





Article

# A General Use QSAR-ARX Model to Predict the Corrosion Inhibition Efficiency of Drugs in Terms of Quantum Mechanical Descriptors and Experimental Comparison for Lidocaine

Carlos Beltran-Perez <sup>1</sup>, Andrés A. A. Serrano <sup>1</sup>, Gilberto Solís-Rosas <sup>1</sup>, Anatolio Martínez-Jiménez <sup>2</sup>, Ricardo Orozco-Cruz <sup>3</sup> , Araceli Espinoza-Vázquez <sup>3,\*</sup> and Alan Miralrio <sup>1,\*</sup> 

<sup>1</sup> Tecnológico de Monterrey, Escuela de Ingeniería y Ciencias, Ave. Eugenio Garza Sada 2501, Monterrey 64849, Mexico; carlos.beltran@tec.mx (C.B.-P.); andres.serrano@tec.mx (A.A.A.S.); gilsolisr@gmail.com (G.S.-R.)

<sup>2</sup> Departamento de Ciencias Básicas, División de CBI (Ciencias Básicas e Ingeniería), Universidad Autónoma Metropolitana, Unidad Azcapotzalco, Área de Física Atómica Molecular Aplicada, San Pablo 180, Ciudad de México 02200, Mexico; amartinez@azc.uam.mx

<sup>3</sup> Unidad Anticorrosión, Instituto de Ingeniería, Universidad Veracruzana, Boca del Río 94292, Mexico; rorozco@uv.mx

\* Correspondence: araespinoza@uv.mx (A.E.-V.); miralrio@tec.mx (A.M.)

**Abstract:** A study of 250 commercial drugs to act as corrosion inhibitors on steel has been developed by applying the quantitative structure-activity relationship (QSAR) paradigm. Hard-soft acid-base (HSAB) descriptors were used to establish a mathematical model to predict the corrosion inhibition efficiency (IE%) of several commercial drugs on steel surfaces. These descriptors were calculated through third-order density-functional tight binding (DFTB) methods. The mathematical modeling was carried out through autoregressive with exogenous inputs (ARX) framework and tested by fivefold cross-validation. Another set of drugs was used as an external validation, obtaining SD, RMSE, and MSE, obtaining 6.76%, 3.89%, 7.03%, and 49.47%, respectively. With a predicted value of IE% = 87.51%, lidocaine was selected to perform a final comparison with experimental results. By the first time, this drug obtained a maximum IE%, determined experimentally by electrochemical impedance spectroscopy measurements at 100 ppm concentration, of about 92.5%, which stands within limits of 1 SD from the predicted ARX model value. From the qualitative perspective, several potential trends have emerged from the estimated values. Among them, macrolides, alkaloids from *Rauwolfia* species, cephalosporin, and rifamycin antibiotics are expected to exhibit high IE% on steel surfaces. Additionally, IE% increases as the energy of HOMO decreases. The highest efficiency is obtained in case of the molecules with the highest  $\omega$  and  $\Delta N$  values. The most efficient drugs are found with  $pK_a$  ranging from 1.70 to 9.46. The drugs recurrently exhibit aromatic rings, carbonyl, and hydroxyl groups with the highest IE% values.

**Keywords:** corrosion inhibition; lidocaine; QSAR; tight binding; ARX model; FROLS algorithm; lidocaine



**Citation:** Beltran-Perez, C.; Serrano, A.A.A.; Solís-Rosas, G.; Martínez-Jiménez, A.; Orozco-Cruz, R.; Espinoza-Vázquez, A.; Miralrio, A. A General Use QSAR-ARX Model to Predict the Corrosion Inhibition Efficiency of Drugs in Terms of Quantum Mechanical Descriptors and Experimental Comparison for Lidocaine. *Int. J. Mol. Sci.* **2022**, *23*, 5086. <https://doi.org/10.3390/ijms23095086>

Academic Editor: Dongho Kim

Received: 19 March 2022

Accepted: 27 April 2022

Published: 3 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

### 1.1. Corrosion Inhibition and QSAR Fundamentals

Amongst the metals, steel is the most used iron alloy for industrial applications [1], such as oil, food, energy, chemical, and construction industries. Being highly ductile, durable, and resistant, steel is highly appreciated for its mechanical properties. Furthermore, the several different alloys obtained at a considerable low-cost increase the variety of properties exhibited. Unfortunately, corrosion is probably the most common phenomenon that leads to weakening metals; this originates from the electrochemical interaction of metallic surfaces with a corrosive environment. Furthermore, the sulfates, oxides, and other compounds produced from these interactions modify the inherent properties of the metal surface, leading to undesired behaviors [2].

From the above, the most common corrosion inhibition strategies are dedicated to steel [3–5], although copper and aluminum alloys are studied to a lesser extent [6–12]. According to the National Association of Corrosion Engineers (NACE), through its “International Measures of Prevention, Application and Economics of Corrosion Technology” study, the Global cost of the damages produced by corrosion in 2013 was US\$2.5 trillion. This massive amount of resources represented 3.4% of the Global Gross Domestic Product (GDP) in 2013 [13]. Thus, providing novel solutions to reduce the undesired effects of corrosion on metals is a global priority.

The most recurrent strategy to reduce the corrosion on metals is to employ corrosion inhibitors on their surface [3,14–16]. A corrosion inhibitor is a substance that, added in small amounts to the metal surface, reduces the action of the corrosive media on the metal by forming a protective film. These organic species could bond strongly to the metal surface through intermolecular interactions, from the inhibitor molecule to neighboring metal atoms [2]. Electrostatic interactions, London dispersion forces, and even covalent bonding could be exhibited. Mostly, these organic compounds contain nitrogen, oxygen, and sulfur. Moreover, organic molecules rich in  $\pi$ -electrons, associated with either triple, double, or conjugated bonds and aromatic rings, are recurrently found to act as corrosion inhibitors [2–5,17].

Several organic compounds acting as corrosion inhibitors have been evaluated in recent years—for instance, plant extracts [14,17–23] and commercial drugs [3–8,10,24–30]. However, the massive amounts of phytochemicals and drugs that can be tested as corrosion inhibitors need intelligent strategies for their study [31–33]. In pharmacology, the evaluation and prediction of biological activities and properties for massive amounts of potential drugs have been assessed for a long time by the quantitative structure-activity/property relationship (QSAR/QSPR) paradigm [34–37]. QSAR/QSPR proposes predicting a given activity/property of quantifiable descriptors. These values can be extracted from existing databases, experiments, theoretical calculations, or simulations. Consequently, data curation, descriptors selection, and mathematical modeling join the problem proposed by the QSAR/QSPR approach [38].

Besides, the QSAR/QSPR modeling in the corrosion field is scarce. The groundbreaking work of Zhang et al. proposes QSAR linear models correlating parameters, obtained mainly by quantum-chemical calculations, such as polarizability, dipole moment, frontier orbital energies, and others, to predict the corrosion inhibition efficiency (IE%) of 18 inhibitors on steel, obtaining average deviations of about 9.82%. Another relevant and quite elegant QSAR linear model is the one reported by Keshavarz and coworkers [39], given the number of nitrogen atoms, amino groups, and other structural parameters. This QSAR model accounts for root mean squared deviation (RMSD), mean absolute error, and maximum errors of about 6.15%, 4.93%, and 12.0%, respectively.

Moreover, recent reports on QSAR/QSPR studies applied to predict IE% of organic corrosion inhibitors on metal surfaces borrow tools from data science. For instance, Liu and coworkers used support vector machine (SVM) models with 11 top descriptors to characterize 20 benzimidazole derivatives [40]. The root mean square error (RMSE) reported is about 4.45%. On the other hand, Ser et al. reported that the corrosion inhibition efficiency of pyridines and quinolines on iron surfaces had been evaluated utilizing machine learning-based QSPR relationships [41]. The authors obtained the mathematical model by genetic algorithm-artificial neural network methods, leading to linear and nonlinear models, with RMSE of about 16.5% and 8.8%, respectively. In this case, the models considered up to nine variables, mainly obtained from DFT calculations.

### 1.2. QSAR Paradigm and HSAB Descriptors

This work proposes modeling the quantitative structure-activity relationship (QSAR) paradigm using several empirical and theoretical descriptors to predict organic molecule corrosion inhibition efficiency. Since these descriptors are almost any quantitative values

that can be measured by experiments or determined by theoretical calculations, this work proposes a set of empirical and theoretical descriptors.

The QSAR descriptors determine a drug's biological activity, looking for specific behaviors. For instance, as proposed by Hansch and Muir, the  $\log P$  octanol-water partition coefficient was correlated with the biological activity  $1/C$  [35]. The  $\log P$  descriptor is massively used to study potential drugs since it determines a substance's concentrations between a hydrophobic phase and a hydrophilic phase. Other QSAR descriptors commonly used in drug design are the solubility coefficient  $\log S$ ,  $pK_a$ , molecular weight, polar surface area, polarizability, and H-bond acceptor and donor counts [35]. Some common QSAR descriptors are defined within Pearson's HSAB [42] theory. These descriptors are based on the vertical ionization energy ( $I$ ) and the vertical electron affinity ( $A$ ). By Koopmans' theorem, both can be calculated by  $I = -E_{\text{HOMO}}$  and  $A = -E_{\text{LUMO}}$ , where  $E_{\text{HOMO}}$  and  $E_{\text{LUMO}}$  are the energies of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO), respectively. According to the HSAB principle, two species will interact easily if either hard or soft [42]. A molecule's global hardness ( $\eta$ ) can be calculated by  $\eta = (I - A)/2$ . Other helpful properties are the absolute electronegativity  $\chi = (I + A)/2$  and the global electrophilicity  $\omega = \mu^2/2\eta$  [43]. The first one determines the "ability" of a given molecule to attract electrons from the environment, whereas the second is related to the energetic stabilization a species gains by obtaining an additional electron.

Lastly, the fraction of electrons transferred ( $\Delta N$ ) is a valuable property to elucidate the behavior of the organic molecule interacting with the metal surface. This value can be calculated by  $\Delta N = (\chi_{\text{Metal}} - \chi_{\text{Inhibitor}})/[2(\eta_{\text{Metal}} + \eta_{\text{Inhibitor}})]$ , where the electronegativities and harnesses of the metal species and the corrosion inhibitor molecule are used. Lukovits and coworkers reported that the higher  $\Delta N$  value, the higher the corrosion inhibition efficiency [44]. Additionally, isosurfaces of frontier molecular orbitals and electrostatic potential arise to describe the active sites and the electrostatic interactions [45].

This work elucidated the contribution to the corrosion inhibition properties in terms of these HSAB descriptors. In addition, the most common QSAR descriptors listed in the previous subsection were studied for completeness. Since DFT calculations require massive computational resources, third-order density-functional tight binding (DFTB) method with Lennard-Jones (LJ) dispersion-correction was used to obtain the quantum chemical descriptors. Thus, the method, hereafter labeled as DFTB3-LJ, was used as implemented in the DFTB+ 21.2 quantum chemistry package [46,47] (more details in Supplementary Materials).

From all the above, this work aims to obtain a QSAR model to predict common commercial drugs to act as high-performance corrosion inhibitors. Thus, this work is divided into four subsections shown as follows. Firstly, the QSAR model is rationalized. Secondly, we describe how the corrosion inhibition efficiency behaves in terms of the descriptors chosen. Then, we discuss the characterization of the high-performance corrosion inhibitors by a quantum-chemical study, and lastly, we provide a theoretical-experimental evaluation of lidocaine as a corrosion inhibitor on steel.

## 2. Theoretical and Experimental Methods

### 2.1. NARMAX System Identification Approach

Nonlinear autoregressive moving averages with exogenous inputs, NARMAX, system identification methodology [48] is used to build models smartly using historical data from the inputs, outputs, and other components. Some relevant applications of these techniques are stock prices [49] and weather [50] prediction, speech recognition [51], pattern classification [52], and aircraft dynamics [53]. As a model, the autoregressive with exogenous inputs (ARX) model is more straightforward (and easier to solve) than other more complex models such as NARMAX. This simplicity is because the ARX model does not require more detailed information about the system to be identified, such as nonlinear elements, which is why it is considered convenient for engineering applications where sufficient data in terms of representativeness is attainable as monitoring or diagnosis [54].

The NARMAX system identification methodology builds first a dictionary or matrix  $D$  composed of the system descriptors and corresponding observations. At that point, the forward regression orthogonal least squares (FROLS) algorithm [55,56] and the error reduction ratio (ERR) estimator [57] process the  $D$  matrix to produce an iterative feature selection that is capable of building a compact but representative model.

The ARX model in the system identification context has been successfully applied in various areas, such as troubleshooting and diagnostics for cooling dehumidifiers [54], fault analysis modeling for variable air volume [58], monitoring of buildings' energy consumption [59], interstitial glucose prediction during human physical activity [60], to predict the global magnetic disturbance in near-Earth space [61], the variability of the Atlantic meridional circulation [62], the Artemia population swimming motion [63], Drosophila photoreceptor responses [64], and EEG signal identification in the neuroscience field [65].

