

**COMPUTATIONAL INSIGHT INTO GRAPHENE FUNCTIONALIZATION FOR
DNA-SEQUENCING APPLICATION: A DFT APPROACH**

A thesis presented to the Department of Materials Science and Engineering,

African University of Science and Technology, Abuja

In Partial Fulfilment of the Requirements for the Degree of

Master of Science

By

ADNAN ALIYU

(41064)

Supervised by:

Dr. Benjamin O. Tayo

Dr. Vitalis C. Anye



African University of Science and Technology (AUST),

www.aust.edu.ng

P.M.B. 681, Garki, Abuja F.C.T. Nigeria

June, 2024.

**COMPUTATIONAL INSIGHT INTO GRAPHENE FUNCTIONALIZATION FOR
DNA-SEQUENCING APPLICATION: A DFT APPROACH.**

By:

Adnan Aliyu

A thesis approved by the Department of Materials Science and Engineering.

RECOMMENDED:

Supervisor, Dr. Benjamin O. Tayo

.....

Co-supervisor, Dr. Vitalis C. Anye

.....

Head, Dept. of Materials Science and Engineering

Approved:

Vice President, Academics

Date:

© Copyright Adnan Aliyu 2024

ALL RIGHTS RESERVED

ABSTRACT

Most diseases such as cancer, gene mutation or infections among humans are due to DNA nucleotides mis-sequence. Deoxyribonucleic acid (DNA) is vital in life science and its sequence detection is imperative in the field of disease diagnosis, forensic sciences, and genomics systems; making materials design for DNA identification very crucial. Two-dimensional materials such as graphene doped with some hetero-atoms have been explored for DNA nucleobase detection, but the role of functional groups remain unclear. This study investigates the influence of functional groups in the discrimination of DNA nucleotides: Adenine(A), Guanine (G), Thymine (T) and Cytosine (C). Herein, we studied how functional groups like carboxylate, nitrile, alcohol, amine, amide, methyl-iodide, aldehyde, methyl-fluoride, ketone, methyl-bromide, methyl-chloride, ester, methyl-acid iodide, ether, methyl-acid fluoride, methyl-acid chloride, methyl-acid bromide and carboxylic acid improve the adsorption capacity of DNA nucleotides onto graphene sheet. The stable configurations of DNA bases adsorbed onto the graphene surface were investigated using spartan student software and density functional theory (DFT) for quantum chemical calculations. The adsorption energies and band gaps were determined during interaction. Our findings reveal that nine (9) of functional groups namely: Methyl-Carboxylic Acid, Methyl-Ester, Methyl-Ketone, Methyl-Ether, Methyl-F, Methyl-I, Acid-F, Amine and Amide are more promising in the DNA sequencing process based on their appreciable quantum chemical parameters such as highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO), chemical hardness and softness as well as electrophilicity index. These functional groups have higher affinity for some specific DNA bases. However, the adsorption energies of functionalized graphene-based material (ranging from -0.5 to -3.0) has underscored the that of pristine graphene-based materials (ranging from -0.1 to -0.3 eV) and this is an indicator that, common practice of doping heteroatoms or semiconductor atoms are not the only reliable approach of enhancing 2D-materials for DNA sequencing application, since functionalized graphene-based materials have the potential of competing with doped non-functionalized nano-sheet and nano-pore graphene-based material. The relative adsorption energies hierarchy of nucleotides obtained agrees with previous findings reported in the literature. Our findings confirm the potential of computational methods to predict functionalized graphene's selectivity in discriminating DNA nucleotides, offering a promising avenue for identifying mutations driving tumour growth, predicting prognosis and guiding targeted therapies tailored to the unique genetic profile of each patients' disease.

Keywords: Graphene, Adsorption, Functional group, DNA Sequencing, DFT

DEDICATION

To my late father, Aliyu Umar for choosing an everlasting mother (Prof. Addakano Bello Umar) with a dogged determination. I am a proud product of your great discipline, love and support.

TABLE OF CONTENTS

TITLE PAGE	i
ABSTRACT.....	iv
DEDICATION	v
Table of Contents	vi
List of tables.....	viii
List of figures.....	ix
ACKNOWLEDGEMENT	xi
List of Appendices	xii
CHAPTER ONE	1
1.0 INTRODUCTION.....	1
1.1 Background of the Study.....	1
1.2 Problem Statement	2
1.3 Aim and Objectives	2
1.4 Scope of the Study.....	2
1.5 Justification	2
CHAPTER TWO	4
2.0 Literature Review.....	4
2.1 DNA Sequencing.....	4
2.1.1 Methods for DNA sequencing	4
2.2 Two Dimensional (2D) Materials for DNA Sequencing	7
2.2.1 Graphene.....	8
2.2.3 Hexagonal boron nitride (h-BN).....	10
2.3 Computational Approach to DNA Sequencing.....	11
2.3.1 Density functional theory (DFT)	11
2.3.2 Computational DNA sequencing	12
2.4 Quantitative structure activity/property relationship (QSAR/QSPR).....	13
2.4.1 Data set division into modelling and prediction sets.....	14
2.4.2 Model development	14
2.4.3 QSAR Model validation	14
CHAPTER THREE	17
3.0 MATERIALS AND METHODOLOGY	17

3.1	Hardware and Software.....	17
3.2	Calculation of Descriptors.....	17
3.3	Data collection.....	18
3.3.1	Dataset division into modelling and prediction sets	18
3.3.2	Model development	18
3.3.3	Data set division into modelling and prediction sets.....	18
3.4	Geometry Optimization.....	19
3.4.1	Functional groups.....	19
3.4.2	Graphene Functionalization and DNA Adsorption.....	20
CHAPTER FOUR.....		24
4.0	Results and Discussion	24
4.1	Quantum Chemical Parameter Study	24
4.2	Quantitative Structure–Property Relationship (QSPR) Studies	27
4.3	QSAR Model Validation.....	28
4.4	Molecular Dynamic Simulation Studies	31
4.5	Comparing the most stable functionalized GNS and pristine.	43
CHAPTER FIVE		45
5.0	CONCLUSION AND RECOMMENDATION.....	45
5.1	Conclusion.....	45
5.2	Recommendation.....	45
References.....		46
Appendices.....		51

