

Relationships between the Photodegradation Reaction Rate and Structural Properties of Polymer Systems

Published as part of *The Journal of Physical Chemistry B* virtual special issue "Machine Learning in Physical Chemistry Volume 2".

Meade Erickson, Gerardo Casañola-Martin, Yulun Han, Bakhtiyor Rasulev,* and Dmitri Kilin*



Cite This: *J. Phys. Chem. B* 2024, 128, 2190–2200



Read Online

ACCESS |



Metrics & More



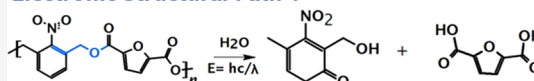
Article Recommendations



Supporting Information

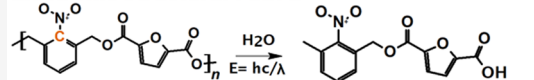
ABSTRACT: The development of reusable polymeric materials inspires an attempt to combine renewable biomass with upcycling to form a biorenewable closed system. It has been reported that 2,5-furandicarboxylic acid (FDCA) can be recovered for recycling when incorporated as monomers into photodegradable polymeric systems. Here, we develop a procedure to better understand the photodegradation reactions combining density functional theory (DFT) based time-dependent excited-state molecular dynamics (TDESMD) studies with machine learning-based quantitative structure–activity relationships (QSAR) methodology. This procedure allows for the unveiling of hidden structural features between active orbitals that affect the rate of photodegradation and is coined InfoTDESMD. Findings show that electrotopological features are influential factors affecting the rate of photodegradation in differing environments. Additionally, statistical validations and knowledge-based analysis of descriptors are conducted to further understand the structural features' influence on the rate of photodegradation of polymeric materials.

Electronic Structural Path 1



Differing Rates of Degradation!

Electronic Structural Path 2



Differing Degradation Products!

1. INTRODUCTION

The application of green chemistry principles is practiced to ensure the responsible use of polymeric materials during the product life cycle and disposal after expected use due to the large amount of polymeric waste accumulating globally.^{1,2} To reduce the pollution of such polymeric wastes, a combination of green chemistry principles and investigative methodologies is necessary.^{3,4} One such combination of green chemistry principles and methodologies is the use of upcycling of renewable biobased compounds where further analysis and experimentation are executed through in-silico methods.^{5–8} Specifically, in-silico approaches can be applied to better determine trends and analyze underutilized data for explanation and guidance for future experimental investigations. This trend analysis can assist in the future processing of polymers throughout, and after, their life cycle in effectively reducing plastic pollutants and benefiting global polymeric waste reduction.

Studies show that the fructose-based compound, 2,5-furandicarboxylic acid (FDCA), has been incorporated as a building block into a photodegradable polymeric system.⁵ This fructose-based building block was able to be degraded via UV-irradiation, later recovered, and finally reused in new polymeric systems after original use.⁵ Recently, the application of in-silico, time-dependent excited-state molecular dynamics (TDESMD), methodology was executed to investigate the

photodegradation process of the FDCA compound in both vacuum and aqueous environments.^{6,9} This allowed for a further understanding of the photodegradation pathways of polymeric chains containing FDCA via interactions with solvent systems. One finds the recovery of biobased building blocks, which corroborated with experimental findings.^{5,6}

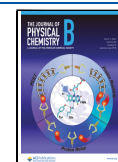
To further enhance the understanding of the photoreactions, here, we focus on relationships between electronic structures and structural properties of the polymeric chains. Various electronic structural information is linked to the photoactivity of compounds and polymers, such as oscillator strength, transition energy, and electron density distribution.^{10,11} Previous studies suggested that UV irradiation influenced electron transfer by changing the electron density distribution from the nitrogen center of diethylenetriamine penta-(methylenephosphonic acid). This effectively increased radical attack during UV irradiation and increased the rate of photodegradation.¹² Brouard et al. have experimentally shown the spin-forbidden dissociation channel following

Received: October 16, 2023

Revised: January 15, 2024

Accepted: January 16, 2024

Published: February 22, 2024



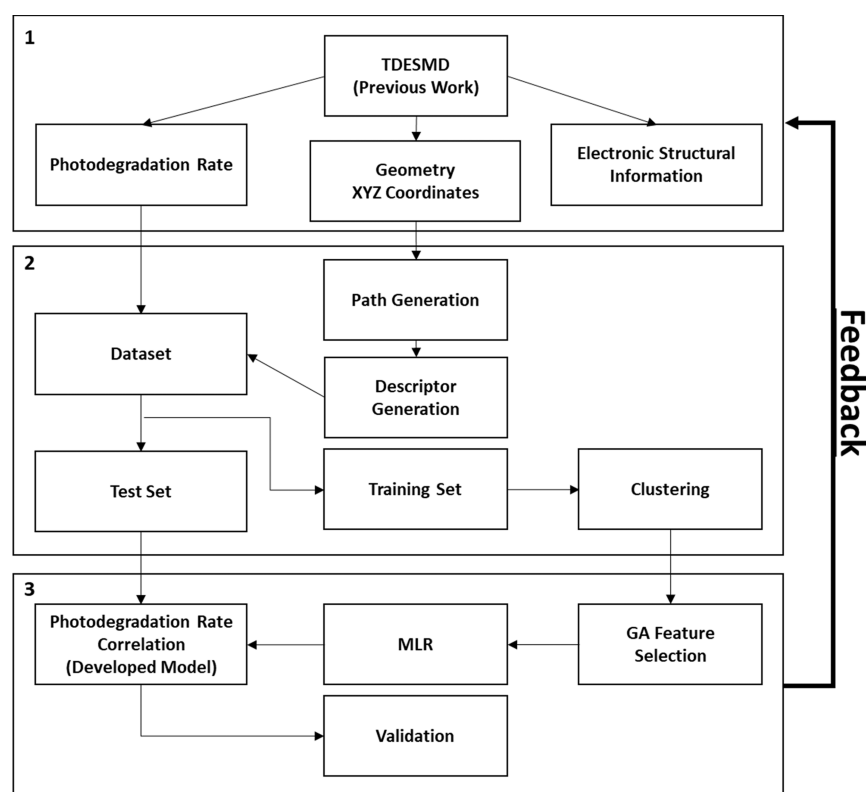


Figure 1. InfoTDESMD flowchart. Section 1 represents preprocessing and data collection. Section 2 represents cheminformatics. Section 3 represents machine learning.

absorption of a 193 nm photon,¹³ where velocity-map imaging has been used to measure spatial anisotropy parameters, photofragmentation velocity distributions, and angular momentum alignment parameters for spin–orbit states of fragments. Their findings show that the electron density distribution peaked perpendicular to the direction of bond breaking. Their findings suggest connections between electron energy distributions and photodegradation channels of N_2O .

Many other groups have used the maximum-entropy method (MEMS) in analyzing the placement of the electron density distribution along paths between atoms.^{14–16} MEM utilizes the electron density information from diffraction data by maximizing the entropy information where it can be used to analyze polymeric systems.^{17–19} In one case, the electron density was distributed closer to the atoms versus the center of the bond. This suggested more ionic behavior and reduced the bond length.¹⁴ The locality of the electron density along the bond lengths, and the information it describes, suggest an influence on factors such as fluorescence efficiency and photocatalytic activity. Both, fluorescence efficiency and photocatalytic activity, relate to the ability of compounds to interact with light and in later stages further degrade due to this perturbation and interaction with light.^{15,16} Although this is a powerful methodology, the analysis is based on electron density distribution between atoms of small molecules where long-range differences in electron density distribution are not able to be accounted for. By using computational methods, it is possible to account for such long-range distributions.

