Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Medicinal Chemistry 39 (2004) 281–289
EUROPEAN JOURNAL OF MEDICINAL CHEMISTRY
www.elsevier.com/locate/ejmech

Short Communication

Structure-based rationalization of antitumor drugs mechanism of action by a MIF approach

Gabriele Cruciani [a,*], Paolo Benedetti [a], Gianluigi Caltabiano [b], Daniele F. Condorelli [b], Cosimo G. Fortuna [b], Giuseppe Musumarra [b]

[a] *Laboratorio di Chemiometria e Chemioinformatica, Dipartimento di Chimica, Università degli Studi di Perugia, Via Elce di Sotto, 8, 06100 Perugia, Italy*
[b] *Dipartimento di Chimica, Università di Catania, Viale A. Doria, 6, 95125 Catania, Italy*

Received 13 October 2003; accepted 22 November 2003

Abstract

Structural patterns for antitumor drugs, evidenced by means of a molecular interaction field (MIF) approach using grid independent descriptors (GRIND), resembled closely those of a previous independent pharmacological classification based on their antitumor mechanism of action. For topoisomerase II inhibitors, antimitotic agents and DNA antimetabolites, systematic structural patterns were evidenced by MIF and the structural features of "outliers" in these classes corresponded to peculiar pharmacological mechanisms of action supported by literature evidences. Alkylating agents and DNA/RNA antimetabolites, interacting with a large variety of targets by different molecular mechanisms, did not exhibit clustering in the structure-based MIF approach. Moreover MIFS were able to point out similarities between drugs which, in spite of apparent dramatic differences in chemical structure, exhibit the same pharmacological behaviour.

Keywords: Antitumor drugs; Cheminformatics; 3D-QSAR; Grid independent descriptors

1. Introduction

The biological and chemical data "explosion" which occurred in the past two decades pointed out the need of bioinformatic and chemoinformatic approaches in handling and extracting information from enormous databases by decoding data signals and relating them to each other.

Effective cancer chemotherapy requires a better understanding of the complex relationships among drug activities, molecular targets and gene expressions in cancer cells from the same tissular origin. We recently reported on the results of multivariate insights into the impressive database available at the National Cancer Institute (NCI) [1], which includes the in vitro activities for common antitumor agents and the molecular targets and gene transcripts expressed in 60 human tumor cell lines. The above statistical approach allowed classification of antitumor agents according to their mechanism of action, to formulate hypotheses on the mechanism of novel anticancer agents, and to relate molecular targets to the in vitro activity of compounds acting by known or unknown mechanisms [2]. Multivariate analysis of the levels for almost 10,000 gene transcripts was able to evaluate the influence of genes on drug efficacy, to assess their selectivity for drugs acting by the same mechanism [3], and to identify either known genes most effective in discriminating different tumoral histotypes (i.e. candidates in the development of new diagnostic tests for cancer detection) or unknown genes deserving high priority in further studies [4]. The latter results point out the advantages of bioinformatics in decoding biological information to adopt shortcuts in genome-based cancer diagnostics and chemotherapy.

In addition to databases analysis, chemoinformatic strategies possess great potentialities in modelling the interactions between biopolymers and ligands. Molecular recognition plays in fact a fundamental role in drug–receptor interactions. Knowledge on the structural and energetic aspects of ligand–macromolecule interactions by molecular modelling allow to formulate hypotheses on drug–receptor molecular interactions and therefore an intelligent selection of structures to be synthesized. In this context, decoding the structure of chemical compounds by means of an appropriate statistical description is a crucial step.

* Corresponding author.
E-mail address: gabri@chemiome.chm.unipg.it (G. Cruciani).

doi:10.1016/j.ejmech.2003.11.013

In the field of drug design, the identification of mathematical descriptions relevant to the pharmacological properties of compounds has been a challenge for decades. Early work of Hansch [5] demonstrated that it is indeed possible to obtain functions correlating mathematical descriptions with the biological properties of compounds. Simple descriptors used in these studies (a few variables representing physicochemical properties) were intuitive and tried to represent what was known at that time about ligand–receptor interactions. Since then, many other mathematical descriptions have been used in drug design, and more specifically in the field of quantitative structure activity relationships (QSAR). Major milestones were the work of Goodford [6] who introduced the concept of molecular interaction field (MIF) and the work of Cramer et al. [7] who introduced the 3D chemical structure into the description of the compounds and hence developed the concept of 3D-QSAR.

However, in the field of 3D-QSAR, the initial alignment of the compounds in the series is widely recognized as one of the most difficult and time-consuming steps. Recently a novel methodology for producing a mathematical description of molecules called grid independent descriptors (GRIND) has been developed [8]. GRIND, easy to obtain even for large series of compounds, were found to be highly relevant with respect to biological properties and therefore are applicable in many different areas of drug design. An important characteristic of these novel descriptors is that they are insensitive to the position and orientation of the molecular structures in the space. Therefore, since the GRIND needs no alignment of compounds, 3D-QSAR analysis requires much less time as compared to standard methodologies. Moreover, the use of these descriptors is not limited to 3D-QSAR, and allows to extend their application in 3D searching, pharmacophore identification, and structure-metabolism relations.

Aim of the present work is to evidence, by means of GRIND descriptors, possible systematic 3D structural patterns for antitumor drugs with known mechanisms in the NCI standard database and/or to rationalise the absence of such grouping. Furthermore, interesting biochemical information might be obtained by comparing the present structure-based classification of compounds with the same mechanism of action, with a previously reported [1] independent pharmacological classification based on in vitro antitumor activities.