LIST OF TABLES

Table	Title	Page
Appendix B:	2D and 3D Adsorption system of non-functionalized graphene -----	53
Table 2.1	Minimum recommended values of validated parameters for generally acceptable QSAR	16
Table 4.1 (a)	Quantum chemical parameters of interacting species.....	26
Table 4.1(b)	Quantum chemical parameters of interacting species	27
Table 4.2	Residuals from model validation for trained and test data.	29
Table 4.3	Adsorption potential of DNA bases relative to E-gap	32
Table 4.4	Adsorption parameters of DNA nucleobases relative to functional groups.	34
Appendix A	2D and 3D structures of functionalized graphene structure with high sequencing potential.....	51
Appendix B	2D and 3D adsorption system of non-functionalized graphene.	53

LIST OF FIGURES

Figure	Title	Page
Figure 2.1	2D materials, which contain various elementary substances (Es) and compounds (Cp) with the elements of B, C, O, Si, Ge, Mo, V, Pb, etc.....	7
Figure 2.2	Graphene: mother of all graphitic forms.....	9
Figure 2.3	Crystal of MoS ₂ composed of vertically stacked, weakly interacting layers held together by van der Waals interactions	10
Figure 2.4	Structural model of a single-layered BN nanotubes made through wrapping of a planar monatomic BN nanosheet.	10
Figure 2.5	DFT algorithm for computing equilibrium geometry and quantized energy levels of a system.	12
Figure 2.6	Four new concepts using graphene nanostructures for DNA sequencing	13
Figure 3.1	Settings for all interacting species (functional groups).....	20
Figure 3.2	Side and Top view of methyl carboxylic acid functionalized graphene nano-structure.	21
Figure 3.3	Model of the functionalized graphene – DNA base adsorption.....	22
Figure 3.4	Optimized graphene and functionalized graphene – DNA base (Guanine, Adenine, Thymine and Cytosine) adsorption system.	23
Figure 4.1	Bar chart of comparison between trained and predicted E-gap for the various functional groups.	30
Figure 4.2	Comparing the adsorption energy and energy gap of the functionalized GNS system with each adsorbed DNA base.....	35
Figure 4.3	Comparing the adsorption energy and energy gap of the Methyl-ketone functionalized GNR system with each adsorbed DNA base.	36
Figure 4.4	Comparing the adsorption energy and energy gap of the Acid-F functionalized GNS system with each adsorbed DNA base.....	37
Figure 4.5	Comparing the adsorption energy and energy gap of the Methyl-F functionalized GNS system with each adsorbed DNA base.....	38
Figure 4.6	Comparing the adsorption energy and energy gap of the Amide functionalized GNS system with each adsorbed DNA base.....	39
Figure 4.7	Comparing the adsorption energy and energy gap of the Amine functionalized GNS system with each adsorbed DNA base.....	40
Figure 4.8	Comparing the adsorption energy and energy gap of the Methyl-I functionalized GNS system with each adsorbed DNA base.....	41
Figure 4.9	Comparing the adsorption energy and energy gap of the ester functionalized GNS system with each adsorbed DNA base.....	42
Figure 4.10	Comparing the adsorption energy and energy gap of the ether functionalized GNS system with each adsorbed DNA base.....	43

Figure 4.11 Comparison of the most stable DNA-sequencing active models and pristine
GNS44

ACKNOWLEDGEMENT

All praises and salutations are due to almighty Allah for his assistance, guidance, clarity and understanding throughout this journey of accomplishing this milestone. Secondly, my sincere gratitude goes to my supervisors, Asst. Prof. Benjamin O. Tayo (University of Central Oklahoma, USA) and Dr. Vitalis Anye, for their great support, guidance and impartation during my MSc journey in AUST. I really appreciate your patience and understanding, especially the sleepless nights you had in order to offer me the needed support during my thesis research/course work. Thank you very much Sir! I would also like to thank Dr. Toyese Oyegoke (A.B.U -Zaria) for putting me through the computational sessions and his continued support to help me become a better computational materials engineer. I would also like to thank Dr. Abdulhakeem Bello and Dr. Usman Abdulmalik for their support and efforts during my stay in AUST. My sincere thanks also go to my everlasting loving mother (Prof. Addakano Bello Umar), your existence gave me more motivation to seek knowledge beyond accumulation of information to change the African narratives. To all AUST faculties and technologists, my gratitude for all your efforts and support. I wish to also thank the Alhaji Abubakar Waziri Bawa for his caring support during my studies. My sincere gratitude goes to my entire family especially members for their continuous support, prayers and for not giving up on me. I appreciate you all! To my roommate Michael Mozie for his support during my studies. Finally, a big thank you to my friends, colleagues and the entire MSc 2023 set, I appreciate all the efforts and constructive contributions we provided each other to attain our goals. May Allah bless you all!

LIST OF APPENDICES

Appendix A: 2D and 3D structures of functionalized graphene structure with high sequencing potential	51
Appendix B: 2D and 3D Adsorption system of non-functionalized graphene	53

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background of the Study

DNA sequencing does not represent only a fundamental code of life, but also provides an enormous opportunity to improve the well-being of mankind by ushering in a new era of precision medicine for curing deadly diseases [1][2]. Technologies available for successful DNA sequencing include: Ionic current detection through a graphene nanopore, tunnelling across a graphene nanogap, In-plane transport of a graphene nanoribbon with a nanopore and Detection methods based on DNA adsorption [3]. Different materials have been explored for the detection of DNA bases with the goal of providing fast and cheap sequencing of longer strands and single bases as accurately as possible [4]. Among these two dimensional (2D) materials, Graphene remains the most widely studied 2D material for DNA sequencing [5][6]. The π - π interactions between graphene and other molecules results to a significant reduction in its band gap making it to have zero band resulting to identification of DNA bases [5][7]. Attempts have been made by researchers to modify graphene for DNA sequencing applications, one of which is functionalization of the material with heteroatoms rich chemical species [8][9], [10]. Using the nanopore technology approach, [11] studied the interaction of DNA bases with nanopores created in finite-size nanoribbons from non-functionalized graphene, phosphorene, and silicene. It was observed that binding energies of DNA bases using nanopores from phosphorene and silicene are similar and generally smaller compared to those from the non-functionalized graphene. On the other hand, [12] studied functionalized graphene nanopores to verify how effective it is for DNA sequencing, the circular symmetric pores are functionalized with hydrogen and a hydroxyl group bonded with carbon atoms of the pore rim. However, the hydrogenated and hydroxylated pores do not show a clear distinction between bases. Therefore, using the adsorption technology this study will focused on exploring the role of other functional

groups in enhancing the adsorption strength of single-base DNA bases relative to graphene nano-structure (GNS).

