorrelations molecular descriptors. The quality of the model was assessed by examining the Wilk's U-statistic, Mahalanobis distance, the percentage of good classification and the proportion between the cases and variables in the equation. Additionally, the model was validated calculating the percentages of good classification in the external prediction series. Compounds in the external prediction series were never used to develop the classification function.

One of the most important steps in computer-aided search of novel Telomerase inhibitors compounds is to design a representative, randomized training and predicting series. With this aim we selected a large and widely variable data set of 546 compounds inhibitors of telomerase with reported activity by interacting with G-quadruplex DNA. The following table shows the different families of compounds belongs to our data base.

**Table 1.** Some kinds of chemical compounds used in this study.

| Family | training | prediction | ref |
|---|---|---|---|
| Quinoline | 8 | 4 | [61,62] |
| Ethidium derivatives | 4 | 1 | [18] |
| Acridine derivatives | 68 | 17 | [16,22,63-67] |
| Pyridine derivatives | 2 | 2 | [27] |
| Quindoline derivatives | 8 | 3 | [68] |
| Acridone derivatives | 17 | 4 | [69] |
| Triazines | 174 | 39 | [19,70 71,72] |
| Anthracene-9,10-dione | 82 | 20 | [14,17,73,74] |
| Catecholic flavonoids | 20 | 7 | [75] |
| Amidofluorenones | 22 | 6 | [15,76] |
| Fluoroquinoanthroxazines | 3 | 1 | [20] |
| Porphyrins | 7 | 3 | [77-79] |

The data set was split in two parts, first group (actives) with concentrations required for 50% inhibition of the telomerase activity ${}^{tel}IC_{50} \leq 1$ μM (230 compounds) and the second group (non-actives) with values of ${}^{tel}IC_{50}$ greater than 1 μM (316 compounds). Some authors report as a potent inhibitor of telomerase that compound with values of ${}^{tel}IC_{50}<5$ μM.[71] For this reason, we used this break point. Further, the whole data was dividing in two parts using k-Means Cluster analysis (k-MCA),[49,51,81-83] training set 80% of the whole data, with 437 (183 compounds with ${}^{tel}IC_{50} \leq 1$ μM and 254 compounds with ${}^{tel}IC_{50} >1$ μM) and test set 20% of the data with 109 compounds (47 compounds with ${}^{tel}IC_{50} \leq 1$ μM values and 62 with ${}^{tel}IC_{50} >1$ μM). Selection of the training and prediction set was carried out taking randomly compounds belonging to each cluster. To ensure a statistically acceptable data partition into several clusters, we took into account the number of members in each cluster and the standard deviation of the variables in the cluster (as low as possible). We also made an inspection of the standard deviation between and within clusters, the respective Fisher ratio and their *p*-level of significance considered to be lower than 0.05.[80,84]

## Results and Discussion

Once performed a representative selection of training set it was used to fit the discriminant function. The model selection was subjected to the principle of parsimony. Then we chose a function with high statistical significance but having few parameters (descriptors) as possible. To derive a discriminant function that permits the classification of chemicals as active (Telomerase Inhibitors) or inactive (non-Telomerase Inhibitors) we use the linear discriminant analysis in which 2D autocorrelations descriptors were used as independent variables. The classification model obtained is given below together with the statistical parameters of the LDA:

$$TI = 7.02 \cdot ATS5v - 6.62 \cdot ATS4p + 6.07 \cdot MATS2m + 3.05 \cdot MATS5e + 3.03 \cdot GATS4m - 5.99 \cdot GATS4e + 1.29 \quad \text{(Eq. 1)}$$