## 2.2. ARX Theoretical Model

The solution method for obtaining the prediction model presented in this work consists of two stages. The first one seeks to place the descriptor data for each drug in the form of a matrix according to the ARX model. In the second part, the descriptors that best explain the IE% are selected in an iterative process to identify the final model.

### 2.2.1. Arrangement of Candidate Terms

In the first step, it is required to arrange the commercial drug's data within rows and their corresponding chemical descriptors data into columns, aiming to produce an arrangement analogous to the ARX mathematical model structure, where only linear terms are included. Hence, the drug's regression representation remains as follows:

$$y_i = \sum_{j=1}^M \beta_j x_{i,j} \forall i \in N \quad (1)$$

The previous series can also be expressed as:

$$y_i = \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_M x_{i,M} \forall i \in N \quad (2)$$

where  $y_i$  is the maximum corrosion inhibition efficiency (IE%) for the  $i$ th commercial drug contained within the database (for  $i = 1, \dots, N$ ),  $x_{i,j}$  is the  $j$ th chemical descriptor value for the  $i$ th commercial drug (for  $j = 1, \dots, M$ ), and  $\beta_1, \dots, \beta_M$  is a set of adjusted weights to be computed once the final model has been identified by the FROLS algorithm.

The goal of generating this series of  $N$  equations is to pack them together to produce an  $N \times 1$  vector array on the left side, namely  $\mathbf{y}$ , and an  $N \times M$  matrix arrangement on the right side, namely  $\mathbf{X}$ , containing  $N$  rows of commercial drugs and  $M$  columns of chemical descriptors. In this stage, each column in  $\mathbf{X}$  represents a model term that is viewed as a *candidate* that can be included (or not) in the final ARX prediction model, as will be explained in the next section.

### 2.2.2. FROLS and ERR Algorithms for Model Structure Selection

The second stage of the solution methodology aims to identify a final ARX prediction model for accurately predicting the corrosion inhibition efficiency. We incorporated the FROLS algorithm to achieve this end, designed to select a subset of candidate terms that best explains the corrosion inhibition efficiencies contained in vector  $\mathbf{y}$ .

In the final model selection, the FROLS algorithm compares each candidate column in  $\mathbf{X}$  with the column vector  $\mathbf{y}$  to incorporate step by step into the model the terms/features that primarily reduce the error prediction of  $\mathbf{y}$ . Such a process continues iteratively until the selected descriptors collectively reach a minimum value for the sum of error reduction ratio (SERR) [48].

Finally, to complete the system's mathematical model, it is necessary to obtain the weights or parameters of the descriptors previously selected by the FROLS algorithm. The

weights are easy to calculate thanks to the fact that the new model has a linear-in-the-parameters representation:

$$\mathbf{y} = \boldsymbol{\beta} \mathbf{Z} \quad (3)$$

where  $\mathbf{y}$  is the response vector,  $\mathbf{Z}$  is the matrix of selected descriptors, and  $\boldsymbol{\beta}$  is the vector of weights obtained from the model. Thus, to compute the vector of linear weights, we only need to clear the vector  $\boldsymbol{\beta}$  as follows:

$$\boldsymbol{\beta} = \mathbf{y} \mathbf{Z}^{-1} \quad (4)$$

### 2.2.3. Cross-Validation

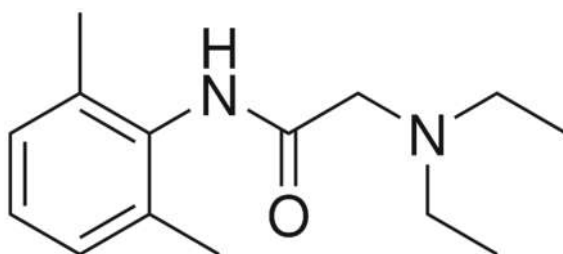
Cross-validation is a standard procedure when the dataset on hand has too few observations for data splitting. In the case of the current study, fivefold cross-validation was employed to seek a reliable model creation; hereafter, such a method is described [66,67]. Firstly, a database of 250 chemical compounds is elaborated as described in Section 3. Secondly, all compounds containing an experimental value of IE% are promoted to the training/validation set. The remaining compounds, those with no available experimental IE%, are set aside. Then, for each ARX/FROLS simulation run, the training/validation set is further randomly divided into five groups. In the first iteration, the first group is treated as a validation subset, and the remaining four groups are designated as the training subset, over which the ARX model and FROLS algorithm are run. The mean squared error (MSE) is computed on the first validation set. This procedure is repeated five times; each time, a different group of observations is treated as a validation subset. This process results in 5 different linear models with five distinct MSE estimates of the test error. The model with the lowest MSE value is promoted to predict the IE% of the remaining compounds.

Other metrics were employed to analyze the model's performance: Mean absolute percentage error (MAPE), standard deviation (SD), mean square error, and root mean square error (RMSE). These metrics, discussed throughout the manuscript, are defined now:  $\text{MAPE} = (1/n) \sum_{i=1}^n |(y_i - \hat{y}_i)/y_i| \times 100\%$ ,  $\text{SD} = [(1/n - 1) \sum_{i=1}^n (\hat{y}_i - y_i)^2]^{1/2}$ ,  $\text{MSE} = (1/n) \sum_{i=1}^n (y_i - \hat{y}_i)^2$ , and  $\text{RMSE} = [(1/n) \sum_{i=1}^n (y_i - \hat{y}_i)^2]^{1/2}$ . Where  $y_i$  is the experimental value of IE% for compound  $i$ ,  $\hat{y}_i$  is the estimated value of IE% for compound  $i$  provided by the model, and  $\bar{y}$  is the average IE% within a sample of compounds of size  $n$ .

## 2.3. Experimental Details

### 2.3.1. Solution Preparation

Different concentrations of the lidocaine compound (Figure 1) were prepared and obtained in an injectable solution from "Farmacias del Ahorro" pharmacy. The initial solution concentration was 0.01 M, dissolved in ethanol, to make later solutions of 0, 10, 20, 50, and 100 parts per million (ppm). The corrosive solution is NaCl 3% (100 mL).



**Figure 1.** Lidocaine molecule experimentally used to validate the ARX model.

### 2.3.2. Electrochemical Evaluation

The standard three-electrode system was used for the electrochemical evaluation at room temperature. The API 5L X70 sample was the working electrode, a saturated Ag/AgCl electrode was the reference electrode, and a graphite bar was the counter electrode. The test sequence was performed on a piece of Gill-AC equipment as follows: (a) Open circuit potential (OCP) was measured for 1800 s; (b) Electrochemical impedance spectroscopy (EIS)

was employed using  $10^{-2}$ – $10^4$  Hz with an amplitude of  $\pm 10$  mV. The exposure area of experimentally used samples was  $0.78 \text{ cm}^2$ . The electrochemical tests were performed in triplicate.

After the EIS measurements, potentiodynamic polarization curves of the inhibitor at different concentrations were performed, measured from  $-300$  mV to  $300$  mV in relation to the open circuit potential (OCP), with the speed of  $60 \text{ mV/min}$  using the ACM Analysis software for data interpretation.

### 2.3.3. Characterization by Atomic Force Microscopy (AFM)

The morphology of the steel sample surface after immersion in the corrosive media, in the presence and absence of the Lidocaine inhibitor, was characterized by atomic force microscopy (AFM) using digital instruments scanning probe microscope with a nanoscope IIIa controller. The AFM was operated in tapping mode using an etched silicon cantilever with a length of  $125 \mu\text{m}$ , with a nominal tip radius of approximately  $10 \text{ nm}$ .

## 3. Results and Discussion

In order to predict the corrosion inhibition efficiency of drugs on steel surfaces, a database with 250 commercial drugs containing common QSAR descriptors and those formulated within Pearson's HSAB theory was used to obtain a linear mathematical model by an ARX analysis. The ARX methodology was also used to exclude the variables that do not contribute to the prediction of IE%. In addition, a more sophisticated method for data analysis, IBM's Watson artificial intelligence [68], will be implemented to compare it with the ARX model and ensure the proper performance of the linear function. Thus, this work is divided into four subsections. Firstly, the model's determination is discussed and compared briefly with the privative AI model. Secondly, the main tendencies exhibited by the linear model are discussed. Then, species predicted as highly efficient corrosion inhibitors are studied and, for extension, correlated with their families. Lastly, the prediction of IE% for lidocaine was compared with its experimental counterpart.

### 3.1. Model Determination

This section explains how an ARX model was estimated to predict the iron corrosion inhibition of different compounds through a system identification methodology based on the FROLS algorithm. First, we narrate the data processing of the stated problem into a linear input-output system, followed by a term selection process to reach the final prediction model.

#### 3.1.1. Data Processing into an ARX Linear System

As stated, system identification aims to find a model that reveals the distinctive elements of a system by processing the history of its interactions with the environment. Here, we took the system to be identified as the corrosion inhibition efficiency on steel, IE%, explained by ten candidate quantum chemical descriptors. We studied the previous interaction through instances collected from 42 commercial chemical substances. The input-output model determination problem implies a suitable selection of the variables (descriptors) present in the following simple ARX linear model:

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \beta_7 x_7 + \beta_8 x_8 + \beta_9 x_9 + \beta_{10} x_{10} \quad (5)$$

where  $y$  is an output term plus ten input terms composed of candidate descriptors, as labeled in Table 1, and a corresponding parameter value or weight  $\beta_i$ . The list of the ten candidate descriptors and their corresponding index number is listed in Table 1. Each drug's molecular weight (MW) was considered a size-dependent parameter. Additionally, the negative base-10 logarithm of the acid dissociation constant of a solution,  $\text{p}K_a = -\log K_a$ , was included to determine the strength of an acid in the solution. Additionally, the octanol-water partition coefficient,  $\log P$ , is a descriptor associated with the concentration of a given substance in the aqueous phase of a two-phase octanol-water mix. Similarly, the

log *S* descriptor is directly related to the water solubility of a substance employing a base-10 logarithm. Besides, the polar surface area (PSA) is the molecular surface associated with heteroatoms and polar hydrogen atoms, giving a quantitative amount related to charge accumulation. In addition, polarizability,  $\alpha$ , denotes the tendency of a particular molecule to acquire an electric dipole moment in the presence of an external electric field. As described previously, energies of HOMO and LUMO orbitals can be related, through Koopman's theorem, to ionization energy and the electron affinity of a given molecular species, respectively. In addition, electrophilicity,  $\omega$ , relates to the change in energy of an electrophile when it comes in contact with a perfect nucleophile, being a measure of the tendency to react between electrophile and nucleophile species. Finally, the fraction of electrons shared,  $\Delta N$ , was selected since it relates to the amount of charge transferred from one species to another (Table 1).

**Table 1.** List of variables included, in the database, used to obtain the QSAR model, symbols, units, and references of their use in QSAR studies for corrosion inhibition. Additionally, parameters and ERR were obtained by the FROLS algorithm after processing the model of Equation (5) for the final model.

$x_i$	Descriptor	Symbol	Units	Reference	Parameter	ERR (%)
$x_1$	Molecular weight	MW	Da	[69,70]	-	-
$x_2$	Acid dissociation constant	pK <sub>a</sub>	-	-	0.5287	0.0600
$x_3$	Octanol-water partition coefficient	log <i>P</i>	-	[71,72]	-	-
$x_4$	Water solubility	log <i>S</i>	-	-	-	-
$x_5$	Polar surface area	PSA	Å <sup>2</sup>	[38,71,72]	-	-
$x_6$	Polarizability	$\alpha$	Å <sup>3</sup>	[72]	-	-
$x_7$	Energy of HOMO	$E_{\text{HOMO}}$	eV	[38,69,71,72]	812.1748	97.8259
$x_8$	Energy of LUMO	$E_{\text{LUMO}}$	eV	[38,69,71,72]	823.4630	0.1034
$x_9$	Electrophilicity	$\omega$	eV	[44,72]	6579.0080	0.0688
$x_{10}$	The fraction of electrons shared	$\Delta N$	-	[44,70,72]	33.1669	1.3933
<b>Sum of ERR</b>						<b>99.4514</b>

Unlike nonlinear models with second or third-order expansions, linear regression models are simpler to solve as they have precisely the same number of variables and model terms. A final critical point in data processing is the determination of the partition of the samples in testing and training data sets. Here, we divided data at random, where 80% of the data was used for training and 20% for validation.

### 3.1.2. Term Selection through FROLS and ERR

In the term selection stage, we consider the linear model in Equation (5) and the training set as entry points to the FROLS algorithm, which can easily detect the most relevant terms in first-order expansions [48]. At the first step, the ERR values of each candidate *m* included in the *D* dictionary, where  $D = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_M\}$ , are determined:

$$\text{ERR}_m = (\mathbf{y}^T \mathbf{p}_m / \mathbf{p}_m^T \mathbf{p}_m)^2 (\mathbf{p}_m^T \mathbf{p}_m) / \mathbf{y}^T \mathbf{y} \quad (6)$$

The ERR value helps determine the final subset of the model's terms because these indicate each one's contribution to the prediction error reduction. The first term selected from the *D* dictionary is always the one with the highest ERR. The descriptor with the highest ERR was  $x_7$ , with an ERR of 97.82% in our testing.