1.2 Problem Statement

DNA Sequencing experiments are usually time-consuming, expensive, and deficient in elucidating the mechanism of the system at the Sub-atomic and molecular levels, Two-dimensional materials such as graphene doped with some hetero-atoms have been explored for DNA nucleobase detection, but the role of functional groups remain unclear.

1.3 Aim and Objectives

The aim of this research is to study the role of functionalized graphene-based materials in the adsorption of DNA bases using density functional theory (DFT) approach. This aim will be achieved through the following objectives:

- I. To determine the quantum chemical parameters of all interacting species in the DNA identification process and develop a predictive model via Quantitative Structure Property Relationship (QSPR) analysis.
- II. To determine the adsorption parameters of DNA identification via molecular dynamics simulation.
- III. Comparing the adsorption strength of the most stable functionalized and pristine.

1.4 Scope of the Study

DNA Sequencing simulation of Graphene-based Materials using some selected functional groups to determine adsorption parameters only with a computer software.

1.5 Justification

DNA Sequencing aids in translating genetic information into clearer results for individuals to make decisions on their genetic prepositions and risks. Therefore, deoxyribonucleic acid (DNA) is vital in life science and its sequence detection is imperative in the field of disease

diagnosis, forensic sciences, and genomics systems; making materials design for DNA identification very crucial.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 DNA Sequencing

Significant progress in biology has been driven by knowledge about deoxyribonucleic acid (DNA). DNA is a molecule composed of deoxyribonucleotides connected by phosphodiester linkages [13]. DNA sequence of living organism is called genome and a section of the DNA coding for a protein is called a gene. DNA sequencing consist of determining sequence of nucleotides of an examined DNA fragment cut out from a genome by restriction enzymes or shut gun approach.

2.1.1 Methods for DNA sequencing

Over time different methods had emerge for sequencing of DNA bases which include: Sanger sequencing, Maxam and Gilbert, Cyclic array sequencing, Sequencing by hybridization, Electrophoresis, Mass spectrometry and nanopore sequencing [14] [15].

2.1.1.1 Sanger sequencing

This method is based on chain elongation termination, which is used in DNA sequencing with the help of polymerase and special nucleotides. The elongation of the chain is terminated by synthetic ddNTPs (dideoxynucleoside triphosphate). Dideoxynucleoside triphosphates are designed for each nucleotide separately [14]. Since each base corresponds to a different ddNTP reaction, the Sanger method of DNA sequencing creates a total of 4 distinct reactions for each base. Novelty in this approach the use of polyacrylamide gels to separate the products or primed synthesis by DNA polymerase in order to increase chain length. The enzyme method for DNA sequencing has been used for genomic research as the main tool to generate the fragments necessary for sequencing, regardless of the sequencing strategy [15]. Two different approaches, shotgun and primer walking sequencing, are the most used for DNA sequencing. Shotgun

sequencing is a random process where DNA is randomly fragmented into smaller pieces, which produces a high level of redundancy there-by increasing the total cost. Second approach represents direct sequencing of unknown DNA where sequence is known. Unknown sequence of DNA is inserted into a vector and amplified [15].

2.1.1.2 Maxam and Gilbert

The DNA sequencing method that was similar to the Sanger and Coulson method in using polyacrylamide gels to resolve bands that terminated at each base throughout the target sequence, but very different in the way that products ending in a specific base were generated [14] . There are three main advantages of the Maxam Gilbert and other chemical methods compared with Sanger's chain termination reaction method. The First advantage is that a fragment can be sequenced from the original DNA fragment, instead of from enzyme copies. Secondly, advantage is no subcloning and no PCR reactions are required [16]. Consequently, for the location of rare bases, the chemical cleavage analysis cannot be replaced by the di-deoxynucleotide terminator method, as the latter analyses the DNA of interest via its complementary sequence, it can, thus, solely providing the sequence information in terms of the four canonical bases. Third advantage relies on the fact that this method is less susceptible to errors with regard to sequencing of secondary structures or enzyme mistakes. Some of the chemical protocols are recognized by different authors as being simple, easy to control, and give clear distinction between the bases [17] [18].

2.1.1.3 Cyclic array sequencing

Platforms performing cyclic array are involved in simultaneously decoding of a two-dimensional array. Main characteristic here is that features are not necessarily separated into individual wells. The Complete process is cyclic in nature because the enzymatic process is applied to interrogate the identity of a single base position for all features in parallel within each cycle [19]. Accuracy of the individual reads. The most popular and widely used platforms

that found commercial use are Roche 454, Illumina, ABI's SOLiD and Pacific Biosciences sequencing. Roche 454 sequencing allows shotgun sequencing of genomes without cloning in *E. coli* or any host cell [20]-[21]. The 454 system was the first next-generation sequencing platform available as a commercial product 454 Genome Sequencer FLX is able to produce 100 Mb sequence with 99.5 % accuracy in over 250 bases length. Illumina sequencing is similar to Roche 454 only difference is that this method uses chain-terminating nucleotides. This machine can analyse more than one billion bases of 75 base reads in a single run [22]-[23].

2.1.1.4 Sequencing by hybridization

The main idea here is to build a two-dimensional grid with k-tuples which is a word of length k. This matrix represents the sequencing chip. DNA is labelled with radioactive material, and each k-tuple present in the sample is hybridized with its reverse complement in the matrix and unhybridized DNA is removed from the matrix [24].

2.1.1.5 Electrophoresis

As a result of the problems, new method for sequencing was invented. Capillary electrophoresis is a fast technique for analysis. This technique can resolve complex mixtures of biopolymers in a high electric field [15].

2.1.1.6 Mass spectrometry has found its place in DNA sequencing.

This method relies on the precise measurement of the masses of DNA fragments. This method shows better results for RNA analysis and it is predicted that this method will not displace conventional methods for most DNA sequence applications [14].

2.1.1.7 Nanopore sequencing

The development of nanopore devices to sequence DNA is an area of intense research that promises exciting results. The essential idea is to make single stranded DNA (ssDNA) pass through (or 'translocate') a nanopore whose diameter is of the order of a few atoms. The

passage of individual bases in the ssDNA is sensed using electrical or optical means [25]. The resulting signal is processed to reveal the identity (i.e., A, C, G or T) of the individual bases as they pass through the nanopore [26].

2.2 Two Dimensional (2D) Materials for DNA Sequencing

2D Materials, in which only one dimension is restricted, are crystalline materials consisting of a single layer of atoms. 2D crystal was thought to be unstable in nature for a long time until 2004 when graphene, a one-atom-thick honeycomb structure composed of carbon atoms, was successfully prepared. Up to now, as schematic shown in Figure 2.1, hundreds of 2D materials have been discovered, and most of them are simple substances from group-IV or binary compounds including metal chalcogenides and complex oxides from group III–V elements [27].

