Article

Synergy of Small Antiviral Molecules on a Black-Phosphorus Nanocarrier: Machine Learning and Quantum Chemical Simulation Insights

Slimane Laref [1,*,†] , **Fouzi Harrou** [2,*,†] , **Bin Wang** [3], **Ying Sun** [2], **Amel Laref** [4], **Taous-Meriem Laleg-Kirati** [2], **Takashi Gojobori** [1] **and Xin Gao** [1]

[1]   Computational Bioscience Research Center (CBRC), King Abdullah University of Science & Technology (KAUST), Thuwal 23955-6900, Saudi Arabia

[2]   A Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia

[3]   Center for Interfacial Reaction Engineering (CIRE), School of Chemical, Biological and Materials Engineering, University of Oklahoma, Norman, OK 73019, USA

[4]   Department of Physics and Astronomy, College of Science, King Saud University, Riyadh 11451, Saudi Arabia

*   Correspondence: slimane.laref@kaust.edu.sa (S.L.); fouzi.harrou@kaust.edu.sa (F.H.)

†   These authors contributed equally to this work.

Abstract: Favipiravir (FP) and Ebselen (EB) belong to a broad range of antiviral drugs that have shown active potential as medications against many viruses. Employing molecular dynamics simulations and machine learning (ML) combined with van der Waals density functional theory, we have uncovered the binding characteristics of these two antiviral drugs on a phosphorene nanocarrier. Herein, by using four different machine learning models (i.e., Bagged Trees, Gaussian Process Regression (GPR), Support Vector Regression (SVR), and Regression Trees (RT)), the Hamiltonian and the interaction energy of antiviral molecules in a phosphorene monolayer are trained in an appropriate way. However, training efficient and accurate models for approximating the density functional theory (DFT) is the final step in using ML to aid in the design of new drugs. To improve the prediction accuracy, the Bayesian optimization approach has been employed to optimize the GPR, SVR, RT, and BT models. Results revealed that the GPR model obtained superior prediction performance with an R^2 of 0.9649, indicating that it can explain 96.49% of the data's variability. Then, by means of DFT calculations, we examine the interaction characteristics and thermodynamic properties in a vacuum and a continuum solvent interface. These results illustrate that the hybrid drug is an enabled, functionalized 2D complex with vigorous thermostability. The change in Gibbs free energy at different surface charges and temperatures implies that the FP and EB molecules are allowed to adsorb from the gas phase onto the 2D monolayer at different pH conditions and high temperatures. The results reveal a valuable antiviral drug therapy loaded by 2D biomaterials that may possibly open a new way of auto-treating different diseases, such as SARS-CoV, in primary terms.

Keywords: MD; machine learning; DFT; inhibitor; black phosphorus; thermodynamic; molecular states; drug vehicles

1. Introduction

The fast development of clinical biomedicine and nanobiotechnology has inspired the current generation of diverse inorganic nanoparticles that provide various pattern modalities as possible substitutions in addressing diseases by synergistic therapy, and principally to treat metastatic pathology [1–3]. Two-dimensional (2D) materials have fascinated extensive consideration due to their exceptional and distinctive properties [4–6]. Encouraged by the great success of graphene, a variety of 2D materials have been deployed for various property-related applications such as energy, environmental science, catalysis, physics, and biomedicine [7–14].

For example, transition metal dichalcogenides (TMDs), nitrides and carbonitrides (MXenes), and hexagonal boron nitride (h-BN), and their derivatives are realized with a monolayer and a few layers, which exhibit weak van der Waals interlayer bonding and strong in-plane covalent bonding [15–17]. Thereby, the biological behavior of Xenes 2D materials is under outstretched investigation to check their biosafety and biocompatibility, such as biodegradability, cytotoxicity, and toxic derivatives [18]. The broad spectra of potential uses and clinical implementation of the Xenes 2D systems in biomedical drug therapy still require proper control of biodegradability [19] and toxicity, as well as an explicit understanding of the biochemical procedures occurring in the drug molecules on 2D nanomaterials and their interfaces. The first aspect is particularly valuable since the understanding of the orientations of molecules with the multi-layered materials can characterize the process of raising biologically friendly materials for gene silencing, scattering complexes, hybrid pharmaceutic ingredients, and biomedicine sensing. In the present study, motivated by the latest developments in 2D nanomaterial synthesis and their encouraging solicitation in anti-tumor drugs [20–22], we conducted a comprehensive investigation in order to use 2D materials as nanocarriers for drug delivery. One such potential pharmacological option for treating various desists is ebselen (EB), while another is favipiravir (FP) [23,24]. Favipiravir has recently been used to treat infections caused by the Ebola virus, and it was found to be safe during clinical studies [25,26]. During the outbreaks of the SARS-CoV-1 and MERS-CoV infections, favipiravir was once more reviewed, and it was shown to be effective against these viruses in animal model systems [27–31]. Ebselen is an organoselenium compound with cytoprotective, anti-inflammatory, and anti-oxidant properties. This compound has previously been investigated to treat multiple diseases, including hearing loss [32] and bipolar syndrome [33]. Ebselen has minimal cytotoxicity [34], and its safety in humans has been estimated in a number of clinical trials [35–38]. Ebselen has the predominates inhibitor targeting the main protease (Mpro) of SARS-CoV-2 (e.g., Mopar is a crucial enzyme of coronaviruses) [24]. Recently, in vitro human cell studies reported that both favipiravir and ebselen could effectively stop the replication of SARS-CoV-2 at an earlier stage [24].

In recent years, the application of machine learning techniques in nanomaterials and pharmaceutical technologies has gained significant attention from researchers. Several studies have used backpropagated neural networks to model various systems [39,40], including Darcy–Forchheimer slip flow models for nanomaterials and ferrofluids [41], non-Fourier heat flux models for transient heat exchange [42], and ternary nanofluid flow between parallel plates [43]. The accuracy of these models was evaluated using reference datasets obtained via the Homotopy Analysis Method and numerical methods. The effects of non-dimensional parameters on the systems were also analyzed. In this paper, by utilizing accelerated molecular dynamics, machine learning, and quantum chemicals based on first-principles density functional theory (DFT), we briefly report the structural stability and thermochemical properties of FP and EB functionalized within a typical phosphorene 2D monolayer. The chemical bonding of FP and EB in 2D black-phosphorus materials is revealed by considering the vacuum and continuum solvent model. However, by calculating the Gibbs free energy changes of each step, we demonstrated that the temperature and pH values would be used to tune the releasing ratio of the antiviral molecules onto the phosphorene sheet. Our results provide new fundamental insights into exploring innovative 2D materials as new nanocarriers for antiviral drug therapy. These atomistic findings open the door to the parametrization of multiscale methodologies, such as contemporary machine learning based on electronic theories, continuum models, Monte Carlo simulations, and coarse-grained methods, to address the interplay between the interaction of drugs/2D monolayers with the cell membrane as a function of their morphology. In addition, we rigorously explored the efficiency of four machine learning models to accelerate the prediction of Hamiltonian and non-bond energy from EB, FP, and FP-EB, namely Bagged Trees, Gaussian Process Regression (GPR), Support Vector Regression (SVR), and Regression Trees (RT). The reason for considering multiple models is to compare and evaluate their performance in accurately predicting the behavior of

the antiviral molecules in the 2D monolayer. We calibrated the machine learning models using Bayesian optimization based on the training data. Results revealed the superior performance of the GPR model by reaching an averaged R^2 of 0.9649. Training efficient and accurate models for approximating DFT is the final step in using machine learning to gain more insight into the design of new drugs.

The remainder of this paper is organized as follows. Section 2 presents the used data and the obtained results. In Section 3, we discuss the pharmachemical implications. Section 4 briefly describes the preliminary materials, including the Mean Squared Error (MSE), the investigated machine learning models, and ab initio calculations. Finally, we offer conclusions in Section 5.

2. Outcomes and Discussion

2.1. Data Description

Molecular dynamics is a powerful approach that enables atoms in the systems to interact with each other based on given parameters, such as temperature, time, pressure, and other conditions. The evolution of molecular simulations can be presented by various terms in the Hamiltonian (kinetic energy, potential energy, interaction energy, free energy). Thereby, the consistency of MD methods combined with ML and other quantum techniques can be considered. The main key in this work is to improve the Hamiltonian and non-bond energy terms from MD simulations as a starting point. The datasets are generated based on the COMPASS III [44] force field of molecular dynamics simulations for small antiviral molecules on BP single layer. By using the ML algorithms mentioned above, the Hamiltonian and the interaction energy are also trained at periodic boundary conditions (PBC) of polycyclic molecules on top of a phosphorene monolayer. Training efficient and accurate models for approximating DFT is the final step in using ML to aid in the design of new drugs.

2.2. Data Analysis

This part is dedicated to assessing the capability of the investigated machine learning models in predicting Hamiltonian and non-bond energy from three different datasets, MLFF-EB, MLFF-FP, and MLFF-FP_EB.

Figure 1 displays the distribution of the Hamiltonian energy from three different datasets. Plotting the probability density can be a useful tool for exploratory data analysis. Generally speaking, it can help visualize a data distribution and gain insights into its properties. It shows the relative frequency of different values and the shape of the distribution, which can identify patterns or anomalies. It can also be used to calculate statistical measures such as mean, variance, and skewness, providing further insight into the properties of the data. From Figure 1, we observe that these datasets are non-Gaussian-distributed. It would challenge traditional prediction methods, such as principal component regression, designed based on the Gaussian assumption of data. Thus, machine learning models designed without assumptions about the data distribution could be promising.