The selection process takes only the remaining unselected terms from dictionary *D* and introduces orthogonal transformations via the Gram–Schmidt algorithm from step two onwards. The orthogonalization process prevents the following candidates from being included from containing information already provided by the descriptors already selected, thus each new term contributes independently to the model's accuracy. Finally, the selection

procedure ends when the error-to-signal ratio (ESR) decreases below a predetermined threshold, where:

$$\text{ESR} = 1 - \sum_{i=1}^{M_0} \text{ERR}_i \leq \rho \quad (7)$$

where  $\rho$  is a very small value, for instance, in this work,  $\rho = 0.005$  and  $M_0$  is the number of unselected candidate variables. The terms included in the final model and the ERR values for each can be consulted in Table 1.

The final calculation involved the estimation of the parameters  $\beta$ , as it is shown in Section 2.2.1, thus the final ARX model, derived from the initial linear formulation in (5), can be stated as follows:

$$\hat{y} = 812.1748 x_7 + 33.1669 x_{10} + 823.4630 x_8 + 6579.0080 x_9 + 0.5287 x_2 \quad (8)$$

Consequently, the linear ARX model obtained accurate and precise results compared to the testing set, as evidenced by the computed mean absolute percentage error and standard deviation. The testing set included aspirin, cephapirin, ascorbic acid, imidazole, trimethoprim, clindamycin, phenobarbital, and doxycycline data, obtaining MAPE, SD, RMSE, and MSE of about 5.18, 2.51, 4.87, and 23.80%, respectively.

In addition, another set of drugs was collected to verify the generalization power of the ARX model. Additional data for streptomycin [73], fexofenadine [74], quinoline [75] N, N-dimethylformamide [76], and mycophenolic acid [25] was used to compute the MAPE, SD, RMSE, and MSE, obtaining 6.76%, 3.89%, 7.03%, and 49.47%, respectively. These values are fully comparable to those obtained with the testing set, pointing to a correct prediction of IE% values for drugs relatively different from those in the original database. IBM's Watson artificial intelligence platform was also used to obtain a private and highly hyperparametrized model. In this case, the five variables ( $\text{p}K_a$ ,  $E_{\text{HOMO}}$ ,  $E_{\text{LUMO}}$ ,  $\omega$ , and  $\Delta N$ ) in the ARX model were included in the AutoAI Watson's routine to fit the experimental IE% values [77]. Additionally, 80% of the data was used for the training set and the remaining 20% for validation. Finally, four different experiments, pipelines, were done by the extra trees regressor algorithm. The model obtained improved the external comparison only to 5.44%, 2.91%, 5.35%, 28.59% for MAPE, SD, RMSE, and MSE, respectively. Thus, the ARX model (obtained through FROLS), a linear function that depends only on five descriptors, obtained results close to a highly hyperparametrized, non-portable, and nonlinear alternative.

Other QSAR models have been proposed recently. Thus, it is pertinent to compare our linear ARX model with those found in the literature regarding standard evaluation metrics. For instance, Quadri and coworkers reported several multiple linear regression and artificial neural network (ANN) models adjusted to twenty pyridazine derivatives. The best ANN model yielded a lower MAPE value, of about 10.2362%, whereas the RMSE and MSE achieved were 10.5637 and 111.5910%, respectively [78]. In comparison, Li and colleagues obtained a QSAR model by a support vector machine approach to predict the performance of benzimidazole derivatives. In this case, the nonlinear model achieved an RMSE of about 6.79% [79]. More recently, they updated the model, improving the RMSE up to 4.45% [40]. In addition, Al-Fakih and coworkers reported QSAR models for furan derivatives, obtained with sparse multiple linear regression using ridge penalty and sparse multiple linear regression using an elastic net, achieving MSE of about 7.75 and 2.34%, respectively [80]. The ARX approach obtained similar metrics compared to those obtained with alternative methods, as reported by other authors.

A brief discussion is included below to identify the extent and perspectives of the ARX linear model. In principle, corrosion inhibition is a multifactorial phenomenon since drug solubility, pH, temperature, concentration, corrosive medium, dynamic conditions, the employed alloy, and even experimental technique used to determine IE% could influence the results [2]. Thus, it is possible to assume that a 5-parameters model such as that introduced above is not enough to catch the variety of conditions experimentally used. However, to the best of our knowledge, these experimental variables are not recurrently



considered, possibly by the scarce IE% values measured with the same experimental design, hindering the formulation of mathematical models. Nevertheless, work conditions are naturally occurring variables that should be considered for robust predictive models.

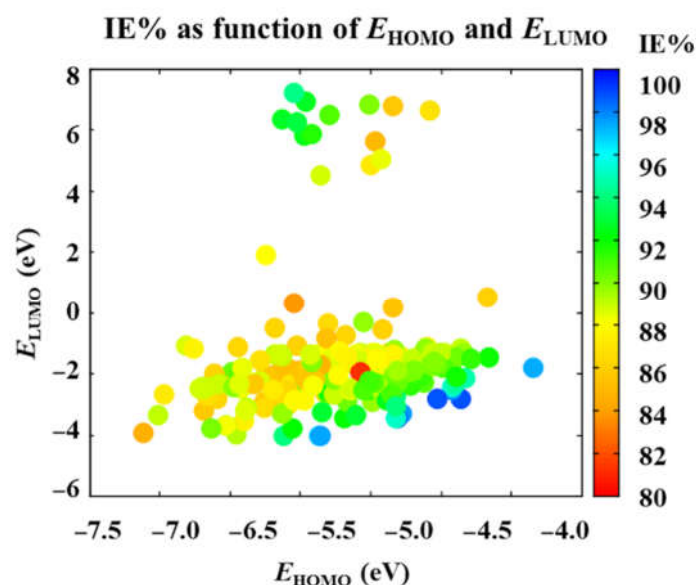
Finally, although linear models to predict IE% are the most common approach [78], nonlinear formulations can also be suitable. A nonlinear version of ARX comes to be the NARMAX model in the current case. According to Gu et al., nonlinear models performed better than linear ones in a study about cortical responses. Nevertheless, the linear terms had larger weights than those in the resulting NARMAX models [81]. In addition, NARMAX approaches are known to identify mathematical models for nonlinear systems, which prevail in nature. This is the case of the solar wind coupling analysis reported by Boynton and coworkers [82] or the peak air pollution levels forecasted by Pisoni et al. [83]. Thus, it is expected that a NARMAX approach to the corrosion inhibition problem may well determine whether the phenomenon is nonlinear in terms of the proposed descriptors. However, as suggested by Boynton, a high number of possible monomials resulting from the polynomial expansion can be a challenging situation. The above stems from the need for a final parsimonious NARMAX model with fewer monomials selected out of a vast majority with no or minimal influence on the phenomenon. In principle, such as in the ARX approach used here, the FROLS algorithm could lead to a small set of monomials within the selected allowable model order [84].

### 3.2. Main Tendencies

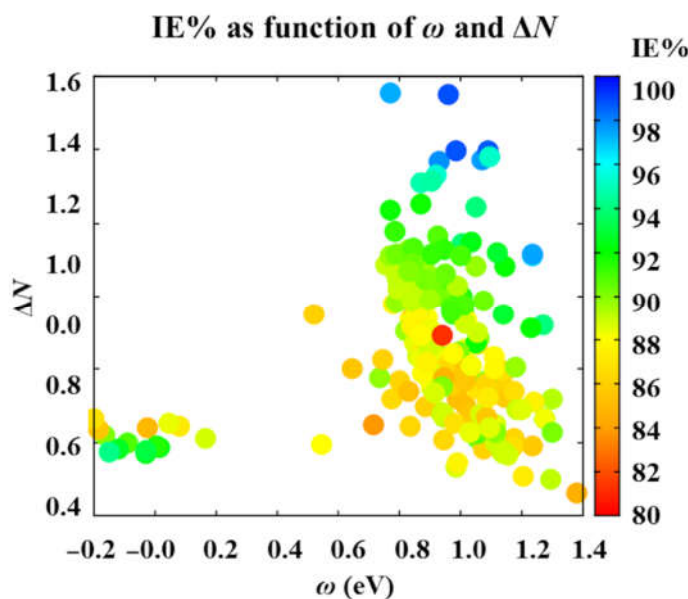
This subsection aims to elucidate the main tendencies exhibited by the predictions, carried out on 250 commercial drugs by the ARX model detailed previously, of the corrosion inhibition efficiency IE%. Thus, five variables contained in the ARX mathematical model ( $E_{\text{HOMO}}$ ,  $E_{\text{LUMO}}$ ,  $\Delta N$ ,  $\omega$ , and  $\text{p}K_{\text{a}}$ ) and molecular weight were used to illustrate the predicted IE% values. In addition, five drugs were excluded from the following analyses due to their unrealistic IE% values predicted above 100%. These species are sulfadiazine (106.31%), methacycline (111.72%), glycine (124.03%), ethosuximide (158.88%), and hexetidine (259.25%).

First of all, Figure 2 shows that corrosion inhibition efficiency IE% increases as the energy of HOMO decreases. Consequently, by Koopman's theorem, the most efficient drug molecules to act as corrosion inhibitors are those with the lowest ionization potential. The above can be rationalized by the necessity of the metal surface atoms to fill their vacant d-orbitals with electrons, coming from the corrosion inhibition molecule in the current case. Consequently, the easier it is to remove valence electrons from the organic molecule, the higher corrosion inhibition performance exhibited.

On the other hand,  $E_{\text{LUMO}}$  splits the IE% values into two sets (Figure 2). The first set of drugs, with  $E_{\text{LUMO}}$  values below 2.0 eV, shows a tendency similar to that exhibited by  $E_{\text{HOMO}}$  since the most efficient corrosion inhibitors are those with the most negative  $E_{\text{LUMO}}$  values and consequently the highest electron affinities. Thus, it is expected that high-performance corrosion inhibitor molecules can catch electrons from the environment and donate them, leading to shared electrons as in covalent interactions. However, the other set of molecules with moderate performance, with  $E_{\text{LUMO}}$  values above 4.5 eV, offers another route to produce the corrosion inhibition effect (Figure 3). Since up to 14 species are obtained with low and even positive  $E_{\text{LUMO}}$  values and intermediate  $E_{\text{HOMO}}$  energies, ranging from  $-6.5$  to  $-5.0$  eV, they are expected to donate charge to the metal surface and handle mostly electrostatic interactions with it. These intermediate efficiency corrosion inhibitors are, in order of increasing IE%, the following ones: cyclopentamine (84.92%), methenamine (85.56%), triethylamine (86.74%), gentamicin (87.12%), mecamlamine (88.32%), kanamycin (88.66%), diethylamine (90.00%), diethanolamine (90.75%), ethambutol (91.46%), amantadine (92.43%), ethanolamine (92.57%), tromethamine (92.59%), tuaminoheptane (92.99%), and ethylamine (94.12%).



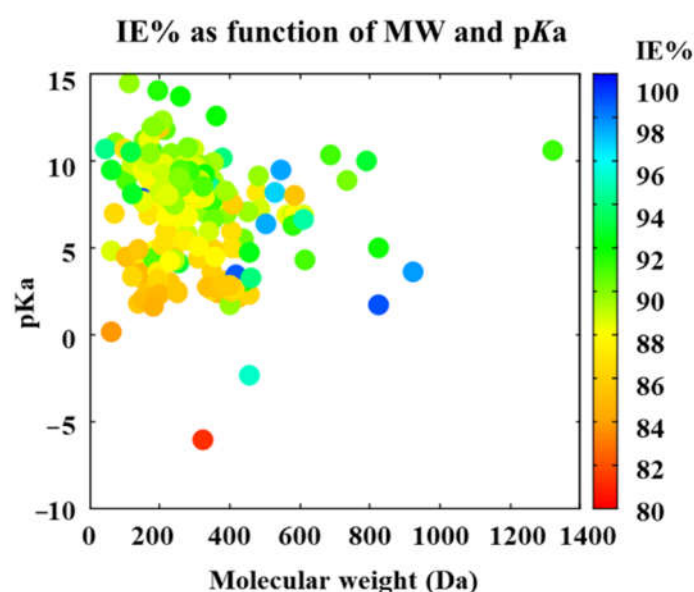
**Figure 2.** Corrosion inhibition efficiency, predicted by the ARX model, as a function of  $E_{\text{HOMO}}$  and  $E_{\text{LUMO}}$ . Variables calculated at the DFTB3-LJ level of theory.



**Figure 3.** Corrosion inhibition efficiency, predicted by the ARX model, as a function of electrophilicity,  $\omega$ , the fraction of electrons transferred,  $\Delta N$ . Variables calculated at the DFTB3-LJ level of theory.