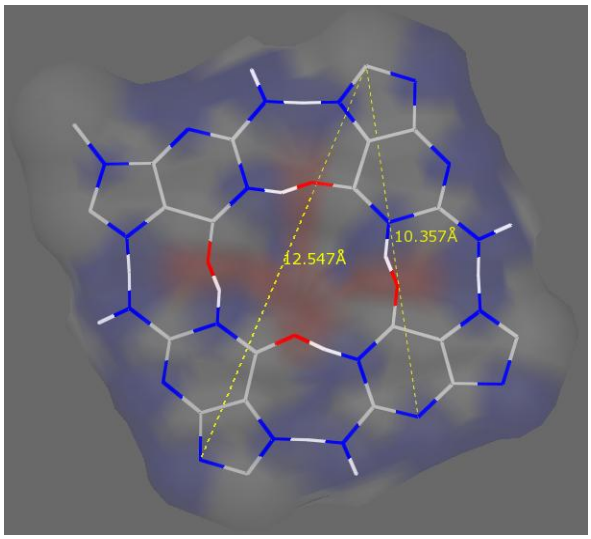
$U = 0.717$ N = 437 F (6,430) = 28.23 $p < 10^{-5}$ $D^2 = 1.61$ $\rho = 26.14$

In this model the coefficient $U$ is the Wilk's statistics, $D^2$ is the squared Mahalanobis distance, and $F$ is the Fisher ratio. The Wilk's $U$-statistics for the overall discrimination can take values on the range from 0 (perfect discrimination) to 1 (no discrimination). For the discrimination of active/inactive compounds studied here, the model classified correctly 93.97 % (152/183) of active and 78.88 % (198/254) of inactive compounds in the training series, for a global good classification of 80.09 %. The percentages of false actives and false inactive compounds in the training series were 20.87 % and 15.84 %, respectively. False actives are those inactive compounds that model classifies as actives, and the false inactive are those actives classified as inactive by the model. In addition, statistical outliers were not detected. The previous statement was based on two facts; all misclassified chemicals (accordingly to posterior probabilities and Mahalanobis's distance) do not rise to model improvement after leaving-out from it. Additionally, k-MCA demonstrates that any group of chemicals did not exist (possible outliers) that differentiate appreciably from the remnant ones.

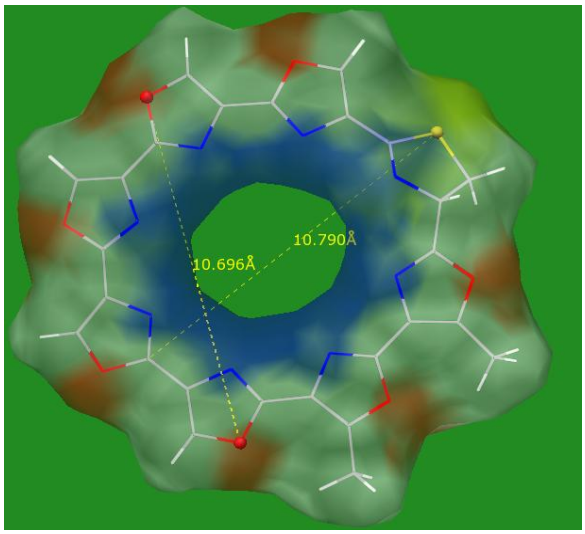
One of the most important criteria for the acceptance or not of a discriminant model, such as model (1), is based on the statistics for the external prediction series. Model 1 classified correctly 80.85 % (38/47) and 77.42 % (48/62) of active and inactive compounds in the prediction series, respectively, which represents an overall predictability of 78.89 %. Four of all compounds at prediction set were considered no classified by a 3.66 %