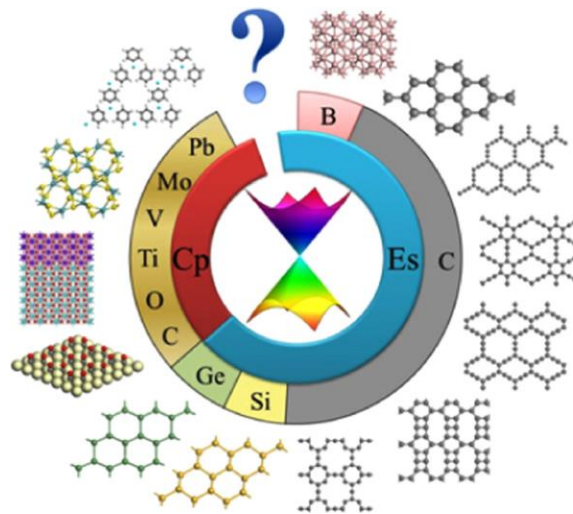


Figure 2.1 2D materials, which contain various elementary substances (Es) and compounds (Cp) with the elements of B, C, O, Si, Ge, Mo, V, Pb, etc.

Due to unique electronic and mechanical properties, 2D materials are expected to have a significant impact on a large variety of applications, ranging from electronics to high performance sensors, gas separation, catalysis, support membranes and so on [27]. As one atomic layer material with comparable thickness to DNA base pair stacking distance of ~ 3.4 Å, 2D material nanopore is a promising device for DNA sequencing. However, there are still

some disadvantages for DNA sequencing by 2D material, especially the detection accuracy is still low. Among all these 2D materials, conductible graphene (known thinnest material), semi-conductible h-BN and MoS₂ (with a large intrinsic bandgap of 1.8 eV and high SNR for DNA sequencing) have aroused many researchers' attentions. Therein, the 2D material made from carbon (graphene), MoS₂ and h-BN are extremely important [28].

2.2.1 Graphene

Graphene, discovered in 2004 has attracted much attention for its unique electrical and mechanical properties. Tremendous research articles and reviews have been published on synthesis and physical properties of graphene [28], [29]. Graphene is one of the first real two-dimensional material (i.e., one atom thick), and it was exfoliated from graphite in nature. Schematically shown in Figure 2.2, graphene could be used to form other type of carbon such as carbon nanotubes, fullerene etc. (. The covalently bonded honeycomb lattice of the graphene sheet shows high stability and unique properties. The electronic mobility of graphene could be as high as $250,000 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. As a alleged zero gap semiconductor, the valence and conduction band in graphene touch with each other in single point [29], [30] . The fragility of a single point contact makes the band structure of graphene highly sensitive to any change, such as external electric fields, mechanical deformations and doping. It is very desirable for sensing applications including DNA [31].

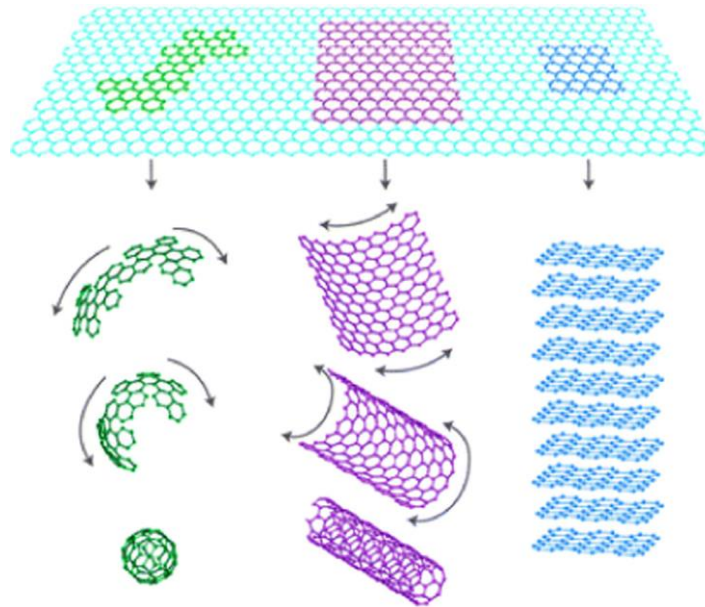


Figure 2.2 Graphene: mother of all graphitic forms.

Reprinted permission from [32]

2.2.2 Molybdenum Disulphide (MoS_2)

Single layer MoS_2 can be extracted from the crystal of MoS_2 composed of vertically stacked, weakly interacting layers held together by van der Waals interactions (Figure 2.3) by using scotch tape [33] or lithium-based intercalation. Unlike pristine graphene with no bandgap, a single layer of MoS_2 has a large intrinsic bandgap of 1.8 eV [34]. Thus, the application of MoS_2 ranges from gas sensor to phototransistors and so on [35]. For example, MoS_2 field effect transistor devices could be fabricated by mechanically exfoliated one layer to few layers MoS_2 nanosheets [36].

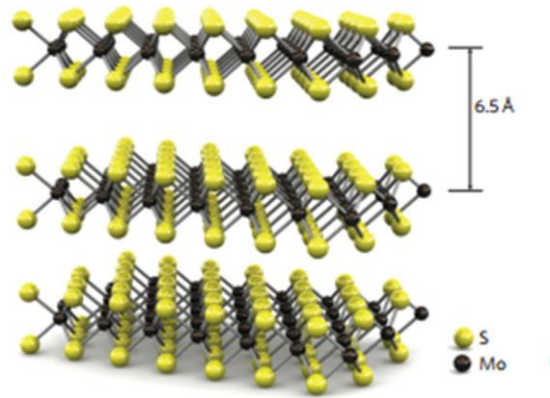


Figure 2.3 Crystal of MoS₂ composed of vertically stacked, weakly interacting layers held together by van der Waals interactions [32], [33].

2.2.3 Hexagonal boron nitride (h-BN)

h-BN is a layered material with a graphite-like structure in which planar networks of BN hexagons are regularly stacked [37], [38]. A high-yield monatomic layer h-BN sheet has been fabricated by using a sonication-centrifugation technique where milligram quantity of h-BN sheets was achievable and ultimately pure [39]. The structure of h-BN sheet is similar to graphene, and they share the same honeycomb lattice structure. However, unlike graphene sheet (surface modification is necessary to open a band gap in graphene), h-BN sheet is a natural semiconductor [40]. As schematically shown in Figure 2.4, the h-BN nanotubes could be rolled from h-BN sheet. The h-BN nanotubes is an electrical insulator, which is really different from metallic carbon nanotubes [39].