In the case of IE% as a function of the electrophilicity and the fraction of electrons shared, the efficiency shows two sets of compounds. The highest IE% values are obtained in the molecules with the highest  $\omega$  and  $\Delta N$  values (see Figure 3). The above can be related to the previous assumptions about how the electrons behave between the corrosion inhibitor molecule and the metal surface. Thus, a highly effective corrosion inhibitor molecule is expected to donate a considerable number of electrons to the metal surface, as denoted by the  $\Delta N$  values calculated for the specific case of an iron surface. Additionally, according to the ARX model, a highly efficient corrosion inhibitor molecule is expected to behave as an electrophile with high power to attract electrons to itself. Besides, the other set of molecules is composed precisely of those drugs named above. Drugs with intermediate performance, with IE% efficiency ranging from 84.92% to 94.12%, cannot donate large amounts of charge to the metal surface and neither to attract it. Thus, the fourteen species with moderate performance are expected to interact by electrostatic interactions.

Lastly, the heat map obtained for IE% as a function of the molecular weight and  $pK_a$  shows more smoothly dispersed values without the two sets obtained previously (see Figure 4). It is noticeable that the highest efficiency is obtained for species with molecular weight ranging from 415 to 823 Da. Additionally, all the high-performance molecules are predicted for those species with positive  $pK_a$ , ranging from 5 to 10. Thus, weak acids are expected to behave as potential corrosion inhibitors, whereas strong acids are not helpful for this application. Although light molecules are not particularly prominent by their corrosion inhibition efficiencies, values above 90% are easily reached. In the case of these light species, with a molecular weight below 400 Da, the IE% increases as the  $pK_a$  increases. Since the commercial drugs must exhibit high corrosion inhibition efficiencies, above 90% according to international standards for industrial applications, the following subsection aims to deep into the species predicted by the ARX models as highly efficient corrosion inhibitors.



**Figure 4.** Corrosion inhibition efficiency, predicted by the ARX model, as a function of molecular weight and  $pK_a$ .

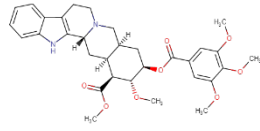
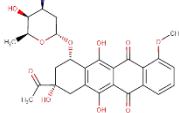
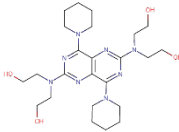
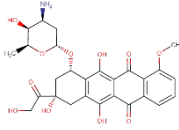
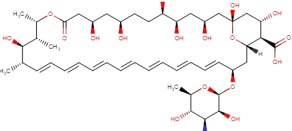
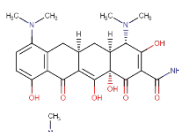
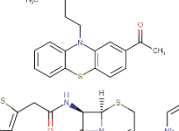
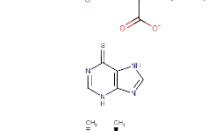
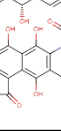
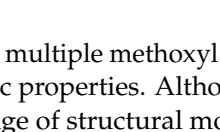
### 3.3. High-Efficiency Corrosion Inhibitors

In order to produce reliable predictions for highly efficient commercial drugs to act as corrosion inhibitors, species with IE% estimated above 95% are studied as follows. The value 95% was chosen to take into account MAPE and SD values computed for the mathematical model of about 5.18% and 2.51%, respectively. This way, it is expected that efficiencies measured under experimental conditions must fall into the regime required for industrial applications, with inhibition efficiencies above 90%. The drugs fulfilling these conditions, shown in Table 2, are minocycline (97.58%), deserpidine (95.29%), daunorubicin (96.67%), dipyridamole (97.28%), doxorubicin (97.40%), amphotericin B (97.55%), acepromazine (97.73%), cephaloridine (98.57%), mercaptopurine (98.66%), and rifampicin (98.71%).

Being weak acids, the most efficient drugs are found with  $pK_a$  ranging from 1.70 to 9.46. Moreover,  $E_{HOMO}$  values are found in a tight range, from  $-5.87$  to  $-4.34$  eV. Conversely,  $E_{LUMO}$  values are exhibited in a broader range, from  $-4.01$  to  $-1.83$  eV. Electrophilicity ranges from 0.77 to 1.23 eV, whereas  $\Delta N$  ranges from 1.12 to 1.55. Thus, all the above is consistent with previous observations about general tendencies. There are no significant similarities among all the high-efficient species; however, some functional groups are recurrently exhibited by these species. In addition, invoking the similarity principle, drugs belonging to the same family are suitable to be presumed as corrosion inhibitors with efficiencies comparable to those obtained for the species shown in Table 2. For instance,

minocycline is a tetracycline antibiotic, with multiple dimethylamino, hydroxyl, and carbonyl groups, obtained as a highly efficient corrosion inhibitor, with an IE% predicted value of 97.58% (Table 2). Other tetracycline antibiotics are presumed to be obtained as suitable corrosion inhibitors. Doxycycline and oxytetracycline are predicted as efficient CIs, with 91.89% and 93.95% IE% values, respectively.

**Table 2.** Descriptor values, common use, and structure of drugs predicted as corrosion inhibitors with efficiency above 95%.

Drug	pK <sub>a</sub>	E <sub>HOMO</sub>	E <sub>LUMO</sub>	ω	ΔN	IE%	Common Use	2D Structure
Deserpidine	6.68	−4.92	−2.42	0.92	1.33	95.29	Antihypertensive and antipsychotic	
Daunorubicin	8.20	−5.87	−4.01	1.23	1.11	96.67	Cancer treatment	
Dipyridamole	6.40	−4.34	−1.83	0.77	1.55	97.28	Anticoagulant	
Doxorubicin	9.46	−5.86	−4.00	1.23	1.12	97.40	Cancer treatment	
Amphotericin B	3.58	−5.27	−3.29	1.07	1.37	97.55	Antibiotic and fungicide	
Minocycline	2.30	−5.32	−3.42	1.09	1.38	97.58	Antibiotic	
Acepromazine	9.30	−4.92	−2.53	0.93	1.37	97.73	Antipsychotic	
Cephaloridine	3.40	−5.31	−3.43	1.09	1.40	98.57	Antibiotic	
Mercaptopurine	7.80	−5.03	−2.83	0.98	1.40	98.66	Cancer treatment	
Rifampicin	1.70	−4.86	−2.81	0.96	1.55	98.71	Antibiotic	

Deserpidine is an example of an ester alkaloid drug, exhibiting multiple methoxyl and carbonyl groups, used for their antihypertensive and antipsychotic properties. Although alkaloids predicted to act as corrosion inhibitors cover a wide range of structural motifs, the most prominent by their structural similarity to deserpidine and predicted IE% value of about 94.55% is reserpine, another alkaloid that can be obtained from *Rauwolfia* species (Table 2). By extension, yohimbine is another alkaloid presumed to act as a suitable corro-

sion inhibitor. Finally, daunorubicin and doxorubicin are two examples of anthracycline class drugs for cancer treatments [85], with predicted IE% values of about 96.67% and 97.40%. Similar to the previous cases, these species exhibit several hydroxyl and carbonyl groups, presumed to be covalent-polar and dative bonds with iron surface atoms, respectively [86,87]. Additionally, multiple  $\pi$ -electrons coming from their aromatic rings can be donated to the iron surface.

Furthermore, amphotericin B is another drug that is predicted to act as a highly efficient corrosion inhibitor, with IE% estimated as 97.55%. This drug belongs to macrolides and the same class of antibiotics, including erythromycin, roxithromycin, azithromycin, and clarithromycin. In particular, Amphotericin B does not contain several aromatic rings, such as those exhibited by the species previously discussed. Thus, this drug, exhibiting eleven hydroxyl groups and two carbonyl ones, is expected to interact mainly with the metal surface by their functional groups. Erythromycin is another macrolide expected to work as a suitable corrosion inhibitor, with IE% estimated as 90.11% (Table 2) according to NRF-005-2009. This decrement in the corrosion inhibition efficiency can be rationalized by the few functional groups in erythromycin capable of interacting with the metal surface compared to amphotericin B.

On the other hand, acepromazine heads the phenothiazine family of antipsychotics, accounting for an estimated corrosion inhibition efficiency of about 97.73%. This molecule's interactions can be explained by the  $\pi$  electrons available in its aromatic rings in addition to a carbonyl group. Other members of the phenothiazine family, promazine and levomepromazine, lack that carbonyl group. These drugs achieve lower corrosion inhibition efficiencies of 91.19 and 91.81% (Table 2). Another relevant drug is cefaloridine, being part of the cephalosporin class, a large group of antibiotics derived from the fungus *acremonium*. Aromatic rings and carbonyl groups are the most relevant structural motifs constituting cefaloridine. However, several cephalosporin drugs were also studied but achieved lower efficiencies. These drugs are cephapirin (85.72%), cephalixin (85.87%), cephalothin (86.06%), cefazolin (86.24%), and cephradine (86.68%). In this case, it is unclear why cefaloridine is so efficient since all cephalosporin exhibits similar functional groups. Thus, it is necessary to deepen the corrosion inhibition effect of cefaloridine and other cephalosporin drugs. According to these measurements, cephapirin (82.5%) and cephalixin (76.9%) achieve lower corrosion inhibition efficiencies in comparison with the predicted values, whereas cephalothin (92.0%), cefazolin (93.9%), and cephradine (95%) are closer to the value predicted for cefaloridine of about 98.57% (see Supplementary Materials). It is plausible that the ARX model does not well describe the cephalosporin family.

Dipyridamole and mercaptopurine are two drugs without relatives in the current study. With high corrosion inhibition efficiencies of about 97.28 and 98.66%, respectively, dipyridamole and mercaptopurine are presumed to strongly interact by their double bonds and polar functional groups. Finally, rifampicin is part of the rifamycins class, a group of antibiotics, with IE% estimated as 98.71% (Table 2). Interestingly, rifampicin is the only of these highly efficient corrosion inhibitor molecules, with an experimental measure of its IE% of about 94.7% (see Supplementary Materials). Furthermore, this species exhibits several functional groups, such as carbonyl and hydroxyl groups and heteroatoms and  $\pi$ -electrons, thus suggesting that rifampicin could interact with the metal surface by covalent, dative bonds, or electrostatic interactions. Thus, the whole rifamycin family could be evaluated as corrosion inhibitors.

The above discussion points out the highly efficient corrosion inhibition properties expected for ten commercial drugs, with IE% values ranging from 95.29 to 98.71% (Table 2). In comparison, other drugs experimentally evaluated exhibited similar or lower corrosion inhibition efficiencies. For instance, losartan, a drug commonly used for hypertension treatment, obtained a maximum IE%, by EIS technique, of only 92.0% [88]. Additionally, salbutamol, a commonly used treatment for asthma, was obtained with a maximum IE%, determined by EIS measurement, of about 84% [89]. Similarly, Anadebe and coworkers collected the maximum IE% achieved by several recently reported commercial drugs,

ranging from 80% to 95% [89]. Drugs applied to carbon steel were tobramycin (80%) [90], metformin (90%) [91], metolazone (92%) [92], and nifedipine (94%) [93]. Additionally, species applied on mild steel surfaces were: dexamethasone (83%) [94], rosuvastatin (90%) [95], ambroxol (94%) [96], and dapsone (95%) [97]. Additionally, Abeng, et al. reported the IE% on carbon steel achieved by moxifloxacin (88.2%), nifedipine (89.6%), metolazone (92.8%), and levofloxacin (94.1%) [98]. Even more, the recently evaluated drugs, mycophenolic acid [25] and fluconazole [24], obtained moderate efficiencies of about 90%. Clearly, all these recently studied drugs obtained values below the threshold assumed in this work. However, other drugs were comparable to those expected for the molecules analyzed in this subsection. This is the case of pyrazinamide, isoniazid, and rifampicin, which achieved maximum IE% values of about 95.86%, 97.89%, and 97.06%, respectively [99]. Thus, the predictions and tendencies discussed could be the object of study and further confirmation.

### 3.4. Experimental Verification

#### (a) Open circuit potential (OCP)

The variations of the OCP without and with the corrosion inhibitor lidocaine reached a steady state at 600 s (Figure 5). It is evident that when a corrosion inhibitor is present, the potential considerably decreases, indicating a drop in the corrosion process.

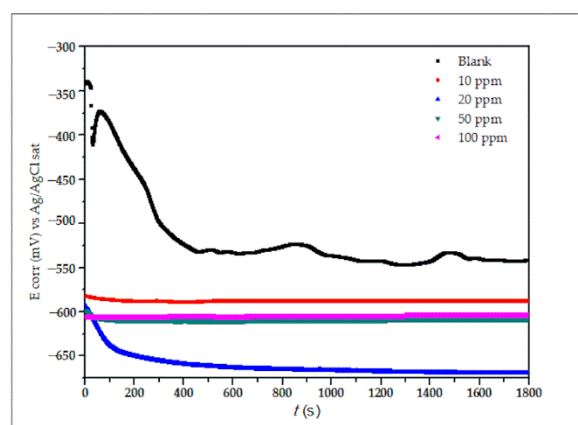


Figure 5. Variation of OCP at different concentrations of Lidocaine in API 5L X70 steel.