2. Methods

MIF were obtained using the program GRID version 21 [9]. GRIND were generated, analysed, and interpreted using the program ALMOND version 3.0 [www.moldiscovery-.com] a software package developed in our group. Computations and graphical display were performed on SGI O2 workstations (MIPS R12000 processor). The process of ligand–receptor interaction has often been represented with the help of the MIF.

If a compound is known to bind a certain receptor, some of the regions defined in its virtual receptor site (VRS) [8] should actually overlap groups of the real receptor site and, therefore, at least a subset of the VRS regions would be relevant for representing the binding properties of the ligand. For the latter statement to be true the VRS must have been obtained from the bioactive conformation of the ligand and the probes used to compute it should represent chemical groups present in the binding site. The molecular descriptors presented in this work are based on the concept of VRS. Basically, GRIND are a small set of variables representing the geometrical relationships between relevant regions of the VRS and as such are independent of the coordinate frame of the space where the MIF is computed. The procedure for obtaining GRIND involves three steps: (i) computing a set of MIF, (ii) filtering the MIF to extract the most relevant regions that define the VRS, and (iii) encoding the VRS into the GRIND variables.

2.1. Computing the MIF

To obtain relevant VRS, the probes used should represent potentially important groups of the binding site. For compounds interacting with proteins, it seems reasonable to use the DRY probe representing hydrophobic interactions, the O probe (carbonyl oxygen) to represent hydrogen bond acceptor groups, and the N1 probe (amide nitrogen) to represent hydrogen bond donor groups. In default mode a grid-spacing of 0.5 Å is used with the grid extending 5 Å beyond a molecule.

2.2. Filtering the MIF

For our purposes, we define the most interesting regions as those characterised by intense favourable (negative) energies of interaction. Unfortunately, a simple contouring procedure would not allow all such regions to be identified, as a region representing a very intense interaction can mask other interactions induced by different parts of the ligand. Instead, we have developed a procedure that extracts these “relevant regions” using an optimisation algorithm, selecting from each MIF a fixed number of nodes optimising a scoring function. This function includes two optimality criteria: the intensity of the field at a node and the mutual node–node distances between the chosen nodes. Therefore, the method extracts from each field a number of nodes (in the order of 150–200) that represent independent, favourable probe–ligand interaction regions. By default, the procedure provides a balanced solution, but sometimes, when the ligands contain charged groups or many polar substituents, the importance of the field values should be decreased, to extract all the recognizable independent regions. Another important parameter of the procedure is the number of extracted nodes. The method uses as default a value of 150, which is enough in most cases, but for compounds with complex structures this number can be increased.

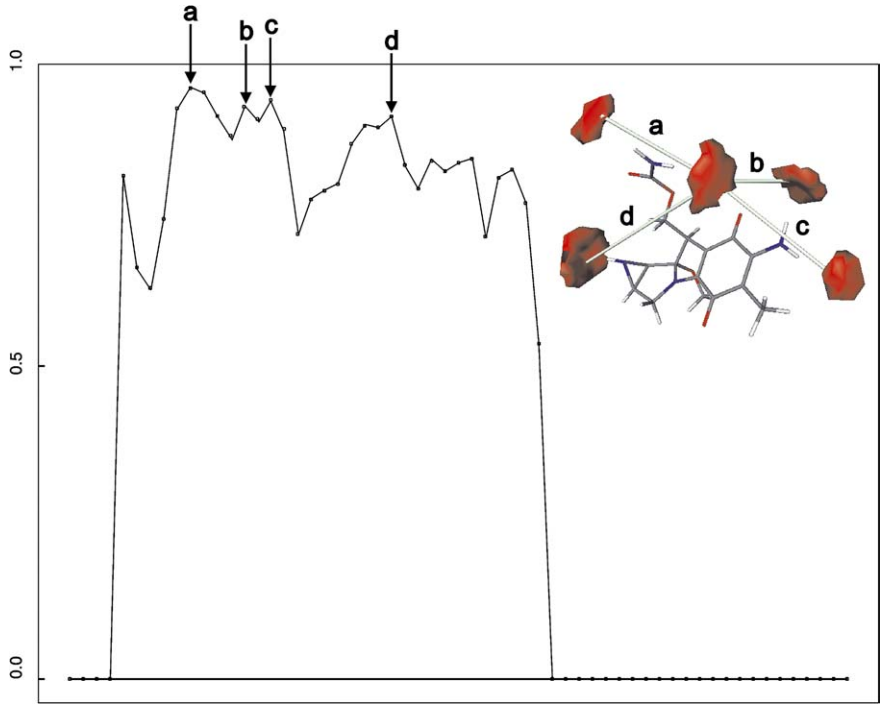


Fig. 1. Correlogram profile obtained for compound **72** using a hydrophobic probe (DRY).