Figure 2a–c depicts the pairwise correlation coefficients between the multivariate data of antiviral drugs on the BP monolayer, MLFF-EB, MLFF-FP, and MLFF-FP_EB, respectively. As the calculation proceeds, there is no change in the temperature profile. This suggests that the model is dynamically stable. We observe that the temperature is absolutely correlated with kinetic energy. Once more, the temperature is uncorrelated with the other variables since it is related to the variations of the potential energy, non-bond energy, and total energy, respectively. Furthermore, it is also observed that in the Hamiltonian, potential energy, kinetic energy, non-bond energy, and total energy are constant throughout the simulation processes. We found a lowering in the kinetic energy that leads to general stability. However, this indicates that the model used in this study shows additional thermal stability of the system. The relaxed geometries are further used for creating a black phosphorus sheet for loading antiviral molecules that characterize the drug-BP based nanocarrier. Accordingly, the optimized ML structures (MLFF-EB, MLFF-FP, and MLFF-FP_EB, at T = 0 K) are in

good agreement with the interaction energies and van der Waals energies found by means of DFT calculations [45]. For more information, please see Tables S1 and S2, respectively. Finally, during the calculations, the entire cell with its external pressure was fixed, and it had a moderate correlation with respect to bonding energy, angle energy, and TPE, as depicted in Figure 2.
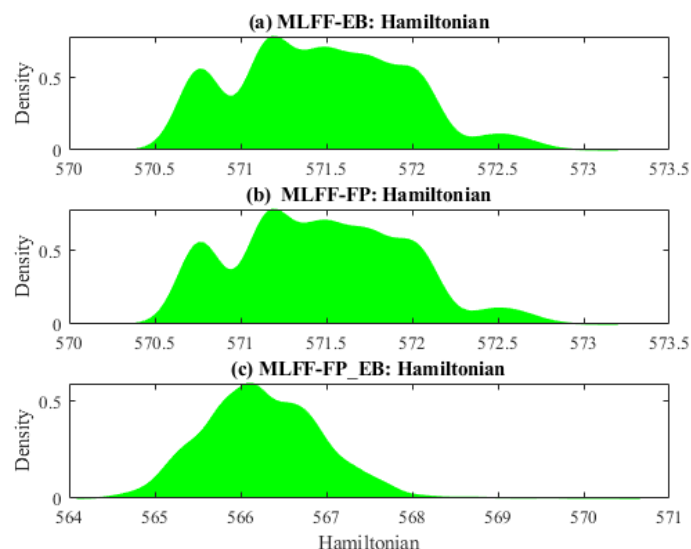

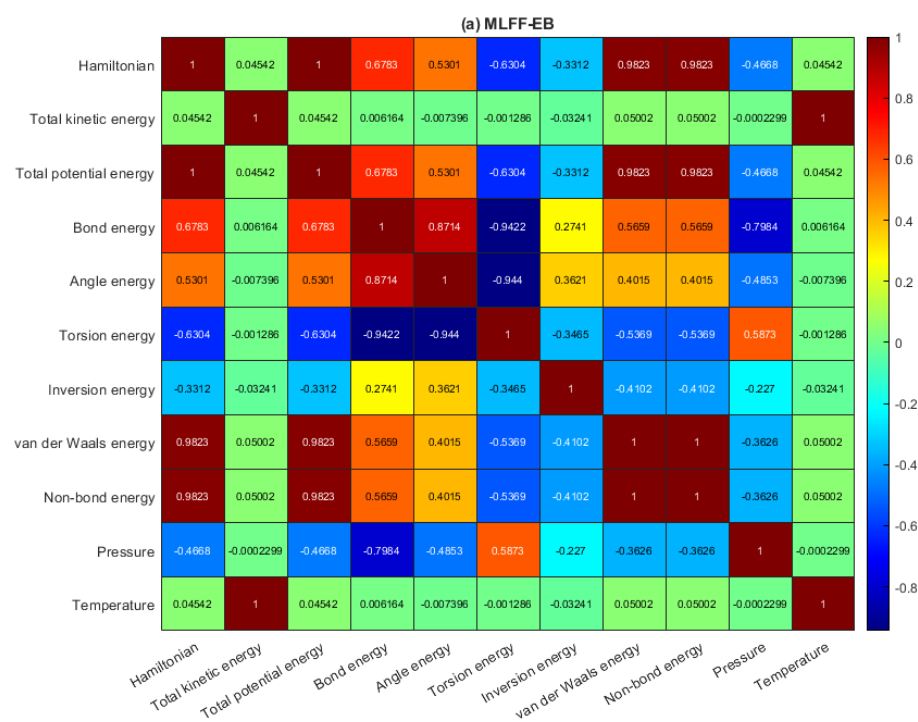**Figure 1.** Distribution of Hamiltonian energy: (**a**) MLFF-EB, (**b**) MLFF-FP, and (**c**) MLFF-FP_EB.
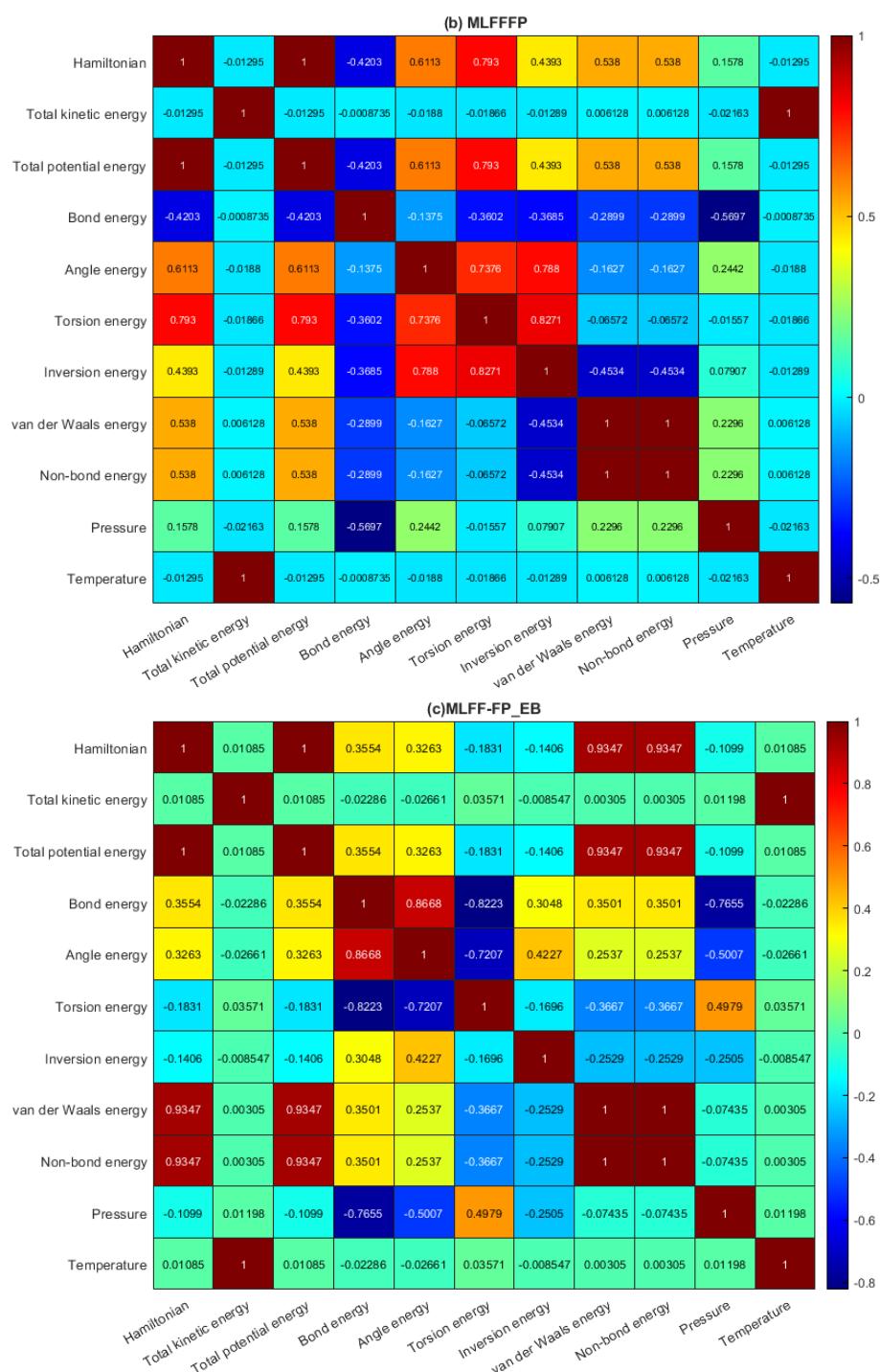

Figure 2. *Cont*.

Figure 2. Heatmap of the correlation matrix: (**a**) MLFF-EB, (**b**) MLFF-FP, and (**c**) MLFF-FP_EB.

2.3. Prediction Results

2.3.1. Hamiltonian Energy Prediction

Three datasets were used to assess the performance of the machine learning models. Each sub-set of data consists of eight input variables (i.e., total kinetic energy, angle energy, torsion energy, inversion energy, van der Waals energy, non-bond energy, pressure, and temperature) and one out-variable (i.e., Hamiltonian energy). Each data set comprises around 2000 samples. In the first experiment, the dataset was divided into two subsets: training composed of 80% and testing with 20%. The k-fold cross-validation technique was considered in constructing these models based on the training data. Specifically, a 5-fold cross-validation procedure was employed in training the investigated models. The Bayesian

optimization approach was employed to optimize the GPR, SVR, RT, and BT models to determine the optimal parameters that minimize the Mean Absolute Error between the actual and predicted Hamiltonian energy using the training data.

Table 1 summarizes the evaluation metrics, RMSE, MAE, R^2, and MAPE, computed from the testing datasets for MLFF-EB, MLFF-FP, and MLFF-FP_EB. Results revealed that the investigated machine learning models exhibited good capability in predicting Hamiltonian energy. We also notice from Table 1 that no single approach uniformly dominates the other models for the three considered datasets. For the MLFF-EB dataset, the SVR model attained the best prediction accuracy with an R^2 of 0.9767. In this case study, the GPR model obtained an accurate prediction with an R^2 of 0.9614 for the MLFF-EB dataset, and the best prediction quality was obtained by the GPR model with an R^2 of 0.9669 when applied to the MLFF-FP_EB dataset.

Table 2 lists the average evaluation metrics per model for Hamiltonian energy prediction. The closest R^2 value to 1 and the lowest RMSE, MAE, and MAPE values characterize the best prediction performance. Results in Table 2 show that all four models performed well in predicting the Hamiltonian energy, with R^2 values ranging from 0.9478 to 0.9649, indicating a good fit between the predicted and actual values. The GPR model had the highest R^2 value (0.9649) and the lowest values for RMSE, MAE, and MAPE, indicating that it performed slightly better than the other models in terms of accuracy. An R^2 value of 0.9649 means that the GPR model can capture 96.49% of the variation in the Hamiltonian energy of the antiviral molecules in the phosphorene monolayer. In addition, the GPR model achieved the lowest RMSE (0.0789) and MAE (0.0521) values among all models, indicating that the model had the smallest average deviation between predicted and actual values. Additionally, the GPR model had the lowest MAPE value (0.0095), indicating that the average percentage error between predicted and actual values was the lowest among all models. Therefore, the GPR model performed the best in terms of all evaluation metrics. This indicates that the model is highly accurate in predicting the Hamiltonian energy and can be used as a reliable tool for designing new antiviral molecules.

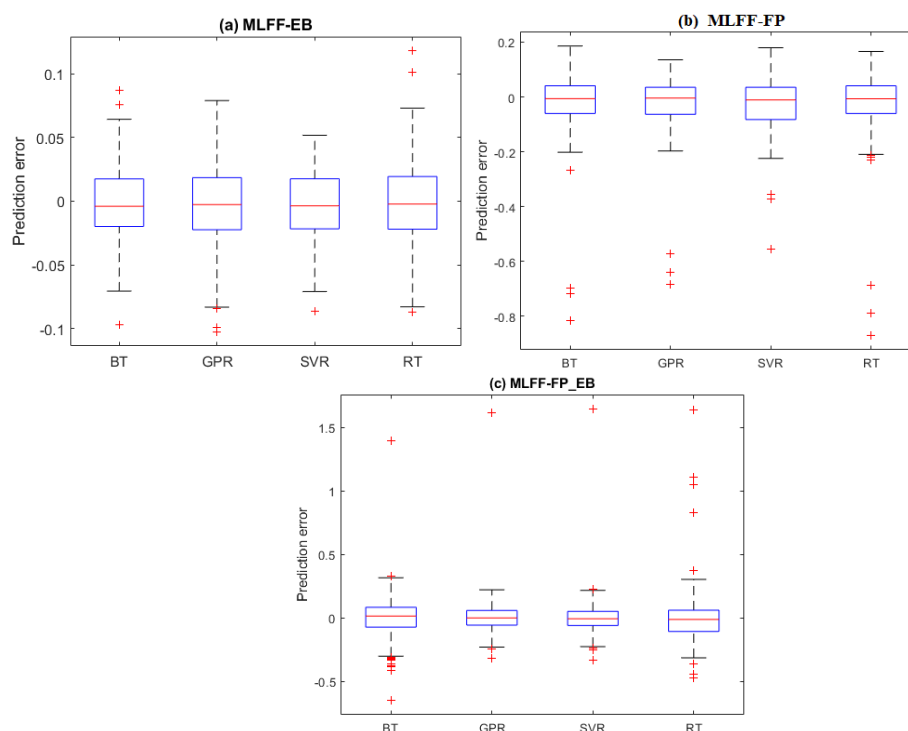**Table 1.** Prediction quality of the Hamiltonian energy using the four optimized models.