Computational analysis of the electron density distribution of compounds and polymeric systems relating to photo activity has focused primarily on using ab initio methodologies.^{20–24} Alekseyev et al. investigated the control of the I^* quantum

yields of CH_3I via transition moments and vibrational states.²⁰ Their work suggests that the electron density distribution at various molecular geometries influences a more efficient excitation of specific fragmentation channels. Li et al. suggested that polymeric systems containing photochromic compounds of 1,4-diket-3, 6-diarylpyrrolo[3, 4-c] pyrroles diketopyrrolo-pyrroles (DPPs) had varying electron density distributions before and after excitation. This varying distribution showed movement from the monomer system to other portions of the compounds affecting the photostability of the material.^{22,23}

The electronic structural information on compounds is dependent on the topological structure, which can be expressed through numerical values by cheminformatics using descriptors.^{25–28} Specifically the application of quantitative structure–activity relationships (QSAR) has been used to investigate the influence that topological and electronic structural information has on photoactivity.^{29–32}

Xiao et al. applied the QSAR methodology to iodinated trihalomethanes with the response variable being the first-order rate constant of photodegradation.³⁰ In their work a three-descriptor regression model was found where the bond strength of carbon–halogen is influential in the rate-determining step. Again, Xiao et al. investigated the photolysis rates of iodoacids where electronic and steric effects of all substituents were used as descriptors and developed a reasonable predictive model with a small data set.³¹ The data correlated well even with fewer than five compounds in the training set. Recently, Wan et al. used QSAR to determine properties that affected the photodegradation of halogenated parabens under simulated sunlight.³² In their work, it was found that C–X bond dissociation energies as well as both electronic and steric effects of halogen substituents. In all cases,

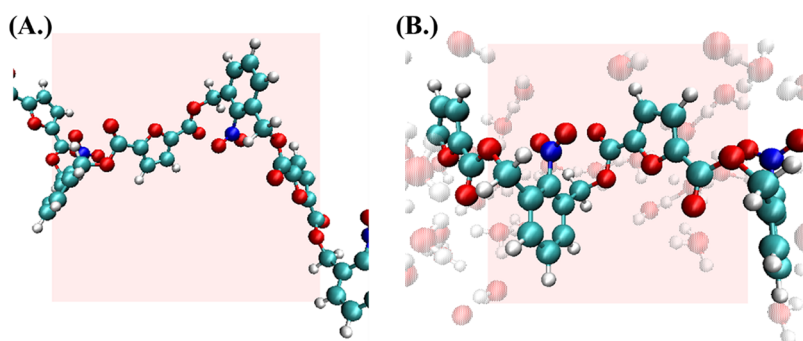


Figure 2. View of the atomistic structures of polymeric chains in (A.) Vacuum system and (B.) Aqueous environment under periodic boundary conditions. The shaded areas indicate the simulation cells. The spheres represent atoms where H (white), C (cyan), N (blue), and O (red).

the use of structural information combined with electronic structural information had combinatorial effects on describing the properties that describe the photodegradation of various compounds. Though none of these studies had computational response values of photodegradation, it is hypothesized that regression models can be developed to explain the connections between electronic and structural information toward the rate of photodegradation.

Here, we combine TDESMD and QSAR (InfoTDESMD) toward the investigation of the photodegradation of FDCA-based polymeric chains. This allows for understanding the relationships between electronic structures and topological information. Specifically, TDESMD calculations were carried out to simulate the photodegradation.³³ Further analysis was conducted using cheminformatics techniques where molecular substructures between the expectation value of the position of active orbitals were captured and used to develop machine learning-based regression models in determining topological factors influencing the rate of photodegradation. The substructures, known as paths, were captured using RDKit, and molecular descriptors were generated using Ochem.eu capabilities with a partnership with alvaDesc.^{34–36} Regression models were developed and analyzed using QSARINS cheminformatics modeling software and package to determine specific properties influencing the photodegradation rate of polymers.³⁷ The analysis of the regression models is interpreted via global interpretation and validated via both statistical and knowledge-based processes.³⁸ Additionally, the increase in mechanistic interoperability of QSAR models, in this case, due to the linearity of MLR methodology, allows for the description and application of information found via model development.

The content throughout this work is organized as follows. The Methods section showcases the computational details via AIMD, DFT, TDESMD, cheminformatics, and regression modeling methodology. The Results section displays the computational findings relating to path development and influential relationships. Regression model development is highlighted, and influential properties are described. In the discussion section, influential properties are compared and discussed compared to other works in detail. Additionally, the similarities and differences between the aqueous and vacuum environment trajectories and their connection with experimental findings are further analyzed. The Conclusions section summarizes findings and highlights the impact on understanding relationships influencing polymer deconstruction and upcycling.

2. METHODS

To illustrate the process implemented, the flowchart shown in Figure 1 is presented. There are three categories including preprocessing and data collection, cheminformatics, and machine learning. Within preprocessing and data collection, computational simulations following TDESMD methodology are recorded. Cheminformatics uses a variety of tools to gather pertinent information to further analyze electronic and topological structural information. The third category of tools is the application of machine learning where regression models are developed and used to give further insight into the photodegradation of polymeric systems.

2.1. Computational Modeling. We consider polymeric chains in vacuum and aqueous systems. The vacuum system has two FDCA-nitrobenzyl units ($C_{28}H_{18}N_2O_{14}$), as shown in Figure 2A. The lattice parameters are $x = 18.544$, $y = 14.149$, and $z = 17.768$ Å. The aqueous system contains one FDCA-nitrobenzyl unit ($C_{14}H_9N_1O_7$) surrounded by 27 H_2O molecules, as shown in Figure 2B. The lattice parameters are $x = 9.580$, $y = 9.887$, and $z = 9.030$ Å. Figure 2 was generated using Visual Molecular Dynamics (VMD) software.³⁹

2.2. Ground State Calculations and TDESMD. First-principles calculations were carried out using DFT with Perdew–Burke–Ernzerhof (PBE)⁴⁰ functional and projected augmented wave (PAW)⁴¹ potentials under periodic boundary conditions. The Kohn–Sham (KS) orbitals,^{42,43} $\phi_i^{KS}(\vec{r}, \{\vec{R}_I\})$, and energies, $\epsilon_i(\{\vec{R}_I\})$, for each nuclear configuration $\{\vec{R}_I\}$ were provided by VASP software.^{44–48} Oscillator strengths, f_{ij} , transition dipoles, $\vec{D}_{ij}$, and absorption spectra are computed on the basis of the orbitals and energies adopting the independent orbital approximations (IOAs). The optimized ground state geometries of the models were used as the starting point for TDESMD calculations.

The initial conditions for TDESMD calculations involving KS orbital pairs, HO- m and LU + n , are selected based on oscillator strengths, which show the probability of absorption of electromagnetic radiation in transitioning between energy levels in molecules.

