

PRÁCTICA II.3
RELACIÓN ESTRUCTURA–ACTIVIDAD II
STRUCTURE–ACTIVITY RELATIONSHIP II

Hugo Cerecetto; Mercedes Gonzalez

Endereço: Dpto. de Química Orgánica. Facultad de Química, UDELAR

hcerecet@bilbo.edu.uy

PROPUESTAS

En estos ejercicios se plantea el trabajo con correlaciones de multivariables.

OBJETIVOS

Obtener y analizar ecuaciones de correlaciones multivariables, entre descriptores fisicoquímicos y actividades biológicas con datos previamente publicados, extrayendo información sobre los requerimientos estructurales para que ciertas familias de compuestos presenten una determinada actividad biológica.

MATERIAL NECESARIO

Programa de análisis estadístico (EXCEL, ORIGIN, etc).

TIEMPO

Cada ejercicio puede llevar 1.5 horas. Depende del tipo de equipamiento. Siendo provechoso realizar una adecuada discusión previa y posterior.

EJERCICIOS

1) Denny W. A. y colaboradores han descrito la síntesis y evaluación biológica de análogos de DACA, reconocida acridina con capacidad intercalante de ADN y actividad inhibitoria de topoisomerasas I y II (*J. Med. Chem.*, 40, 1919-1929, **1997**). Los compuestos son preparados a través de rutas sintéticas clásicas y se evalúan contra tres líneas celulares humanas de leucemia (líneas Jurkat: JL_C, JL_A y JL_D, la primera nativa y las dos restantes mutantes). En la Tabla 1 se resumen las estructuras y los valores de IC₅₀ para la línea JL_C.

En un intento por conocer las exigencias estructurales para una óptima actividad biológica (actividad antitumoral), los autores realizan un estudio de correlación entre descriptores fisicoquímicos y bioactividad.

Para ello se seleccionan los derivados 5-sustituido (**2-11**), los derivados 6-sustituido (**12-18**) y los derivados 7-sustituido junto con DACA (**1, 19-30**) y se estudian las correlaciones existentes entre $\log_{10}(\text{IC}_{50} \text{ JL}_C)$ y descriptores estéricos (MR), de lipofilia (π) y electrónicos (σ_p y σ_m), como los indicados en la Tabla 1. Obteniendo los siguientes resultados:

Compuestos	Correlación con una variable	Incorporación de otra variable	Observaciones
Derivados 5-sustituido	Correlación aceptable entre $\log_{10}(\text{IC}_{50} \text{ JL}_C)$ y MR	Al incorporar σ_p mejora el r de la ecuación, mostrando el término electrónico una menor contribución	Los autores deciden no investigar la existencia de correlación entre la actividad y las variables independientes MR y π juntas
Derivados 6-sustituido	No se encuentran correlaciones significativas		Utilizan σ_m como parámetro electrónico
Derivados 7-sustituido y DACA	Muy buena correlación entre $\log_{10}(\text{IC}_{50} \text{ JL}_C)$ y MR	Al incorporar σ_p no se observa una mejora en el r de la ecuación	Los autores descartan del análisis a los derivados 23 y 30 , por no mostrar un adecuado ajuste

Nº	R	IC ₅₀ JL _C (nM)
1, DACA	H	580
2	5-Me	46
3	5-Et	137
4	5-iPr	2500
5	5-Ph	1665
6	5-OMe	111
7	5-F	103
8	5-Cl	4.2
9	5-Br	26
10	5-NMe ₂	5710
11	5-CF ₃	75
12	6-Me	625
13	6-OMe	580
14	6-F	455
15	6-Cl	397
16	6-Br	446
17	6-NMe ₂	212
18	6-CF ₃	1350
19	7-Me	1350
20	7-Et	1840
21	7-iPr	7000
22	7-tBu	4390
23	7-Ph	240
24	7-OMe	1590
25	7-OH	508
26	7-F	424
27	7-Cl	753
28	7-Br	756
29	7-NMe ₂	2480
30	7-CF ₃	6390