2.3. Encoding the VRS into GRIND

As mentioned above, GRIND encodes the geometrical relationships between the VRS regions in such a way that they are no longer dependent upon their positions in the 3D space. Basically, the encoding is an auto- and cross-correlation transform [8]. The procedure works on the filtered nodes extracted by the previous step and computes the product of the interaction energy for each pair of nodes. The results of the products are handled according to the distance between the nodes. A discrete number of categories, each one representing a small rank of distances, are considered. In regular autocorrelation analysis, all computed terms are summed, and the result characterises each category. In our approach, only the highest product is stored, while others are discarded. This important difference is responsible for the "reversibility" properties of GRIND. A sum cannot be reverted to all its terms but the nodes producing the maximum product can be stored in the computer memory and traced back when necessary. Accordingly, this method is called maximum auto- and cross-correlation (MACC) or, more specifically, MACC-2 [8]. The values obtained from the analysis can be represented directly in correlogram plots, where the products of the node–node energies are reported vs. the distance separating the nodes. One single energy value is obtained for each of the categories considered and represents a small distance range. Fig. 1 shows the correlogram obtained for adriamycin (compound **72**), with hydrophobic probe (DRY). Every peak in the correlogram indicates that the VRS contains two regions separated by a distance corresponding to the abscissa of the peak.

2.3.1. Molecular description: GRID force field

The GRID program [6,10,11] was used to describe the molecular structures. GRID is a computational procedure for detecting energetically favourable binding sites on molecules. The program calculates the interactions between the molecule and a probe group which is moved through a regular grid of points in a region of interest around the target molecule and, at each point, the interaction energy between the probe and the target molecule is calculated as the sum of Lennard–Jones (E_{LJ}), hydrogen bond (E_{HB}) electrostatic interactions (E_{EL}) and, for specific probes, entropic contribution (E_{ENT}):

$$E_{x,y,z} = \sum_{i=1}^{N} E_{LJ} + \sum_{i=1}^{N} E_{HB} + \sum_{i=1}^{N} E_{EL} + \sum_{i=1}^{N} E_{ENT}$$

GRID contains a table of parameters to describe each type of atom occurring in each of the ligand molecules. These parameters define the strength of the Lennard–Jones, hydrogen bond and electrostatic interactions made by an atom and are used in order to evaluate the energy functions. GRID probes are very specific. They give precise spatial information, and this specificity and sensitivity are an advantage since the probes may then be representative of the important chemical groups present in the active site provided that the statistical method used for the analysis can distinguish between different types of interactions.

3. Results and discussion

In the present work 85 compounds belonging to five classes each including drugs with the same known mecha-

nism of action according to the NCI database [1] (**A**, alkylating agents, compounds **1–30; D**, DNA antimetabolites, compounds **31–46; M**, antimitotic agents, compounds **47–57; R**, RNA/DNA antimetabolites, compounds **58–70; T**, topoisomerase II inhibitors, compounds **71–85**) were considered (see Table 1). 2D structures for selected drugs discussed in detail below are reported in Scheme 1.

As compared to those previously classified in Ref. [2], six compounds were excluded from class **A** due to the presence of metal atoms, one (Morpholino-ADR) not included in class **T** (its 3D structure was not available) and five compounds (**50, 51, 52, 54** and **56**), recently classified as antimitotic agents[1], added in order to increase the number of drugs in class **M**. Compounds in classes **R** and **D** are the same as in Ref. [2].

The 3D structures of compounds **1–85** were imported, in Mol-file, from the NCI database and coded as ALMOND descriptors following the procedure described in the Section 2.

PCA afforded a 5 PC model explaning 75.0% of variance, 42.8% 1st PC, 13.8% 2nd PC, 9.0% 3rd PC, 4.6% 4th PC and 4.8% 5th PC.

Fig. 2 allows to detect those *X* variables or chemical descriptors that have a strong direct impact on chemical-clusterisation. The variable DRY-25 represents for instance an interaction between hydrophobic GRID-nodes separated by approximately 12 Å. In our series, topoisomerase II inhibitors are all characterised by having these interactions at the reported distance (see Fig. 2a).

The variable O-22 represents an interaction between ligand–H-bond donor regions separated by approximately 11 Å. They appear in the space formed by alcholic or imidazolic groups (see Fig. 2b). Finally, variable N1-33 is characteristic for topoisomerase II inhibitors, and it represents interactions between an aromatic and-or aliphatic hydroxyl groups in the 3D space (see Fig. 2c) occourring at 16 Å distance.

Fig. 3, the scores plot depicting the data structure elucidated by three PC (already explaining 66% of variance) appears a simplified graphical representation for an immediate evaluation of 3D structural MIF results, an appropriate basis for comparisons and discussions in relation to available pharmacological knowledge.

Topoisomease II inhibitors (class **T**, colour purple) belong to a tight cluster with positive PC1 and PC2 values. Antimitotic agents (class **M**, colour green) exhibit a more dispersed clustering characterised by positive PC2 values, while DNA antimetabolites (class **D**, colour red) are charactarised by negative PC2 values. For alkylating agents (class **A**, colour black) and RNA/DNA antimetabolites (class **R**, colour blue) no clear grouping can be evidenced in Fig. 3.

It is noteworthy that structure-based grouping in Fig. 3 parallels pharmacological classification based on the antitumor mechanism of action; such analogies are particular evident for drugs belonging to classes **T, M** and **D**.