Previous work describes the TDESMD procedure in detail.⁴⁹ Briefly, periodic excitations and de-excitations of the model are executed to simulate light-matter interaction and allow photoreactions to occur through the TDESMD process. This interaction is expressed in Fock matrix $\hat{F}$ and the total energy in eq 1,

$$\hat{F} = \hat{F}^{KS} + \hat{V}^{NA}(t) - \widehat{M} \cdot \vec{e}(t) \quad (1)$$

where $\hat{F}^{KS}$ is Fock matrix in KS orbitals, Ω is the laser field frequency, $\widehat{M}$ stands for transition dipole operator, $\vec{e}(t) = \vec{e} \cdot \cos(\Omega t)$ is the electric field of laser, and $\hat{V}^{NA}(t)$ represents nonadiabatic couplings.^{50,51}

Throughout the TDESMD methodology $|\hat{V}^{NA}| < |\widehat{M} \cdot \vec{e}|$ is assumed. The density operator $\hat{\rho}$ is used to express that the electronic degrees of freedom obey the equation of motion. $\hat{\rho}$, shown in eq 2, is in the terms of the Redfield superoperator $\widehat{\mathcal{R}}$ and Liouville–von Neumann superoperator $\widehat{\mathcal{L}}$,

$$\left(\widehat{\mathcal{L}} + \widehat{\mathcal{R}}\right)\hat{\rho} = \frac{d}{dt}\hat{\rho} = -\frac{i}{\hbar}[\hat{F}, \hat{\rho}] + \left(\frac{d\rho_{ij}}{dt}\right)_{\text{diss}} \quad (2)$$

where $\left(\frac{d\rho_{ij}}{dt}\right)_{\text{diss}}$ is the electronic dissipative transition.^{52,53}

Following the Kohn–Sham self-consistent procedure, the total density is computed, which is then used to determine the energy gradient and forces imposed on the nuclei in eq 3.

$$\frac{d^2}{dt^2}\hat{R}_i(t) = \frac{\hat{F}_i(t)}{M_i} \quad (3)$$

2.3. Rate of Photodegradation: Response Variable.

The rate (k) of photodegradation is calculated by taking the inverse of the time (t) at which the first covalent bond breakage occurs without recombination within the trajectory, as eq 4.^{54,55} Bond breakage can be defined as covalently bonded interatomic distances exceeding $|\vec{R}_i(t) - \vec{R}_j(t)| \geq 1.7$ (Å) to one another where the atoms do not return to an interatomic distance of <1.7 (Å).

$$k = \frac{1}{t} \quad (4)$$

2.4. Path Development. Steps for the topological structural path development process are listed in sequential order.

1. Molecular partial charges of the polymers were calculated from the TDESMD process where electron density distributions were calculated and visualized.⁶

2. To help simplify the analysis process, the expectation value of positions of the active orbitals (EVPAOs) were calculated from the electron density distributions. Active orbitals indicate that two orbitals, one HO- m and one LU + n , are involved in the excitation and deexcitation of the electrons. These calculations were conducted using eqs 5–7, where i represents the index of orbital $\phi_i^{KS}(\vec{r}, \{\vec{R}_I\})$, of interest and corresponding partial charge density $\rho_i(x, y, z) = |\phi_i^{KS}(\vec{r}, \{\vec{R}_I\})|^2$, $\langle X \rangle$ is the center of the orbital for the x -Cartesian projection, $\langle Y \rangle$ is the center of the orbital for the y -Cartesian projection, and $\langle Z \rangle$ is the center of the orbital for the z -Cartesian projection.

$$\langle X \rangle_i = \left\langle \int_{Z_1}^{Z_2} \int_{Y_1}^{Y_2} \int_{X_1}^{X_2} \rho_i(x, y, z) dx dy dz \times x \right\rangle \quad (5)$$

$$\langle Y \rangle_i = \left\langle \int_{Z_1}^{Z_2} \int_{Y_1}^{Y_2} \int_{X_1}^{X_2} \rho_i(x, y, z) dx dy dz \times y \right\rangle \quad (6)$$

$$\langle Z \rangle_i = \left\langle \int_{Z_1}^{Z_2} \int_{Y_1}^{Y_2} \int_{X_1}^{X_2} \rho_i(x, y, z) dx dy dz \times z \right\rangle \quad (7)$$

The EVPAOs act as singular points to which XYZ coordinate information is retained from the electronic density distributions. A 2D structure is shown in Figure 3 to represent

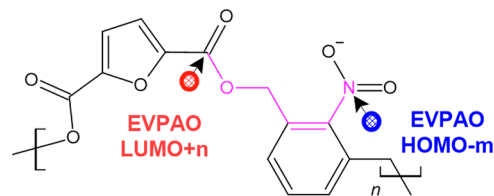


Figure 3. Example of EVPAO atom distance and path determination.

the EVPAO location, closest atoms, and path development. The blue and red circles with patterns in the center represent the HOMO- m and LUMO + n EVPAOs locations, respectively. The black arrows point toward the closest atom of the EVPAOs, determined by Euclidean distances. The pink bonds represent the path between the two closest atoms of the EVPAOs and are the structures analyzed in this work.

3. The closest non-hydrogen atoms of the FDCA-nitrobenzyl compound to the EVPAOs are selected as the two atoms at which electron densities would be most likely located in the XYZ coordinate space. The determination of the closest atoms to EVPAO was executed with geometry-optimized ground state systems at time zero. The two atoms closest to the EVPAOs are used as the terminal atoms of the structure of interest.

4. Between the two atoms, the shortest path is found using RDKit's shortest path determination functions.³⁴ These shortest paths are discussed from now on as "paths."

5. The paths are then extracted as separate molecular fragments, which represent chemical compounds. The paths retain both electronic structural information and allow for topological structural information to be generated in a numerical fashion. If both EVPAOs were found to be closest to the same indexed atom, then the singular atom was used as the path for that pair of electron orbitals.

2.5. Descriptor Generation. The paths, shown in Tables 1 and 2, were generated as conventional molecule files, and then converted to *.mol files using OpenBabel,⁵¹ where the optimization of the paths occurred during conversion from XYZ coordinates to *.mol files and using the software

Table 1. Path and Photodegradation Related Information in the Aqueous Environment

initial excitation	D_{ij} (Å)	walk distance	rate, k (ps ⁻¹)
HO-30→LU	2.37	5	0.728
HO-21→LU	1.70	1	0.663
HO-13→LU + 1	1.47	1	0.000
HO-12→LU	2.90	5	0.488
HO-6→LU	2.97	6	0.389
HO-2→LU	2.56	5	0.446
HO-2→LU + 2	1.24	4	0.723
HO-1→LU	1.87	1	0.000
HO-1→LU + 3	0.22	0	0.691
HO-1→LU + 4	0.50	0	0.698
HO→LU	2.08	1	0.000

Table 2. Paths and Photodegradation Related Information in the Vacuum Environment

initial excitation	D_{ij} (Å)	walk distance	rate, k (ps ⁻¹)
HO-21→LU + 2	4.96	9	0.000
HO-21→LU + 8	4.39	9	0.674
HO-20→LU + 3	2.57	6	0.000
HO-19→LU + 5	1.03	2	0.000
HO-18→LU + 6	1.15	0	0.000
HO-17→LU	6.85	15	0.000
HO-15→LU + 8	1.36	2	0.000
HO-14→LU + 1	1.52	1	0.000
HO-13→LU	5.31	6	0.000
HO-13→LU + 2	0.89	2	0.766
HO-13→LU + 8	0.69	2	0.000
HO-11→LU + 25	4.10	8	0.000
HO-10→LU + 2	1.93	7	0.000
HO-9→LU	4.39	4	0.000
HO-8→LU + 7	1.52	0	0.000
HO-8→LU + 24	1.91	2	0.000
HO-7→LU + 3	3.64	7	0.729
HO-3→LU + 4	2.95	2	0.000
HO-3→LU + 8	3.78	4	0.597
HO-2→LU + 3	3.34	7	0.742
HO-2→LU + 6	1.61	2	0.000
HO-2→LU + 9	3.74	7	0.683
HO-1→LU + 1	4.58	7	1.157
HO-1→LU + 3	6.23	11	0.000
HO-1→LU + 9	6.76	11	0.755
HO→LU	12.2	18	0.000
HO→LU + 2	6.41	10	0.000
HO→LU + 3	1.47	0	0.604
HO→LU + 9	0.91	0	0.736

optimization options.⁵¹ the structural information from the paths was encoded as numerical values, by means of the molecular descriptors, in the ochem.eu online platform where the descriptors were generated with alvaDesc.^{34,35} AlvaDesc generated approximately 3800 descriptors per path with different categories that include among others: 2D and 3D matrix-based descriptors, 2D autocorrelations, constitutional indices, charge-based descriptors, 0D, 2D, and 3D descriptors, molecular properties, and topological descriptors. The molecular descriptors with high correlation and noninformative information were discarded based on constant value, near constant, and pair correlation criteria higher than 0.95. Additionally, data gathered from the previous source including transition energy and oscillator strength were added to the data set as electronic structural descriptors.⁶