Descriptores fisicoquímicos				
R	MR	π	σ_p	σ_m
H	1.0	0	0	0
Me	5.7	0.56	-0.17	-0.07
Et	10.3	1.02	-0.15	-0.07
i-Pr	15.0	1.53	-0.15	-0.07
Ph	25.4	1.96	-0.01	0.06
Ome	7.9	-0.02	-0.27	0.12
F	0.9	0.14	0.06	0.34
Cl	6.0	0.71	0.23	0.37
Br	8.9	0.86	0.23	0.39
NMe ₂	15.5	0.18	-0.83	-0.15
CF ₃	5.0	0.88	0.54	0.43
TBu	19.6	1.98	-0.20	-0.10
OH	2.8	-0.67	-0.37	0.12

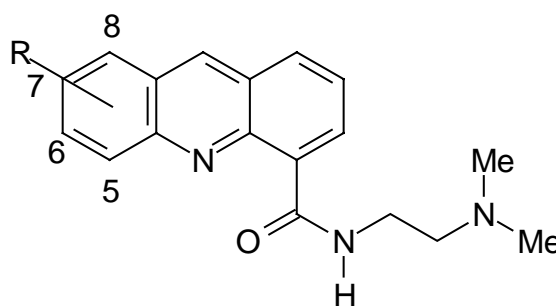


Tabla 1

- 1.1. ¿Porque los autores utilizan los valores de $\log_{10}(\text{IC}_{50} \text{ JL}_C)$ para las correlaciones, en lugar de usar directamente los valores de $\text{IC}_{50} \text{ JL}_C$?
- 1.2. Indique la ecuación entre la variable dependiente y MR para los derivados 5-sustituido. ¿Cuál es la ecuación cuando se incorpora σ_p en el análisis? ¿Porqué la contribución del factor electrónico es menor? ¿Porqué los autores no investigan conjuntamente MR y π ?
- 1.3. En el análisis de los derivados 6-sustituido, ¿cuál es el motivo por el que utilizan σ_m como descriptor electrónico (en lugar de usar σ_p como en los restantes análisis)?

1.4. Indique la ecuación entre la variable dependiente y MR para los derivados 7-sustituido. ¿Cuál es la ecuación cuando se incorpora σ_p en el análisis?

RESOLUCIÓN

1.1. Al tratarse de valores de $IC_{50} JLC$ dispares (por ejemplo para los derivados 5-sustituido varían desde 5710 en el caso del derivado **10** a 4.2 en el derivado **8**) con el uso de \log_{10} se consigue un conjunto de datos biológicos homogéneos en la variable dependiente (3.76 y 0.62 para los derivados **10** y **8**, respectivamente).

1.2. La contribución del efecto electrónico es menor ya que el coeficiente de este descriptor en la ecuación posee un error estándar mayor que el correspondiente coeficiente de MR.

MR y π no se incluyen conjuntamente en un análisis QSAR porque existe una alta colinealidad entre ambos descriptores ($r=0.68$), mientras que MR y σ_p resultan ortogonales ($r=0.38$).

1.3. El motivo es el interés en conocer como influye el efecto electrónico de los sustituyentes sobre el nitrógeno heterocíclico de la acridina y de esta forma saber la influencia en propiedades tales como pKa.

Mientras que los sustituyentes en los derivados 5 y 7-sustituido están en posiciones *orto* y *para* respecto a este heteroátomo, en los derivados 6-sustituido están en posición *meta*.

De ahí que en ciertos casos (los primeros) se use el descriptor electrónico σ_p y en el otro (el segundo) el descriptor σ_m .

2) Tratando de confirmar que en el mecanismo de citotoxicidad de derivados de 1,2,5-oxadiazol *N*-óxido participa la posición “bencílica” de dichos compuestos (Boiani y col., *Eur. J. Med. Chem.*, 36, 771-782, **2001**, por ejemplo como se muestra en la figura 1), se plantea un estudio teórico de correlaciones entre descriptores fisicoquímicos de los sustituyentes del sistema heterocíclico y la actividad.