***Structural Interpretation***

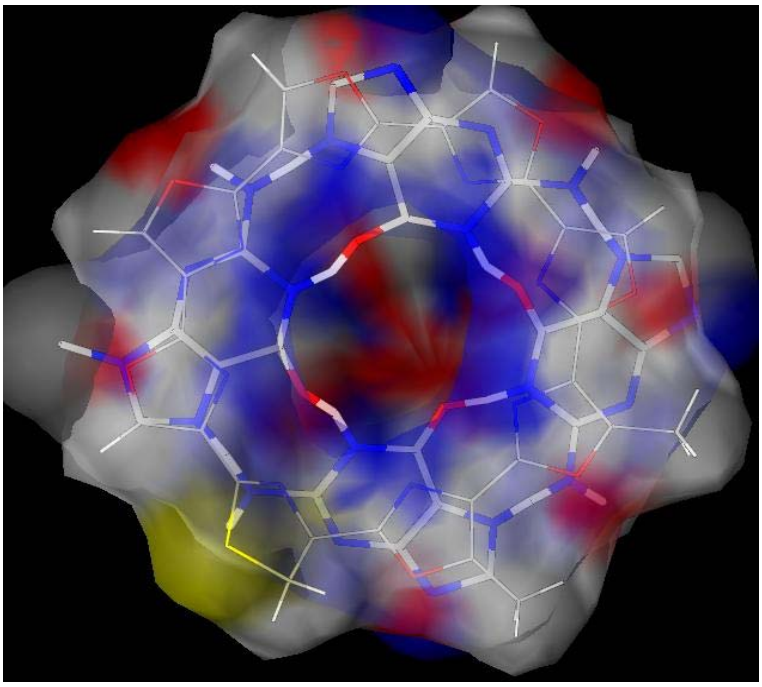
The factors that stabilize the G-quartet structure are known. Some of them such as hydrophobic and electrostatic interactions and hydrogen binding are related with the stabilization of the DNA structure. Molecules that acting by stabilization to G-quartet structure they must display a plane core, hydrophobic **fig 1a**, that allow the interaction of the molecule with guanines coplanar bases that they form the G-quartet structure. The positive contribution of ATS5v variable suggests that hydrophobic interactions play a fundamental role to distances of 5Å. The effect of compounds with high Van der Walls volume values to distances of 5Å increase the activity, a classic example is the telomestatin **fig 1b** , a natural product with one of the smaller values of $^{tel}IC_{50}$=0.05 µM.[25] Figure 2a shows the union of oxazole groups, with high polarizability values to this distance, which is in correspondence to the variable. Also this compound can interact correctly with G-quartet structure allowing the hydrophobic interactions (**fig1 c**).



A




B

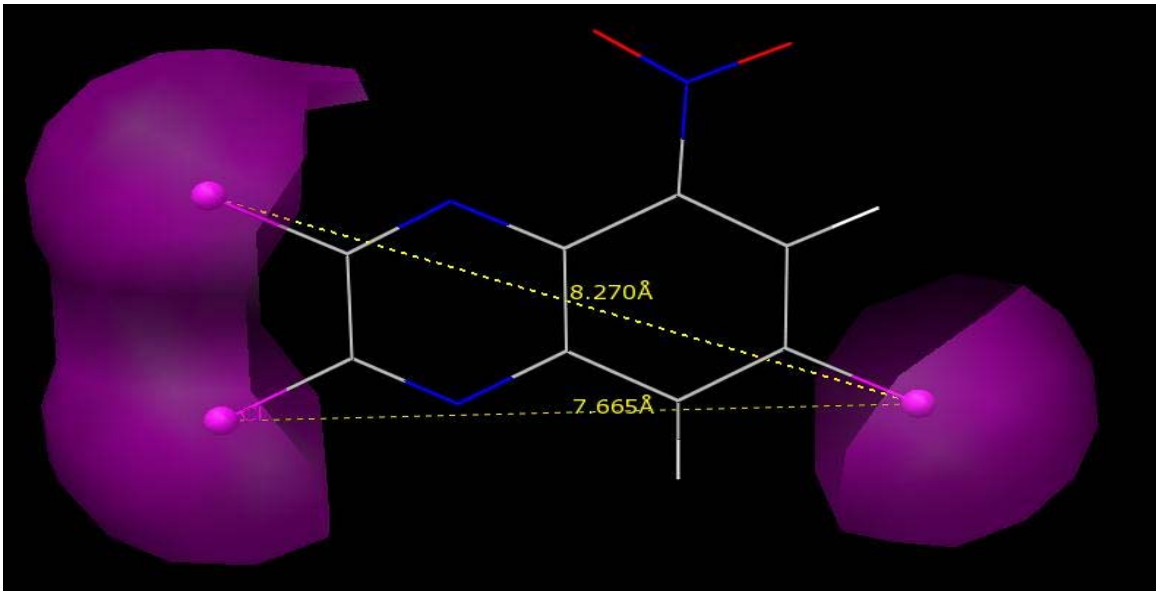


C

**Figure 1**. a) G-quadruplex structure. b) Telomestatin. c) Hydrophobic interactions between Telomestatin and G-quadruplex structure.