2.6. QSAR: Regression Model Development and Validation. A data set of all path descriptors, electronic structural information, and photodegradation rate values was developed where the quantitative structure–activity relationship (QSAR) methodology was applied. QSAR is generally used as a tool for determining the structural factors that allow for prediction of investigated response value.^{28,56–59} Due to the computational response values and the complexity of both electronic and topological structural information, the QSAR methodology is used to highlight correlations that influence the rate of photodegradation rather than developing predictions.

In this work, genetic algorithm (GA) and multilinear regression (MLR) analysis methods were used to develop

the multilinear regression (MLR) models.⁶⁰ This algorithm was emphasized since the descriptors chosen by GA can be discussed as influential in a linear fashion to the rate of photodegradation.⁵² GA has been used to address numerous problems relating to computation processing time as well as reduce to number of descriptors in final models.^{53,54} MLR models were developed using QSARINS software.³⁷

In this work, several QSAR models were developed having the range of 1–2 descriptors. These developing model techniques were followed by statistical analysis. The evaluation specified the comparison of the squared correlation coefficient for training and test sets, R^2_{train} and R^2_{ext} respectively. The calculation of R^2 is presented in eq 8.

$$R^2 = \frac{\sum_{i=1}^n (y_i^{\text{obs}} - y_i^{\text{pred}})^2}{\sum_{i=1}^n (y_i^{\text{obs}} - \bar{y}^{\text{obs}})^2} \quad (8)$$

The variables for eq 8 are as follows: y_i^{obs} is the experimental (observed) value of the property for the *i*th/*j*th compound; y_i^{pred} is the predicted value for the *i*th/*j*th compound; $\bar{y}$ is the mean experimental value of the property in the training set while $\hat{y}$ is the mean experimental value of the validation set.

A training set of 10 and 26 paths for aqueous and vacuum environments, respectively, was used for training the regression models. A test set of 3 and 1 for vacuum and aqueous environments, respectively. A test set is used for validation purposes, where a training set may hold unforeseen biases that can be expressed on a test set. Due to some paths having the same chemical structure, it was necessary to have nonrepetitive compounds in the test set for more robust predictive measures. Several one and two descriptor regression models were developed. The applicability domain (AD) was calculated by the leverage approach. This method allows to verify the models' predictive reliability.^{58,59}

To visualize the AD of the models, the Williams plot was used. The Williams plot axes are the standardized cross-validated residuals (Std. Resid.), *Y*-axis, versus leverage (*Hat i/ i*) values, *X*-axis. Std. Resid. clearly depict the response outliers, *Y*-outliers, and HAT depicts the structurally influential compound outliers, *X*-outliers, in a model. Once regression models were developed, the selection was based on their statistical performance: (1) highest R^2_{train} value and highest R^2_{ext} value. This was to show the highest correlation between structural information descriptors and photodegradation rate. (2) A low correlation, < 0.6, between descriptors was presented. (3) Williams plot analysis, to make sure the data set was robust enough to incorporate all paths, all paths needed to be within the AD of the presented Williams plots.

An example of a two-descriptor regression model is presented in eq 9, where the d_1 and d_2 are the first and second descriptors selected from the GA, respectively. C_1 and C_2 are the coefficients of the descriptors determined by MLR. C_3 is the intercept and y is the response variable, in this case, the rate of photodegradation.

$$y = C_1 \times d_1 + C_2 \times d_2 + C_3 \quad (9)$$

3. RESULTS

The initial excitation of active orbital pairs, the Euclidean distance in Angstrom (Å), the walk distance, and the rate of photodegradation k (ps⁻¹), are shown in Tables 1 and 2 for both aqueous and vacuum environments, respectively. The atoms that were found to be in the structural path can be

simplified to a path length where the number of atoms is shown as walk distance. For discussion purposes, the rate of photodegradation, which was previously reported, is shown. Some entries have a path distance of 0 atoms where both expectation values of orbitals were nearest to the same exact indexed atom. The developed structural paths between active orbitals are shown in Tables S1 and S2 for the aqueous and vacuum environments, respectively. Chemical equations are shown for representative paths in Schemes S1 and S2.

The information presented shows that similar paths may have differing rates of degradation. For example, three orbital pairs, HO-30 $\rightarrow$ LU, HO-12 $\rightarrow$ LU, and HO-2 $\rightarrow$ LU have the same path but differ in rates. A rare situation when the same paths provide different rates prompts a necessity for the inclusion of an additional descriptor in the model. One hypothesizes that such descriptor may include the value of computed transition energy and or higher order spatial features of the involved orbitals, such as second-order momenta, related with spatial delocalization of the involved orbitals, for example, $\langle (X - \langle X \rangle_i)^2 \rangle_i$ or inverse participation ratio. Orbital pairs were generally present in regions near the center of the system as presented in Figure S1, where yellow spheres are representative of the expectation values of orbitals.

3.1. Model Development and Validation. The data set was used to develop models for both environments. The best models are shown in eqs 10 and 11 for aqueous and vacuum environments, respectively.

$$K_{\text{Aqueous}} = (-0.816) \times \text{ChiA_Coulomb} + 0.889 \quad (10)$$

Model parameters were $R^2 = 0.667$; $R = 0.817$; for Aqueous system

$$K_{\text{vacuum}} = (-0.150) \times \text{MATs7s} + 0.298 \quad (11)$$

Model parameters were $R^2 = 0.687$; $R = 0.829$; for vacuum system. The descriptors and their definitions are listed in Table 3.

Table 3. Descriptors Information in Models (10) and (11)

environment	descriptor	definition	descriptor group
aqueous	ChiA_Coulomb	average Randic-like index from Coulomb matrix	3D matrix-based descriptors
vacuum	MATs7s	Moran coefficient of lag 7 weighted by intrinsic state	2D autocorrelations

The statistical analysis was conducted where model parameters were reported above, showing statistical significance and reasonable levels of error. R^2 is reported since it indicates the variation in data explained by the relationship between the independent variable, descriptors, and the dependent variable, rate of photodegradation. For ease of discussion, it is the statistical value used to measure the predictability of the independent variable onto the dependent variable. R^2 , the squared correlation coefficient, represents the correlation between the independent variables and the dependent variable. Both are used to discuss predictability and correlation.