Table 1
NCI standard database antitumor drugs considered in the present study [a]

Entry	NCI Code	Class [b]	Name
1	102627	**A**	Yoshi-864
2	132313	**A**	Dianhydrogalactitol
3	135758	**A**	Piperazinedione
4	142982	**A**	Hycanthone
5	167780	**A**	Asaley
6	172112	**A**	Spirohydantoin mustard
7	178248	**A**	Chlorozotocin
8	182986	**A**	AZQ
9	25154	**A**	Pipobroman
10	26980	**A**	Mitomycin C
11	296934	**A**	Teroxirone
12	3088	**A**	Chlorambucil
13	329680	**A**	Hepsulfam
14	338947	**A**	Clomesone
15	344007	**A**	Piperazine, 1(2-chloroethyl)-4-(3-chloropropyl)-dihydrochloride
16	34462	**A**	Uracil nitrogen mustard
17	348948	**A**	Cyclodisone
18	353451	**A**	Mitozolamide
19	357704	**A**	Cyanomorpholinodoxorubicin
20	409962	**A**	BCNU
21	56410	**A**	Porfiromycin
22	6396	**A**	Thio-tepa
23	73754	**A**	Fluorodopan
24	750	**A**	Busulfan
25	762	**A**	Nitrogen mustard
26	79037	**A**	CCNU
27	8806	**A**	Melphalan
28	95441	**A**	Methyl CCNU
29	95466	**A**	PCNU
30	9706	**A**	Triethylenemelamine
31	107392	**D**	5-HP
32	118994	**D**	Inosine glycodialdehyde
33	127716	**D**	5-aza-2′-deoxycytidine
34	145668	**D**	Cyclocytidine
35	1895	**D**	Guanazole
36	27640	**D**	2′-deoxy-5-fluorouridine
37	303812	**D**	Aphidicolin glycinate
38	32065	**D**	Hydroxyurea
39	330500	**D**	Macbecin II
40	51143	**D**	Pyrazoloimidazole
41	63878	**D**	Ara-C
42	71261	**D**	Beta-TGDR
43	71851	**D**	Alpha-TGDR
44	752	**D**	Thioguanine
45	755	**D**	Thiopurine
46	95678	**D**	3-HP
47	125973	**M**	Taxol
48	153858	**M**	Mytansine
49	332598	**M**	Rhizoxin
50	33410	**M**	Colchicine derivative
51	361792	**M**	Thiocolchicine
52	406042	**M**	Allocolchicine
53	49842	**M**	Vinblastine sulfate
54	608832	**M**	Taxol derivative
55	67574	**M**	Vincristine sulfate
56	757	**M**	Colchicine

(continued on next page)

Table 1
(continued)

Entry	NCI Code	Class [b]	Name
57	83265	**M**	Trityl cysteine
58	102816	**R**	5-azacytidine
59	126771	**R**	Dichlorallyl lawsone
60	139105	**R**	Baker's soluble antifol
61	143095	**R**	Pyrazofurin
62	148958	**R**	Ftorafur(pro-drug)
63	153353	**R**	L-alanosine
64	163501	**R**	Acivicin
65	19893	**R**	5-fluorouracil
66	224131	**R**	*N*-(phosphonoacetyl)-L-aspartate(PALA)
67	264880	**R**	5,6-dihydro-5-azacytidine
68	352122	**R**	Trimetrexate
69	368390	**R**	Brequinar
70	740	**R**	Methotrexate
71	122819	**T**	VM-26
72	123127	**T**	Doxorubicin
73	141540	**T**	VP-16
74	164011	**T**	Rubidazone
75	249992	**T**	*m*-AMSA
76	267469	**T**	Deoxydoxorubicin
77	268242	**T**	*N,N*-dibenzyl daunomycin
78	269148	**T**	Menogaril
79	301739	**T**	Mitoxantrone
80	308847	**T**	Amonafide
81	337766	**T**	Bisantrene HCL
82	349174	**T**	Oxanthrazole
83	355644	**T**	Anthrapyrazole derivative
84	366140	**T**	Pyrazoloacridine
85	82151	**T**	Daunorubicin

[a] 3D structures available on: http://dtp.nci.nih.gov.
[b] **A**, alkylating agents; **D**, DNA antimetabolites; **M**, antimitotic agents; **R**, RNA/DNA antimetabolites; **T**, topoisomerase II inhibitors.

Systematic structural patterns and eventually the presence of "outliers" in Fig. 3 are discussed below for each class of drugs in relation to their pharmacological behaviour.

*3.1. Class **T** (topoisomerase II inhibitors, compounds **71–85**)*

Fig. 3 shows a tight cluster for compounds **72, 73, 76, 79, 82** and **83**, characterised by both PC1 and PC2 positive values and the corresponding correlograms evidence significant energetic values for the N1 probe in variables 27–41, indicating distant H-bonding acceptor groups. High energetic values for dry probe, consistent with the presence of a single dry group (anthraquinone), can be obtained for short distances. Accordingly **74** and **78** do not belong to the above tight cluster, the former due to the additional presence of a phenyl ring distant from the principal dry group and the latter due to the steric hindrance of two OH which make the dry less accessible.

Topoisomerase II (TOP2) poisons interfere with the breakage/reunion reaction of TOP2 resulting in DNA cleavage. However, the mechanism of action of compounds **74** (Rubidazone) and **78** (menogaril) does not differ significantly from that of other members of class **T**. Menogaril is a semisynthetic anthracycline that is less cardiotoxic than doxorubicin in a preclinical model: using purified mammalian topoisomerases, menogaril was shown to poison topoisomerase II but not topoisomerase I [12]. Rubidazone (or zorubicin) belongs to the group of the anthracyclines and acts as topoisomerase I inhibitor. Although both molecules are characterised by original metabolic and pharmacokinetic features it is not possible to envisage any correlation with structural properties.