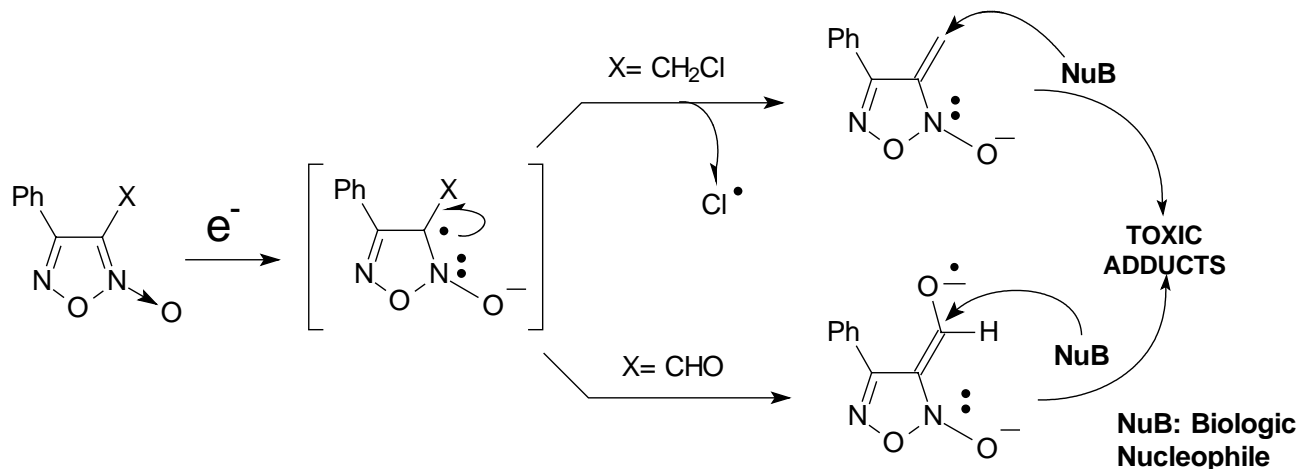


Figura 1

Boiani y col. describen la síntesis y evaluación biológica de diferentes derivados. Los compuestos son preparados a través de rutas sintéticas clásicas y se evalúan por su citotoxicidad selectiva hacia condiciones de hipoxia frente a la línea celular V79 (fibroblastos de hámster chino). En la Tabla 1 se resumen algunas de las estructuras y los valores de actividad. Para los estudios de relación estructura-actividad las actividades biológicas se categorizan como se indica en la tabla.

Para realizar el estudio de correlaciones se adjuntan, en la Tabla 2, los descriptores electrónicos de Swain-Lupton (constante de campo, F , y constante de resonancia, R), de Hammett (σ_p), los descriptores hidrofóbicos (π), los descriptores estéricos (MR) y la variable indicadora I_{NO} , que describe la presencia (1) o ausencia (0) de la función N -óxido en la molécula.

	Nº	R	n	SF_{hipox}^a	SF_{foxia}^b	$SF_{hipox_cat}^c$
	1	CHO	1	0	0	0.0
	2	CN	1	0	0	0.0
	3	CH ₂ Cl	1	2	0	0.5
	4	CH=NNHC(S)NH ₂	1	54	53	1.0
	5	CH ₂ OH	1	63	55	1.0
	6	CH ₂ Cl	0	68	68	1.5
	7	CH=NOH	1	70	35	1.5

8	CH=CHCO ₂ Et	1	96	72	2.0
9	CH ₂ OH	0	96	96	2.0
10	CHO	0	98	94	2.0

Tabla 1. (a) Fracción de supervivencia en hipoxia a 20 μ M. **(b)** Fracción de supervivencia en oxia a 20 μ M. **(c)** Categorizado según: 0.0= 0-1 % de *SFhipox*, 0.5= 2-34 % de *SFhipox*, 1.0= 35-64 % de *SFhipox*, 1.5= 65-95 % de *SFhipox*, 2.0= 96-100 % de *SFhipox*.