#### (b) Concentration effect of lidocaine by EIS

The equivalent electric circuits employed to fit the experimental data of the Nyquist diagrams are shown in Figure 6. A Randles circuit in the case of the sample of Figure 6a was used without inhibitor (Blank). A parallel circuit with two constant phase elements (Figure 6b) for the samples with inhibitor.

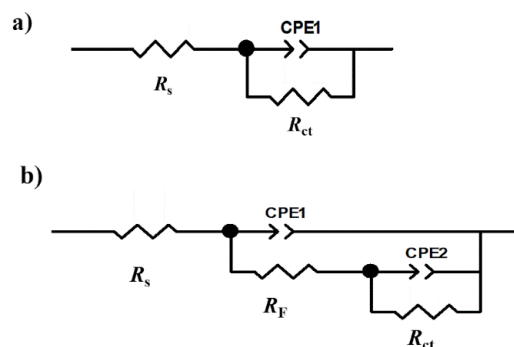


Figure 6. Equivalent electrical circuits. (a) A Randles circuit in the case of the sample was used without inhibitor (Blank). (b) A parallel circuit with two constant phase elements for the samples with inhibitor.

Where  $R_s$  is the solution resistance,  $R_{ct}$  is the charge transfer resistance,  $CPE_{inh}$  is the constant phase element of the inhibitor, and  $CPE_{Rct}$  is the constant phase element associated with the double layer. The inhibitor efficiency IE% was calculated by the equation [100]:

$$IE\% = 100 [(R_{p \text{ blank}}^{-1} - R_{p \text{ inh}}^{-1})/R_{p \text{ blank}}^{-1}] \quad (9)$$

where  $R_{p \text{ blank}}^{-1}$  is the polarization resistance of blank and  $R_{p \text{ inh}}^{-1}$  is the polarization resistance of sample with inhibitor.

The polarization resistance ( $R_p$ ) was calculated with:

$$R_p = R_{ct} + R_F \quad (10)$$

where  $R_{ct}$  is charge transference resistance and  $R_F$  film resistance, in  $\Omega \text{ cm}^2$ .

The electrochemical double-layer capacitance ( $C_{dl}$ ) was calculated through the next equation [101]:

$$C_{dl} = Y_0^{1/n} (R_s^{-1} + R_{ct}^{-1})^{(n-1)/n} \quad (11)$$

where  $Y_0$  is the constant phase element,  $R_s$  is the solution resistance ( $\Omega \text{ cm}^2$ ), and  $R_{ct}$  is the charge transfer resistance ( $\Omega \text{ cm}^2$ ).

For the description of a frequency-independent phase shift between an applied AC potential and its current response, a constant phase element ( $CPE$ ) is used, defined in the impedance representation as:

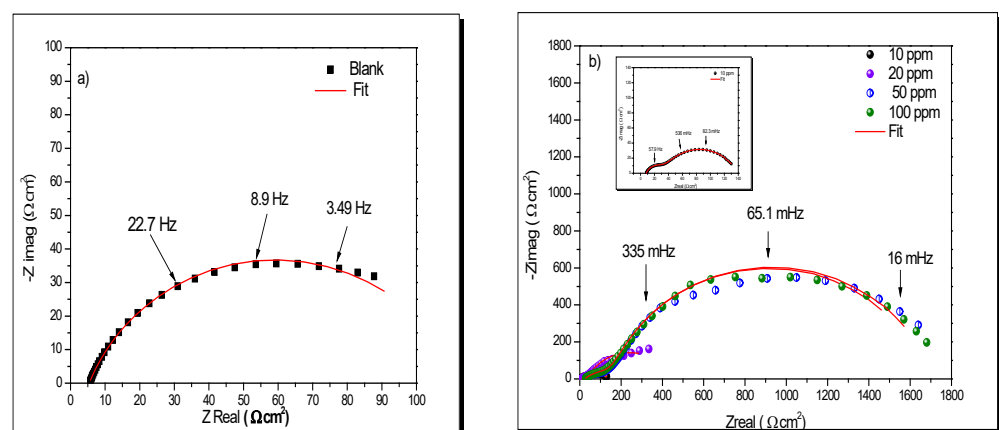
$$Z_{CPE} = Y_0^{-1}(j\omega)^{-n} \quad (12)$$

where  $Y_0$  is the  $CPE$  constant,  $n$  is the  $CPE$  exponent that can be used as a gauge of the heterogeneity or roughness of the surface,  $j = -1$  is an imaginary number, and  $\omega$  is the angular frequency in  $\text{rad s}^{-1}$ . Depending on  $n$ ,  $CPE$  can represent a resistance ( $Z_{CPE} = R$ ,  $n = 0$ ), a capacitance ( $Z_{CPE} = C$ ,  $n = 1$ ), and a Warburg impedance ( $Z_{CPE} = W$ ,  $n = 0.5$ ), or inductance ( $Z_{CPE} = L$ ,  $n = -1$ ). The correct equation to convert  $Y_0$  into  $C_F$  is given by [102]

$$C_F = Y_0 (\omega')^{n-1} \quad (13)$$

where  $C_F$  is the film capacitance ( $\mu\text{F}/\text{cm}^2$ ) and  $\omega'$  is the angular frequency at which  $Z_{\text{real}}$  is maximum.

Figure 7a shows the Nyquist diagram of API 5L X52, called "blank". A semicircle that is not entirely close was observed, attributed to charge transfer resistance.



**Figure 7.** Nyquist plots (a) without inhibitor and (b) different concentrations of lidocaine in API 5L X70 immersed in NaCl 3%.

On the other hand, the Nyquist diagram in Figure 7b in all concentrations showed two processes: one attributed to inhibitor film and the charge transfer resistance [103].

Table 3 shows the electrochemical parameters obtained after fitting with the equivalent electric circuits of Figure 6. At 100 ppm, the highest value of  $R_{ct}$  was reached ( $1522 \Omega \text{ cm}^2$ ).

This behavior could be attributed to the high adsorption process of the organic compound on the metallic surface. On the other hand, the  $C_{dl}$  and  $C_F$  values decreased because the surface is protected by the inhibitor at the metal/solution interface [104].

**Table 3.** Electrochemical parameters at different concentrations of lidocaine in API 5L X70 immersed in NaCl 3%.

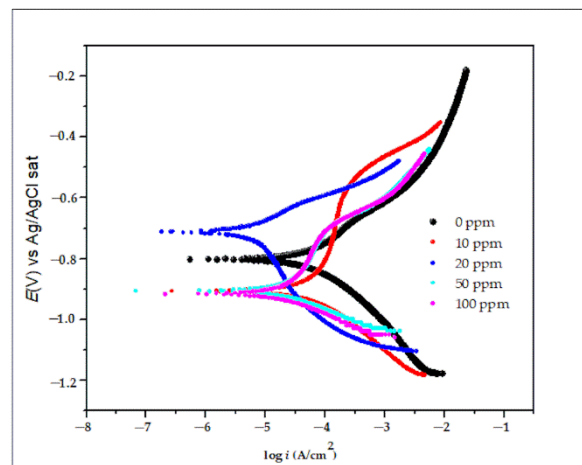
C (ppm)	$R_s$ ( $\Omega \text{ cm}^2$ )	$n$	$C_{dl}$ ( $\mu\text{F}/\text{cm}^2$ )	$R_{ct}$ ( $\Omega \text{ cm}^2$ )	$C_F$ ( $\mu\text{F}/\text{cm}^2$ )	$n_2$	$R_{mol}$ ( $\Omega \text{ cm}^2$ )	$R_{total}$ ( $\Omega \text{ cm}^2$ )	IE% (%)
0	6	0.800	2960	127	-	-	-	-	-
10	8.24	0.80	181.3	102.00	4034.0	0.8	28.70	130.70	3.2
20	10.53	0.77	187.5	404.10	622.2	0.52	337.90	742.00	83.0
50	24.66	0.85	90.3	1493.00	40.7	0.49	151.70	1644.70	92.3
100	24.29	0.84	51.9	1522.00	26.0	0.48	157.00	1679.00	92.5

(c) Polarization curves

The polarization curves of the API 5L X70 steel, immerse in 3% NaCl with and without lidocaine, are shown in Figure 8. The polarization parameters are enlisted in Table 4: corrosion potential ( $E_{corr}$ ), current density ( $i_{corr}$ ), tafel slopes ( $b_a$  and  $b_c$ ), and the inhibition efficiency (IE%) was determined by:

$$IE\% = [1 - i_{corr}/i_{corr \text{ blank}}] \times 100 \tag{14}$$

where  $i_{corr}$  and  $i_{corr \text{ blank}}$  are the current density with and without inhibitor.



**Figure 8.** Polarization curves of different concentrations of lidocaine in API 5L X70 steel immersed in 3% NaCl.

**Table 4.** Polarization parameters for lidocaine in API 5L X70 immersed in NaCl 3%.

C (ppm)	$E_{corr}$ (mV) vs. Ag/AgCl sat	$i_{corr}$ ( $\mu\text{A}/\text{cm}^2$ )	$b_a$ (mV/dec)	$-b_c$ (mV/dec)	IE% (%)
0	-804.7	67.4	159.5	173	-
10	-909.7	65.0	146.6	161.5	3.4
20	-709.6	4.9	104.5	204.1	92.6
50	-907.7	7.4	170.5	60.3	89.0
100	-916.5	8.2	187.8	68.2	87.4

Figure 8 shows that the curve shifts to the left due to the corrosion current density decreasing when the concentration increases due to the adsorption of the inhibitor.

The electrochemical parameters by this technique show that the corrosion current density ( $i_{corr}$ ) value decreased in the presence of the lidocaine inhibitor, being attributed to



the protected metal surface (Table 4). This phenomenon implies that the inhibitor might suppress the anodic reaction of the metal dissolution and the detachment of cathodic hydrogen [105], while the inhibition efficiency by this technique shows similar results to the other technique (EIS), and the best concentration was 20 ppm with 92.6% to protect the metal surface. Finally, according to corrosion potential ( $E_{\text{corr}}$ ), at a concentration of 20 ppm, the lidocaine behaved as an anodic inhibitor, while at 10, 50, and 100 ppm, the behavior was cathodic.

(d) Adsorption process

The corrosion inhibition displaces the water molecules and replaces them with the inhibitor molecules on the metal surface.

Nevertheless, the superficial coverage ( $\theta$ ) for the different lidocaine concentrations as corrosion inhibitors in this system was evaluated by EIS using IE%:

$$\theta = (1/100) \text{ IE\%} \quad (15)$$

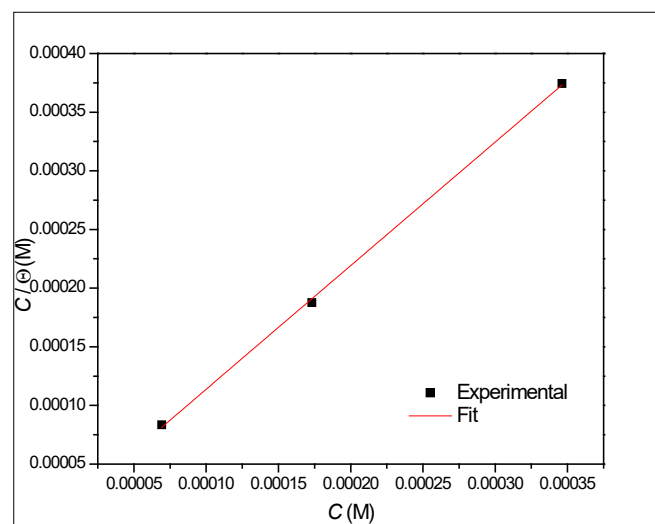
Using the Langmuir isotherm, a good fit is obtained, and according to the value of the free energy of adsorption of Gibbs (Equation (17)), the combined process continues [106].

$$C/\theta = k_{\text{ads}}^{-1} + C \quad (16)$$

$$\Delta G_{\text{ads}}^0 = -55.5 RT \ln k_{\text{ads}} \quad (17)$$

where  $C$  is the concentration,  $\theta$  is the coating coverage,  $k_{\text{ads}}$  is the adsorption constant,  $R$  is the ideal gas constant, and  $T$  is the temperature. In some reports, the  $\Delta G_{\text{ads}}^0$  will indicate which adsorption mechanism follows the organic compound; if it is lower than  $-20$  kJ/mol, it can be considered a physisorption process. If it is higher than  $-40$  kJ/mol ( $\Delta G_{\text{ads}}^0 > -40$  kJ/mol), then it is a chemisorption process, but if it is in the middle of  $-20$  kJ/mol and  $-40$  kJ/mol, the type of process that is taking place is called “combined”.

Figure 9 shows a good fit, having a correlation coefficient ( $R^2$ ) of 0.9996. The equation obtained is  $C/\theta = 1.0534 C + 8 \times 10^{-6}$ , obtaining a  $\Delta G_{\text{ads}}^0 = -38.39$  kJ/mol due to a combined type process.