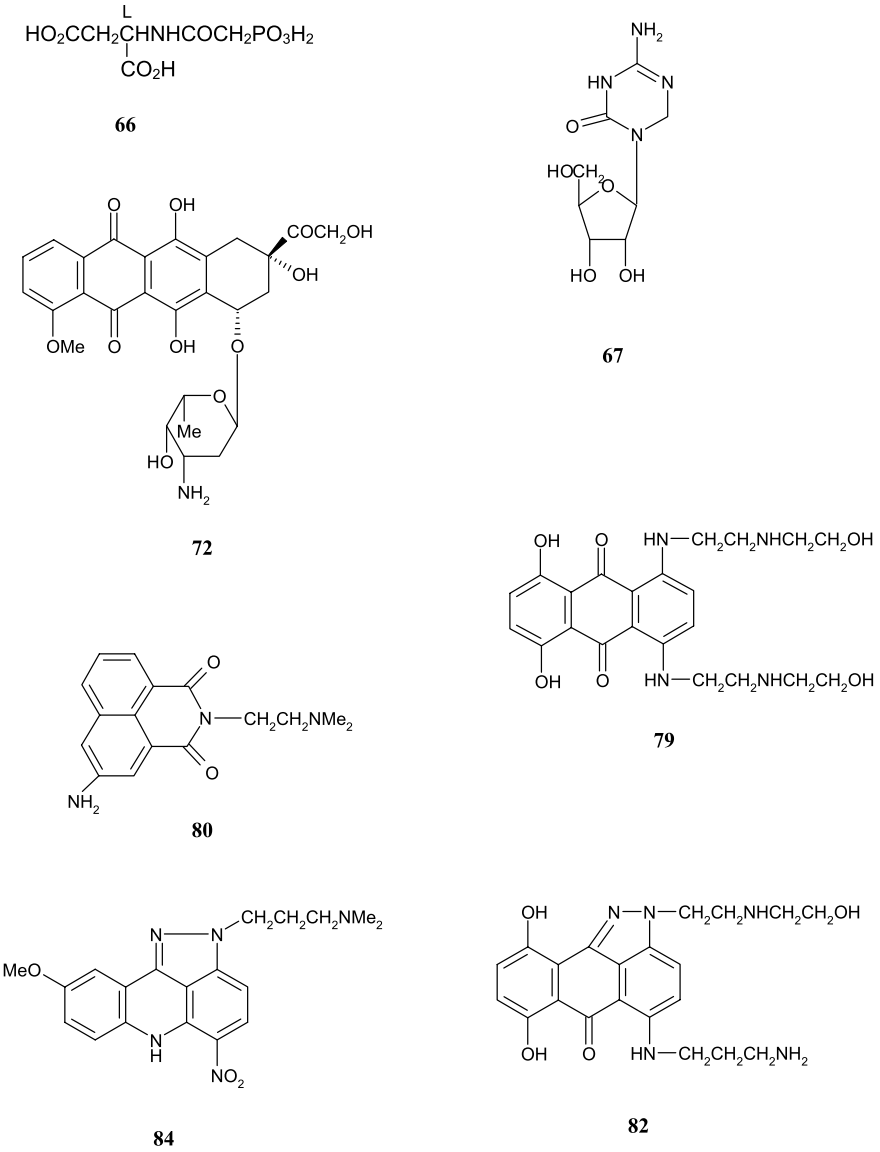
Drugs **75, 80, 81** and **84**, characterised by positive PC3 values, are away from the **T** cluster due to the absence of OH groups (strong H-bonding donors).

Wang et al. [13] have recently reported the existence of two different classes (ATP-sensitive and -insensitive) TOP2 poisons that can be identified due to their differential sensitivity to the ATP-bound conformation of TOP2. In the presence of 1 mM ATP or the nonhydrolyzable analog adenosine 5′-(beta, gamma-imino)triphosphate, TOP2-mediated DNA cleavage induced by ATP-sensitive TOP2 poisons (e.g. doxorubicin, etoposide, mitoxantrone) was stimulated 30–100-fold, whereas DNA cleavage induced by ATP-insensitive TOP2 poisons (e.g. amonafide, batracylin, and menadione) was only slightly (less than threefold) affected. In addition, ADP was shown to strongly antagonize TOP2-mediated DNA cleavage induced by ATP-sensitive but not ATP-insensitive TOP2 poisons. Interestingly, amonafide (**80**) is far away from the **T** cluster which includes mainly ATP-sensitive TOP2 inhibitors. Another drug located far from the **T** cluster, is pyrazoloacridine (**84**: PZA) an investigational nucleic acid binding agent that inhibits the activity of topoisomerases I and II through a mechanism distinct from other topoisomerase poisons [14]. PZA is an investigational anticancer agent with a tetracyclic structure containing a 9-methoxy substitution and a potentially reducible 5-nitro ring substituent [15]. PZA exerts broad cytotoxicity against tumor cell lines in vitro and in vivo [16–18]. Furthermore, PZA retains activity against several drug-resistant cancer phenotypes including cells that are resistant to etoposide and doxorubicin through deficient topoisomerase II-mediated DNA damage, and cells that are camptothecin-resistant because of low levels of topoisomerase I [17,19]. PZA exerts cytotoxicity by virtue of its ability to bind to nucleic acids, which interferes with RNA synthesis, DNA synthesis, and DNA repair [19,20]. It interferes with the normal function of both topoisomerase I and II through a mechanism that is distinct from commercially available topoisomerase I and II-targeting agents.