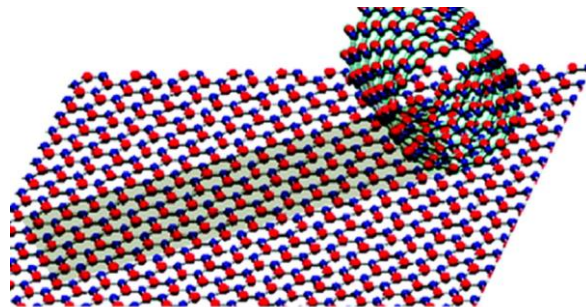


Figure 2.4 Structural model of a single-layered BN nanotubes made through wrapping of a planar monatomic BN nanosheet.

2.3 Computational Approach to DNA Sequencing

2.3.1 Density functional theory (DFT)

Density Functional Theory (DFT) is certainly the most widely applied ab initio approach for modelling the ground-state (GS) properties of molecules, clusters and solids. Interestingly this model was originally viewed as unable to describe chemical bonds, but continuous improvements of the exchange–correlation functionals – the only approximated term in DFT, has allowed DFT to become a blockbuster in both chemistry and physics. The most fundamental parameter in DFT is the electron density $\rho(\mathbf{r})$, in terms of which all the chemical quantities are expressed [41]. The structural parameters calculated through the electronic density ($\rho(\mathbf{r})$ concept) compare well with the parameters calculated by the Schrodinger equation in terms of the single-electron wavefunction (Ψ), where the latter can be applied [42]. Since the theory is simpler than classical quantum mechanics, which relies upon solving the many-bodied Schrodinger equation with poorly-known electron correlation factors, interest has grown in understanding the structure, properties, reactivity, and dynamics of atoms, molecules and clusters using DFT, and large systems involving many atoms or molecules are now simulated on a regular basis. In the field of reaction chemistry, DFT overcomes the limitations of wave mechanics, by reducing the complexity (number of degrees of freedom) of the problem by dealing with electron density functionals, each of which comprise the wave functions for many electrons, and it is emerging as a unique approach for the study of reaction mechanisms, including defining the configurations and energetics of transition states[41], [42]. DFT calculation progresses with an algorithm in trying to solve an approximate schrodinger equation as shown in Figure 2.5

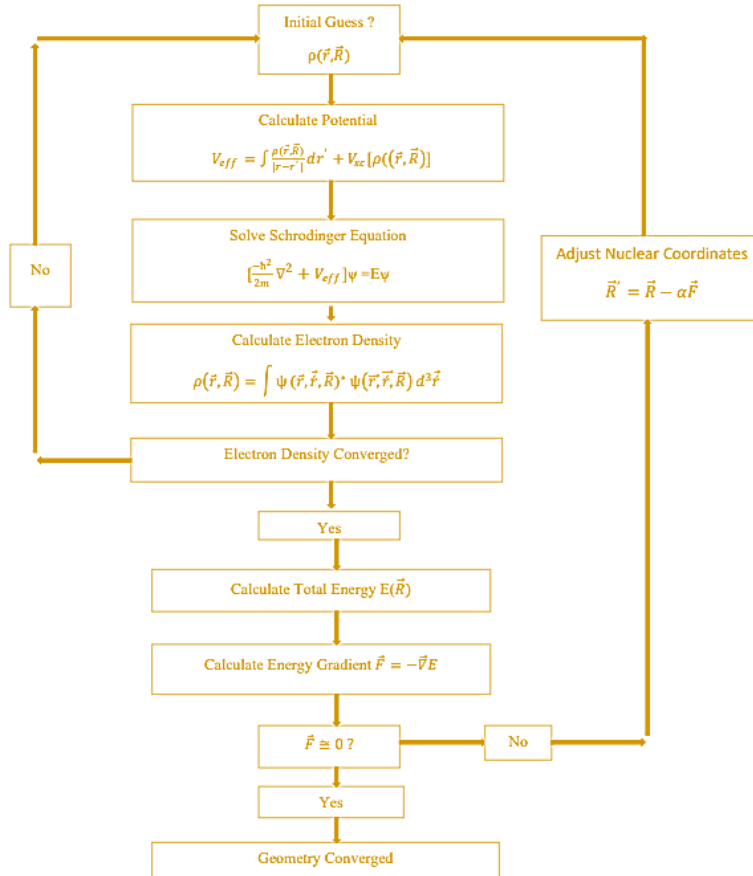


Figure 2.5 DFT algorithm for computing equilibrium geometry and quantized energy levels of a system.

2.3.2 Computational DNA sequencing

There are four (4) different concepts used in the study of DNA sequencing using 2D materials which are:

- a. Ionic current detection through a graphene nanopore.
- b. Tunnelling across a graphene nanogap.
- c. In-plane transport of a graphene nanoribbon with a nanopore.
- d. Detection methods based on DNA adsorption.

An illustration of these four (4) DNA sequencing concepts is presented in Figure 2.6.

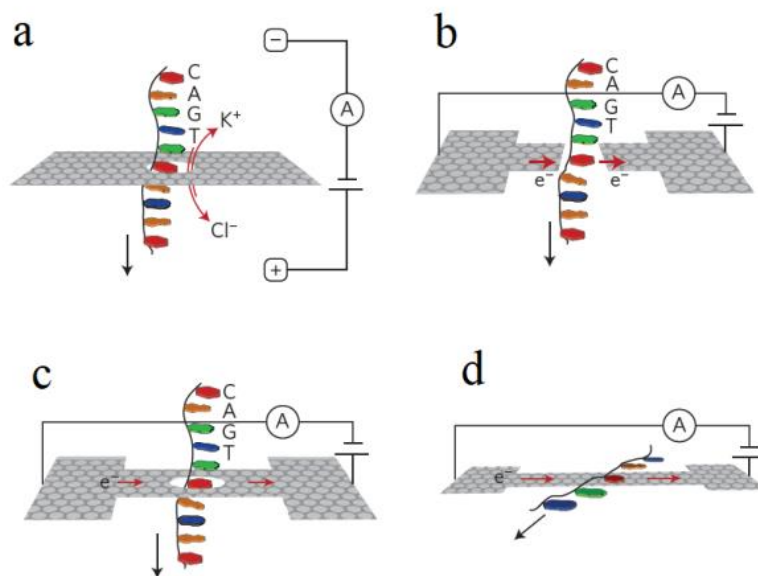


Figure 2.6 Four new concepts using graphene nanostructures for DNA sequencing [43].

a, Detection of changes in the ionic current through a nanopore in a graphene membrane due to the passage of a DNA molecule. **b**, Modulations of a tunnelling current through a nanogap between two graphene electrodes due to presence of a DNA molecule. **c**, Variations in the in-plane current through a graphene nanoribbon due to traversal of a DNA molecule. **d**, Changes in a graphene current due to the physisorption of DNA bases onto the graphene.