The models were analyzed further by investigating the correlation plot between the computationally determined rate of degradation and the predicted values of the rate of degradation. Figure 4A shows the correlation plot for the aqueous environment. Orbital pairs are shown as individual

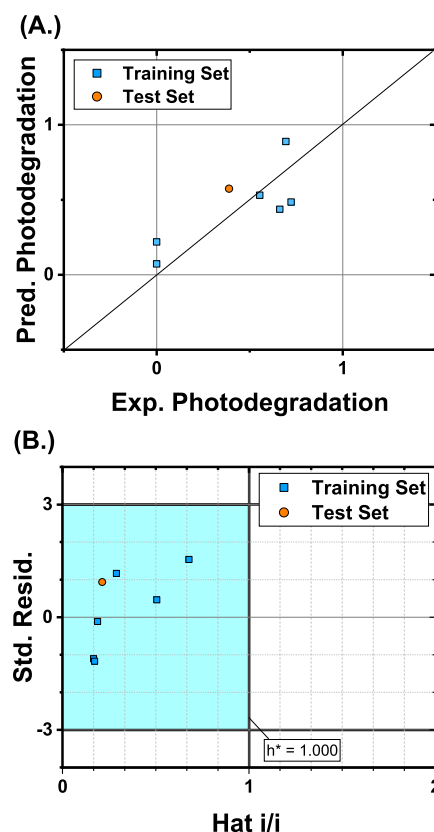


Figure 4. Aqueous environment model: correlation plot (A) and Williams plot (B).

dots in both the training and test sets. Further validation of our model, to make sure the structures were within reason to include in our data set, the use of applicability domain of the Williams plot. This plot statistically determined that structures were within a prescribed structural domain for model development. The applicability domain, blue region of Figure 4B, is bounded by standard deviations of the structural differences determined by QSARINS software.³⁷ If beyond the 3-standard deviation, the structure would be outside of the applicability domain and not representative to be used in our data set. The X-axis of Figure 4B is the measure of the potential contribution a compound has on the fitted compounds collectively labeled as (*Hat i/i*). It is useful in determining if compounds are outliers regarding their descriptor values, which therefore may be excessively influencing the regression results. It can be seen here that all compounds are within the applicability domain. The training set refers to the paths that were used to develop the model presented in eq 10. While prediction set refers to chosen paths that were used as “test compounds” to test the developed model for application outside of the training paths.

Similarly, for the vacuum environment, the correlation plot was developed and is shown in Figure 2A. The Williams plot for the vacuum environment is shown in Figure 2B where there were no compounds outside of the 3 standard residual thresholds. The training set for Figure 2 refers to the paths that were used to develop the model presented in eq 11.

4. DISCUSSION

4.1. Aqueous System Model Performance and Descriptor Analysis. It can be seen that the correlation

plot of Figure 4A shows all compounds within relative agreement being close to the line $y = x$. No visual outliers were present indicating the statistical aqueous environment model predicts compounds' rate of photodegradation relatively well. The Williams Plot of the model based on aqueous environment calculations in Figure 4A shows no outliers, and all compounds are within the applicability domain. The R^2 value of 0.667 gives a moderate level of predictability. In statistical terms, converting R^2 to R gives a value of 0.817 which suggests a high level of correlation between the $ChiA_Coulomb$ descriptor and the rate of photodegradation, where the descriptor of $ChiA_Coulomb$, has pertinent information regarding the electronic and structural aspects affecting the photodegradation rate of the polymer system. To understand the connection between the $ChiA_Coulomb$ descriptor and the rate of photodegradation, a descriptive analysis of the descriptor and a knowledge-based validation are conducted.

The Coulomb matrix is a global representation of the electrostatic forces between the nuclei. It is calculated using eq 12 and has been used in determining the atomic charges of molecules.^{61,62}

Off-diagonal elements correspond to the Coulomb repulsion between atoms i and j , while diagonal elements encode a polynomial fit of atomic energies to nuclear charge, where the nuclear charge of differing atoms is Z_i and Z_j , the nuclear positions via Cartesian coordinates between two nuclear charges is R_{ij} .

$$M_{ij}^{Coulomb} = \begin{cases} 0.5Z_i^{2.4} & \text{for } i = j \\ \frac{Z_i Z_j}{R_{ij}} & \text{for } i \neq j \end{cases} \quad (12)$$

Although electrostatic effects have not been directly linked to influencing the photodegradation rate via intramolecular interactions, it is known that intermolecular interactions affect the stability of polymers during photodegradation. This descriptor attempts to describe the influence of the solvent system on the photodegradation of the polymer material.

From our previous work, we suggested that proton transfer of the solvent system onto the polymer material influenced the photodegradation process to develop degradation products that matched wet lab experiments.⁶ With this in mind, we can expand on the pertinent information on the descriptor by highlighting electrostatic effects and proton transfer on the rate of photodegradation. It is important to mention that the intrinsic state has been used previously to successfully model the rate of hydrogen abstraction process which is another form of proton transfer.⁶³

The electrostatic effect, also known as the dipole–dipole interaction, between the H_2O molecules and the polymer material has a large influence on the photodegradation rate. The influence is believed to lead to changes in the electronic structural localization during the photodegradation process following TDESMD methodology. This interaction, combined with energy inputted via photoluminescence during the photodegradation process, allows proton transfer to occur. The proton transfer then exacerbates the interaction due to chemical bonding, debonding, and nuclear motion throughout the trajectory. The $ChiA_Coulomb$ descriptor, and its insight relating to the electrostatic interaction, aligns with our previous

works' explanation of solvent proton transfer influencing the photodegradation rate.

4.2. Vacuum System Model Performance and Descriptor Analysis. When viewing the correlation plot of the vacuum environment, Figure 5A, it is shown that the model

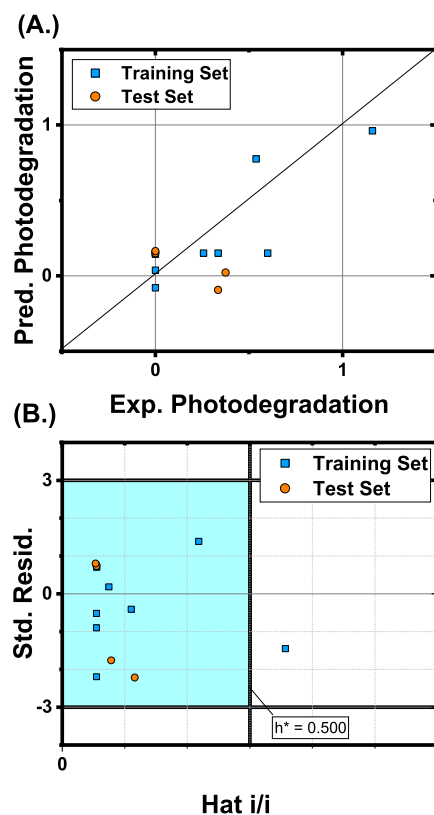


Figure 5. Vacuum environment model: correlation plot (A) and Williams plot (B).

adequately predicts the compound rate of photodegradation relative to the computational-experimentally found values. The data set used to train and test the model falls within the AD of the Williams plot shown in Figure 5B. Since all compounds, except for one, are within the AD we can say our model can adequately predict compounds that are structurally similar. The one compound that is beyond the 0.500 $HATS\ i/i$ threshold is still within the 3 Std deviations (3σ). The residual threshold indicates minor structural differences that do not influence the predictability of the model to a great extent. These two figures combined highlight the validation tools used to adequately critique the model for gathering unseen knowledge via statistical methods.

The R^2 value of 0.687 is a moderate level of predictability. In statistical terms, converting R^2 to R gives a value of 0.829 which suggests a high level of correlation between the $MATS7s$ descriptor and the rate of photodegradation. To understand the connection between the $MATS7s$ descriptor and the rate of photodegradation, a descriptive analysis of the descriptor and a knowledge-based validation is conducted.