The above considerations, pointing out literature evidences for the peculiarities of **80** and **84** on their mechanism of action, parallels and corroborates structure-based peculiarities evidenced by the MIF approach.

*3.2. Class **M** (antimitotic agents, compounds **47–57**)*

Class **M**, including well known antitumor agents such as taxol (**47**) vinblastine (**53**) and vincristine (**55**), is character-



Scheme 1

ised by positive PC1 and PC2 PCA scores values. These molecules show H-bonding acceptor network. The H-bond interaction energy for these molecules is by far the most intense. For example, their capabilities as H-bond acceptors are twice as strong as those produced by **D** class. Moreover **M** class molecules are quite large and report hydrophobic interaction at large distances. For example, taxol shows two hydrophobic regions at 22 Å distance and vinblastine at 21 Å distance. The above feature is characteristic for compounds in this class.

*3.3. Class **D** (DNA antimetabolites, compounds **31–36**)*

Apart from a few exceptions, most drugs in class **D** (**31, 33, 34, 35, 36, 42, 43** and **46**) are characterised by negative PC2 values and low DRY probe values for variables higher than 20, indicating the absence of distant hydrophobic groups (often only one of them is present). Low energy values for probe N1 for variables higher than 31 (i.e. close H acceptor groups) and higher values for probe O for variables up to 35 (i.e. H donor not very distant from each other) are consistent with few functional groups close to each other. Distant from the **D** drug cluster we find **40** and **45** with much lower N1 and O probe values, **37** and **38** with no dry and **39** with extended dry mainly due to an aliphatic long chain (with PC2 similar to that of **40**, but clearly discriminated by PC1 values).

The peculiarities evidenced by the MIF approach for **37, 38, 40** and **45** can be supported and rationalized on the basis of their pharmacological mechanism of action.

DNA antimetabolites which do not cluster with other drugs in the **D** class do not share the same molecular targets: compound **40** (pyrazoloimidazolo) and compound **38** (hydroxyurea) are considered ribonucleotide reductase inhibitors [21], compound **45** (thiopurine) is mainly an inhibitor of de novo purine synthesis, and **37** is a derivative of aphidicolin, a fungal derived tetracyclic diterpene antibiotic known to be an inhibitor of DNA polymerases alpha and delta [22].

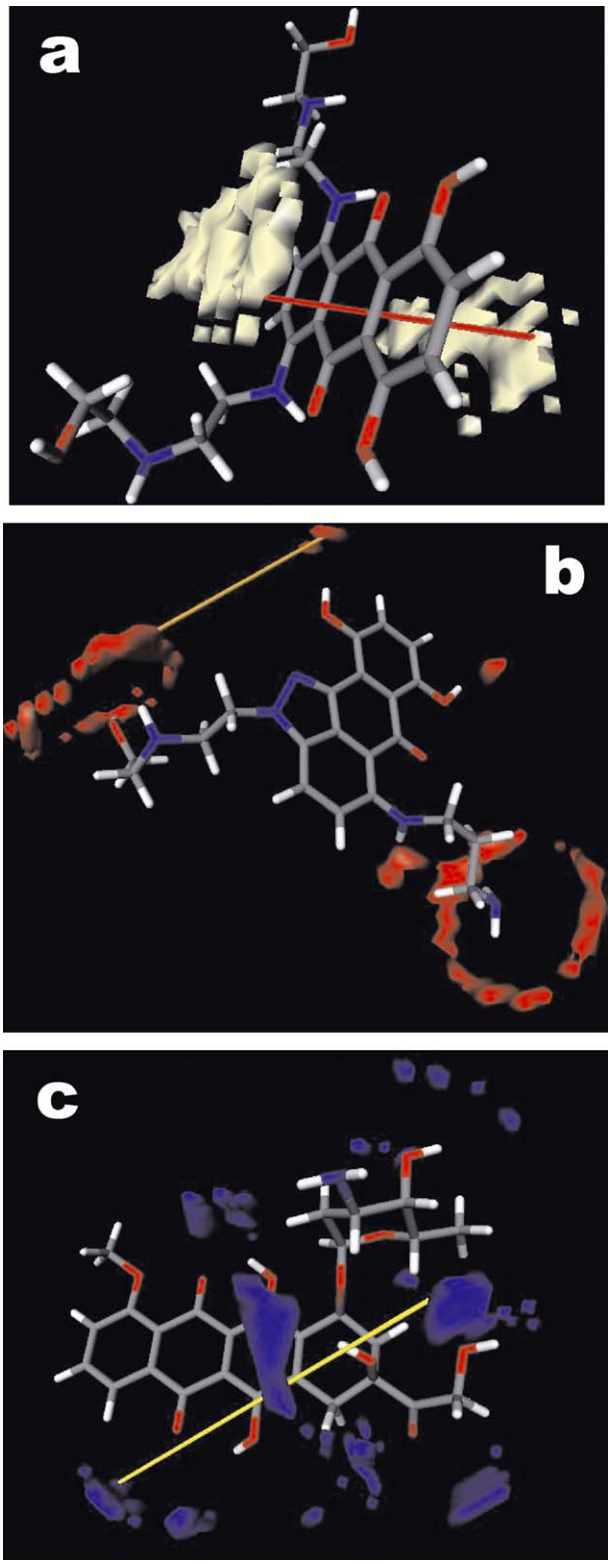


Fig. 2. Isocontour plots of GRID MIF computed around topoisomerase II inhibitors: (a) with a Dry probe (compound **79**); (b) with a hydrogen bond donor probe, O (compound **82**) and (c) with a hydrogen bond acceptor probe, N1 (compound **72**).