2.4 Quantitative structure activity/property relationship (QSAR/QSPR)

Quantitative structure-activity relationships (QSAR) have been a subject of intense interest in the field of medicinal chemistry in determining the molecular structure as well as elucidating the electronic structure and reactivity, and to a small extent in the field of corrosion. However, in recent years, a quantitative structure-activity relationship (QSAR) has aroused many researchers' interests in the studies of corrosion inhibitors [44]. The success of the QSAR approach can be explained by the insights offered for the structural determination of chemical properties, and the possibility of estimating the properties of the new chemical compounds without any need for them to be synthesized and tested [45]. QSAR utilizes linear regression of statistical analysis to build mathematical equation models, which could elucidate the relationship for molecular structures of the compounds with their potential biological activities. QSAR will also help to create a preliminary hypothesis regarding the mechanism of action by investigating the compounds on a particular biological system. Through this approach, it is

assumed that the compounds that fit in a QSAR model are acting with the same mechanism of action [44], [46].

2.4.1 Data set division into modelling and prediction sets

The data set was divided into two sets, the modelling and prediction set. The modelling set is used in developing the model, it contains seventy percent (70%) of the entire data set. While the test set which constitutes the remaining thirty percent (30%) of the whole data set was not used in the construction of the model but to ascertain the predictive ability of the model. This partitioning ensures that a similar principle can be employed for the activity prediction of the test set. Kennard-Stone Algorithm will be applied for dividing Dataset into a modelling and test set [47].

2.4.2 Model development

In QSAR studies the identification and selection of descriptors which provide maximum information in activity variations and have minimum co-linearity are important. Therefore, a genetic function algorithm (GFA) improves the model accuracy in the selection of proper descriptors. Multiple Linear Regression was used on the modelling set to show the relationship between the dependent variable Y (%IE) and independent variable X (molecular descriptors). In regression analysis, the contingent mean of the dependent variable (%IE) Y relies on (descriptors) X [48].

2.4.3 QSAR Model validation

Validation of a QSAR model's stability and predictive ability is another key step in QSAR modelling. Different statistical parameters have been used for the validation of the suitability of the built model for the prediction of the anti-cancer activity of the studied compounds. This includes the square of the correlation coefficient (R^2) this describes the fraction of the total variation attributed to the model. The closer the value of R^2 is to 1.0, the better the regression

and equation explain the Y variable. R^2 is the most commonly used internal validation indicator and is expressed as in equation (1):

$$R^2 = 1 - \frac{\sum(Y_{exp} - Y_{pred})^2}{\sum(Y_{exp} - Y_{m\ training})^2} \quad (1)$$

Where, Y_{exp} ; Y_{pred} ; $Y_{m\ training}$ are the experimental property, the predicted property and the mean experimental property of the samples in the training set, respectively [49]. The minimum recommended value for this parameter as shown in Table 1.

Adjusted R^2 (R^2_{adj}): R^2 value varies directly with the increase in the number of regressors i.e., descriptors, thus, R^2 cannot be a useful measure for the goodness of model fit. Therefore, R^2 is adjusted for the number of explanatory variables in the model. The adjusted R^2 is defined as in equation (2):

$$R^2_{adj} = 1 - (1 - R^2) \frac{N-1}{N-P-1} = \frac{(N-1)R^2 - P}{N-P+1} \quad (2)$$

Where p = number of independent variables in the model and N = sample size. The minimum recommended value for this parameter as shown in Table 2.1.

Cross-validation coefficient parameter (Q^2_{cv}), which is the most commonly used internal validation indicator and is expressed as in equation (3):

$$Q^2_{cv} = 1 - \frac{\sum(Y_{pred} - Y_{exp})^2}{\sum(Y_{exp} - Y_{m\ ntrng})^2} \quad (3)$$

where Y_{exp} is the experimental activity, Y_{pred} is the predicted activity, and $Y_{m\ training}$ is the mean of the experimental activity of the validation set. However, it should be noted that a high Q^2_{cv} does not necessarily mean high predictability of the developed model. In other words, the high value of Q^2_{cv} is a necessary condition, but not sufficient for a developed model to have high predictability [50].

Therefore, validation through an external prediction set (R^2_{test}), is a very paramount parameter that is used to test the external predictive ability of a QSAR model. The R^2_{test} value is calculated by equation (4):

$$R^2_{\text{test}} = 1 - \frac{\sum(Y_{\text{pred}} - Y_{\text{exp}})^2}{\sum(Y_{\text{exp}} - Y_{\text{mntrng}})^2} \quad (4)$$

where Y_{exp} is the experimental activity, Y_{pred} is the predicted activity, and Y_{mntrng} is the mean of the experimental activity of the validation set.

Table 2.1 Minimum recommended values of validated parameters for generally acceptable QSAR [48]-[50].

S/N	Symbol	Name	Value
1	R^2	Coefficient of determination	≥ 0.6
2	Q^2	Cross-validation coefficient	< 0.5
3	$R^2 - Q^2$	Difference between R^2 and Q^2	≤ 0.3
4	N_{test}	Minimum number of external test set	≥ 5
5	R^2	Coefficient of determination	≥ 0.6

CHAPTER THREE

3.0 MATERIALS AND METHODOLOGY

3.1 Hardware and Software

The hardware and software that was used include Acer Nitro-5 Intel (R) Core (TM) i5-8300H CPU 8.00GB RAM @ 2.30GHz and 4.00GHz turbo speed processor, 4GB (GTX1050) dedicated graphics card, on Windows 11 64-bit Operating system, × 64-based processor, Spartan Student v9.0.3 software and Microsoft Office Excel 2020.