For the descriptor in the vacuum model, $MAT7S$, the S standing for “weighted by intrinsic state” can be described and interpreted. The intrinsic state is calculated by eq 13.⁶⁴

Table 4. Connection between MATS7s and Photodegradation Rate–Vacuum

HOMO – LUMO Pair(s)	SMILES	photodegradation rate	MATS7s	correlation
HO-22→LU + 2, HO-22→LU + 8	ONC=CCOCCOC	0.335	−0.8226	−0.362
HO-0→LU + 2	CCOCC=CCOCC	0	−0.776	
HO-1→LU + 3, HO-1→LU + 9	OCCOCC=CCOCC	0.375	−0.4345	N/A
HO-11→LU + 25	NC=CCOCCOC	0	−0.3838	
HO-17→LU + 0	ONC=CCOCCOCCOCC=CN	0	−0.02167	N/A
HO-8→LU + 7	C	0.335	0	
HO-0→LU + 3				N/A
HO-0→LU + 9				
HO-18→LU + 6				N/A
HO-14→LU + 1	O=N	0	0	
HO-3→LU + 4	OCC	0	0	N/A
HO-19→LU + 5				
HO-15→LU + 8	COC	0.256667	0	N/A
HO-13→LU + 2				
HO-13→LU + 8				N/A
HO-8→LU + 24	CNO	0	0	
HO-2→LU + 6				N/A
HO-9→LU + 0	OCC=CN	0	0	
HO-3→LU + 8	OCCOC	0.6	0	N/A
HO-20→LU + 3	NC=CCOCC	0	0	
HO-13→LU + 0				N/A
HO-0→LU + 0	CCOCC=CCOCCOCCOCC=CN	0	0.04669	
HO-10→LU + 2	ONC=CCOCC	0.5375	2.112	0.942
HO-7→LU + 3				
HO-2→LU + 3				N/A
HO-2→LU + 9				
HO-1→LU + 1	OCCOCC=CN	1.16	2.743	N/A

$$\text{Intrinsic State} = \frac{\left(\frac{2}{N}\right)^2 \times \delta_i^v + 1}{\delta_i} \quad (13)$$

N = principal quantum number (2 for C, N, and O atoms), δ_i is the connectivity vertex degree of a graph (number of σ bonds), and δ_i^v = valence vertex degree of i th atom (number of valence electrons).

To describe the equation in a differing manner, the intrinsic state of an atom can be thought of as the ratio of lone pair electrons and π bonds to σ bonds for the considered atom. More available lone pair electrons and π bonds suggest the possible availability of electrons for intramolecular interactions along the specified path.⁶⁴ To connect to our work, this intramolecular availability allows for the movement of electrons needed for excitation and deexcitation assuming the electron moves across the shortest path of the compound. This aligns with our work well due to the general assumptions made during path development.

The intrinsic state includes information about the electronic structural properties of compounds while also including the connectivity ratio of links to nodes, following the graph theory. Others have suggested that molecular connectivity descriptors generally have a dominant nature even in single-descriptor statistical models.⁶³ This dominance is highlighted in our work, where it can describe topological information (paths) and electronic properties (electron density distribution), which relate heavily to the excitation and deexcitation process for photodegradation to occur.

An interesting finding is, shown in Table 4, that when the MATS7s descriptor has a value above 0, the photodegradation rate is heavily correlated 0.942, where the higher the

descriptor, the higher the photodegradation rate. This suggests that the more lone pair or π bond electrons allow for a higher rate of photodegradation. Additionally, the higher magnitude of either positive or negative MATS7s values relates back to a higher correlation of the photodegradation rate. This finding is only found in compounds that have 7 atoms or larger due to the 7-atom minimum requirement of the descriptor.

The availability of lone pair electrons and π bond electrons is believed to allow for the movement of electrons across the molecule to and from the active orbitals. This information is insightful and aligns with the theory that electrons move across the molecule from deexcited states to excited states and back.

5. CONCLUSIONS

In summary, the relation of the expectation value of position for active orbitals to the photodegradation rate of FDCA polymeric chains was analyzed by a novel process (InfoTDESMD) combining DFT-based TDESMD calculations with cheminformatics techniques. By analyzing the paths between the expectation value of charge density localization for orbital pairs involved in TDESMD calculations, one identifies a data set with several fragments and over 3000 descriptors per fragment. These paths containing both electronic structural and topological information are collected and analyzed by applying various software packages. The applied InfoTDESMD procedure suggests that the molecular structures' paths involved in an excitation between HO- m to LU + n had an influence on the rate of photodegradation. The values of various descriptors depend on the systems. The descriptor used to model the rate of photodegradation for an aqueous environment suggested a strong influence by neighboring solvent factors including electrostatic effects. In the vacuum

environment, it was found that the rate of photodegradation was influenced heavily by the intrinsic state of paths with 7 or more atoms. This suggests that available lone pair electrons and π bonds allow for the enhanced rate of photodegradation.

The application of InfoTDESMD allowed for insight into the connection between topological information and the rate of photodegradation. InfoTDESMD provides opportunities to investigate photoreactions for polymeric systems with various computational techniques faster than original TDESMD simulations, keeping similar precision of predicting power. Our calculations will further provide insights and knowledge of uninvestigated properties of TDESMD photodegraded systems to find correlations and predictions between the photoactivity and structure of compounds.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.3c06854>.

Example of converting orbital pairs to paths; simulation cell details; theoretical methods; computational details; and additional degradation results from TDESMD trajectories (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Bakhtiyor Rasulev – Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, North Dakota 58105, United States; orcid.org/0000-0002-7845-4884; Email: bakhtiyor.rasulev@ndsu.edu

Dmitri Kilin – Department of Chemistry and Biochemistry, North Dakota State University, Fargo, North Dakota 58105, United States; orcid.org/0000-0001-7847-5549; Email: dmitri.kilin@ndsu.edu

Authors

Meade Erickson – Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, North Dakota 58105, United States

Gerardo Casañola-Martin – Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, North Dakota 58105, United States

Yulun Han – Department of Chemistry and Biochemistry, North Dakota State University, Fargo, North Dakota 58105, United States

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpcb.3c06854>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the use of resources at the Center for Computationally Assisted Science and Technology (CCAST) at North Dakota State University, which was made possible in part by National Science Foundation (NSF) MRI Award No. 2019077. We also thank Aaron Forde and Amirhadi Alesadi for collective discussion and editing. The authors thank the DOE BES NERSC facility for computational resources, allocation award “Computational Modeling of Photo-catalysis and Photo-induced Charge Transfer Dynamics on Surfaces”, supported by the Office of Science of the DOE under Contract DE-AC02-

05CH11231. M.E. and B.R. thank the Department of Coatings and Polymeric Materials at North Dakota State University for providing the Lowell Wood funding. The authors thank Paola Gramatica for generously providing a free license for the QSARINS software. D.K. acknowledges the support of the National Science Foundation under Grant CHE-1944921. D.K. thanks David Micha, Oleg Prezhdo, Sergei Tretiak, Svetlana Kilina, Andrei Kryjevski, Benjamin Levine, Evert Jan Meijer, P. Stanley May, Mary T. Berry, and Sivaguru Jayaraman for inspiring discussions.