MLFF-FP				
Methods	R^2	RMSE	MAE	MAPE
BT	0.9743	0.0266	0.0217	0.0049
GPR	0.9663	0.0304	0.0238	0.0053
SVR	0.9767	0.0253	0.0208	0.0047
RT	0.9674	0.0299	0.0238	0.0053
MLFF-EB				
Methods	R^2	RMSE	MAE	MAPE
BT	0.9513	0.1001	0.0655	0.0115
GPR	0.9614	0.0891	0.0606	0.0106
SVR	0.9447	0.1067	0.0809	0.0142
RT	0.9501	0.1014	0.0654	0.0114
MLFF-FP_EB				
Methods	R^2	RMSE	MAE	MAPE
BT	0.9489	0.1459	0.1011	0.0179
GPR	0.9669	0.1173	0.0718	0.0127
SVR	0.9665	0.1182	0.0710	0.0125
RT	0.9259	0.1756	0.1109	0.0196

Table 2. Averaged evaluation metrics per model for Hamiltonian energy prediction.

	R^2	**RMSE**	**MAE**	**MAPE**
BT	0.9582	0.0909	0.0628	0.0114
GPR	0.9649	0.0789	0.0521	0.0095
SVR	0.9626	0.0834	0.0576	0.0105
RT	0.9478	0.1023	0.0667	0.0121

It is worth examining the prediction errors from the four investigated models (i.e., BT, GPR, SVR, and RT). The prediction error refers to the deviation between the true Hamiltonian energy measurements and their predicted values. Figure 3a–c displays the boxplots of the prediction errors for each model based on testing datasets from the MLFF-EB, MLFF-FP, and MLFF-FP_EB datasets, respectively. Visually, we observed that the prediction errors of the GPR and SVR models are concentrated around zero, indicating better prediction accuracy than the BT and RT models. Furthermore, it can also be seen that the GPR model with narrower boxes and whiskers reaches superior performance compared to the SVR for the MLFF-FP and MLFF-FP_EB datasets (Figure 3). Moreover, SVR and GPR models show relatively comparable performance under the MLFF-EB dataset (Figure 3).

As a visual illustration, the scatter graphs and plots of actual and predicted Hamiltonian energy of MLFF-EB, MLFF-FP, and MLFF-FP_EB, when applying the trained GPR model to testing data are depicted in Figure 4. To simplify visual readability, we presented only the results from the best model, and the results from the other models were omitted. We observe that the predicted values of Hamiltonian energy obtained by the GPR models are fairly close to the actual data of MLFF-EB, MLFF-FP, and MLFF-FP_EB, which indicates the promising performance of the GPR model.

Figure 3. Prediction Distribution of prediction errors of the four models based on testing datasets: (**a**) MLFF-EB, (**b**) MLFF-FP, and (**c**) MLFF-FP_EB.
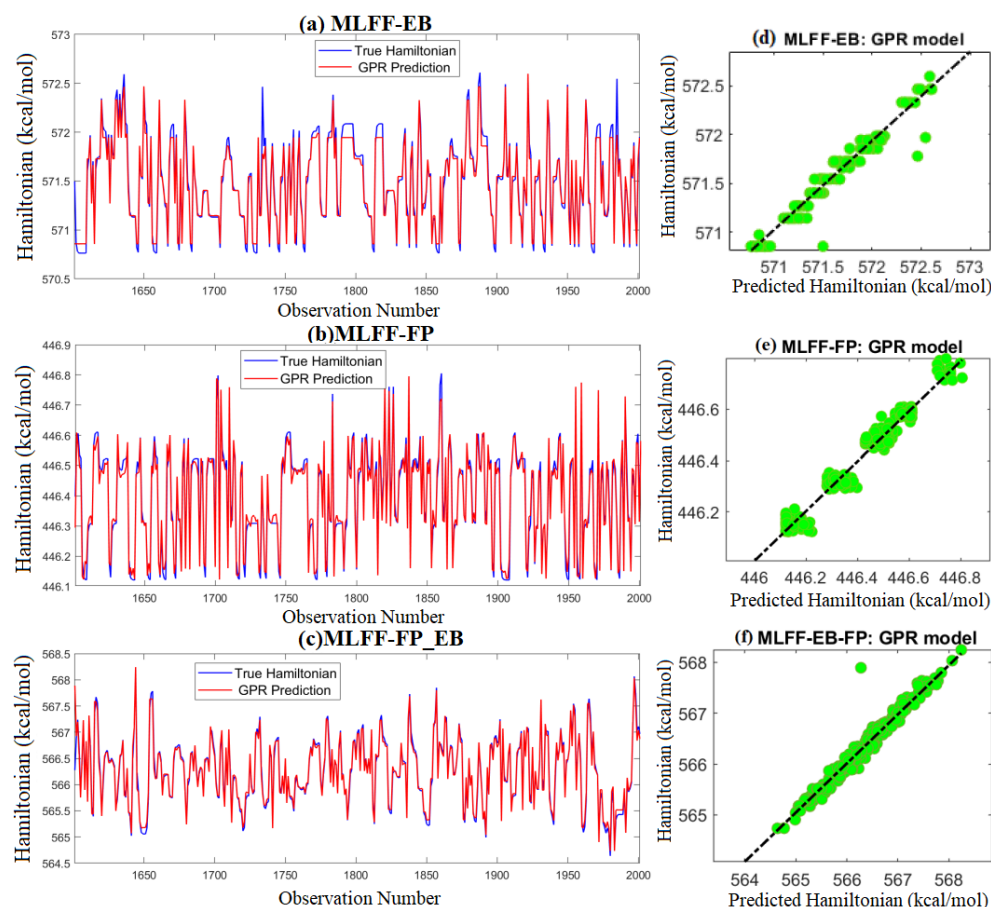
Figure 4. Predicted Hamiltonian energy from the GPR model based on test data: (**Top**) MLFF-EB, (**Middle**) MLFF-FP, and (**Bottom**) MLFF-FP_EB.

2.3.2. Non-Bond Energy Prediction

In the second experiment, the main key objective was to predict and improve the non-bond energy for MLFF-EB, MLFF-FP, and MLFF-FP_EB based on seven input variables (i.e., total kinetic energy, angle energy, torsion energy, inversion energy, van der Waals energy, pressure, and temperature) related to ab initio results [45]. Towards this end, we constructed four machine learning models (BT, SVR, GPR, and RT) based on a 5-fold cross-validation with the training dataset. The parameters of each model have been optimized using the Bayesian optimization approach in training. The trained models have been employed for non-bond energy based on testing datasets. Then, the statistical metrics (R^2, RMSE, MAE, and MAPE) are computed to compare the prediction performance of the investigated machine learning models (Table 3). For the three datasets, the GPR approach dominates the other approaches (i.e., BT, SVR, and RT) in terms of the four performance evaluation metrics (Table 3). Regarding the MLFF-FP, the four models provide relatively comparable performance with an R^2 around 0.97; the GPR model slightly outperformed the other models with an R^2 of 0.9792. Considering the MLFF-EB data, the GPR model obtained the best prediction performance for MLFF-EB and MLFF-FP_EB with an R^2 of 0.9735 and 0.9725, respectively. Table 4 lists the averaged evaluation metrics per model for non-bond energy prediction. The GPR model performed the best among all the models with an R^2 value of 0.9751, indicating a very good fit between the predicted and actual non-bond energies. The GPR model also had the lowest RMSE and MAE values, indicating that the predicted values were very close to the actual values. The MAPE value for the GPR model was 0.9579%, the lowest among all models, indicating that the model had the lowest percentage error in prediction. Therefore, based on these evaluation metrics, the GPR model was the best performer among all the models for non-bond energy prediction.

Table 3. Prediction quality of the non-bond energy.

	R^2	RMSE	MAE	MAPE
MLFF-FP				
BT	0.9772	0.0279	0.0218	0.2466
GPR	0.9792	0.0266	0.0208	0.2348
SVR	0.9772	0.0279	0.0211	0.2382
RT	0.9775	0.0277	0.0221	0.2502
MLFF-EB				
BT	0.9685	0.0758	0.0493	0.4496
GPR	0.9735	0.0695	0.0468	0.4277
SVR	0.9705	0.0734	0.0492	0.4503
RT	0.9729	0.0703	0.0486	0.4448
MLFF-FP_EB				
BT	0.9648	0.1253	0.0843	2.5297
GPR	0.9725	0.1107	0.0732	2.2111
SVR	0.9710	0.1138	0.0771	2.3273
RT	0.9689	0.1178	0.0795	2.3512