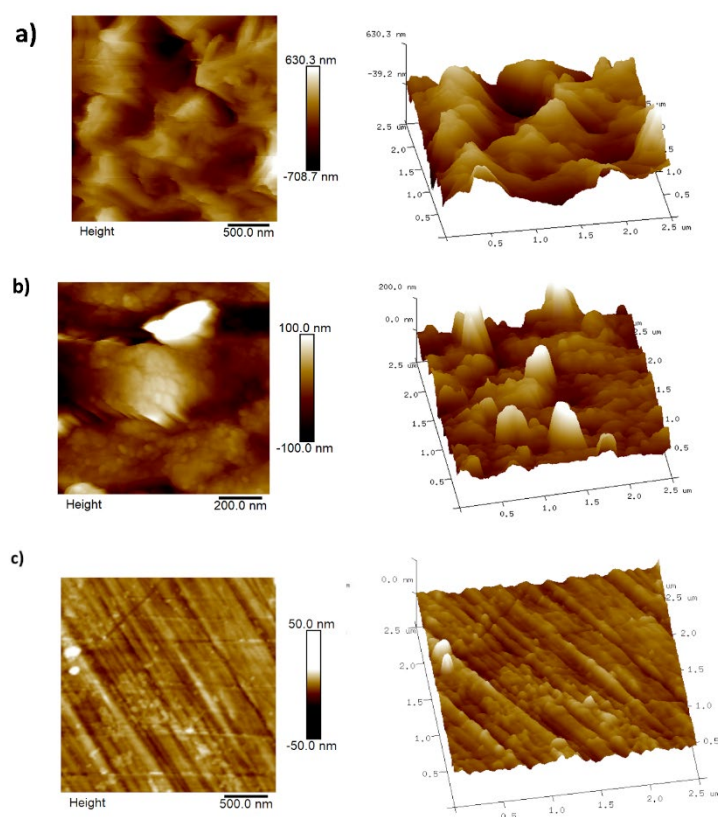


**Figure 9.** Langmuir isotherm at different concentrations of lidocaine in API 5L X70 steel immersed in NaCl 3%.

(e) AFM analysis

Figure 10 shows AFM images recorded on the surfaces of steel samples, and Table 5 shows the roughness values; both Ra (the mean roughness or arithmetic average of the

absolute values of the roughness profile ordinates) and  $R_q$  (root mean square roughness or the root mean square average of the roughness profile ordinates) are reported.



**Figure 10.** AFM images (2D and 3D formats) were recorded on the surface of API 5L X70 steel samples after 24 h immersion in NaCl 3%: (a) in the absence and (b) in the presence of 50 ppm lidocaine. The image in (c) corresponds to the as-polished steel sample not immersed in the corrosive media.

**Table 5.** Roughness values calculated from the AFM images shown in Figure 10.

AFM image	$R_a$ (nm)	$R_q$ (nm)
a	142	181
b	30.5	45
c	3.4	4.3

After exposure to the corrosive media for 24 h in the absence (Figure 10a) and in the presence of 50 ppm of lidocaine (Figure 10b), it can be noted that the roughness values of the steel sample that was not protected with lidocaine are notorious compared with that of the sample that was protected with lidocaine or that which was not exposed to the corrosive media, Figure 10c, such notable rugosity is due to the different corrosion products formed in each case.

#### 4. Conclusions

A QSAR relationship constructed through a linear ARX model was used to predict the corrosion inhibition efficiency of 250 commercial drugs on steel. The ARX treatment found the five most important descriptors to predict IE%, reducing by half the number of variables used in the linear model. These variables, obtained mostly from quantum chemical calculations of gas-phase molecules at the DFTB3-LJ level, are the  $pK_a$ , electrophilicity, HOMO and LUMO energies, and the fraction of electrons transferred to bulk iron.

The ARX model obtained a MAPE and SD of about 5.18 and 2.51%, respectively, compared to the testing set. Another five drugs not included in the original database were used as an external validation set for which the computed MAPE and SD were approximately 6.76% and 3.89%, respectively, thus confirming the previous predictions. In addition, a fivefold model, obtained by IBM's Watson AI extra trees regressor algorithm, was used to compare it with the ARX one. In this case, IBM's Watson model improved the external comparison only to 5.44% and 2.91% for MAPE and SD, respectively. Thus, the linear ARX model is competitive compared to highly hyperparametrized and privative alternatives.

Overall, there are several tendencies of IE% as a function of the selected variables. For instance, IE% increases as the energy of HOMO decreases. Additionally, the highest IE% values are obtained in the case of the molecules with the highest  $\omega$  and  $\Delta N$  values. The most efficient drugs are found with  $pK_a$  ranging from 1.70 to 9.46. The drugs recurrently exhibit aromatic rings, carbonyl, and hydroxyl groups with the highest IE% values. Ten drugs are predicted with IE% above 95%—those are: minocycline (97.58%), deserpidine (95.29%), daunorubicin (96.67%), dipyridamole (97.28%), doxorubicin (97.40%), amphotericin B (97.55%), acepromazine (97.73%), cephaloridine (98.57%), mercaptopurine (98.66%), and rifampicin (98.71%). Alkaloids from *Rauwolfia* species, macrolides, cephalosporin, and rifamycin antibiotics are expected to exhibit high IE% on steel surfaces.

Lastly, lidocaine was predicted and experimentally tested for the first time. At 100 ppm concentration, the lidocaine showed IE% of about 92.5% using electrochemical impedance spectroscopy and 87.4% by polarization curves; in comparison, the ARX model predicted 87.51%. The thermodynamic analysis showed that the lidocaine follows a mixed adsorption process in the API 5L X70 surface. This linear model proposed to deepen the use of commercial and reused drugs to act as corrosion inhibitors on steel.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23095086/s1>.

**Author Contributions:** Conceptualization, A.E.-V., C.B.-P., A.A.A.S. and A.M.; methodology, A.E.-V., C.B.-P., A.A.A.S. and A.M.; software, C.B.-P. and A.M.; validation, A.E.-V., C.B.-P., A.A.A.S. and A.M.; formal analysis, A.E.-V., C.B.-P., A.A.A.S., R.O.-C., A.M.-J. and A.M.; investigation, A.E.-V., C.B.-P., A.A.A.S., G.S.-R., R.O.-C., A.M.-J. and A.M.; resources, A.E.-V., C.B.-P., R.O.-C., A.M.-J. and A.M.; data curation, A.E.-V., C.B.-P., A.A.A.S., G.S.-R., R.O.-C., A.M.-J. and A.M.; writing—original draft preparation A.E.-V., C.B.-P., A.A.A.S. and A.M.; writing—review and editing, A.E.-V., C.B.-P., A.A.A.S., G.S.-R., R.O.-C., A.M.-J. and A.M.; supervision, A.E.-V. and A.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** The APC was funded by Tecnológico de Monterrey through the grants for scientific papers publication fund.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** A.E.-V., A.M., and R.O.-C. wish to acknowledge the SNI for the distinction of its membership and the stipend received. A.E.-V. and R.O.-C. express their gratitude to the Instituto de Ingeniería of the Universidad Veracruzana (UV). A.M. thanks to Laboratorio de Supercomputo del Bajío and Hasso Plattner Institute for the supercomputer resources received.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. *Mineral Commodity Summaries 2021*; Mineral Commodity Summaries: Reston, VA, USA, 2021; p. 200.
2. Sastri, V.S. *Green Corrosion Inhibitors: Theory and Practice*; John Wiley and Sons: Hoboken, NJ, USA, 2011.
3. Shamnamol, G.K.; Sreelakshmi, K.P.; Ajith, G.; Jacob, J.M. Effective Utilization of Drugs as Green Corrosion Inhibitor—A Review. In *AIP Conference Proceedings*; AIP Publishing LLC.: Puducherry, India, 2020; p. 070006.

4. Sharma, S.; Kumar, A. Recent Advances in Metallic Corrosion Inhibition: A Review. *J. Mol. Liq.* **2021**, *322*, 114862. [[CrossRef](#)]
5. Nduma, R.C.; Fayomi, O.S.I.; Nkiko, M.O.; Inegbenebor, A.O.; Udoye, N.E.; Onyisi, O.; Sanni, O.; Fayomi, J. Review of Metal Protection Techniques and Application of Drugs as Corrosion Inhibitors on Metals. *IOP Conf. Ser. Mater. Sci. Eng.* **2021**, *1107*, 012023. [[CrossRef](#)]
6. Elsaoud, A.A.; Mabrouk, E.M.; Seyam, D.F.; El-Etre, A. Recyclization of Expired Megavit Zinc (MZ) Drug as Metallic Corrosion Inhibitor for Copper Alloy C10100 in Nitric Acid Solution. *J. Bio-Tribo-Corros.* **2021**, *7*, 64. [[CrossRef](#)]
7. Li, H.; Zhang, S.; Tan, B.; Qiang, Y.; Li, W.; Chen, S.; Guo, L. Investigation of Losartan Potassium as an Eco-Friendly Corrosion Inhibitor for Copper in 0.5 M H<sub>2</sub>SO<sub>4</sub>. *J. Mol. Liq.* **2020**, *305*, 112789. [[CrossRef](#)]
8. Tasić, Ž.Z.; Petrović Mihajlović, M.B.; Radovanović, M.B.; Simonović, A.T.; Antonijević, M.M. Experimental and Theoretical Studies of Paracetamol as a Copper Corrosion Inhibitor. *J. Mol. Liq.* **2021**, *327*, 114817. [[CrossRef](#)]
9. Fateh, A.; Aliofkhaezrai, M.; Rezvanian, A.R. Review of Corrosive Environments for Copper and Its Corrosion Inhibitors. *Arab. J. Chem.* **2017**, *13*, 481–544. [[CrossRef](#)]
10. Su, P. Expired Drug Theophylline as Potential Corrosion Inhibitor for 7075 Aluminium Alloy in 1M NaOH Solution. *Int. J. Electrochem. Sci.* **2020**, *15*, 1412–1425. [[CrossRef](#)]
11. Chaubey, N.; Singh, V.K.; Quraishi, M.A. Papaya Peel Extract as Potential Corrosion Inhibitor for Aluminium Alloy in 1 M HCl: Electrochemical and Quantum Chemical Study. *Ain. Shams. Eng. J.* **2018**, *9*, 1131–1140. [[CrossRef](#)]
12. Nathiya, R.S.; Perumal, S.; Murugesan, V.; Raj, V. Evaluation of Extracts of Borassus Flabellifer Dust as Green Inhibitors for Aluminium Corrosion in Acidic Media. *Mater. Sci. Semicond. Process.* **2019**, *104*, 104674. [[CrossRef](#)]
13. Koch, G.; Varney, J.; Thompson, N.; Moghissi, O.; Gould, M.; Payer, J. International Measures of Prevention, Application, and Economics of Corrosion Technologies Study. *NACE Int.* **2016**, *216*, 2–3.
14. Miralrio, A.; Espinoza Vázquez, A. Plant Extracts as Green Corrosion Inhibitors for Different Metal Surfaces and Corrosive Media: A Review. *Processes* **2020**, *8*, 942. [[CrossRef](#)]
15. Kadhim, A.; Al-Amiery, A.A. Corrosion Inhibitors. A Review. *Int. J. Corros. Scale Inhib.* **2021**, *10*, 54–67. [[CrossRef](#)]
16. Papavinasam, S. Corrosion Inhibitors. In *Uhlig's Corrosion Handbook*; Revie, R.W., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2011; pp. 1021–1032.
17. Salleh, S.Z.; Yusoff, A.H.; Zakaria, S.K.; Taib, M.A.A.; Abu Seman, A.; Masri, M.N.; Mohamad, M.; Mamat, S.; Ahmad Sobri, S.; Ali, A.; et al. Plant Extracts as Green Corrosion Inhibitor for Ferrous Metal Alloys: A Review. *J. Clean. Prod.* **2021**, *304*, 127030. [[CrossRef](#)]
18. Majd, M.T.; Ramezanzadeh, M.; Bahlakeh, G.; Ramezanzadeh, B. Probing Molecular Adsorption/Interactions and Anti-Corrosion Performance of Poppy Extract in Acidic Environments. *J. Mol. Liq.* **2020**, *304*, 112750. [[CrossRef](#)]
19. Kalaiselvi, K.; Chung, I.-M.; Kim, S.-H.; Prabakaran, M. Corrosion Resistance of Mild Steel in Sulphuric Acid Solution by *Coreopsis Tinctoria* Extract: Electrochemical and Surface Studies. *Anti-Corros. Methods Mater.* **2018**, *65*, 408–416. [[CrossRef](#)]
20. Saxena, A.; Prasad, D.; Haldhar, R.; Singh, G.; Kumar, A. Use of Saraca Ashoka Extract as Green Corrosion Inhibitor for Mild Steel in 0.5 M H<sub>2</sub>SO<sub>4</sub>. *J. Mol. Liq.* **2018**, *258*, 89–97. [[CrossRef](#)]
21. Aysel, B.; Meltem, D. The Use of Papaver Somniferum L. Plant Extract as Corrosion Inhibitor. *Prot. Met. Phys. Chem. Surf.* **2019**, *55*, 1182–1194. [[CrossRef](#)]
22. Alrefaee, S.H.; Rhee, K.Y.; Verma, C.; Quraishi, M.A.; Ebeso, E.E. Challenges and Advantages of Using Plant Extract as Inhibitors in Modern Corrosion Inhibition Systems: Recent Advancements. *J. Mol. Liq.* **2021**, *321*, 114666. [[CrossRef](#)]
23. Baitule, P.K.; Victoria, S.N.; Manivannan, R. Review on Assessment of Corrosion of Mild Steel in Alkaline Environment by Using Plant Extract. *IOP Conf. Ser. Mater. Sci. Eng.* **2021**, *1057*, 012012. [[CrossRef](#)]
24. Espinoza-Vázquez, A.; Rodríguez-Gómez, F.J.; Negrón-Silva, G.E.; González-Olvera, R.; Ángeles-Beltrán, D.; Palomar-Pardavé, M.; Miralrio, A.; Castro, M. Fluconazole and Fragments as Corrosion Inhibitors of API 5L X52 Steel Immersed in 1M HCl. *Corros. Sci.* **2020**, *174*, 108853. [[CrossRef](#)]
25. Rodríguez-Gómez, F.J.; Valdelamar, M.P.; Vazquez, A.E.; Del Valle Perez, P.; Mata, R.; Miralrio, A.; Castro, M. Mycophenolic Acid as a Corrosion Inhibitor of Carbon Steel in 3% Wt. NaCl Solution. An Experimental and Theoretical Study. *J. Mol. Struct.* **2019**, *1183*, 168–181. [[CrossRef](#)]
26. Haruna, K.; Saleh, T.A.; Quraishi, M.A. Expired Metformin Drug as Green Corrosion Inhibitor for Simulated Oil/Gas Well Acidizing Environment. *J. Mol. Liq.* **2020**, *315*, 113716. [[CrossRef](#)]
27. Ma, X. Electrochemical Studies of Expired Drug (Formoterol) as Oilfield Corrosion Inhibitor for Mild Steel in H<sub>2</sub>SO<sub>4</sub> Media. *Int. J. Electrochem. Sci.* **2020**, *15*, 1964–1981. [[CrossRef](#)]
28. Farahati, R.; Mousavi-Khoshdel, S.M.; Ghaffarinejad, A.; Behzadi, H. Experimental and Computational Study of Penicillamine Drug and Cysteine as Water-Soluble Green Corrosion Inhibitors of Mild Steel. *Prog. Org. Coat.* **2020**, *142*, 105567. [[CrossRef](#)]
29. Abdollahi, F.; Foroughi, M.M.; Shahidi Zandi, M.; Kazemipour, M. Electrochemical Investigation of Meloxicam Drug as a Corrosion Inhibitor for Mild Steel in Hydrochloric and Sulfuric Acid Solutions. *Prog. Color Color. Coat.* **2020**, *13*, 155–165. [[CrossRef](#)]
30. El-Haddad, M.N.; Fouda, A.E.S. Evaluation of Curam Drug as an Ecofriendly Corrosion Inhibitor for Protection of Stainless Steel-304 in Hydrochloric Acid Solution: Chemical, Electrochemical, and Surface Morphology Studies. *J. Chin. Chem. Soc.* **2021**, *68*, 826–836. [[CrossRef](#)]
31. Gece, G. Drugs: A Review of Promising Novel Corrosion Inhibitors. *Corros. Sci.* **2011**, *53*, 3873–3898. [[CrossRef](#)]