■ REFERENCES

- (1) Anastas, P.; Eghbali, N. Green Chemistry: Principles and Practice. *Chem. Soc. Rev.* **2010**, 39 (1), 301–312.
- (2) Geyer, R.; Jambeck, J. R.; Law, K. L. Production, Use, and Fate of All Plastics Ever Made. *Sci. Adv.* **2017**, 3 (7), No. e1700782.
- (3) Maertens, A.; Plugge, H. Better Metrics for “Sustainable by Design”: Toward an in Silico Green Toxicology for Green(Er) Chemistry (Vol 6, Pg 1999, 2018). *ACS Sustain. Chem. Eng.* **2018**, 6 (4), 5662.
- (4) Crawford, S. E.; Hartung, T.; Hollert, H.; Mathes, B.; van Ravenzwaay, B.; Steger-Hartmann, T.; Studer, C.; Krug, H. F. Green Toxicology: A Strategy for Sustainable Chemical and Material Development. *Environ. Sci. Eur.* **2017**, 29, 16.
- (5) Rajendran, S.; Raghunathan, R.; Hevus, I.; Krishnan, R.; Ugrinov, A.; Sibi, M. P.; Webster, D. C.; Sivaguru, J. Programmed Photodegradation of Polymeric/Oligomeric Materials Derived from Renewable Bioresources. *Angew. Chemie - Int. Ed.* **2015**, 54 (4), 1159–1163.
- (6) Erickson, M.; Han, Y. L.; Rasulev, B.; Kilin, D. Molecular Dynamics Study of the Photodegradation of Polymeric Chains. *J. Phys. Chem. Lett.* **2022**, 13 (19), 4374–4380.
- (7) Singathi, R.; Raghunathan, R.; Krishnan, R.; Kumar Rajendran, S.; Baburaj, S.; Sibi, M. P.; Webster, D. C.; Sivaguru, J. Towards Upcycling Biomass-Derived Crosslinked Polymers with Light. *Angew. Chem., Int. Ed.* **2022**, 61 (31), No. e202203353.
- (8) Zhu, Y. Q.; Romain, C.; Williams, C. K. Sustainable Polymers from Renewable Resources. *Nature* **2016**, 540 (7633), 354–362.
- (9) Han, Y.; Meng, Q.; Rasulev, B.; May, P. S.; Berry, M. T.; Kilin, D. S. Photofragmentation of the Gas-Phase Lanthanum Isopropylcyclopentadienyl Complex: Computational Modeling vs Experiment. *J. Phys. Chem. A* **2015**, 119 (44), 10838–10848.
- (10) Wang, F.; Li, F. L.; Xu, M. M.; Yu, H.; Zhang, J. G.; Xia, H. T.; Lang, J. P. Facile Synthesis of a Ag(I)-Doped Coordination Polymer with Enhanced Catalytic Performance in the Photodegradation of Azo Dyes in Water. *J. Mater. Chem. A* **2015**, 3 (11), 5908–5916.
- (11) Seki, S.; Yoshida, Y.; Tagawa, S.; Asai, K. Electronic Structure of Radical Anions and Cations of Polysilanes with Structural Defects. *Macromolecules* **1999**, 32 (4), 1080–1086.
- (12) Kuhn, R.; Jensch, R.; Bryant, I. M.; Fischer, T.; Liebsch, S.; Martienssen, M. The Influence of Selected Bivalent Metal Ions on the Photolysis of Diethylenetriamine Penta(Methylenephosphonic Acid). *Chemosphere* **2018**, 210, 726–733.
- (13) Brouard, M.; Clark, A. P.; Vallance, C.; Vasyutinskii, O. S. Velocity-Map Imaging Study of the O(P-3)+N-2 Product Channel Following 193 Nm Photolysis of N₂O. *J. Chem. Phys.* **2003**, 119 (2), 771–780.
- (14) Rajkumar, M.; Arunpandian, M.; Leeladevi, K.; Arunachalam, S.; Rameshkumar, P. Fabrication of Pebble Stone-like Nanostructure: Focus on Photocatalysis, Photoluminescence and Electron Density Distribution Analysis. *Physica* **2021**, 620, No. 413222.
- (15) Sivaganesh, D.; Saravanakumar, S.; Sivakumar, V.; Sasikumar, S.; Gopal, J. N.; Kalpana, S.; Rajajeyaganthan, R. Sm³⁺ Induced-SrWO₄ Phosphor: Analysis of Photoluminescence and Photocatalytic Properties with Electron Density Distribution Studies. *J. Mater. Sci. Electron.* **2020**, 31 (11), 8865–8883.
- (16) Parvathi, L. T.; Arunpandian, M.; Sivaganesh, D.; Nagarajan, E. R.; Karuthapandian, S. Flower Decorated Rod-like Pd @ MnO₂

Nanocomposite: Focus on Photocatalysis, Rietveld Refinement Analysis and Electron Density Distribution Analysis. *Physica* **2022**, *625*, No. 413475.

(17) Takata, M.; Nishibori, E.; Sakata, M. Charge Density Studies Utilizing Powder Diffraction and MEM. Exploring of High T_c Superconductors, C-60 Superconductors and Manganites. *Zeitschrift Fur Krist.* **2001**, *216* (2), 71–86.

(18) Collins, D. M. ELECTRON-DENSITY IMAGES FROM IMPERFECT DATA BY ITERATIVE ENTROPY MAXIMIZATION. *Nature* **1982**, *298* (5869), 49–51.

(19) VoigtMartin, I. G. The Impact of New Microscopical Methods on the Investigation of Polymers 0.2. Electron Crystallography and High Resolution Imaging. *Acta Polym.* **1996**, *47* (8), 311–322.

(20) Alekseyev, A. B.; Liebermann, H. P.; Buenker, R. J. An Ab Initio Study of the CH₃I Photodissociation. II. Transition Moments and Vibrational State Control of the I-* Quantum Yields. *J. Chem. Phys.* **2007**, *126* (23), 234103.

(21) Qureshi, M.; Khan, Z.; Manoharan, S. S. Excited State Energy Distributions in Stereo Selective White Light Emission of 1,2-Dibenzthiazolyl Ethylene Isomers. *Chem. Phys. Lett.* **2009**, *473* (1–3), 184–188.

(22) Li, Y. Z.; Cui, J. G.; Zhao, J. N.; Liu, J. L.; Song, P.; Ma, F. C. Effect of Electronic Acceptor Segments on Photophysical Properties of Low-Band-Gap Ambipolar Polymers. *Sci. World J.* **2013**, *2013*, No. 890215.

(23) Martin, R. L. Natural Transition Orbitals. *J. Chem. Phys.* **2003**, *118* (11), 4775–4777.

(24) Makki, H.; Adema, K. N. S.; Peters, E. E. A. J. F.; Laven, J.; van der Ven, L. G. J.; van Benthem, R. A. T. M. R.; de With, G. A Simulation Approach to Study Photo-Degradation Processes of Polymeric Coatings. *Polym. Degrad. Stab.* **2014**, *105*, 68–79.

(25) Engel, T. Basic Overview of Chemoinformatics. *J. Chem. Inf. Model* **2006**, *46*, 2267–2277.

(26) Karuth, A.; Alesadi, A.; Xia, W.; Rasulev, B. Predicting Glass Transition of Amorphous Polymers by Application of Cheminformatics and Molecular Dynamics Simulations. *Polymer (Guildf)*. **2021**, *218*, No. 123495.

(27) Karuth, A.; Casanola-Martin, G.; Lystrom, L.; Sun, W.; Kilin, D.; Kilina, S.; Rasulev, B. Combined Machine Learning, Computational and Experimental Analysis of the Iridium(III) Complexes with Red to Near-IR Emission. *J. Phys. Chem. Lett.* **2024**, *15*, 471.

(28) Erickson, M. E.; Ngongang, M.; Rasulev, B. A Refractive Index Study of a Diverse Set of Polymeric Materials by QSPR with Quantum-Chemical and Additive Descriptors. *Molecules* **2020**, *25* (17), 3772.