The poor solubility of aphidicolin led to the development of aphidicolin glycinate **37**, a water soluble ester currently in clinical trials.

*3.4. Classes **A** and **R** (alkylating agents, RNA/DNA antimetabolites)*

Drugs classified as alkylating agents (class **A**) and RNA/DNA antimetabolites (class **R**) do not exhibit significant clustering in Fig. 3.

Eterogenous groups, such as DNA/RNA antimetabolites, show a large dispersion in structural parameters. This group includes inhibitors of cytosine DNA methyltransferase (**67**: 5,6-dihydro-5′-azacytidine; **58**: 5-azacytidine), inhibitors of dihydroorotate dehydrogenase, an enzyme of the pyrimidine biosynthesis (**69**: brequinar), inhibitors of RNA synthesis and of thymidilate synthase (**65**: fluoruracil), and various antifolic agents. Such a large variety of molecular mechanisms and targets precludes a clear identification of common structural elements.

Previous multivariate classification based on the in vitro antitumor activities (expressed as $\log GI_{50}$) pointed out that most alkylating agents belong also to the class of RNA/DNA antimetabolites [2]. The above finding can be expected, as that of alkylating agents is the only antitumor mechanism of action based merely on the drug chemical reactivity rather than on the behaviour in complex biochemical pathways. Therefore, similarities between alkylating agents and RNA/DNA antimetabolites, whose activities are mainly mediated by a block in DNA based processes such as transcription and replication, are not surprising. It is again noteworthy that the above findings, first pointed out by a pharmacologically-based classification [2], are now confirmed by the present approach based only on the chemical structure.

The MIF approach, considering interactions with probes, is able to point out similarities between drugs which, in spite of apparent dramatic differences in chemical structure (e.g. **66** and **67**), exhibit the same pharmacological behaviour (e.g. RNA/DNA antimetabolites).

The presence of a specific RNA/DNA mechanism is difficult to be identified and might be evidenced by more specific "probes" such as gene transcripts. Unfortunately, gene products suggested by a bioinformatic approach [3] to be selective for the **R** class (GC 11121, GC17689 and GC18564), are still unknown and therefore deserve high priority in further studies.

Strict structural analogies between **A** and **R** drugs can be observed in Fig. 3 for spirohydantoin mustard (**6**), uracil nitrogen mustard (**16**), CCNU (**26**), methyl CCNU (**28**), dichlorallyl lawsone (**59**) and ftorafur (pro-drug) (**62**) all characterised by CH_2Cl.

4. Conclusions

In relation to the aim of the present work, systematic structural patterns for antitumor drugs were evidenced by means of GRIND descriptors. It is noteworthy that above structure-based grouping corresponds to an independent pharmacological classification based on their antitumor

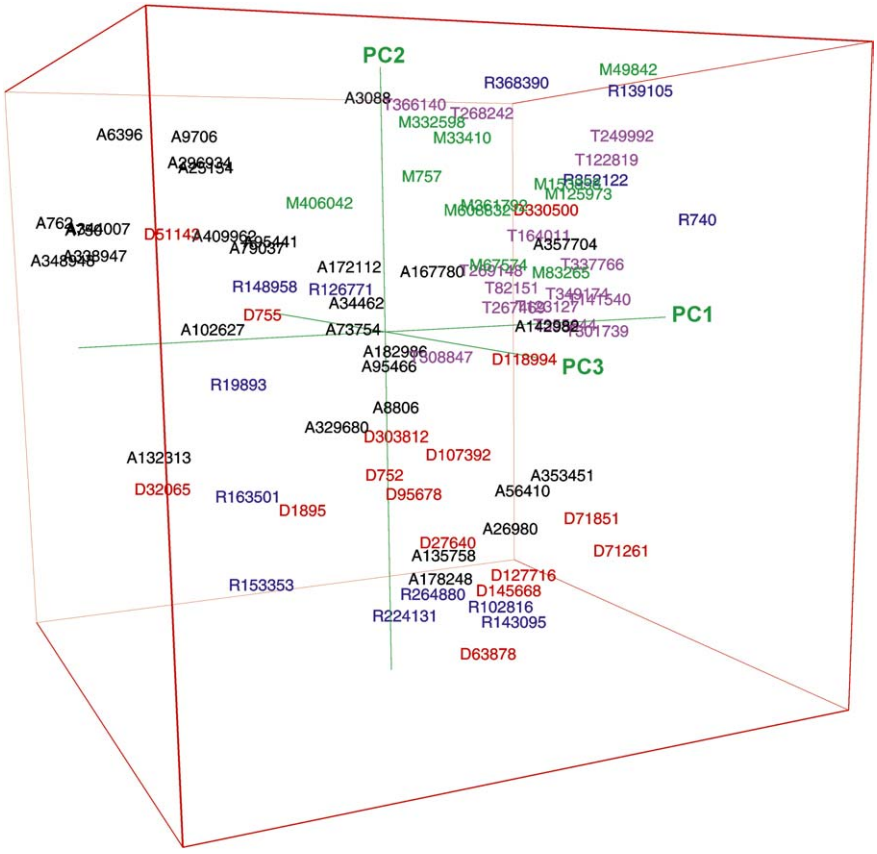


Fig. 3. Three components PCA scores plot for compounds **1–85**. Alkylating agents (coloured black, compounds **1–30**), DNA antimetabolites (coloured red, compounds **31–46**), antimitotic agents (coloured green, compounds **47–57**), DNA/RNA antimetabolites (coloured blue, compounds **58–70**), topoisomerase inhibitors II (coloured purple, compounds **71–85**).

mechanism of action for topoisomerase II inhibitors **T**, antimitotic agents **M** and DNA antimetabolites **D**.