The equation 1 shows a positive contribution of MATS5e. This descriptor is able to explain the contribution of electronegative atoms on radius of 5Å such as O and S atoms, although is not possible to relate this kind of interaction with factors related with the the stabilization of G-quartet structure, at this distance.

Also the variable ATS4p has a negative contribution to the inhibition of telomerase. This descriptor is related with polarizability at 4Å of distance. At this distance the hydrophobic interaction not playing a main role yet. For example the quinoxalines (See figure 2a) has three chlorine atoms, and their big size and high polarizability it is in a ratio near to 4Å. This compound has a $IC_{50} > 1$ μM.[85] The same occurs with GATS4e variable, which contributes of negative form to the activity. The results shown in the fig 2a and 2b evidence this performance. In the second case (fig 2b), the acridone has six atoms of fluorine to this distance and this compound has $IC_{50}>139$ μM[69]



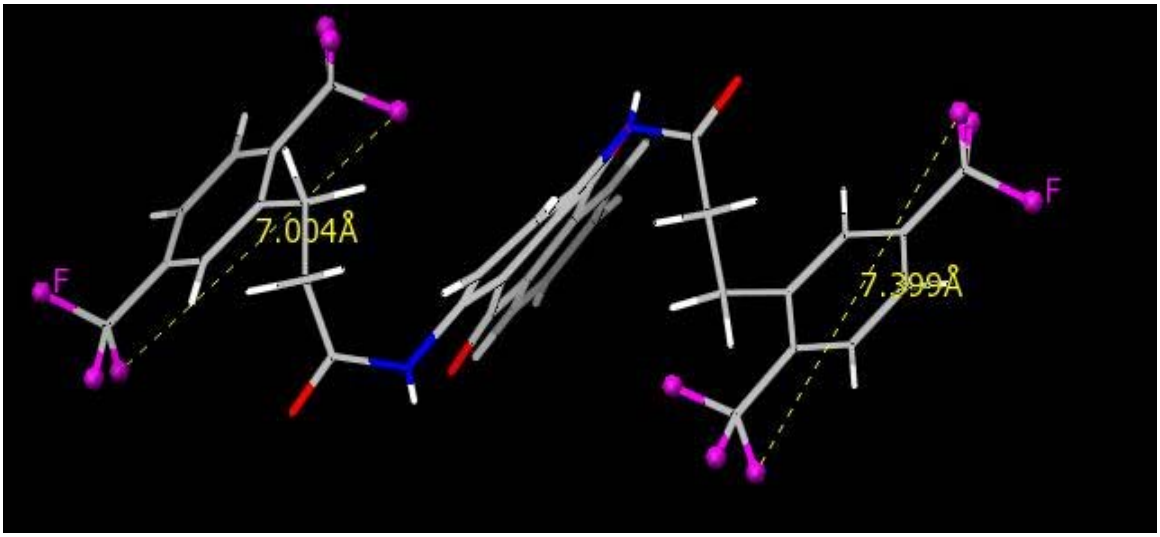



Figure2. a) Quinoxaline structure. b) Acridone

**Concluding Remarks**

Despite some criticism, there is an increasing necessity of topological-indices-based QSAR models in order to rationalize the drug discovery process. In this sense, the 2D-autocorrelations approach has been extended to the discovery of novel drugs leads,[86,87] but the QSAR models developed with these descriptors have been used only for reduced or homologous series of compounds. Consequently, the model capacity to predict the activity for different structural features is decreased. In the present paper the 2D-autocorrelation approach has been used to obtain good predictive linear models in order to account for TI activity. Thence, we can assert that the 2D- autocorrelation descriptors may be used as an efficient alternative to massive screening of drugs.

**Acknowledgments**

The authors acknowledge the Cuban Higher Education Ministry (R&D project number 6.181-2006) for financial support.

**References**


(1) Blackburn, E. H. *Cell* **1994**, *77*, 621-3.