Nº	R	F	R	σ_p	π	MR	I _{NO}
1	CHO	0.31	0.13	0.42	-0.65	6.88	1
2	CN	0.51	0.19	0.66	-0.57	6.33	1
3	CH ₂ Cl	0.10	0.03	0.12	0.17	10.49	1
4	CH=NNHC(S)NH ₂	0.46	-0.02	0.40	-0.27	29.92	1
5	CH ₂ OH	0.00	0.00	0.00	-1.03	7.19	1
6	CH ₂ Cl	0.10	0.03	0.12	0.17	10.49	0
7	CH=NOH	0.25	-0.13	0.10	-0.38	10.28	1
8	CH=CHCO ₂ Et	0.24	-0.19	0.03	0.86	27.21	1
9	CH ₂ OH	0.00	0.00	0.00	-1.03	7.19	0
10	CHO	0.31	0.13	0.42	-0.65	6.88	0

Tabla 2

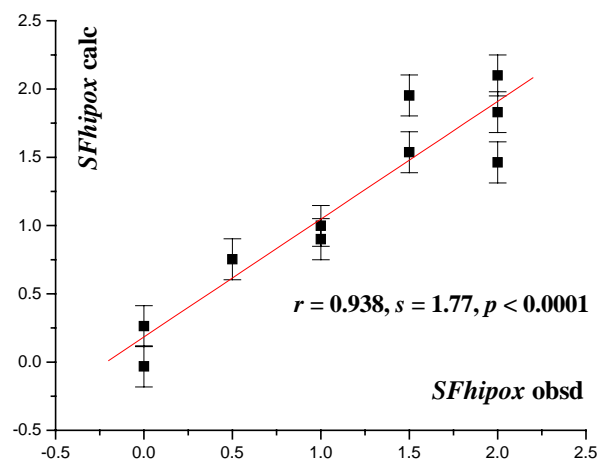
Analice, la relación entre la citotoxicidad en hipoxia (categorizada como se muestra en la tabla 1) y las propiedades fisicoquímicas de los sustituyentes **R** para los derivados mostrados. ¿El mecanismo de citotoxicidad en hipoxia de los derivados de 1,2,5-oxadiazol *N*-óxido puede ser descrito como se muestra en la Figura 1?

RESOLUCIÓN

No existen correlaciones (mono o multivariable) estadísticamente válidas entre la actividad biológica estudiada y los descriptores **F**, σ_p , π , ó **MR**.

Sin embargo, es posible obtener una correlación válida entre la actividad y las variables independientes **R** y I_{NO} conjuntamente.

Ecuación 1 ***Sfhipox_cat* = 2.1 (\pm 0.2) – 1.2 (\pm 0.2) I_{NO} – 4.9 (\pm 0.9) **R****
 $n = 10, r = 0.938, F = 25.44, p < 0.0006$



Nº	SFhipox	SFhipox calc (± 0.1) ^a
1	0.0	0.2
2	0.0	-0.1
3	0.5	0.7
4	1.0	1.0
5	1.0	0.9
6	1.5	1.9
7	1.5	1.6
8	2.0	1.9
9	2.0	2.1
10	2.0	1.4

(a) Calculada usando Ec. 1.

Según la Ec. 1, el mecanismo de acción puede ser descrito como se muestra en la figura.

Ya que de acuerdo a la ecuación es necesaria la presencia de la función N-óxido para que disminuya la fracción de supervivencia en hipoxia (término en variable INO), lo que estaría en concordancia con la etapa 1 de monoelectronación mostrada en la figura. Y, por otro lado, la ecuación muestra que a mayor valor de R, constante de resonancia, mayor citotoxicidad. Esto estaría de acuerdo con la etapa de deslocalización del radical hacia la posición "bencílica" tal como se indica en la figura.