3.2 Calculation of Descriptors

In this work, quantum chemical calculations were carried out with a Spartan Student v9.0.3 Wave-Function programming Package. Quantum chemical descriptors generated were the energy of the highest occupied molecular orbital (E-HOMO), the energy of the lowest unoccupied molecular orbital (E-LUMO), dipole moment, polarizability and some others were calculated such as energy band gap ($\Delta E = E_{\text{LUMO}} - E_{\text{HOMO}}$), global electronic chemical potential (μ), chemical softness (σ), chemical hardness (η) and solvation energy using appropriate relations (Equations 1–5) as previously reported in literature [42], [47], [51].

$$\Delta E = E_{\text{HOMO}} - E_{\text{LUMO}} \quad (5)$$

$$\eta = -\frac{1}{2}(E_{\text{HOMO}} - E_{\text{LUMO}}) \quad (6)$$

$$\sigma = \frac{1}{\eta} \quad (7)$$

$$\omega = \frac{\mu^2}{2\eta} \quad (8)$$

$$E_{\text{ads}} = E_{\text{graphene+DNA base}} - E_{\text{graphene}} - E_{\text{DNA base}} \quad (9)$$

Where: E_{ads} = Adsorption energy, $E_{\text{graphene+DNA base}}$ = Energy of the adsorption system, E_{graphene} = Energy of graphene and $E_{\text{DNA base}}$ = Energy of the DNA base

3.3 Data collection

Eighteen (18) primary carbon functional groups as potential pristine graphene were collected from the quantum chemical parameters calculation carried out using spartan student v9.0.3 which was used for this present study upon applying equations (5 - 8).

3.3.1 Dataset division into modelling and prediction sets

The data set was divided into two sets, the modelling and prediction set. The modelling set is used in developing the model, it contains seventy percent (70%) of the entire data set which are: Carboxylate, Carboxylic acid, Nitrile, Alcohol, Ester, Aldehyde, Ketone, Ether, Methyl-F, Methyl-Cl, Methyl-Br, Methyl-I, Acid-F. While the test set which constitutes the remaining thirty percent (30%) of the whole data set was not used in the construction of the model but to ascertain the predictive ability of the model and those functional groups are: Acid-Cl, Acid-Br, Acid-I, Amine, Amide.

3.3.2 Model development

In QSAR studies the identification and selection of descriptors which provide maximum information in activity variations and have minimum co-linearity are important. Therefore, a genetic function algorithm (GFA) improves the model accuracy in the selection of proper descriptors. Multiple Linear Regression was used on the modelling set to show the relationship between the dependent variable Y (E_{ads}) and independent variable X (polarizability, ovality, polar surface area, PSA., dipole moment, solvation energy). In regression analysis, the contingent mean of the dependent variable (E_{ads}) Y relies on (polarizability, ovality, polar surface area, PSA, dipole moment, solvation energy) X.

3.3.3 Data set division into modelling and prediction sets

The data set was divided into two sets, the modelling and prediction set. The modelling set is used in developing the model, it contains seventy percent (70%) of the entire data set. While

the test set which constitutes the remaining thirty percent (30%) of the whole data set was not used in the construction of the model but to ascertain the predictive ability of the model. This partitioning ensures that a similar principle can be employed for the activity prediction of the test set. Kennard-Stone Algorithm will be applied for dividing Dataset into a modelling and test set.

In order to assess the predictive ability and to check the statistical significance of the developed models, the proposed models were applied to predict the values of %IE of an external set that was not used in the development of the model. The predictive powers of the developed regression models on the training set were evaluated by predicted values of the prediction set.

3.4 Geometry Optimization

3.4.1 Functional groups

The 2D structure of each of the functional groups (Methyl Amine, Methyl Amide, Methyl ketone, Methyl Ether, Methyl-I, Methyl Alcohol, Methyl Aldehyde, Methyl Acid-I, Methyl Ester, Methyl-Br, Methyl Carboxylic, Methyl Acid-Br, Methyl Carboxylate, Methyl-Cl, Methyl Acid-F, Methyl Nitrile, Methyl-F) was drawn using spartan student v9.0.3 and were converted to the 3D structure. The structures were cleaned by minimizing and checking using a density function theory (DFT) option on Spartan student, so as to remove all strain from the structure of the molecule. Additionally, this will guarantee a well-defined and stable conformer relationship within the compounds in the study. Geometry optimization was set at the ground state utilizing the density functional theory (DFT) at the Becke88 three-parameter hybrid exchange potentials with Lee-Yang-Parr correlation potential (B3LYP) level of theory and for the basis set 6-31G* was selected. The detailed of the calculation is presented in figure 7 for carboxylate functional of which same settings was made for all other functions. The QSAR option was checked to enable the computation of some of the functional group structural

properties predictors such as: solvation energy, polarizability, dipole moment, ovality, LogP and polar surface area (PSA).