32. Vaszilcsin, N.; Ordodi, V.; Borza, A. Corrosion Inhibitors from Expired Drugs. *Int. J. Pharm.* **2012**, *431*, 241–244. [[CrossRef](#)]
33. Verma, C.; Ebenso, E.E.; Bahadur, I.; Quraishi, M.A. An Overview on Plant Extracts as Environmental Sustainable and Green Corrosion Inhibitors for Metals and Alloys in Aggressive Corrosive Media. *J. Mol. Liq.* **2018**, *266*, 577–590. [[CrossRef](#)]
34. Neves, B.J.; Braga, R.C.; Melo-Filho, C.C.; Moreira-Filho, J.T.; Muratov, E.N.; Andrade, C.H. QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery. *Front. Pharmacol.* **2018**, *9*, 1275. [[CrossRef](#)]
35. Selassie, C.; Verma, R.P. History of Quantitative Structure-Activity Relationships. In *Burger's Medicinal Chemistry and Drug Discovery*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2010; p. bmc001.
36. Hu, S.; Chen, P.; Gu, P.; Wang, B. A Deep Learning-Based Chemical System for QSAR Prediction. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 3020–3028. [[CrossRef](#)] [[PubMed](#)]
37. Toropov, A.A.; Toropova, A.P. QSPR/QSAR: State-of-Art, Weirdness, the Future. *Molecules* **2020**, *25*, 1292. [[CrossRef](#)] [[PubMed](#)]
38. Khaled, K.F. Modeling Corrosion Inhibition of Iron in Acid Medium by Genetic Function Approximation Method: A QSAR Model. *Corros. Sci.* **2011**, *53*, 3457–3465. [[CrossRef](#)]
39. Keshavarz, M.H.; Esmailpour, K.; Golikand, A.N.; Shirazi, Z. Simple Approach to Predict Corrosion Inhibition Efficiency of Imidazole and Benzimidazole Derivatives as Well as Linear Organic Compounds Containing Several Polar Functional Groups: Simple Approach to Predict Corrosion Inhibition. *Zeitschrift für anorganische und allgemeine Chemie* **2016**, *642*, 906–913. [[CrossRef](#)]
40. Liu, Y.; Guo, Y.; Wu, W.; Xiong, Y.; Sun, C.; Yuan, L.; Li, M. A Machine Learning-Based QSAR Model for Benzimidazole Derivatives as Corrosion Inhibitors by Incorporating Comprehensive Feature Selection. *Interdiscip. Sci. Comput. Life Sci.* **2019**, *11*, 738–747. [[CrossRef](#)]
41. Ser, C.T.; Žuvela, P.; Wong, M.W. Prediction of Corrosion Inhibition Efficiency of Pyridines and Quinolines on an Iron Surface Using Machine Learning-Powered Quantitative Structure-Property Relationships. *Appl. Surf. Sci.* **2020**, *512*, 145612. [[CrossRef](#)]
42. Pearson, R.G. Hard and Soft Acids and Bases. *J. Am. Chem. Soc.* **1963**, *85*, 3533–3539. [[CrossRef](#)]
43. Chattaraj, P.K.; Giri, S. Electrophilicity Index within a Conceptual DFT Framework. *Annu. Rep. Sect. C Phys. Chem.* **2009**, *105*, 13. [[CrossRef](#)]
44. Lukovits, I.; Kálmán, E.; Zucchi, F. Corrosion Inhibitors—Correlation between Electronic Structure and Efficiency. *CORROSION* **2001**, *57*, 3–8. [[CrossRef](#)]
45. Bulat, F.A.; Toro-Labbé, A.; Brinck, T.; Murray, J.S.; Politzer, P. Quantitative Analysis of Molecular Surfaces: Areas, Volumes, Electrostatic Potentials and Average Local Ionization Energies. *J. Mol. Model.* **2010**, *16*, 1679–1691. [[CrossRef](#)]
46. Aradi, B.; Hourahine, B.; Frauenheim, T. DFTB+, a Sparse Matrix-Based Implementation of the DFTB Method. *J. Phys. Chem. A* **2007**, *111*, 5678–5684. [[CrossRef](#)] [[PubMed](#)]
47. Hourahine, B.; Aradi, B.; Blum, V.; Bonafé, F.; Buccheri, A.; Camacho, C.; Cevallos, C.; Deshayé, M.Y.; Dumitrică, T.; Dominguez, A.; et al. DFTB+, a Software Package for Efficient Approximate Density Functional Theory Based Atomistic Simulations. *J. Chem. Phys.* **2020**, *152*, 124101. [[CrossRef](#)] [[PubMed](#)]
48. Billings, S.A. *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*; John Wiley & Sons, Ltd.: Chichester, UK, 2013.
49. Yassin, I.M.; Abdul Khalid, M.F.; Herman, S.H.; Pasya, I.; Wahab, N.A.; Awang, Z. Multi-Layer Perceptron (MLP)-Based Nonlinear Auto-Regressive with Exogenous Inputs (NARX) Stock Forecasting Model. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2017**, *7*, 1098. [[CrossRef](#)]
50. Chadalawada, J.; Havlicek, V.; Babovic, V. A Genetic Programming Approach to System Identification of Rainfall-Runoff Models. *Water Resour. Manag.* **2017**, *31*, 3975–3992. [[CrossRef](#)]
51. Cheah, K.W.; Ahmad, N.A. Fuzzy Recursive Least-Squares Approach in Speech System Identification: A Transformed Domain LPC Model. *Int. J. Electr. Comput. Eng. IJECE* **2017**, *7*, 842. [[CrossRef](#)]
52. Pal, S.K.; Wang, P.P. *Genetic Algorithms for Pattern Recognition*, 1st ed.; Pal, S.K., Wang, P.P., Eds.; CRC Press: Boca Raton, FL, USA, 2017.
53. Harris, J.; Arthurs, F.; Henrickson, J.V.; Valasek, J. Aircraft System Identification Using Artificial Neural Networks with Flight Test Data. In Proceedings of the 2016 International Conference on Unmanned Aircraft Systems (ICUAS, IEEE), Arlington, VA, USA, 7–10 June 2016; pp. 679–688.
54. Gao, Y.; Liu, S.; Li, F.; Liu, Z. Fault Detection and Diagnosis Method for Cooling Dehumidifier Based on LS-SVM NARX Model. *Int. J. Refrig.* **2016**, *61*, 69–81. [[CrossRef](#)]
55. Chen, S.; Billings, S.A.; Luo, W. Orthogonal Least Squares Methods and Their Application to Non-Linear System Identification. *Int. J. Control* **1989**, *50*, 1873–1896. [[CrossRef](#)]
56. Billings, S.A.; Korenberg, M.J.; Chen, S. Identification of Non-Linear Output-Affine Systems Using an Orthogonal Least-Squares Algorithm. *Int. J. Syst. Sci.* **1988**, *19*, 1559–1568. [[CrossRef](#)]
57. Korenberg, M.; Billings, S.A.; Liu, Y.P.; McILROY, P.J. Orthogonal Parameter Estimation Algorithm for Non-Linear Stochastic Systems. *Int. J. Control* **1988**, *48*, 193–210. [[CrossRef](#)]
58. Yoshida, H.; Kumar, S. Development of ARX Model Based Off-Line FDD Technique for Energy Efficient Buildings. *Renew. Energy* **2001**, *22*, 53–59. [[CrossRef](#)]
59. Antonucci, D.; Filippi Oberegger, U.; Pasut, W.; Gasparella, A. Building Performance Evaluation through a Novel Feature Selection Algorithm for Automated Arx Model Identification Procedures. *Energy Build.* **2017**, *150*, 432–446. [[CrossRef](#)]