Hugo Cerecetto; Mercedes Gonzalez

hcerecet@bilbo.edu.uy

El trabajo en los laboratorios de Medicinal Chemistry debe mantener altos estándares de precaución y buen uso.

El manejo de instrumental eléctrico, la utilización de calor, el material de cristal y los disolventes no presentan un especial problema, si se siguen las instrucciones del supervisor.

Este documento ha sido supervisado por el Prof. Hugo Cerecetto and Mercedes González (hcerecet@fq.edu.uy; megonzal@fq.edu.uy) quien informa que no existen problemas específicos de seguridad en la realización de este ejercicio, incluyendo toxicidad, inflamabilidad y explosión, ni cualquier otro destacable, dentro de lo usual en un laboratorio de Medicinal Chemistry.

Se agradecerá comunicar al Editor cualquier posible incidencia.

EXERCISE II.3

STRUCTURE–ACTIVITY RELATIONSHIP II

Hugo Cerecetto and Mercedes Gonzalez

Endereço: Dpto. de Química Orgánica. Facultad de Química, UDELAR

E-mail: hcerecet@fq.edu.uy & megonzal@fq.edu.uy

PROPOSAL

In these exercises, the work is planned with multivariable correlations.

OBJECTIVES

To obtain and analyze equation with multivariable correlation, between physicochemical descriptors and biological activities with data already published, extracting information about the structural requirements for certain families of compounds present certain biological activity.

MATERIAL

Statistic analysis program

TIME

Each exercise could take 1.5 h, depending on the type of equipment. A final discussion of the work could be very profitable.

EXERCISES

1) W. A. Denny and coworkers described the synthesis and biological evaluation of DACA analogs, known acridine with capacity to intercalate to DNA and with topoisomerases I and II activity (*J. Med. Chem.*, 40, 1919–1929, **1997**). The compounds were prepared through classical synthetic routes and are evaluated against three leukemia human cell lines (lines Jurkat: JL_C, JL_A, and JL_D, the first native and the other two mutants). Table 1 lists the structures and the IC₅₀ values for the JL_C line.

Intending to know the structural requirements for an optimal biological activity (antitumor activity), the authors made a correlation study between physicochemical descriptors and bioactivity.

For that, the 5-substituted derivatives (**2–11**), the 6-substituted derivatives (**12–18**), and the 7-substituted derivatives together with DACA (**1, 19–30**) were selected and the correlation between log₁₀(IC₅₀ JL_C) and steric (MR), lipophilic

(π) and electronic (σ_p and σ_m) descriptors were studied (Table 1). The following results were obtained.

Compounds	Correlation with one variable	Incorporation of other variable	Observations
5-substituted derivatives	Acceptable correlation between $\log_{10}(\text{IC}_{50} \text{JL}_C)$ and MR	Incorporating σ_p , the r increase, showing the electronic term a lower contribution	The authors decided to not investigate the existence of correlation between the activity and the independent variables MR and π together
6-substituted derivatives	No significant correlation was found		σ_m was studied as electronic descriptor
7-substituted derivatives and DACA	A very good correlation between $\log_{10}(\text{IC}_{50} \text{JL}_C)$ and MR	Incorporating σ_p , no increase in r was observed	The authors discard the analysis of derivatives 23 and 30 , not showing an adequate adjust.

Nº	R	IC ₅₀ JL _C (nM)
1, DACA	H	580
2	5-Me	46
3	5-Et	137
4	5- <i>i</i> Pr	2500
5	5-Ph	1665
6	5-OMe	111
7	5-F	103
8	5-Cl	4.2
9	5-Br	26
10	5-NMe ₂	5710
11	5-CF ₃	75
12	6-Me	625
13	6-OMe	580
14	6-F	455
15	6-Cl	397
16	6-Br	446
17	6-NMe ₂	212
18	6-CF ₃	1350
19	7-Me	1350
20	7-Et	1840
21	7- <i>i</i> Pr	7000
22	7- <i>t</i> Bu	4390
23	7-Ph	240
24	7-OMe	1590
25	7-OH	508
26	7-F	424
27	7-Cl	753
28	7-Br	756
29	7-NMe ₂	2480
30	7-CF ₃	6390