(2) Bryan, T. M.; Cech, T. R. *Curr Opin Cell Biol* **1999**, *11*, 318-24.

(3) Harley, C. B.; Villeponteau, B. *Curr Opin Genet Dev* **1995**, *5*, 249-55.

(4) Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. *Science* **1994**, *266*, 2011-5.

(5) Smogorzewska, A.; de Lange, T. *Annu Rev Biochem* **2004**, *73*, 177-208.

(6) Griffith, J. D.; Comeau, L.; Rosenfield, S.; Stansel, R. M.; Bianchi, A.; Moss, H.; de Lange, T. *Cell* **1999**, *97*, 503-14.

(7) Duan, W.; Rangan, A.; Vankayalapati, H.; Kim, M. Y.; Zeng, Q.; Sun, D.; Han, H.; Fedoroff, O. Y.; Nishioka, D.; Rha, S. Y.; Izbicka, E.; Von Hoff, D. D.; Hurley, L. H. *Mol Cancer Ther* **2001**, *1*, 103-20.

(8) Li, G. Z.; Eller, M. S.; Firoozabadi, R.; Gilchrest, B. A. *Proc Natl Acad Sci U S A* **2003**, *100*, 527-31.

(9) Davis, J. T. *Angew Chem Int Ed Engl* **2004**, *43*, 668-98.

(10) Mergny, J. L.; Riou, J. F.; Mailliet, P.; Teulade-Fichou, M. P.; Gilson, E. *Nucleic Acids Res* **2002**, *30*, 839-65.

(11) Shay, J. W.; Bacchetti, S. *Eur J Cancer* **1997**, *33*, 787-91.

(12) Han, H.; Hurley, L. H.; Salazar, M. *Nucleic Acids Res* **1999**, *27*, 537-42.

(13) Rossetti, L.; Franceschin, M.; Bianco, A.; Ortaggi, G.; Savino, M. *Bioorg Med Chem Lett* **2002**, *12*, 2527-33.

(14) Perry, P. J.; Gowan, S. M.; Reszka, A. P.; Polucci, P.; Jenkins, T. C.; Kelland, L. R.; Neidle, S. *J Med Chem* **1998**, *41*, 3253-60.

(15) Perry, P. J.; Read, M. A.; Davies, R. T.; Gowan, S. M.; Reszka, A. P.; Wood, A. A.; Kelland, L. R.; Neidle, S. *J Med Chem* **1999**, *42*, 2679-84.

(16) Harrison, R. J.; Cuesta, J.; Chessari, G.; Read, M. A.; Basra, S. K.; Reszka, A. P.; Morrell, J.; Gowan, S. M.; Incles, C. M.; Tanious, F. A.; Wilson, W. D.; Kelland, L. R.; Neidle, S. *J Med Chem* **2003**, *46*, 4463-76.

(17) Read, M. A.; Wood, A. A.; Harrison, J. R.; Gowan, S. M.; Kelland, L. R.; Dosanjh, H. S.; Neidle, S. *J Med Chem* **1999**, *42*, 4538-46.

(18) Koeppel, F.; Riou, J. F.; Laoui, A.; Mailliet, P.; Arimondo, P. B.; Labit, D.; Petitgenet, O.; Helene, C.; Mergny, J. L. *Nucleic Acids Res* **2001**, *29*, 1087-96.

(19) Riou, J. F.; Guittat, L.; Mailliet, P.; Laoui, A.; Renou, E.; Petitgenet, O.; Megnin-Chanet, F.; Helene, C.; Mergny, J. L. *Proc Natl Acad Sci U S A* **2002**, *99*, 2672-7.

(20) Kim, M. Y.; Duan, W.; Gleason-Guzman, M.; Hurley, L. H. *J Med Chem* **2003**, *46*, 571-83.

(21) Caprio, V.; Guyen, B.; Opoku-Boahen, Y.; Mann, J.; Gowan, S. M.; Kelland, L. M.; Read, M. A.; Neidle, S. *Bioorg Med Chem Lett* **2000**, *10*, 2063-6.