Figure 5a–c shows the boxplots of the prediction errors for the four models based on testing datasets (MLFF-EB, MLFF-FP, and MLFF-FP_EB). Visually, we observe that the boxplots are around zero, which means that the prediction quality is good and the prediction errors are small. Additionally, it can be seen that the GPR model provides relatively slightly less prediction errors than the other models for the three considered datasets.
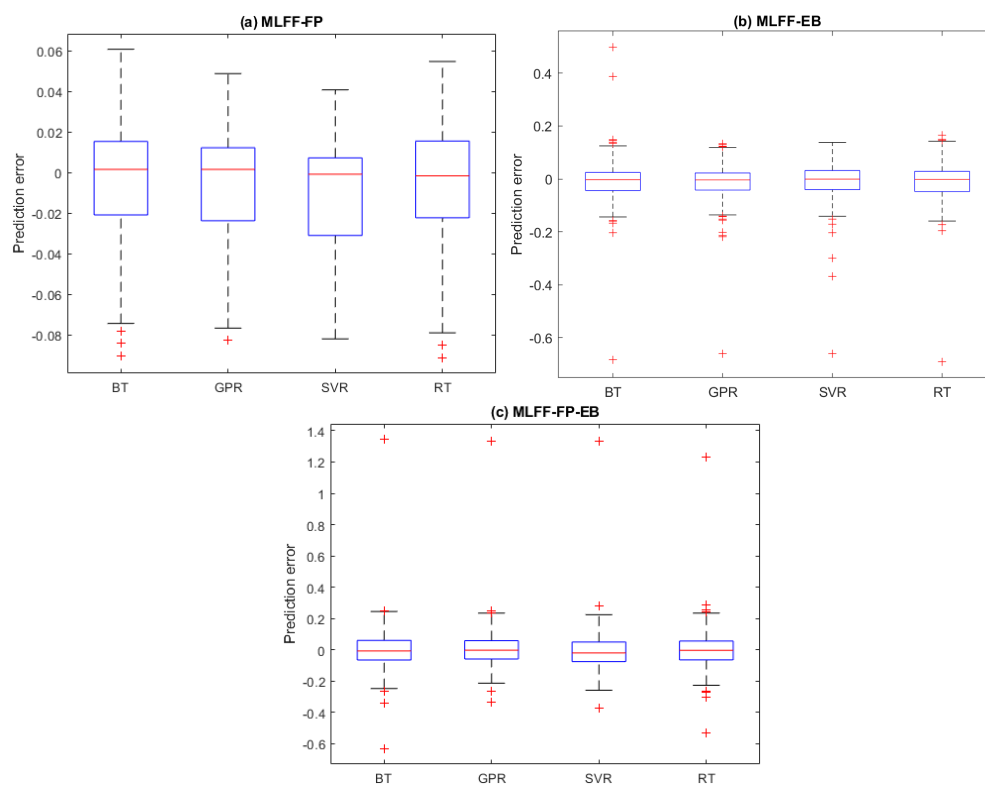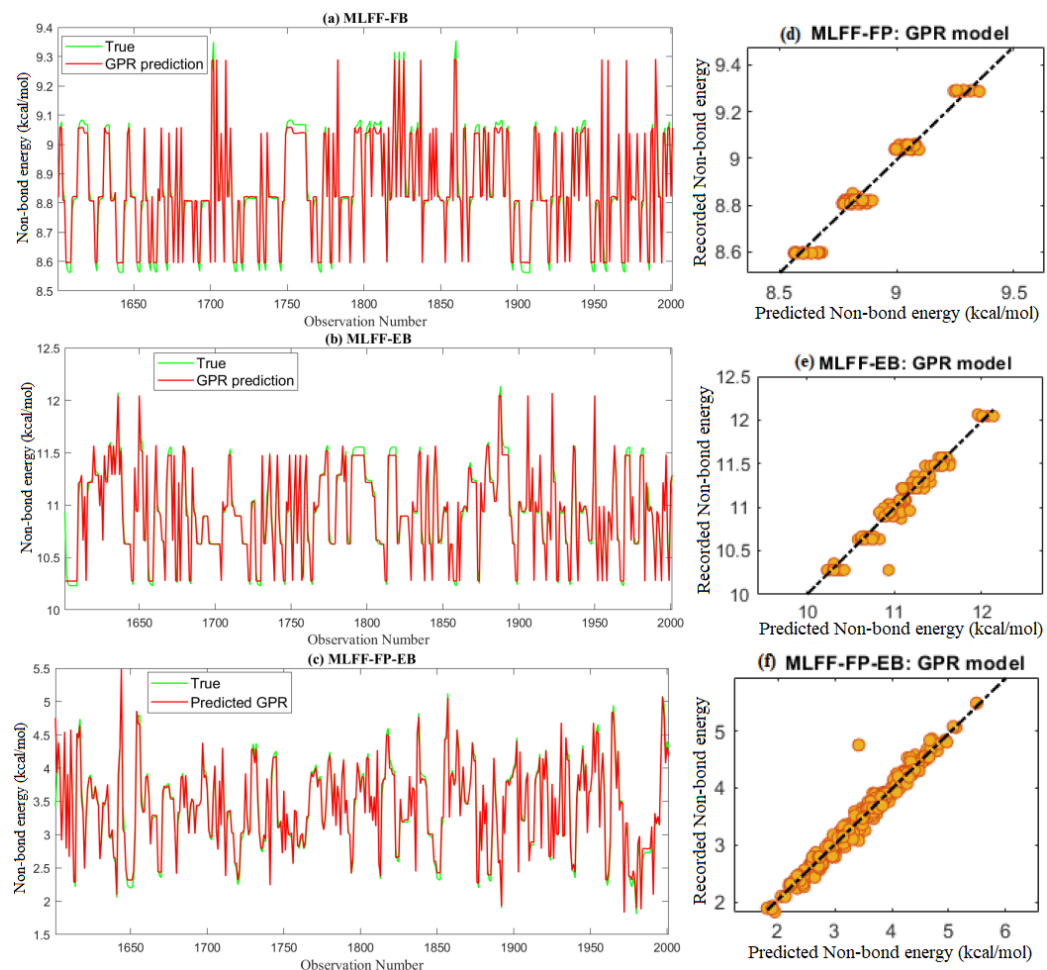

Figure 5. Prediction errors of the prediction models from the testing datasets: (**a**) MLFF-EB, (**b**) MLFF-FP, and (**c**) MLFF-FP_EB.

Table 4. Averaged evaluation metrics per model for non-bond energy prediction.

	R^2	RMSE	MAE	MAPE
BT	0.9701	0.0763	0.0518	1.0753
GPR	0.9751	0.0689	0.0469	0.9579
SVR	0.9729	0.0717	0.0491	1.0053
RT	0.9731	0.0719	0.0501	1.0154

The actual MD results and the predicted non-bond energy from the four models based on testing datasets are depicted in Figure 6a–c. We found that the predicted values of non-bond energy from the GPR model closely followed the measured data. Figure 6d–f shows the scatter plot of the measured and predicted values of non-bond energy from the GPR model. A scatter plot is a useful tool to visually compare a dataset's predicted and measured data. In Figure 6d–f, the colored dots represent the data points, where the *x*-axis represents the predicted values and the *y*-axis represents the measured non-bond energy values. If the dots are closely clustered around the diagonal line, the model performs well and makes accurate predictions. Thus, from Figure 6d–f, we observe that the predictions are close to the measured values.

Figure 6. Predicted non-bond energy from the GPR model based on test data.

3. Pharmachemical Implications

3.1. Drug Release

3.1.1. The pH Sensitivity

The 2D biomaterials are extensively applied in thermotherapy by tuning the drug–virus interactions. It is important to assign an effective nanocarrier for drug therapy;

similarly, it is well-noted that the close link concerning virus and plasma antigen levels of tissue type (plasminogen) has a substantial clinical impact [46]. Accordingly, as a result of defection tissue, the normal and infected tissue should have different pH numbers. As a result, the stability of different respiratory diseases is the maximum at a pH close to 6, so the application of the voltages also affects the pH of the plasma tissue. The voltage induced is widely used in physical therapy to treat a variety of diseases. However, the pH measurement is simulated in the solvent environment by placing altered charged ions. On the other hand, the solid–solvent interface provides an alternative challenge for biomaterial simulation due to the large number of degrees of freedom hosted by the liquid states. For this purpose, we used an implicit solvation model, where the inhibitor molecule/2D biomaterial is treated at the level of a quantum mechanic state, and the solvent is treated as a continuum model described by the relative permittivity. This implicit solvation model is found to mimic the correct experimental solvation energy in various solvents [47].

To predict the pH sensitivity of the binding characteristics, we follow the method in Ref. [47] by adding or removing electrons number of valence states, expressed by the modification of the work function and the applied voltage, which has been demonstrated to provide consistent results with the experiments [48]. The adsorption energy of neutral FP, EB, and FP + EB on the BP single layer is -2.61, -5.41, and -8.86 meV/Å^2, respectively; inclusion of 0.1 extra electrons will promote the energy of the system and cause a lowering of the binding energy. For example, adding half electrons in FP, EB, and FP + EB on BP surface changes the energy to 4.02, -1.95, and 2.82 meV/Å^2, respectively, and the desorption of the hybrid FP + EB drugs becomes spontaneous. This scheme allows us to frequently control the chemical desorption by introducing fractional charge at the interface under the Poisson–Boltzmann approximation. In particular, pH-dependent adsorption depicted in Figure 7 suggests that larger adsorption energy is obtained in a polar environment with a higher pH value. Thus, when they are expected to release pharmaceutical drugs in a pointing position, an application of a low pH value with a smaller positive voltage would be appropriate.
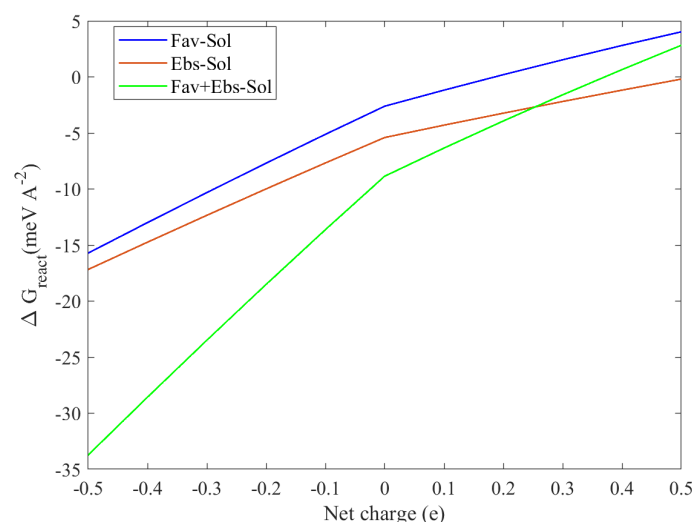

Figure 7. The calculated adsorption energy process as a function of net charge at the solvent phase of the BP sheet with favipiravir, ebselen, and favipiravir + ebselen hybrid drugs. The minus red and plus blue sign colors represent the difference between pH < 0 and pH > 0 values, respectively.

3.1.2. Thermotherapy Properties

The electronic and pH-influencing properties of antiviral drugs were described above. The key character of chemical adsorption nature related to temperature dependence is the next step in this investigation. Correspondingly, as represented in Figure 8, the temperature indeed alters the adsorption energy of the molecules at the BP monolayer, either at vacuum or solvent level. It can be well-noticed that by considering the vibrational correction, despite

the fact that the entropy change is probably dominated by translational entropy rather than vibration, it shifts the adsorption to more positive or less negative numbers at a higher temperature. The zero-point energy at lower temperature values (0–100 K) has a minor effect on the energy (between 1–2 meV/Å^2 for FP, EB, and FP + EB, respectively). On the other hand, increasing temperature from 150 to 420 K causes shifts of up to 4.1 meV/Å^2 for FP, EB, and FP + EB on the BP-vacuum interface with the incorporation of the corrected vibrational free energy. Equivalent trends have been found by adding the continuum solvent level to these antiviral-molecules/BP 2D-system. Surprisingly, the adsorption remains stable up to 450 K, at the vacuum phase for hybrid FP + EB on the BP sheet. The thermotherapy is vigorous, and the releasing rate of the antiviral molecules will be highly enhanced due to the large drop in the adsorption energy at higher temperatures.