Physicochemical descriptors				
R	MR	π	σ_p	σ_m
H	1.0	0	0	0
Me	5.7	0.56	-0.17	-0.07
Et	10.3	1.02	-0.15	-0.07
<i>i</i> -Pr	15.0	1.53	-0.15	-0.07
Ph	25.4	1.96	-0.01	0.06
Ome	7.9	-0.02	-0.27	0.12
F	0.9	0.14	0.06	0.34
Cl	6.0	0.71	0.23	0.37
Br	8.9	0.86	0.23	0.39
NMe ₂	15.5	0.18	-0.83	-0.15
CF ₃	5.0	0.88	0.54	0.43
TBu	19.6	1.98	-0.20	-0.10
OH	2.8	-0.67	-0.37	0.12

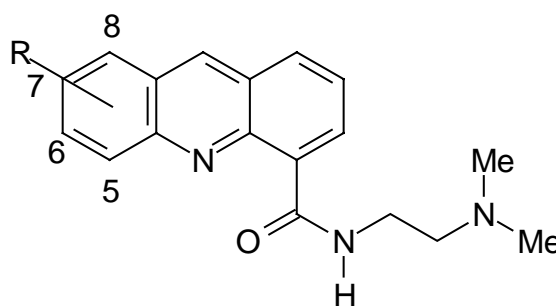


Table 1

- 1.5. Why do the authors use the values of $\log_{10}(\text{IC}_{50} \text{ JL}_C)$ for the correlations, instead of using the $\text{IC}_{50} \text{ JL}_C$ values directly?
- 1.6. Show the equation between the dependent variable and MR for the 5-substituted derivatives. Which is the equation when σ_p is incorporated in the analysis? Why is the electronic factor lower? Why do the authors not investigate MR and π together?
- 1.7. Analyzing the 6-substituted derivatives, what is the reason for using σ_m as electronic descriptor (instead of using σ_p like in the other analysis)?

- 1.8. Show the equation obtained between the dependent variable and MR for the 7-substituted derivatives. Which is the equation when σ_p is incorporated in the analysis?

RESOLUTION

1.1. Because the $IC_{50} JL_C$ values are very unequal (i.e., 5-substituted derivatives vary from 5710 for derivative **10** to 4.2 for derivative **8**) with the \log_{10} homogenous biological data are obtained (3.76 and 0.62 for derivatives **10** and **8**, respectively).

1.2. The contribution of the electronic effect is lower because the descriptor coefficient in this equation has a standard error greater than for the corresponding coefficient of MR.

MR and π are not included together in the QSAR analysis because a high colinearity exists between both descriptors ($r = 0.68$), while MR and σ_p result in orthogonales ($r = 0.38$).

1.3. It is interesting to know how the electronic effect of the substituents affects the heterocyclic nitrogen of the acridine and thus investigate the effect in properties such as pK_a .

While the substituents in the 5- and 7-substituted derivatives are in the *ortho* and *para* position with respect to the heteroatom, in the 6-substituted derivatives they are in the *meta* position.

So in certain cases (the first one) the electronic descriptor σ_p is used and in the other (the second) the σ_m descriptor is used.

2) Trying to confirm that the cytotoxic mechanism of 1,2,5-oxadiazole *N*-oxides involves the “benzylic” position of such compounds (Boiani et al., *Eur. J. Med. Chem.*, 36, 771–782, **2001**, i.e., as shown in Fig. 1), we plan a theoretical study of correlations between the physicochemical descriptors of the substituents of the heterocyclic system and the activity.