(22) Mergny, J. L.; Lacroix, L.; Teulade-Fichou, M. P.; Hounsou, C.; Guittat, L.; Hoarau, M.; Arimondo, P. B.; Vigneron, J. P.; Lehn, J. M.; Riou, J. F.; Garestier, T.; Helene, C. *Proc Natl Acad Sci U S A* **2001**, *98*, 3062-7.

(23) Teulade-Fichou, M. P.; Carrasco, C.; Guittat, L.; Bailly, C.; Alberti, P.; Mergny, J. L.; David, A.; Lehn, J. M.; Wilson, W. D. *J Am Chem Soc* **2003**, *125*, 4732-40.

(24) Gowan, S. M.; Heald, R.; Stevens, M. F.; Kelland, L. R. *Mol Pharmacol* **2001**, *60*, 981-8.

(25) Shin-ya, K.; Wierzba, K.; Matsuo, K.; Ohtani, T.; Yamada, Y.; Furihata, K.; Hayakawa, Y.; Seto, H. *J Am Chem Soc* **2001**, *123*, 1262-3.

(26) De Cian, A.; Delemos, E.; Mergny, J. L.; Teulade-Fichou, M. P.; Monchaud, D. *J Am Chem Soc* **2007**, *129*, 1856-7.

(27) Pennarun, G.; Granotier, C.; Gauthier, L. R.; Gomez, D.; Hoffschir, F.; Mandine, E.; Riou, J. F.; Mergny, J. L.; Mailliet, P.; Boussin, F. D. *Oncogene* **2005**, *24*, 2917-28.

(28) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *Nucleic Acids Res* **2006**, *34*, 2723-35.

(29) Parkinson, G. N.; Lee, M. P.; Neidle, S. *Nature* **2002**, *417*, 876-80.

(30) Rosu, F.; De Pauw, E.; Guittat, L.; Alberti, P.; Lacroix, L.; Mailliet, P.; Riou, J. F.; Mergny, J. L. *Biochemistry* **2003**, *42*, 10361-71.

(31) Kubinyi, H. *QSAR: Hansch Analysis and Related Approaches.* New York., 1993.

(32) Devillers, J.; Domine, D. *Environ. Res.* **1997**, *7*, 195–232.

(33) Moreau, G.; Broto, P. *Nouv. J. Chim.* **1980**, *4*, 359–360.

(34) Moreau, G.; Broto, P. *Nouv. J. Chim.* **1980**, *4*.

(35) Wagener, M.; Sadowski, J.; Gasteiger, J. *J. Am. Chem. Soc.* **1995**, *117*, 7769–7775.

(36) Devillers, J., . ; Gordon and Breach Science 1999, p 595–612.

(37) Todechini, R.; Consonni, V. *Handbook of Molecular Descriptors*; Mannhold, R., Kubinyi, H., Timmerman, H., Eds ed. Wiley-ECH, 2000; Vol. 11.

(38) Moreau, G.; Broto, P. *Nouv J Chim* **1980**, *4*, 757-764.

(39) Moreau, G.; Broto, P. *Nouv J Chim* **1980**, *4*, 359-360.

(40) Broto, P.; Moreau, G.; Vandicke, C. *Eur J Med Chem* **1984**, *19*, 79-84.

(41) Broto, P.; Moreau, G.; Vandicke, C. *Eur J Med Chem* **1984**, *19*, 66-70.

(42) Broto, P.; Moreau, G.; Vandicke, C. *Eur J Med Chem* **1984**, *19*, 71-78.

(43) Saiz-Urra, L.; Gonzalez, M. P.; Fall, Y.; Gomez, G. *Eur J Med Chem* **2007**, *42*, 64-70.

(44) González, M. P.; Morales, A. H. *J Comput Aided Mol Des* **2003**, *17*, 665-72.

(45) González, M. P.; Dias, L.; Morales, A. H. *Polymer* **2004**, *15*, 5353-5359.

(46) González, M. P.; Morales, A. H.; Gonzalez-Diaz, H. *Polymer* **2004**, *45*, 2073-2079.

(47) González, M. P.; Morales, A. H.; Molina, R.; García, J. R. *Polymer* **2004**, *45*, 2773-2779.