In general, the drug will have a larger release rate beyond 350 K, and it might be due to the height's thermal deviation at a larger temperature. In the ex vivo evaluation process, the phosphorene-loaded drugs can be used as an anti-defected cell drug that has a pertinent act in the thermal therapeutic [45]. The temperature could upturn to around 450 K within a few minutes. The excellent therapeutic efficiency of these drugs in the black phosphorus nanocarrier is firmly associated to the large ability to target drug-release and the hyperthermia of the infected cell. Further, these findings suggest that the hybrid FP + EB therapy of low nanocarrier-based delivery systems can be activated effectively [48]. This comprehensive investigation will contribute to optimizing virus growth and storage conditions, facilitating the molecular characterization of this important pathogen. The interest of accommodating those inhibitors in close contact with the phosphorene sheet as a new delivery scheme is that *it could continuously control the desorption free energy* of the aromatic molecules by applying an electric field, together with the presence of external stimuli to tune their pH environment, and changing their temperature.
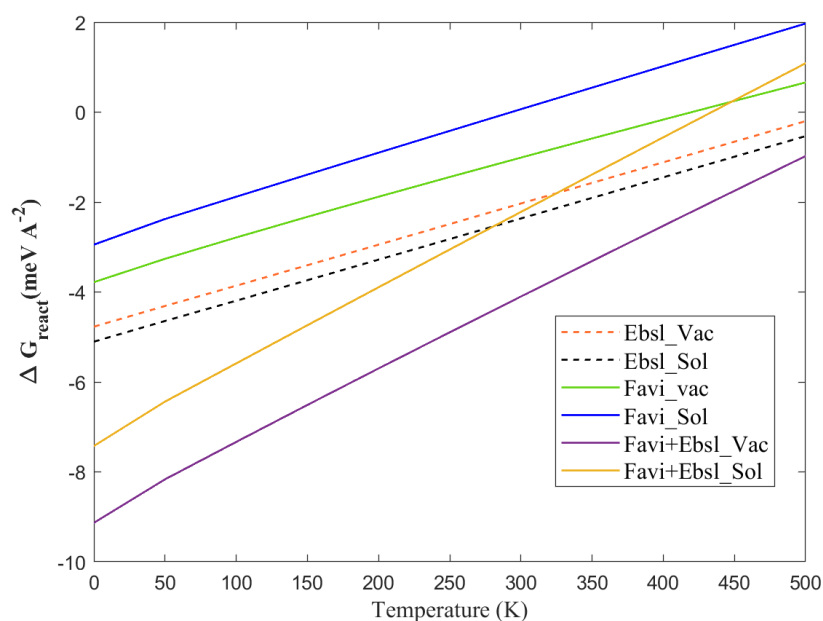

Figure 8. The free Gibbs adsorption energy process as a function of temperature at the vacuum and solvent phases of the BP sheet with favipiravir, ebselen, and favipiravir+ebselen hybrid drugs.

4. Materials and Methods

This section is dedicated to molecular dynamics simulations and machine learning models deployed for this study, which is related to describing nanocarrier loaded drugs prediction. ML approaches are effective methods for accelerating DFT relaxations and require significantly fewer iterations (DFT calculations) for convergence. It has shown excellent transferability along FP, EB, and hybrid FP and EB molecules physisorbed on a BP single layer, and closely describes the ab initio interaction energy of the entire system. This

is an encouraging initiative toward a common class of force-field and ab initio approaches combined with ML for drug molecules on ultra-thin film crystalline 2D material.

4.1. Molecular Dynamics Simulations

Molecular dynamics (MD) simulations based on atomic force fields can accurately simulate the non-covalent geometry of a large number of molecular systems from non-equilibrium systems to thermodynamic states at the atomic level. In the present study, the current data were obtained by means of molecular dynamics simulations, as implemented in Forcite software [44]. To properly adsorb FP and EB drugs on a BP single layer, we used a (5 × 4) supercell model of black phosphorus containing 80 phosphorus atoms for the nanocarrier. A vacuum region of 20 Å is introduced perpendicular to the BP film to prevent unphysical periodic effects. A Hamiltonian including kinetic energy, potential energy, interaction energy, and free energy as a function of the change of temperature and pressure was simulated at 350 K and atmospheric pressure, respectively. We used an NVT canonical ensemble based on the Nosé thermostat. During the MD simulations, we adopted a time step of 0.3 fs for an ensemble of 100 ps of calculations to reach the thermodynamic states and accurate MD calculations.

4.2. Gaussian Process Regressor

GPR models, which are within nonparametric kernel-driven learning models, showed extended modeling ability for handling nonlinear prediction problems because of their nonlinear approximation capabilities [49–51]. They are well-known for their ability and flexibility in modeling small-sized data with a Gaussian or non-Gaussian distribution, as well as providing uncertainty measures on predictions [52,53]. Moreover, the GPR model's benefit is its capacity to provide a confidence interval to evaluate the reliability of the predicted values, making it widely used in numerous applications [54]. The GPR models have been utilized in various applications, including renewable energy systems monitoring [50,55], COVID-19 spread prediction [56], and spatio-temporal $PM_{2.5}$ prediction [57].

To introduce GPR, we have to consider the response **y** of a function **f** at the input **x**, which is expressed as [58],

$$\mathbf{y}_i = f(\mathbf{x}_i) + \varepsilon_i. \tag{1}$$

Here, $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is an additive noise, and $f(x)$ is considered as a random variable. It is worth pointing out that the uncertainty on f could decrease significantly by the observation of the function's output at different input points.

The function $f(\mathbf{x})$ is considered following a Gaussian process. Accordingly, y_i follows a joint Gaussian distribution [58]:

$$\mathbf{y} = [y_1, y_2, \ldots y_n]^\top \sim \mathcal{N}(\mathbf{m}(\mathbf{x}), \mathbf{K} + \sigma^2 \mathbf{I}), \tag{2}$$

where $\mathbf{m}(\mathbf{x}) = [m(x_1), m(x_2), \ldots m(x_n)]^\top$ refers to the vector of mean values $m(\cdot)$, **I** denotes the identity matrix, and **K** refers to the $n \times n$ covariance matrix with (i, j)th element $\mathbf{K}_{ij} = k(x_i, x_j)$. Within the GPR framework, $\mathbf{K}_{ij} = k(x_i, x_j)$ is termed the kernel function [52,58]. Numerous kernels have been used in the literature, including the Linear kernel, Rational Quadratic (RQ) kernel, Squared Exponential (SE) kernel, Matern 5/2 (M52) kernel, and Exponential (Exp) kernel. The two commonly used kernels (i.e., linear and squared exponential) are expressed respectively as [52]:

$$k_{\text{Lin}}(x_i, x_j) = \theta_1^2 + \theta_2^2 (x_i - \theta_3)(x_j - \theta_3). \tag{3}$$

$$k_{\text{SE}}(x_i, x_j) = \theta_1 \exp\left(\frac{(x_i - x_j)^2}{\theta_2}\right). \tag{4}$$

To optimize the GPR model in the training stage, we need to determine the kernel parameters that enable maximizing the following likelihood [58,59].

$$\boldsymbol{\theta}_{\text{opt}} = \arg \max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}), \tag{5}$$

where $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots]$ are the parameters of the kernel (also called hyperparameters), the mean values $m(.)$ are taken to be zero, and

$$L(\boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{K} + \sigma^2 \mathbf{I}|}} \exp\left(-\frac{1}{2}(\mathbf{y}^\top (\mathbf{K} + \sigma^2 \mathbf{I})\mathbf{y})\right). \tag{6}$$

Many efforts have been reported in the literature to calibrate the hyper-parameters by using grid search or random search. Here, Bayesian optimization will be adopted to fine-tune the GPR hyperparameters by maximizing the marginal likelihood in (5) with respect to θ [60].

We assume x_* is a new test input, then the estimated mean and variance associated with $\hat{y}_* = f(x_*) = f_*$ are [58,59]:

$$\hat{y}_* = \mathbf{k}_*^\top (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{y}, \tag{7}$$

and

$$\Sigma_* = k_{**} - \mathbf{k}_*^\top (\mathbf{K} + \sigma^2 \mathbf{I})^{-1} \mathbf{k}_*. \tag{8}$$

respectively. Then, $\mathbf{y}_*$ follows a conditional distribution which has the form:

$$y_* | \mathbf{y} \sim \mathcal{N}(\hat{y}_*, \Sigma_*), \tag{9}$$

where $K = k(X, X)$, $K_{**} = k(X_*, X_*)$, and $K_* = k(X, X_*)$ are the covariance matrices computed based on the training data, the testing set, and both training and test sets, respectively.

Overall, GPR-based prediction involves two main steps. The first step is to estimate the hyperparameters, which include the kernel and noise variance, in the GPR model by maximizing the negative log marginalized likelihood based on the training data. The second step is to compute the predictive posterior distribution of GPR for new inputs, X_*, by using the mean of the predictive distribution in Equation (8) as the predicted value. For more details about the GPR model, see Refs. [49,50,58,61].

4.3. Bayesian Optimization Procedure

As discussed before, we need to calibrate hyperparameters during training to obtain good prediction performance when constructing a GPR model. Fine-tuning a machine learning model entails modifying its hyperparameters to optimize performance on a particular task. Optimal hyperparameter configuration has a direct impact on the model's performance. Grid Search, Random Search, and Bayesian Optimization are popular methods for hyperparameter tuning in machine learning. Grid Search is a simple but computationally expensive method that exhaustively searches all possible combinations of hyperparameters within a predefined range. While Grid Search can guarantee to find the optimal combination of hyperparameters, it can be very time-consuming and computationally expensive, especially when the number of hyperparameters and their possible values are high [62]. Random Search is another method for hyperparameter tuning that randomly samples a set of hyperparameters from a predefined range. Unlike Grid Search, Random Search does not guarantee finding the optimal combination of hyperparameters, but it is generally faster and more computationally efficient than Grid Search [62]. On the other hand, Bayesian Optimization (BO) is a more advanced method that uses probabilistic models and previous evaluations to guide the Search [63]. It can quickly find good combinations of hyperparameters with fewer evaluations than Grid Search and Random Search.

This study adopts Bayesian optimization to find the optimal values of GPR's hyperparameters [63]. Recently, the BO approach has been frequently employed to fine-tune hyperparameters in machine learning methods. This could be attributed to its design using Gaussian processes and Bayesian inference to reach global optimization [60]. Importantly, the BO procedure considers the previous evaluations to determine the hyperparameter set and then evaluates the next step, which enables reducing the time of the optimization procedure [64]. Moreover, the BO procedure can optimize functions without closed form [65]. The optimization process based on the BO procedure needs fewer iterations compared to a grid search optimization.

The main idea of the BO procedure consists of building a probabilistic proxy model for the objective function, employing prior experiments' results as training data [66,67]. Essentially, the proxy model (e.g., the Gaussian process) is not expensive to calculate and enables getting pertinent information on where we should evaluate the true loss function to reach suitable results. Eventually, we consider now the problem of adjusting m hyperparameters $\mathbf{P} = \mathbf{p}_1, \ldots, \mathbf{p}_m$. To this end, the aim is to determine

$$\mathbf{P}^* = \arg \min_{\mathbf{P}} \mathbf{g}(\mathbf{P}|\{(x_i, y_i)\}_{i=1}^n), \tag{10}$$

where $\mathbf{g}$ denotes the the objective function [56]. The optimization procedure is controlled by an appropriate acquisition function that decides the next set of hyperparameters to be evaluated [68].