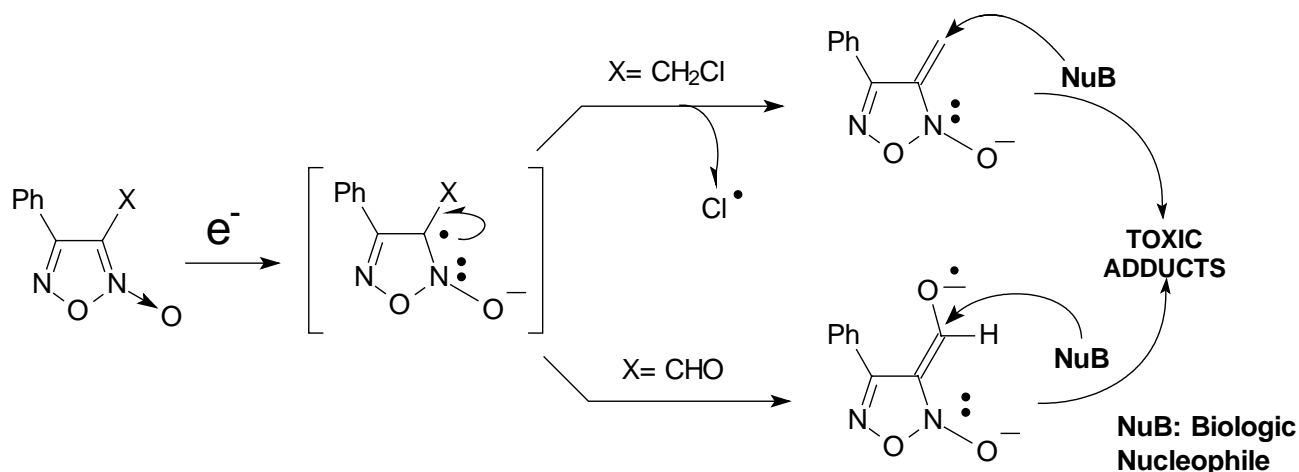


Fig. 1

Boiani and colleagues describe the synthesis and biological evaluation of different derivatives. The compounds were prepared through classical synthetic routes, and their selective cytotoxic activity in hypoxia was determined against the cell line V79 (fibroblast of Chinese hamster). In Table 1, some structures and activity values are gathered. For the structure–activity relationship studies, the biological activities are categorized as indicated in the table.

For the development of the correlation studies, the electronic descriptors of Swain-Lupton (field constant, F , and resonance constant, R), of Hammett (σ_p), the hydrophobic descriptor (π), the steric descriptor (MR), and the indicator variable I_{NO} , that describe the presence (1) or absence (0) of the N -oxide in the molecule are presented in Table 2.

	N^o	R	n	SF_{hipox}^a	SF_{oxia}^b	$SF_{hipox_cat}^c$
	1	CHO	1	0	0	0.0
	2	CN	1	0	0	0.0
	3	CH ₂ Cl	1	2	0	0.5
	4	CH=NNHC(S)NH ₂	1	54	53	1.0
	5	CH ₂ OH	1	63	55	1.0
	6	CH ₂ Cl	0	68	68	1.5
	7	CH=NOH	1	70	35	1.5
	8	CH=CHCO ₂ Et	1	96	72	2.0
	9	CH ₂ OH	0	96	96	2.0
	10	CHO	0	98	94	2.0

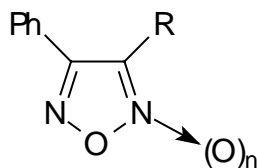


Table 1 (a) Survival fraction in hypoxia at 20 μM . **(b)** Survival fraction in oxia at 20 μM . **(c)** Category as: 0.0 = 0–1 % of *SFhipox*, 0.5 = 2–34 % of *SFhipox*, 1.0 = 35–64 % of *SFhipox*, 1.5 = 65–95 % of *SFhipox*, 2.0 = 96–100 % of *SFhipox*.