(48) González, M. P.; Toropov, A. A.; Duchowicz, P. R.; Castro, E. A. *Molecules* **2004**, *9*, 1019-1033.

(49) González, M. P.; Morales, A. H.; Cabrera, M. A. *Bioorg Med Chem* **2005**, *13*, 1775-81.

(50) González, M. P.; Suarez, P. L.; Fall, Y.; Gomez, G. *Bioorg Med Chem Lett* **2005**, *15*, 5165-5169.

(51) González, M. P.; Terán, C.; Fall, Y.; Diaz, L. C.; Morales, A. H. *Polymer* **2005**, *46*, 2783 - 2790.

(52) González, M. P.; Teran, C.; Fall, Y.; Teijeira, M.; Besada, P. *Bioorg Med Chem* **2005**, *13*, 601-8.

(53) Bauknecht, H.; Zell, A.; Bayer, H.; Levi, P.; Wagener, M.; Sadowski, J.; Gasteriger, J. *J. Chem. Inform. Comput. Sci.* **1996**, *36*, 1205–1213.

(54) Moran, P. A. P. *Biometrika* **1950**, *37*, 17–23.

(55) Geary, R. F. *Incorp. Stat* **1954**, *5*, 115–145.

(56) Todeschini, R.; Consonni, V.; Maui, A.; Pavan, M.; Version 5.4 ed.; Talete srl: 2006.

(57) Perez, M. A.; Sanz, M. B.; Torres, L. R.; Avalos, R. G.; González, M. P.; Diaz, H. G. *Eur J Med Chem* **2004**, *39*, 905-16.

(58) Ramos de Armas, R.; Gonzalez Diaz, H.; Molina, R.; González, M. P.; Uriarte, E. *Bioorg Med Chem* **2004**, *12*, 4815-22.

(59) Morales, A. H.; Cabrera Perez, M. A.; González, M. P.; Ruiz, R. M.; Gonzalez-Diaz, H. *Bioorg Med Chem* **2005**, *13*, 2477-88.

(60) González, M. P.; Morales, A. H.; Collado, I. G. *Mol Divers* **2006**, *10*, 109-18.

(61) Cookson, J. C.; Dai, F.; Smith, V.; Heald, R. A.; Laughton, C. A.; Stevens, M. F.; Burger, A. M. *Mol Pharmacol* **2005**, *68*, 1551-8.

(62) Heald, R. A.; Modi, C.; Cookson, J. C.; Hutchinson, I.; Laughton, C. A.; Gowan, S. M.; Kelland, L. R.; Stevens, M. F. *J Med Chem* **2002**, *45*, 590-7.

(63) Incles, C. M.; Schultes, C. M.; Kempski, H.; Koehler, H.; Kelland, L. R.; Neidle, S. *Mol Cancer Ther* **2004**, *3*, 1201-6.

(64) Harrison, R. J.; Gowan, S. M.; Kelland, L. R.; Neidle, S. *Bioorg Med Chem Lett* **1999**, *9*, 2463-8.

(65) Schultes, C. M.; Guyen, B.; Cuesta, J.; Neidle, S. *Bioorg Med Chem Lett* **2004**, *14*, 4347-51.

(66) Moore, M. J.; Schultes, C. M.; Cuesta, J.; Cuenca, F.; Gunaratnam, M.; Tanious, F. A.; Wilson, W. D.; Neidle, S. *J Med Chem* **2006**, *49*, 582-99.

(67) Martins, C.; Gunaratnam, M.; Stuart, J.; Makwana, V.; Greciano, O.; Reszka, A. P.; Kelland, L. R.; Neidle, S. *Bioorg Med Chem Lett* **2007**, *17*, 2293-8.

(68) Zhou, J. L.; Lu, Y. J.; Ou, T. M.; Zhou, J. M.; Huang, Z. S.; Zhu, X. F.; Du, C. J.; Bu, X. Z.; Ma, L.; Gu, L. Q.; Li, Y. M.; Chan, A. S. *J Med Chem* **2005**, *48*, 7315-21.