Figure 9 depicts the basic idea of the BO procedure to optimize the GPR model during the training phase. Specifically, the mean squared error (MSE) between the actual molecular dynamics data and the GPR predictions is computed at each iteration. The optimized model is obtained once the MSE converges to a small value, close to zero.
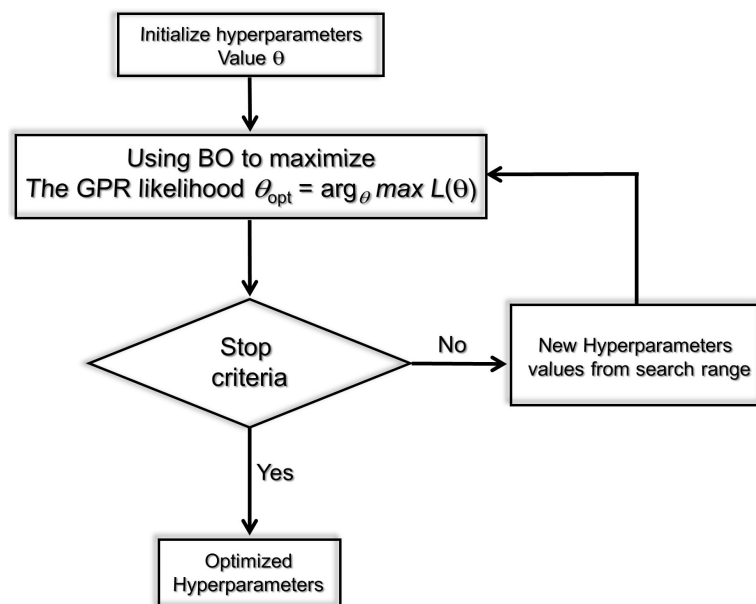

Figure 9. Schematic drawing of the main idea of the BO-based optimized GPR strategy.

4.4. SVR Models

Here, we briefly introduce another widely used approach, the SVR model, a flexible data-based approach with good learning capacity via kernel tricks. The key idea underlying the SVR consists in mapping the train data to a higher dimensional space and conducting linear regression in that space. This SVR model is known to be efficient in dealing with nonlinear regression by using the kernel trick, which enables mapping the input features into high-dimensional feature spaces [69,70]. Moreover, the relevant concept used in designing the SVR model lies in structural risk minimization. It is demonstrated that SVR

provides satisfactory performance with limited samples [71]. Thus, VVR models have been broadly exploited in numerous applications, such as solar irradiance prediction [72], wind power prediction [51], and anomaly detection [50]. This study will use optimized SVR via Bayesian optimization for comparison.

4.5. Bagged Tree Model

Likewise, we discuss another important machine learning model called the bagged tree (BT), also known as bootstrap aggregating [73]. The essence of this model is based on merging the benefit of the bagging technique and decision trees to enhance prediction quality. Importantly, in the BT model, bootstrap sampling is employed to generate numerous samples from the original dataset. Next, multiple distinct decision trees are built, and their outputs are aggregated to obtain the final output [74]. Hence, the prediction quality of the decision trees will be improved, the error will be reduced, and the overfitting issue in individual trees will be bypassed significantly [75,76].

4.6. Evaluation Metrics

In this study, we assess the accuracy of the predicting models using three metrics: root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and R^2.

- RMSE: Root Mean Squared Error measures the differences between predicted and actual values. It is calculated by taking the square root of the average of the squared differences between the predicted and actual values.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2},$$ (11)

- MAPE: Mean Absolute Percentage Error is a measure of the accuracy of a model in predicting values. It is calculated as the average of the absolute percentage difference between predicted and actual values.

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \%,$$ (12)

- MAE: Mean Absolute Error is a measure of the average magnitude of the errors in a set of predictions, without considering their direction. It is calculated as the average of the absolute differences between predicted and actual values.

$$MAE = \frac{\sum_{t=1}^{n} |y_t - \hat{y}_t|}{n},$$ (13)

- R-squared: R^2 is a statistical measure that represents the proportion of the variance for a dependent variable that is explained by an independent variable or variables in a regression model. It is also known as the coefficient of determination, and its value ranges between 0 and 1.

$$R^2 = \frac{\sum_{t=1}^{n} [(y_t - \bar{y}) \cdot (\hat{y}_t - \bar{y})]^2}{\sqrt{\sum_{t=1}^{n} (y_t - \bar{y})^2} \cdot \sqrt{\sum_{t=1}^{n} (\hat{y}_t - \bar{y})^2}},$$ (14)

where y_t is the number of Hamiltonian data, $\hat{y}_t$ is its corresponding forecasted Hamiltonian and non-bond energy, and n is the number of records. Importantly, a higher R^2 value indicates a better fit of the model to the data. For example, if the R^2 value is between 0.7 and 0.9, it is generally considered to indicate a good fit. An R^2 higher than 0.9 indicates a very good fit of the model to the data. It means that the model can explain a large proportion of the variability in the data. on the other hand, a lower MAPE value indicates better model prediction accuracy. Lower RMSE, MAE, and MAPE values would imply better precision

and prediction quality. Generally, a MAPE value of less than 10% is considered good, while a value greater than 10% is considered poor. However, the threshold for an acceptable MAPE value can vary depending on the application and the specific domain. Therefore, it is important to consider the context and specific requirements when interpreting the MAPE value.

4.7. Prediction Framework

This work investigates the feasibility of machine learning methods to predict the Hamiltonian and non-bond energy terms from MD simulations. The prediction methodology employed in this study is depicted in Figure 10. We compared four machine learning methods, including BT, SVR, GPR, and RT. The prediction framework consists of two key steps: model construction and model prediction. The data were first preprocessed to a zero mean and unit variance by subtracting the mean, μ, and dividing by the standard deviation, σ.

$$\mathbf{y}_s = (\mathbf{y} - \mu)/\sigma. \tag{15}$$

After obtaining the predicted values, we reversed this step. The normalized data were then split into training and testing sets. The training set was used to construct the machine-learning models. Here, we employ Bayesian optimization to determine the optimal parameters of the machine learning methods. We applied a five-fold cross-validation technique in training the investigated models. After that, the previously constructed models were then utilized to predict the Hamiltonian (or non-bond energy) in the testing step. The model's accuracy was checked by comparing measured data to predicted data via the score indicators: R^2, RMSE, MAE, and MAPE. An R^2 closer to 1 and lower RMSE, MAE, and MAPE values reflect an accurate prediction.
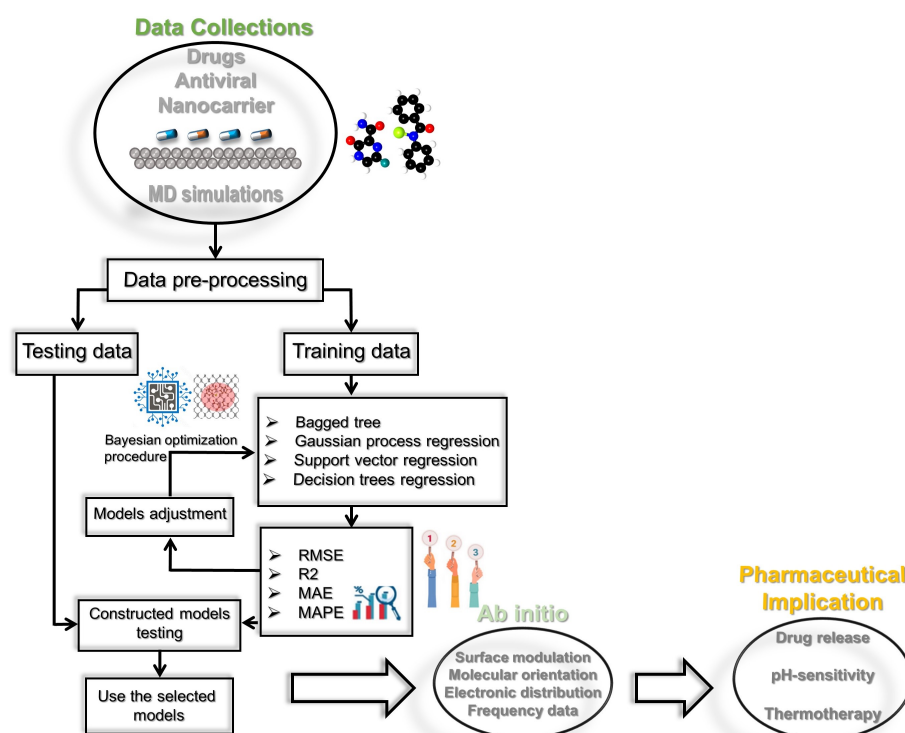

Figure 10. Illustration of the used prediction framework and general steps in antiviral drug discovery for small molecules on black phosphorus, developing quantitative structure–activity relationship models to optimize lead structures for pharmaceutical implication.

4.8. Ab Initio Calculations

One of the most popular approaches in quantum chemistry is density functional theory (DFT) that computes a wide variety of properties of almost any kind of atomistic system, such as molecules, crystals, surfaces, interfaces, and even biomaterial carriers when combined with machine learning. Thus, as the next step of this study, all DFT calculations were carried out using the Vienna Ab Initio Simulation Package (VASP 5.4) [77,78]. All electrons with a projected augmented wave (PAW) formalism were used to model the electron–ion interactions [79]. The exchange–correlation functional contribution to the total energy was modeled using the generalized gradient approximation (GGA). In addition to the modified Perdew–Burke–Ernzerhof (PBE) [79], for treating the van der Waals (vdW), interaction corrections through the D3–BJ approach were incorporated in all calculations [80,81], as established by Grimme et al. [82,83]. We used a (5×4) supercell model of a black phosphorus monolayer containing 80 phosphorus atoms for the nanocarrier to adsorb FP and EB drugs. A vacuum region of 20 Å was introduced perpendicular to the BP film to prevent unphysical periodic effects. Given the large size of the supercell, numerical integration was employed over the Brillouin zone using a $3 \times 3 \times 1$ Γ-centered k-point grid [84]. A denser $5 \times 5 \times 1$ mesh was applied during the self-consistent field (SCF) for electronic and frequency calculations. The pseudo-wave functions were expanded in a plane-wave basis set with an energy cut-off of 450 eV. The accuracy was ensured by adopting energy convergence criteria of 1×10^{-5} eV. Therefore, structural relaxation was performed using the conjugate gradient technique until the force on each atom was below 0.01 eV/Å. This approach has been successfully employed in the past to describe fullerene adsorption on single-layer graphene [6].