N ^o	R	F	R	σ_p	π	MR	I _{NO}
1	CHO	0.31	0.13	0.42	-0.65	6.88	1
2	CN	0.51	0.19	0.66	-0.57	6.33	1
3	CH ₂ Cl	0.10	0.03	0.12	0.17	10.49	1
4	CH=NNHC(S)NH ₂	0.46	-0.02	0.40	-0.27	29.92	1
5	CH ₂ OH	0.00	0.00	0.00	-1.03	7.19	1
6	CH ₂ Cl	0.10	0.03	0.12	0.17	10.49	0
7	CH=NOH	0.25	-0.13	0.10	-0.38	10.28	1
8	CH=CHCO ₂ Et	0.24	-0.19	0.03	0.86	27.21	1
9	CH ₂ OH	0.00	0.00	0.00	-1.03	7.19	0
10	CHO	0.31	0.13	0.42	-0.65	6.88	0

Table 2

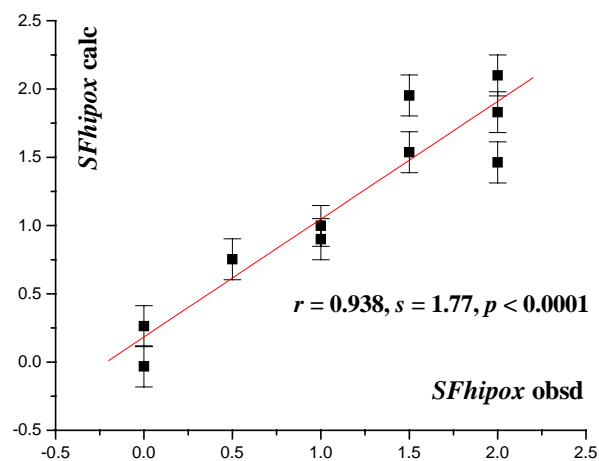
Analyze the relationship between the cytotoxicity in hypoxia and the physicochemical properties of the substituent R for the shown derivatives. Could the cytotoxicity in hypoxia of the 1,2,5-oxadiazole *N*-oxide derivatives mechanism be described as shown in Fig. 1?

RESOLUTION

No statistically significant relationships (mono or multivariable) between the biological activity studied and the descriptors *F*, σ_p , π , or MR were obtained.

But, it was possible to obtain a valid correlation between the activity and the independent variables R and I_{NO} together.

Equation 1 ***Sfhipox_cat* = 2.1 (± 0.2) – 1.2 (± 0.2) I_{NO} – 4.9 (± 0.9) R**
 $n = 10, r = 0.938, F = 25.44, p < 0.0006$



Nº	<i>SFhipox</i>	<i>SFhipox calc</i> (±0.1) ^a
1	0.0	0.2
2	0.0	-0.1
3	0.5	0.7
4	1.0	1.0
5	1.0	0.9
6	1.5	1.9
7	1.5	1.6
8	2.0	1.9
9	2.0	2.1
10	2.0	1.4

(a) Calculated using eq. 1.

According to eq. 1, the mechanism of action can be described as shown in the figure.

Because according to the equation the presence of the *N*-oxide group is necessary for lowering the survival fraction in hypoxia (term in variable I_{NO}), this is in agreement with the first step of mono electron addition shown in Fig. I. On the other hand, the equation shows that with a bigger value of *R*, resonance constant, a greater cytotoxicity is observed. This is in agreement with the delocalization of the radical through the benzylic position as shown in the figure.

Hugo Cerecetto
hcerecet@fq.edu.uy

High standards in safety measures should be maintained in all work carried out in Medicinal Chemistry Laboratories. The handling of electrical instruments, heating elements, glass materials, dissolvents and other inflammable materials does not present a problem if the supervisor's instructions are carefully followed.

This document has been supervised by Prof. Hugo Cerecetto and Mercedes González (hcerecet@fq.edu.uy; megonzal@fq.edu.uy) who has informed that no special risk (regarding toxicity, inflammability, explosions), outside of the standard risks pertaining to a Medicinal Chemistry laboratory exist when performing this exercise.

If your exercise involves any "special" risks, please inform the editor.