The peculiarities of "outliers" evidenced by MIF, e.g. for **80** (amonafide) and **84** (pyrazoloacridine) in topoisomerase II inhibitors and for **37, 38, 40** and **45** in DNA antimetabolites are supported by literature evidences and can be rationalized on the basis of their peculiar pharmacological mechanisms of action.

Drugs classified as alkylating agents (class **A**) and RNA/DNA antimetabolites (class **R**), interacting with a large variety of targets by different molecular mechanisms, did not exhibit significant clustering in the MIF approach. The above finding, already pointed out by a pharmacologically-based classification [2], is confirmed here by the analysis of GRIND descriptors based only on the chemical structure. Moreover MIFS, considering interactions with selected probes, were able to point out similarities between drugs which, in spite of apparent dramatic differences in chemical structure, exhibit the same pharmacological behaviour.

As demonstrated by this paper, the use of GRIND descriptors, requiring much less time as compared to standard methodologies, is not limited to 3D-QSAR, and the very fact that they do not require a molecular superimposition allows them to be applied in a number of different fields such as in 3D searching, pharmacophore identification, and structure-metabolism relations.

Acknowledgements

We wish to thank the University of Catania and MIUR (Rome) for partial financial support of this work.

References

[1] Database resources currently available on the World Wide Web: http://www.dtp.nci.nih.gov/.
[2] G. Musumarra, D.F. Condorelli, A.S. Costa, M. Fichera, J. Comp. Aided Mol. Design 15 (2001) 219–234.
[3] G. Musumarra, D.F. Condorelli, S. Scire, A.S. Costa, Biochem. Pharmacol. 62 (2001) 547–553.
[4] G. Musumarra, V. Barresi, D.F. Condorelli, S. Scire, J. Biol. Chem. 384 (2003) 321–327.
[5] C. Hansch, Acc. Chem. Res. 2 (1969) 232–239.
[6] P.J. Goodford, J. Med. Chem. 28 (1985) 849–857.
[7] R.D. Cramer III, D.E. Patterson, J. Bunce, J. Am. Chem. Soc. 110 (1988) 5959–5967.
[8] M. Pastor, G. Cruciani, I. McLay, S. Pickett, S. Clementi, J. Med. Chem. 43 (2000) 3233–3243.
[9] GRID v. 21 Molecular Disovery Ltd. (http://www.moldiscovery.com).
[10] D.N.A. Boobbyer, P.J. Goodford, P.M. Mcwhinnie, R.C. Wade, J. Med. Chem. 32 (1989) 1083–1094.
[11] R.C. Wade, K.J. Clerk, P.J. Goodford, J. Med. Chem. 36 (1993) 140–147.
[12] S.P. Sim, B. Gatto, C. Yu, A.A. Liu, T.K. Li, D.S. Pilch, E.J. LaVoie, L.F. Liu, Biochemistry 36 (1997) 13285–13291.

[13] H. Wang, Y. Mao, N. Zhou, T. Hu, T.S. Hsieh, L.F. Liu, J. Biol. Chem. 19 (2001) 15990–15995.

[14] J.L. Grem, N. Harold, B. Keith, A.P. Chen, V. Kao, C.H. Takimoto, J.M. Hamilton, J. Pang, M. Pace, G.B. Jasser, M.G. Quinn, B.P. Monahan, Clin. Cancer. Res. 7 (2002) 2149–2156.

[15] J.S. Sebolt, S.V. Scavone, C.D. Pinter, K.L. Hamelehle, D.D. Von Hoff, R.C. Jackson, Cancer Res. 47 (1987) 4299–4304.

[16] J.S. Sebolt, M. Havlick, K. Hamelehle, J. Nelson, R. Jackson, Cancer Chemother. Pharmacol. 24 (1989) 219–224.

[17] R.C. Jackson, J.S. Sebolt, J.L. Shillis, W.R. Leopold, Cancer Investig. 8 (1990) 39–47.

[18] P. LoRusso, A.J. Wozniak, L. Polin, D. Capps, W.R. Leopold, L.M. Werbel, L. Biernat, M.E. Dan, T.H. Corbett, Cancer Res. 50 (1990) 4900–4905.

[19] J.L. Grem, P.M. Politi, S.L. Berg, N.M. Benchekroun, M. Patel, F.M. Balis, B.K. Sinha, W. Dahut, C.J. Allegra, Biochem. Pharmacol. 51 (1996) 1649–1659.

[20] A.A. Adjei, M. Charron, E.K. Rowinsky, P.A. Svingen, J. Miller, J.M. Reid, J. Sebolt-Leopold, M.M. Ames, S.H. Kaufmann, Clin. Cancer Res. 4 (1998) 683–691.

[21] M. Matsumoto, T. Tihan, J.G. Cory, Cancer Chemother. Pharmacol. 5 (1990) 323–329.

[22] P.J. O'Dwyer, J.D. Moyerm, M. Suffness, S.D. Harrison Jr, R. Cysyk, T.C. Hamilton, J. Plowman, Cancer Res. 3 (1994) 724–729.