5. Conclusions

We have systematically unveiled the interaction characteristics of inhibitor drugs on the surface of a phosphorene sheet by using accelerated molecular dynamics simulations, machine learning, and density functional theory calculations. Importantly, four machine learning models (i.e., BT, GPR, SVR, and RT) were employed to predict Hamiltonian and non-bond energy in EB, FP, and FP + EB. Results indicate the promising prediction performance of machine learning models. Specifically, the optimized GPR model obtained the best performance in this case study compared to the other investigated models. Specifically, the GPR model consistently outperformed the other three models (BT, SVR, and RT) for both Hamiltonian and Non-Bond energy predictions. In particular, the GPR model achieved the highest R^2 value of 0.9649 for Hamiltonian energy prediction and 0.9751 for Non-Bond energy prediction. Additionally, the GPR model had the lowest RMSE, MAE, and MAPE values for both energy predictions, indicating its high accuracy and precision in predicting the energy values. Therefore, the GPR model can be considered the best-performing model among the tested machine-learning models in this study. Further, regarding DFT calculations, we carried these out in a vacuum and an aqueous phase to evaluate the binding ability and thermochemical properties for active drug delivery. We found that the hybrid drugs are physisorbed and enable functionalization of the 2D phosphorene, and they show robust thermostability. All the drug molecules displayed strong van der Waals interactions when combined with the phosphorene sheet. Moreover, the calculated pH values and Gibbs free energy changes at different temperatures suggest that the FP and EB enabled the release of the drug from the 2D monolayer at high temperatures and different pH conditions. These findings would point to drug/2D biosystems as a new and potentially effective way to steer drug vehicle delivery in thermotherapy.

T = 0 K. Table S2: ML, and PBE–D3(BJ) van der Waals energy prediction of MLFF– EB, MLFF–FP, and MLFF–FP_EB adsorption states in (eV) and (kcal/mol) at T = 0 K.

Author Contributions: Conceptualization, S.L. and F.H.; methodology, S.L., A.L. and F.H.; software, S.L., F.H. and A.L.; validation, S.L., B.W., Y.S. and F.H.; formal analysis, S.L. and T.G.; investigation, T.G. and X.G.; resources, S.L. and F.H.; data curation, S.L. and F.H.; writing—original draft preparation, S.L., B.W., T.-M.L.-K. and F.H.; writing—review and editing, S.L. and F.H.; visualization, T.-M.L.-K. and T.G.; supervision, X.G. and T.G.; project administration, T.G.; funding acquisition, T.G. and X.G. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data and results associated with this article are available at https://github.com/SLIM23-CBRC/MD-ML-Hamiltonian (accessed on 1 April 2023).

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

References

1. Kotagiri, N.; Sudlow, G.P.; Akers, W.J.; Achilefu, S. Breaking the depth dependency of phototherapy with Cerenkov radiation and low-radiance-responsive nanophotosensitizers. *Nat. Nanotechnol.* **2015**, *10*, 370–379. [CrossRef] [PubMed]
2. Hong, G.; Lee, J.C.; Robinson, J.T.; Raaz, U.; Xie, L.; Huang, N.F.; Cooke, J.P.; Dai, H. Multifunctional in vivo vascular imaging using near-infrared II fluorescence. *Nat. Med.* **2012**, *18*, 1841–1846. [CrossRef] [PubMed]
3. Yang, K.; Feng, L.; Shi, X.; Liu, Z. Nano-graphene in biomedicine: Theranostic applications. *Chem. Soc. Rev.* **2013**, *42*, 530–547. [CrossRef] [PubMed]
4. Ferrari, M. Cancer nanotechnology: Opportunities and challenges. *Nat. Rev. Cancer* **2005**, *5*, 161–171. [CrossRef] [PubMed]
5. Lin, H.; Chen, Y.; Shi, J. Insights into 2D MXenes for versatile biomedical applications: Current advances and challenges ahead. *Adv. Sci.* **2018**, *5*, 1800518. [CrossRef]
6. Laref, S.; Asaduzzaman, A.; Beck, W.; Deymier, P.A.; Runge, K.; Adamowicz, L.; Muralidharan, K. Characterization of graphene–fullerene interactions: Insights from density functional theory. *Chem. Phys. Lett.* **2013**, *582*, 115–118. [CrossRef]
7. Backes, C.; Abdelkader, A.M.; Alonso, C.; Andrieux-Ledier, A.; Arenal, R.; Azpeitia, J.; Balakrishnan, N.; Banszerus, L.; Barjon, J.; Bartali, R.; et al. Production and processing of graphene and related materials. *2D Mater.* **2020**, *7*, 022001. [CrossRef]
8. Gao, G.; Gao, W.; Cannuccia, E.; Taha-Tijerina, J.; Balicas, L.; Mathkar, A.; Narayanan, T.; Liu, Z.; Gupta, B.K.; Peng, J.; et al. Artificially stacked atomic layers: Toward new van der Waals solids. *Nano Lett.* **2012**, *12*, 3518–3525. [CrossRef]
9. Laref, A.; Alshammari, N.; Laref, S.; Luo, S. Surface passivation effects on the electronic and optical properties of silicon quantum dots. *Sol. Energy Mater. Sol. Cells* **2014**, *120*, 622–630. [CrossRef]
10. Novoselov, K.S.; Fal'ko, V.I.; Colombo, L.; Gellert, P.; Schwab, M.; Kim, K. A roadmap for graphene. *Nature* **2012**, *490*, 192–200. [CrossRef]
11. Wick, P.; Louw-Gaume, A.E.; Kucki, M.; Krug, H.F.; Kostarelos, K.; Fadeel, B.; Dawson, K.A.; Salvati, A.; Vázquez, E.; Ballerini, L.; et al. Classification framework for graphene-based materials. *Angew. Chem. Int. Ed.* **2014**, *53*, 7714–7718. [CrossRef] [PubMed]
12. Laref, A.; Alsagri, M.; Alay-e Abbas, S.M.; Barakat, F.; Laref, S.; Huang, H.; Xiong, Y.; Yang, J.; Wu, X. Impact of phosphorous and sulphur substitution on Dirac cone modification and optical behaviors of monolayer graphene for nano-electronic devices. *Appl. Surf. Sci.* **2019**, *489*, 358–371. [CrossRef]
13. Smith, A.T.; LaChance, A.M.; Zeng, S.; Liu, B.; Sun, L. Synthesis, properties, and applications of graphene oxide/reduced graphene oxide and their nanocomposites. *Nano Mater. Sci.* **2019**, *1*, 31–47. [CrossRef]
14. Laref, A.; Alsagri, M.; Alay-e Abbas, S.M.; Laref, S.; Huang, H.; Xiong, Y.; Yang, J.; Khandy, S.A.; Rai, D.P.; Varshney, D.; et al. Electronic structure and optical characteristics of AA stacked bilayer graphene: A first principles calculations. *Optik* **2020**, *206*, 163755. [CrossRef]
15. Huang, X.; Tang, S.; Mu, X.; Dai, Y.; Chen, G.; Zhou, Z.; Ruan, F.; Yang, Z.; Zheng, N. Freestanding palladium nanosheets with plasmonic and catalytic properties. *Nat. Nanotechnol.* **2011**, *6*, 28–32. [CrossRef]
16. Anasori, B.; Lukatskaya, M.R.; Gogotsi, Y. 2D metal carbides and nitrides (MXenes) for energy storage. *Nat. Rev. Mater.* **2017**, *2*, 1–17. [CrossRef]
17. Laref, S.; Wang, B.; Inal, S.; Al-Ghamdi, S.; Gao, X.; Gojobori, T. A Peculiar Binding Characterization of DNA (RNA) Nucleobases at MoOS-Based Janus Biosensor: Dissimilar Facets Role on Selectivity and Sensitivity. *Biosensors* **2022**, *12*, 442. [CrossRef] [PubMed]
18. Ma, B.; Martín, C.; Kurapati, R.; Bianco, A. Degradation-by-design: How chemical functionalization enhances the biodegradability and safety of 2D materials. *Chem. Soc. Rev.* **2020**, *49*, 6224–6247. [CrossRef]
19. Liu, H.; Du, Y.; Deng, Y.; Peide, D.Y. Semiconducting black phosphorus: Synthesis, transport properties and electronic applications. *Chem. Soc. Rev.* **2015**, *44*, 2732–2743. [CrossRef]
20. Shen, L.; Li, B.; Qiao, Y. Fe_3O_4 nanoparticles in targeted drug/gene delivery systems. *Materials* **2018**, *11*, 324. [CrossRef]

21. Rahimi, R.; Solimannejad, M. BC$_3$ graphene-like monolayer as a drug delivery system for nitrosourea anticancer drug: A first-principles perception. *Appl. Surf. Sci.* **2020**, *525*, 146577. [CrossRef]

22. Hashemzadeh, H.; Raissi, H. Covalent organic framework as smart and high efficient carrier for anticancer drug delivery: A DFT calculations and molecular dynamics simulation study. *J. Phys. D Appl. Phys.* **2018**, *51*, 345401. [CrossRef]

23. Liu, C.; Zhou, Q.; Li, Y.; Garner, L.V.; Watkins, S.P.; Carter, L.J.; Smoot, J.; Gregg, A.C.; Daniels, A.D.; Jervey, S.; et al. Research and development on therapeutic agents and vaccines for COVID-19 and related human coronavirus diseases. *ACS Cent. Sci.* **2020**, *6*, 315–331 . [CrossRef] [PubMed]

24. Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, *582*, 289–293. [CrossRef]

25. Scavone, C.; Brusco, S.; Bertini, M.; Sportiello, L.; Rafaniello, C.; Zoccoli, A.; Berrino, L.; Racagni, G.; Rossi, F.; Capuano, A. Current pharmacological treatments for COVID-19: What's next? *Br. J. Pharmacol.* **2020**, *177*, 4813–4824. [CrossRef]

26. Sreekanth Reddy, O.; Lai, W.F. Tackling COVID-19 using remdesivir and favipiravir as therapeutic options. *ChemBioChem* **2021**, *22*, 939–948. [CrossRef]

27. Guy, R.K.; DiPaola, R.S.; Romanelli, F.; Dutch, R.E. Rapid repurposing of drugs for COVID-19. *Science* **2020**, *368*, 829–830. [CrossRef] [PubMed]

28. Dong, L.; Hu, S.; Gao, J. Discovering drugs to treat coronavirus disease 2019 (COVID-19). *Drug Discov. Ther.* **2020**, *14*, 58–60. [CrossRef]

29. Delang, L.; Abdelnabi, R.; Neyts, J. Favipiravir as a potential countermeasure against neglected and emerging RNA viruses. *Antivir. Res.* **2018**, *153*, 85–94. [CrossRef]

30. Furuta, Y.; Komeno, T.; Nakamura, T. Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. *Proc. Jpn. Acad. Ser. B* **2017**, *93*, 449–463. [CrossRef]

31. Furuta, Y.; Gowen, B.B.; Takahashi, K.; Shiraki, K.; Smee, D.F.; Barnard, D.L. Favipiravir (T-705), a novel viral RNA polymerase inhibitor. *Antivir. Res.* **2013**, *100*, 446–454. [CrossRef] [PubMed]

32. Lynch, E.; Kil, J. Development of ebselen, a glutathione peroxidase mimic, for the prevention and treatment of noise-induced hearing loss. *Semin. Hear.* **2009**, *30*, 047–055 . [CrossRef]

33. Singh, N.; Halliday, A.C.; Thomas, J.M.; Kuznetsova, O.V.; Baldwin, R.; Woon, E.C.; Aley, P.K.; Antoniadou, I.; Sharp, T.; Vasudevan, S.R.; et al. A safe lithium mimetic for bipolar disorder. *Nat. Commun.* **2013**, *4*, 1–7. [CrossRef]

34. Renson, M.; Etschenberg, E.; Winkelmann, J. 2-Phenyl-1,2-benzisoselenazol-3(2H)-one Containing Pharmaceutical Preparations and Process for the Treatment of Rheumatic Diseases. U.S. Patent 4,352,799, 5 October 1982 .