60. Romero-Ugalde, H.M.; Garnotel, M.; Doron, M.; Jallon, P.; Charpentier, G.; Franc, S.; Huneker, E.; Simon, C.; Bonnet, S. ARX Model for Interstitial Glucose Prediction during and after Physical Activities. *Control Eng. Pract.* **2019**, *90*, 321–330. [CrossRef]
61. Ayala Solares, J.R.; Wei, H.-L.; Boynton, R.J.; Walker, S.N.; Billings, S.A. Modeling and Prediction of Global Magnetic Disturbance in Near-Earth Space: A Case Study for  $K_p$  Index Using NARX Models: MODELING AND PREDICTION OF  $K_p$  INDEX. *Space Weather* **2016**, *14*, 899–916. [CrossRef]
62. Ayala-Solares, J.R.; Wei, H.-L.; Bigg, G.R. The Variability of the Atlantic Meridional Circulation since 1980, as Hindcast by a Data-Driven Nonlinear Systems Model. *Acta Geophys.* **2018**, *66*, 683–695. [CrossRef]
63. Rashid, M.T.; Frasca, M.; Ali, A.A.; Ali, R.S.; Fortuna, L.; Xibilia, M.G. Nonlinear Model Identification for Artemia Population Motion. *Nonlinear Dyn.* **2012**, *69*, 2237–2243. [CrossRef]
64. Luna, C.; Coca, D. An Empirical Model of Drosophila Photoreceptor-LMC Network. *BMC Neurosci.* **2015**, *16*, 47. [CrossRef]
65. Li, Y.; Cui, W.-G.; Guo, Y.-Z.; Huang, T.; Yang, X.-F.; Wei, H.-L. Time-Varying System Identification Using an Ultra-Orthogonal Forward Regression and Multiwavelet Basis Functions With Applications to EEG. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *29*, 2960–2972. [CrossRef] [PubMed]
66. Berk, R.A. *Statistical Learning from a Regression Perspective*; Springer Texts in Statistics; Springer International Publishing: Cham, Switzerland, 2020.
67. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer Texts in Statistics; Springer: New York, NY, USA, 2013; Volume 103.
68. Chen, Y.; Elenee Argentinis, J.; Weber, G. IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research. *Clin. Ther.* **2016**, *38*, 688–701. [CrossRef]
69. Ogunyemi, B.T.; Latona, D.F.; Adejoro, I.A. Molecular Modeling and Quantitative Structure–Property Relationships (QSPRs) of Purine Derivatives as Corrosion Inhibitor in Acid Medium. *Sci. Afr.* **2020**, *8*, e00336. [CrossRef]
70. Yesudass, S.; Olasunkanmi, L.O.; Bahadur, I.; Kabanda, M.M.; Obot, I.B.; Ebenso, E.E. Experimental and Theoretical Studies on Some Selected Ionic Liquids with Different Cations/Anions as Corrosion Inhibitors for Mild Steel in Acidic Medium. *J. Taiwan Inst. Chem. Eng.* **2016**, *64*, 252–268. [CrossRef]
71. Jalgham, R.T.T. Theoretical, Monte Carlo Simulations and QSAR Studies on Some Triazole Derivatives as Corrosion Inhibitors for Mild Steel in 1 M HCl. *ES Energy Environ.* **2021**, *13*, 37–49. [CrossRef]
72. Obot, I.B. Experimental, DFT and QSAR Models for the Discovery of New Pyrazines Corrosion Inhibitors for Steel in Oilfield Acidizing Environment. *Int. J. Electrochem. Sci.* **2020**, *15*, 9066–9080. [CrossRef]
73. Shukla, S.K.; Singh, A.K.; Ahamad, I.; Quraishi, M.A. Streptomycin: A Commercially Available Drug as Corrosion Inhibitor for Mild Steel in Hydrochloric Acid Solution. *Mater. Lett.* **2009**, *63*, 819–822. [CrossRef]
74. Ahamad, I.; Prasad, R.; Quraishi, M.A. Experimental and Theoretical Investigations of Adsorption of Fexofenadine at Mild Steel/Hydrochloric Acid Interface as Corrosion Inhibitor. *J. Solid State Electrochem.* **2010**, *14*, 2095–2105. [CrossRef]
75. Ebenso, E.E.; Obot, I.B.; Murulana, L.C. Quinoline and Its Derivatives as Effective Corrosion Inhibitors for Mild Steel in Acidic Medium. *Int. J. Electrochem. Sci.* **2010**, *5*, 1574–1586.
76. Martínez Jiménez, P.; Espinoza Vázquez, A.; Rodríguez Gómez, F.J. Evaluation of N,N-Dimethylformamide as Corrosion Inhibitor on API 5L X70 Using 3% NaCl. *MRS Adv.* **2019**, *4*, 2391–2399. [CrossRef]
77. IBM Watson Studio—AutoML—IBM AutoAI. Available online: <https://www.ibm.com/in-en/cloud/watson-studio/autoai> (accessed on 9 February 2022).
78. Quadri, T.W.; Olasunkanmi, L.O.; Akpan, E.D.; Fayemi, O.E.; Lee, H.-S.; Lgaz, H.; Verma, C.; Guo, L.; Kaya, S.; Ebenso, E.E. Development of QSAR-Based (MLR/ANN) Predictive Models for Effective Design of Pyridazine Corrosion Inhibitors. *Mater. Today Commun.* **2022**, *30*, 103163. [CrossRef]
79. Li, L.; Zhang, X.; Gong, S.; Zhao, H.; Bai, Y.; Li, Q.; Ji, L. The Discussion of Descriptors for the QSAR Model and Molecular Dynamics Simulation of Benzimidazole Derivatives as Corrosion Inhibitors. *Corros. Sci.* **2015**, *99*, 76–88. [CrossRef]
80. Al-Fakih, A.M.; Algamal, Z.Y.; Lee, M.H.; Abdallah, H.H.; Maarof, H.; Aziz, M. Quantitative Structure-Activity Relationship Model for Prediction Study of Corrosion Inhibition Efficiency Using Two-Stage Sparse Multiple Linear Regression: QSAR Study Using Two-Stage SMLR. *J. Chemom.* **2016**, *30*, 361–368. [CrossRef]
81. Gu, Y.; Yang, Y.; Dewald, J.P.A.; van der Helm, F.C.T.; Schouten, A.C.; Wei, H.-L. Nonlinear Modeling of Cortical Responses to Mechanical Wrist Perturbations Using the NARMAX Method. *IEEE Trans. Biomed. Eng.* **2021**, *68*, 948–958. [CrossRef]
82. Boynton, R.J.; Balikhin, M.A.; Billings, S.A.; Sharma, A.S.; Amariutei, O.A. Data Derived NARMAX Dst Model. *Ann. Geophys.* **2011**, *29*, 965–971. [CrossRef]
83. Pisoni, E.; Farina, M.; Carnevale, C.; Piroddi, L. Forecasting Peak Air Pollution Levels Using NARX Models. *Eng. Appl. Artif. Intell.* **2009**, *22*, 593–602. [CrossRef]
84. Boynton, R.; Balikhin, M.; Wei, H.-L.; Lang, Z.-Q. Applications of NARMAX in Space Weather. In *Machine Learning Techniques for Space Weather*; Elsevier: Amsterdam, the Netherlands, 2018; pp. 203–236.
85. Kukula-Koch, W.A.; Widelski, J. Alkaloids. In *Pharmacognosy*; Elsevier: Amsterdam, the Netherlands, 2017; pp. 163–198.
86. Nunomura, N.; Sunada, S. Density Functional Theory Study of The Interaction Of Hydroxyl Groups With Iron Surface. *Arch. Metall. Mater.* **2015**, *60*, 931–933. [CrossRef]
87. Lepetit, C.; Maraval, V.; Canac, Y.; Chauvin, R. On the Nature of the Dative Bond: Coordination to Metals and beyond. The Carbon Case. *Coord. Chem. Rev.* **2016**, *308*, 59–75. [CrossRef]

88. Qiang, Y.; Guo, L.; Li, H.; Lan, X. Fabrication of Environmentally Friendly Losartan Potassium Film for Corrosion Inhibition of Mild Steel in HCl Medium. *Chem. Eng. J.* **2021**, *406*, 126863. [[CrossRef](#)]
89. Anadebe, V.C.; Nnaji, P.C.; Onukwuli, O.D.; Okafor, N.A.; Abeng, F.E.; Chukwuike, V.I.; Okoye, C.C.; Udoh, I.I.; Chidiebere, M.A.; Guo, L.; et al. Multidimensional Insight into the Corrosion Inhibition of Salbutamol Drug Molecule on Mild Steel in Oilfield Acidizing Fluid: Experimental and Computer Aided Modeling Approach. *J. Mol. Liq.* **2022**, *349*, 118482. [[CrossRef](#)]
90. Abeng, F.E.; Anadebe, V.C.; Idim, V.D.; Edim, M.M. Anti-Corrosion Behaviour of Expired Tobramycin Drug on Carbon Steel in Acidic Medium. *S. Afr. J. Chem.* **2020**, *73*, 125–130.
91. Onyechu, I.B.; Abdel-Azeim, S.; Chauhan, D.S.; Quraishi, M.A. Electrochemical and Computational Insights on the Application of Expired Metformin Drug as a Novel Inhibitor for the Sweet Corrosion of C1018 Steel. *ACS Omega* **2021**, *6*, 65–76. [[CrossRef](#)]
92. Abeng, F.E.; Ikpi, M.E.; Anadebe, V.C.; Emori, W. Metolazone Compound as Corrosion Inhibitor for API 5L X-52 Steel in Hydrochloric Acid Solution. *Bull. Chem. Soc. Ethiop.* **2020**, *34*, 407–418. [[CrossRef](#)]
93. Ikpi, M.E.; Abeng, F.E. Electrochemical and Quantum Chemical Investigation on Adsorption of Nifedipine as Corrosion Inhibitor at API 5L X-52 Steel/HCL ACID Interface. *Arch. Metall. Mater.* **2020**, *65*, 125–131. [[CrossRef](#)]
94. Anadebe, V.C.; Onukwuli, O.D.; Abeng, F.E.; Okafor, N.A.; Ezeugo, J.O.; Okoye, C.C. Electrochemical-Kinetics, MD-Simulation and Multi-Input Single-Output (MISO) Modeling Using Adaptive Neuro-Fuzzy Inference System (ANFIS) Prediction for Dexamethasone Drug as Eco-Friendly Corrosion Inhibitor for Mild Steel in 2 M HCl Electrolyte. *J. Taiwan Inst. Chem. Eng.* **2020**, *115*, 251–265. [[CrossRef](#)] [[PubMed](#)]
95. Gholamhosseinzadeh, M.R.; Aghaie, H.; Zandi, M.S.; Giahi, M. Rosuvastatin Drug as a Green and Effective Inhibitor for Corrosion of Mild Steel in HCl and H<sub>2</sub>SO<sub>4</sub> Solutions. *J. Mater. Res. Technol.* **2019**, *8*, 5314–5324. [[CrossRef](#)]
96. Geethamani, P.; Narmatha, M.; Dhanalakshmi, R.; Aejitha, S.; Kasthuri, P.K. Corrosion Inhibition and Adsorption Properties of Mild Steel in 1 M Hydrochloric Acid Medium by Expired Ambroxol Drug. *J. Bio-Tribo-Corros.* **2019**, *5*, 16. [[CrossRef](#)]
97. Singh, P.; Chauhan, D.S.; Chauhan, S.S.; Singh, G.; Quraishi, M.A. Chemically Modified Expired Dapsone Drug as Environmentally Benign Corrosion Inhibitor for Mild Steel in Sulphuric Acid Useful for Industrial Pickling Process. *J. Mol. Liq.* **2019**, *286*, 110903. [[CrossRef](#)]
98. Abeng, F.E.; Ikpi, M.E.; Ushie, O.A.; Anadebe, V.C.; Nyong, B.E.; Obeten, M.E.; Okafor, N.A.; Chukwuike, V.I.; Nkom, P.Y. Insight into Corrosion Inhibition Mechanism of Carbon Steel in 2 M HCl Electrolyte by Eco-Friendly Based Pharmaceutical Drugs. *Chem. Data Collect.* **2021**, *34*, 100722. [[CrossRef](#)]
99. Dahiya, S.; Pahuja, P.; Lgaz, H.; Chung, I.-M.; Lata, S. Advanced Quantum Chemical and Electrochemical Analysis of Ravage Drugs for Corrosion Inhibition of Mild Steel. *J. Adhes. Sci. Technol.* **2019**, *33*, 1066–1089. [[CrossRef](#)]
100. Saraswat, V.; Yadav, M. Improved Corrosion Resistant Performance of Mild Steel under Acid Environment by Novel Carbon Dots as Green Corrosion Inhibitor. *Colloids Surf. Physicochem. Eng. Asp.* **2021**, *627*, 127172. [[CrossRef](#)]
101. Verma, C.; Abdellattif, M.H.; Alfantazi, A.; Quraishi, M.A. N-Heterocycle Compounds as Aqueous Phase Corrosion Inhibitors: A Robust, Effective and Economic Substitute. *J. Mol. Liq.* **2021**, *340*, 117211. [[CrossRef](#)]
102. Rodrigues, F.A.d.S.; Gonçalves, Y.M.H.; Horta, B.A.C.; Santos, I.d.S.; Silva, B.V.; D'Elia, E. Experimental and Theoretical Studies of Isonitrosoacetanilides Derivatives as Corrosion Inhibitors for Mild Steel in 1 Mol L<sup>-1</sup> HCl. *J. Mol. Struct.* **2021**, *1245*, 131256. [[CrossRef](#)]
103. Berdimurodov, E.; Kholikov, A.; Akbarov, K.; Guo, L. Experimental and Theoretical Assessment of New and Eco-Friendly Thioglycoluril Derivative as an Effective Corrosion Inhibitor of St2 Steel in the Aggressive Hydrochloric Acid with Sulfate Ions. *J. Mol. Liq.* **2021**, *335*, 116168. [[CrossRef](#)]
104. Srivastava, V.; Salman, M.; Chauhan, D.S.; Abdel-Azeim, S.; Quraishi, M.A. (E)-2-Styryl-1H-Benzo[d]Imidazole as Novel Green Corrosion Inhibitor for Carbon Steel: Experimental and Computational Approach. *J. Mol. Liq.* **2021**, *324*, 115010. [[CrossRef](#)]
105. Espinoza Vázquez, A.; López Reséndiz, L.A.; Figueroa, I.A.; Rodríguez Gómez, F.J.; Figueroa, M.; Ángeles Beltrán, D.; Castro, M.; Miralrio, A. Corrosion Inhibition Assessment on API 5L X70 Steel by Preussomerin G Immersed in Saline and Saline Acetic. *J. Adhes. Sci. Technol.* **2021**, *35*, 873–899. [[CrossRef](#)]
106. Toghan, A.; Gadow, H.S.; Dardeer, H.M.; Elabbasy, H.M. New Promising Halogenated Cyclic Imides Derivatives as Potential Corrosion Inhibitors for Carbon Steel in Hydrochloric Acid Solution. *J. Mol. Liq.* **2021**, *325*, 115136. [[CrossRef](#)]