(29) Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; et al. QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, *57* (12), 4977–5010.

(30) Xiao, Y. J.; Fan, R. L.; Zhang, L. F.; Yue, J. Q.; Webster, R. D.; Lim, T. T. Photodegradation of Iodinated Trihalomethanes in Aqueous Solution by UV 254 Irradiation. *Water Res.* **2014**, *49*, 275–285.

(31) Xiao, Y. J.; Zhang, L. F.; Zhang, W.; Lim, K. Y.; Webster, R. D.; Lim, T. T. Comparative Evaluation of Iodoacids Removal by UV/Persulfate and UV/H₂O₂ Processes. *Water Res.* **2016**, *102*, 629–639.

(32) Wan, D.; Chen, Y.; Su, J.; Liu, L.; Zuo, Y. G. Ultraviolet Absorption Redshift Induced Direct Photodegradation of Halogenated Parabens under Simulated Sunlight. *Water Res.* **2018**, *142*, 46–54.

(33) Chen, J.; Hochstatter, A. M.; Kilin, D.; May, P. S.; Meng, Q.; Berry, M. T. Photofragmentation of Gas-Phase Lanthanide Cyclopentadienyl Complexes: Experimental and Time-Dependent Excited-State Molecular Dynamics. *Organometallics* **2014**, *33* (7), 1574–1586.

(34) Landrum, G.; Tosco, P. *RdKit*, 2023.

(35) Sushko, I.; Novotarskyi, S.; Korner, R.; Pandey, A. K.; Rupp, M.; Teetz, W.; Brandmaier, S.; Abdelaziz, A.; Prokopenko, V. V.; Tanchuk, V. Y.; et al. Online Chemical Modeling Environment (OCHEM): Web Platform for Data Storage, Model Development and

Publishing of Chemical Information. *J. Comput. Aided. Mol. Des.* **2011**, *25* (6), 533–554.

(36) Mauri, A. AlvaDesc: A Tool to Calculate and Analyze Molecular Descriptors and Fingerprints. In *Ecotoxicological Qsars*; Roy, K., Ed.; Humana Press Inc: Totowa, 2020; pp. 801–820.

(37) Gramatica, P.; Chirico, N.; Papa, E.; Cassani, S.; Kovarich, S. QSARINS: A New Software for the Development, Analysis, and Validation of QSAR MLR Models. *J. Comput. Chem.* **2013**, *34* (24), 2121–2132.

(38) Polishchuk, P. Interpretation of Quantitative Structure–Activity Relationship Models: Past, Present, and Future. *J. Chem. Inf. Model.* **2017**, *57* (11), 2618–2639.

(39) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.

(40) Perdew, J. P.; Burke, K.; Ernzerhof, M.; Perdew, John P.; Kieron Burke, M. E.; Perdew, John P.; Matthias Ernzerhof, K. B. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865.

(41) Holzwarth, N. A. W.; Tackett, A. R.; Matthews, G. E. A Projector Augmented Wave (PAW) Code for Electronic Structure Calculations, Part I: Atompaw for Generating Atom-Centered Functions. *Comput. Phys. Commun.* **2001**, *135* (3), 329–347.

(42) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140* (4A), A1133–A1138.

(43) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), B864–B871.

(44) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), B864–B871.

(45) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B Condens Matter* **1993**, *47* (1), 558–561.

(46) Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metal-Amorphous-Semiconductor Transition in Germanium. *Phys. Rev. B Condens Matter* **1994**, *49* (20), 14251–14269.

(47) Furthmüller, G. K. J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169.

(48) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6*, 15–50.

(49) Chen, J.; Meng, Q.; Stanley May, P.; Berry, M. T.; Kilin, D. S. Time-Dependent Excited-State Molecular Dynamics of Photodissociation of Lanthanide Complexes for Laser-Assisted Metal-Organic Chemical Vapour Deposition. *Mol. Phys.* **2014**, *112* (3–4), 508–517.

(50) Rabi, I. I.; Ramsey, N. F.; Schwinger, J. Use of Rotating Coordinates in Magnetic Resonance Problems. *Rev. Mod. Phys.* **1954**, *26* (2), 167–171.

(51) Rabi, I. I. Space Quantization in a Gyration Magnetic Field. *Phys. Rev.* **1937**, *51* (8), 652–654.

(52) Nguyen, T. S.; Nanguneri, R.; Parkhill, J. How Electronic Dynamics with Pauli Exclusion Produces Fermi–Dirac Statistics. *J. Chem. Phys.* **2015**, *142* (13), 134113.

(53) Nguyen, T. S.; Parkhill, J. Nonadiabatic Dynamics for Electrons at Second-Order: Real-Time TDDFT and OSCE₂. *J. Chem. Theory Comput* **2015**, *11* (7), 2918–2924.

(54) Meng, Q.; May, P. S.; Berry, M. T.; Kilin, D. Sequential Hydrogen Dissociation from a Charged Pt₁₃H₂₄ cluster Modeled By ab Initio Molecular Dynamics. *Int. J. Quantum Chem.* **2012**, *112* (24), 3896–3903.

(55) Meng, Q.; Chen, J.; Kilin, D. Proton Reduction at Surface of Transition Metal Nanocatalysts. *Mol. Sim.* **2015**, *41* (1–3), 134–145.

(56) Daghighi, A.; Casanola-Martin, G. M.; Timmerman, T.; Milenkovic, D.; Lucic, B.; Rasulev, B. In Silico Prediction of the Toxicity of Nitroaromatic Compounds: Application of Ensemble Learning QSAR Approach. *Toxics* **2022**, *10* (12), 746.

(57) Nantasenamat, C.; Isarankura-Na-Ayudhya, C.; Naenna, T.; Prachayasittikul, V. A practical Overview of Quantitative Structure–Activity Relationship. *EXCLI J.* **2009**, *8*, 74–88.

(58) Muratov, E. N.; Bajorath, J.; Sheridan, R. P.; Tetko, I. V.; Filimonov, D.; Poroikov, V.; Oprea, T. I.; Baskin, I. I.; Varnek, A.; Roitberg, A.; et al. QSAR without Borders. *Chem. Soc. Rev.* **2020**, 49 (11), 3525–3564.

(59) Pagar, R. R.; Musale, S. R.; Pawar, G.; Kulkarni, D.; Giram, P. S. Comprehensive Review on the Degradation Chemistry and Toxicity Studies of Functional Materials. *Acs Biomater. Sci. Eng.* **2022**, 8 (6), 2161–2195.

(60) Walters, W. P.; Goldman, B. B. Feature Selection in Quantitative Structure-Activity Relationships. *Curr. Opin. Drug Discovery Dev.* **2005**, 8 (3), 329–333.

(61) Rupp, M.; Tkatchenko, A.; Muller, K. R.; von Lilienfeld, O. A. Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning. *Phys. Rev. Lett.* **2012**, 108 (5), No. 058301.

(62) Montavon, G.; Hansen, K.; Fazil, S.; Rupp, M.; Biegler, A.; Ziehe, A.; Tkatchenko, A.; Lilienfeld, A.; Müller, K. R. Learning Invariant Representations of Molecules for Atomization Energy Prediction. In *Advances in Neural Information Processing Systems 25*; Weinberger, F., Ed.; Curran Associates Inc., 2012; pp. 440–448.

(63) Pogliani, L. Modeling with Molecular Pseudoconnectivity Descriptors. A Useful Extension of the Intrinsic I-State Concept. *J. Phys. Chem. A* **2000**, 104 (39), 9029–9045.

(64) Todeschini, V. C. R.; Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*; WILEY-VCH, 2000; Vol. 11.