(69) Harrison, R. J.; Reszka, A. P.; Haider, S. M.; Romagnoli, B.; Morrell, J.; Read, M. A.; Gowan, S. M.; Incles, C. M.; Kelland, L. R.; Neidle, S. *Bioorg Med Chem Lett* **2004**, *14*, 5845-9.

(70) Mailliet, P.; Riou, J. F.; Mergny, J. L.; Laoui, A.; Lavelle, F.; Petitgenet, O.; Aventis Pharma S.A.: France, 2003; Vol. US 6642964 B1.

(71) Mailliet, P.; Riou, J. F.; Alasia, M.; T., C.; Doerflinger, G.; Mergny, J. L.; Laoui, A.; Petitgenet, O.; Renou, E.; Aventis Pharma S.A.: France, 2005; Vol. US 6858608 B2, p 26.

(72) Mailliet, P.; Laoui, A.; Riou, J. F.; Doerflinger, G.; Mergny, J. L.; Hamy, F.; Caulfield, T.; Aventis Pharma S.A.: France, 2005.

(73) Perry, P. J.; Reszka, A. P.; Wood, A. A.; Read, M. A.; Gowan, S. M.; Dosanjh, H. S.; Trent, J. O.; Jenkins, T. C.; Kelland, L. R.; Neidle, S. *J Med Chem* **1998**, *41*, 4873-84.

(74) Huang, H. S.; Chou, C. L.; Guo, C. L.; Yuan, C. L.; Lu, Y. C.; Shieh, F. Y.; Lin, J. J. *Bioorg Med Chem* **2005**, *13*, 1435-44.

(75) Menichincheri, M.; Ballinari, D.; Bargiotti, A.; Bonomini, L.; Ceccarelli, W.; D'Alessio, R.; Fretta, A.; Moll, J.; Polucci, P.; Soncini, C.; Tibolla, M.; Trosset, J. Y.; Vanotti, E. *J Med Chem* **2004**, *47*, 6466-75.

(76) Neidle, S.; Harrison, R. J.; Reszka, A. P.; Read, M. A. *Pharmacol Ther* **2000**, *85*, 133-9.

(77) Shi, D. F.; Wheelhouse, R. T.; Sun, D.; Hurley, L. H. *J Med Chem* **2001**, *44*, 4509-23.

(78) Weelhouse, R. T.; Hurley, L. H.; Regents of the University of Texas system: United States of America, 2000; Vol. 6087493.

(79) Maraval, A.; Franco, S.; Vialas, C.; Pratviel, G.; Blasco, M. A.; Meunier, B. *Org Biomol Chem* **2003**, *1*, 921-7.

(80) Dillon, W. R.; Goldstein, M. *Multivariate analysis: Methods and applications*; John Wiley and Sons. Inc: N. Y., 1984.

(81) González, M. P.; Dias, L. C.; Helguera, A. M.; Rodriguez, Y. M.; de Oliveira, L. G.; Gomez, L. T.; Diaz, H. G. *Bioorg Med Chem* **2004**, *12*, 4467-75.

(82) González, M. P.; Gonzalez Diaz, H.; Molina Ruiz, R.; Cabrera, M. A.; Ramos de Armas, R. *J Chem Inf Comput Sci* **2003**, *43*, 1192-9.

(83) Molina, E.; Diaz, H. G.; González, M. P.; Rodriguez, E.; Uriarte, E. *J Chem Inf Comput Sci* **2004**, *44*, 515-21.

(84) Gago, F. B. *Métodos computacionales de modelado molecular y diseño de fármacos*; Universidad de Alcalá de Henares ed. Madrid, 2002.

(85) Kim, J. H.; Kim, J. H.; Lee, G. E.; Kim, S. W.; Chung, I. K. *Biochem J* **2003**, *373*, 523-9.

(86) González, M. P.; Caballero, J.; Morales, A. H.; Garriga, M.; González, G.; Fernandez, M. *Bull Math Biol* **2006**, *68*, 735-51.

(87) Saiz-Urra, L.; González, M. P.; Teijeira, M. *Bioorg Med Chem* **2007**, *15*, 3565-71.