35. Kil, J.; Lobarinas, E.; Spankovich, C.; Griffiths, S.K.; Antonelli, P.J.; Lynch, E.D.; Le Prell, C.G. Safety and efficacy of ebselen for the prevention of noise-induced hearing loss: A randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet* **2017**, *390*, 969–979. [CrossRef] [PubMed]

36. Masaki, C.; Sharpley, A.L.; Cooper, C.M.; Godlewska, B.R.; Singh, N.; Vasudevan, S.R.; Harmer, C.J.; Churchill, G.C.; Sharp, T.; Rogers, R.D.; et al. Effects of the potential lithium-mimetic, ebselen, on impulsivity and emotional processing. *Psychopharmacology* **2016**, *233*, 2655–2661. [CrossRef] [PubMed]

37. Chen, L.; Gui, C.; Luo, X.; Yang, Q.; Günther, S.; Scandella, E.; Drosten, C.; Bai, D.; He, X.; Ludewig, B.; et al. Cinanserin is an inhibitor of the 3C-like proteinase of severe acute respiratory syndrome coronavirus and strongly reduces virus replication in vitro. *J. Virol.* **2005**, *79*, 7095–7103. [CrossRef]

38. Carmo, A.; Pereira-Vaz, J.; Mota, V.; Mendes, A.; Morais, C.; da Silva, A.C.; Camilo, E.; Pinto, C.S.; Cunha, E.; Pereira, J.; et al. Clearance and persistence of SARS-CoV-2 RNA in patients with COVID-19. *J. Med Virol.* **2020**, *92*, 2227–2231. [CrossRef]

39. Khan, M.I.; Shoaib, M.; Zubair, G.; Kumar, R.N.; Prasannakumara, B.; Mousa, A.A.A.; Malik, M.; Raja, M.A.Z. Neural artificial networking for nonlinear Darcy–Forchheimer nanofluidic slip flow. *Appl. Nanosci.* **2022**, 1–20. [CrossRef]

40. Shoaib, M.; Kausar, M.; Khan, M.I.; Zeb, M.; Gowda, R.P.; Prasannakumara, B.; Alzahrani, F.; Raja, M.A.Z. Intelligent backpropagated neural networks application on Darcy-Forchheimer ferrofluid slip flow system. *Int. Commun. Heat Mass Transf.* **2021**, *129*, 105730. [CrossRef]

41. Raja, M.A.Z.; Shoaib, M.; Zubair, G.; Khan, M.I.; Gowda, R.P.; Prasannakumara, B.; Guedri, K. Intelligent neuro-computing for entropy generated Darcy–Forchheimer mixed convective fluid flow. *Math. Comput. Simul.* **2022**, *201*, 193–214. [CrossRef]

42. Varun Kumar, R.; Alsulami, M.; Sarris, I.; Prasannakumara, B.; Rana, S. Backpropagated Neural Network Modeling for the Non-Fourier Thermal Analysis of a Moving Plate. *Mathematics* **2023**, *11*, 438. [CrossRef]

43. Sharma, R.P.; Madhukesh, J.; Shukla, S.; Prasannakumara, B. Numerical and Levenberg–Marquardt backpropagation neural networks computation of ternary nanofluid flow across parallel plates with Nield boundary conditions. *Eur. Phys. J. Plus* **2023**, *138*, 63. [CrossRef]

44. Akkermans, R.L.; Spenley, N.A.; Robertson, S.H. COMPASS III: Automated fitting workflows and extension to ionic liquids. *Mol. Simul.* **2021**, *47*, 540–551. [CrossRef]

45. Laref, S.; Wang, B.; Gao, X.; Gojobori, T. Computational Studies of Auto-Active van der Waals interaction Molecules on Ultra-thin Black-Phosphorus Film. *Molecules* **2023**, *28*, 681. [CrossRef]

46. Zuo, Y.; Warnock, M.; Harbaugh, A.; Yalavarthi, S.; Gockman, K.; Zuo, M.; Madison, J.A.; Knight, J.S.; Kanthi, Y.; Lawrence, D.A. Plasma tissue plasminogen activator and plasminogen activator inhibitor-1 in hospitalized COVID-19 patients. *Sci. Rep.* **2021**, *11*, 1–9. [CrossRef]

47. Cheng, T.; Wang, L.; Merinov, B.V.; Goddard, W.A., III. Explanation of dramatic pH-dependence of hydrogen binding on noble metal electrode: Greatly weakened water adsorption at high pH. *J. Am. Chem. Soc.* **2018**, *140*, 7787–7790. [CrossRef]

48. Ou, W.; Byeon, J.H.; Thapa, R.K.; Ku, S.K.; Yong, C.S.; Kim, J.O. Plug-and-play nanorization of coarse black phosphorus for targeted chemo-photoimmunotherapy of colorectal cancer. *ACS Nano* **2018**, *12*, 10061–10074. [CrossRef]

49. Xie, Y.; Zhao, K.; Sun, Y.; Chen, D. Gaussian processes for short-term traffic volume forecasting. *Transp. Res. Rec.* **2010**, *2165*, 69–78. [CrossRef]

50. Harrou, F.; Saidi, A.; Sun, Y.; Khadraoui, S. Monitoring of photovoltaic systems using improved kernel-based learning schemes. *IEEE J. Photovoltaics* **2021**, *11*, 806–818. [CrossRef]

51. Lee, J.; Wang, W.; Harrou, F.; Sun, Y. Wind power prediction using ensemble learning-based models. *IEEE Access* **2020**, *8*, 61517–61527. [CrossRef]

52. Williams, C.K.; Rasmussen, C.E. Gaussian processes for regression. In *Advances in Neural Information Processing Systems 8, Proceedings of the 1995 Conference, Denver, CO, USA, 27–30 November 1995*; MIT Press: Cambridge, MA, USA , 1996.

53. MacKay, D.J. *Gaussian Processes—A Replacement for Supervised Neural Networks?*; Cambridge University: Cambridge, UK , 1997.

54. García-Nieto, P.J.; García-Gonzalo, E.; Puig-Bargués, J.; Duran-Ros, M.; de Cartagena, F.R.; Arbat, G. Prediction of outlet dissolved oxygen in micro-irrigation sand media filters using a Gaussian process regression. *Biosyst. Eng.* **2020**, *195*, 198–207. [CrossRef]

55. Alkesaiberi, A.; Harrou, F.; Sun, Y. Efficient wind power prediction using machine learning methods: A comparative study. *Energies* **2022**, *15*, 2327. [CrossRef]

56. Alali, Y.; Harrou, F.; Sun, Y. A proficient approach to forecast COVID-19 spread via optimized dynamic machine learning models. *Sci. Rep.* **2022**, *12*, 1–20. [CrossRef] [PubMed]

57. Yang, W.; Deng, M.; Xu, F.; Wang, H. Prediction of hourly $PM_{2.5}$ using a space-time support vector regression model. *Atmos. Environ.* **2018**, *181*, 12–19. [CrossRef]

58. Schulz, E.; Speekenbrink, M.; Krause, A. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions. *J. Math. Psychol.* **2018**, *85*, 1–16. [CrossRef]

59. Seeger, M. Gaussian processes for machine learning. *Int. J. Neural Syst.* **2004**, *14*, 69–106. [CrossRef]

60. Nguyen, V.H.; Le, T.T.; Truong, H.S.; Le, M.V.; Ngo, V.L.; Nguyen, A.T.; Nguyen, H.Q. Applying Bayesian Optimization for Machine Learning Models in Predicting the Surface Roughness in Single-Point Diamond Turning Polycarbonate. *Math. Probl. Eng.* **2021**, *2021*, 1–16 . [CrossRef]

61. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; MIT Press: Cambridge, MA, USA , 2012.

62. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.

63. Protopapadakis, E.; Voulodimos, A.; Doulamis, N. An investigation on multi-objective optimization of feedforward neural network topology. In Proceedings of the 2017 8th International Conference on Information, Intelligence, Systems & Applications (IISA), Larnaca, Cyprus, 27–30 August 2017; pp. 1–6.

64. Shahriari, B.; Swersky, K.; Wang, Z.; Adams, R.P.; De Freitas, N. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* **2015**, *104*, 148–175. [CrossRef]

65. Snoek, J.; Larochelle, H.; Adams, R.P. Practical bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems 25 (NIPS 2012): 26th Annual Conference on Neural Information Processing Systems 2012*; Morgan Kaufmann Publishers, Inc.: Burlington, MA, USA, 2012.

66. Cui, J.; Tan, Q.; Zhang, C.; Yang, B. A novel framework of graph Bayesian optimization and its applications to real-world network analysis. *Expert Syst. Appl.* **2021**, *170*, 114524. [CrossRef]

67. Nikolaidis, P.; Chatzis, S. Gaussian process-based Bayesian optimization for data-driven unit commitment. *Int. J. Electr. Power Energy Syst.* **2021**, *130*, 106930. [CrossRef]

68. Springenberg, J.T.; Klein, A.; Falkner, S.; Hutter, F. Bayesian optimization with robust Bayesian neural networks. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 4134–4142.

69. Yu, P.S.; Chen, S.T.; Chang, I.F. Support vector regression for real-time flood stage forecasting. *J. Hydrol.* **2006**, *328*, 704–716. [CrossRef]

70. Hong, W.C.; Dong, Y.; Chen, L.Y.; Wei, S.Y. SVR with hybrid chaotic genetic algorithms for tourism demand forecasting. *Appl. Soft Comput.* **2011**, *11*, 1881–1890. [CrossRef]

71. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [CrossRef]

72. Lee, J.; Wang, W.; Harrou, F.; Sun, Y. Reliable solar irradiance prediction using ensemble learning-based models: A comparative study. *Energy Convers. Manag.* **2020**, *208*, 112582. [CrossRef]

73. Zhang, Y.; Haghani, A. A gradient boosting method to improve travel time prediction. *Transp. Res. Part C Emerg. Technol.* **2015**, *58*, 308–324. [CrossRef]

74. Bauer, E.; Kohavi, R. An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Mach. Learn.* **1999**, *36*, 105–139. [CrossRef]

75. Harrou, F.; Saidi, A.; Sun, Y. Wind power prediction using bootstrap aggregating trees approach to enabling sustainable wind power integration in a smart grid. *Energy Convers. Manag.* **2019**, *201*, 112077. [CrossRef]

76. Ribeiro, M.H.D.M.; dos Santos Coelho, L. Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. *Appl. Soft Comput.* **2020**, *86*, 105837. [CrossRef]

77. 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. [CrossRef]

78. Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169. [CrossRef] [PubMed]

79. Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **1999**, *59*, 1758. [CrossRef]

80. Perdew, J.P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865. [CrossRef] [PubMed]

81. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [CrossRef]

82. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465. [CrossRef]

83. Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463–1473. [CrossRef]

84. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799. [CrossRef]