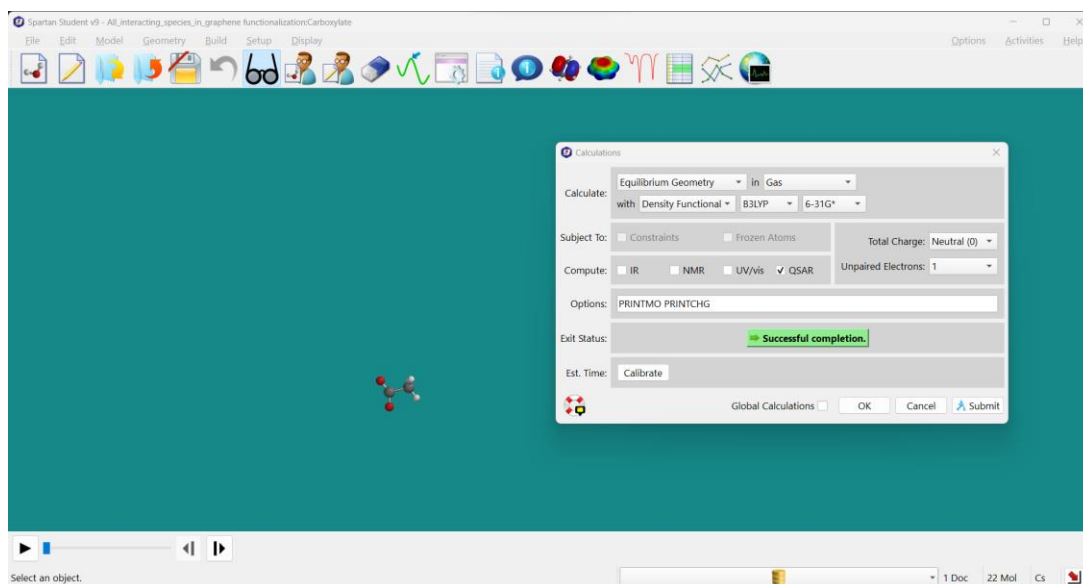


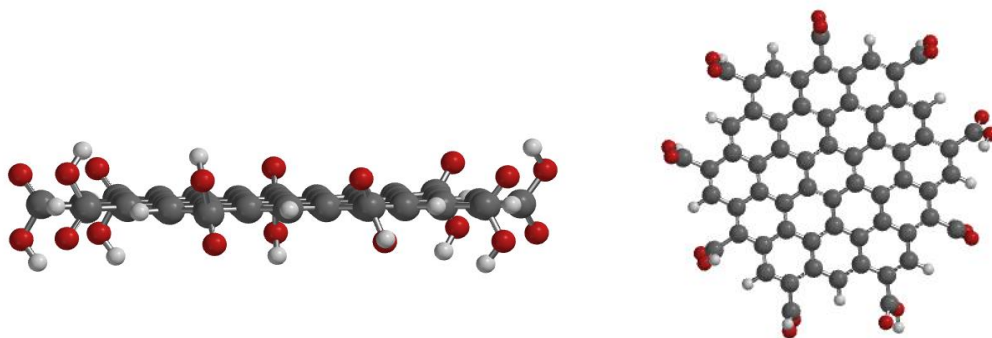
Figure 3.1 Settings for all interacting species (functional groups).

The figure above is an illustration of spartan student v9 user interface used in this work, the small pop-up box is for setting the DFT calculation parameters, method and type of molecular interaction.

3.4.2 Graphene Functionalization and DNA Adsorption.

3.4.2.1 Graphene functionalization

A 3D 12Å diameter single-atom layer graphene nanosheet structure was developed using spartan student and alternately bonded to participating functional groups relative to the four (4) DNA bases as shown in Figure 3.2.



(a). Side view

(b). Top View

Figure 3.2 Side and Top view of methyl carboxylic acid functionalized graphene nanostructure.

Same functionalization pattern was carried out for relative to all functional groups, the detail is presented in appendices.

3.4.2.2 Adsorption of DNA bases on functionalized graphene.

Physisorption of DNA bases was carried out on functionalized graphene with an adsorption distance of 3\AA forming an adsorption system. This system was optimized to calculate its equilibrium geometry energy which form the basis for determining the adsorption energy using equation 9 above.

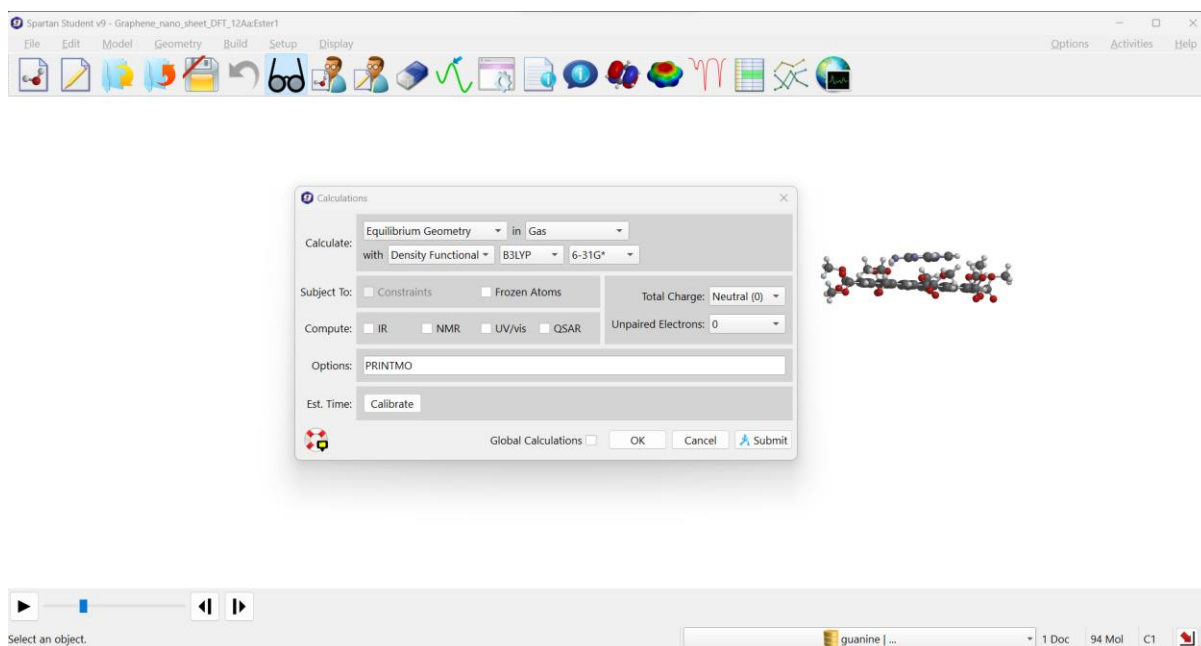
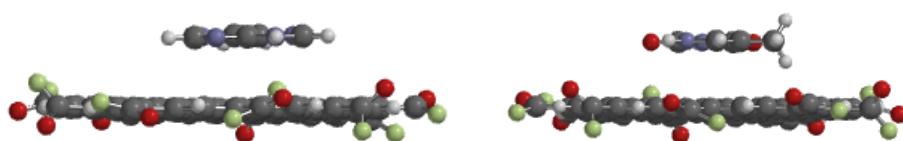
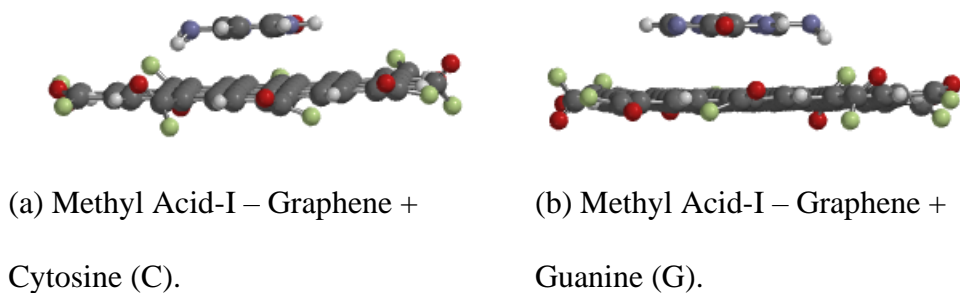


Figure 3.3 Model of the functionalized graphene – DNA base adsorption.

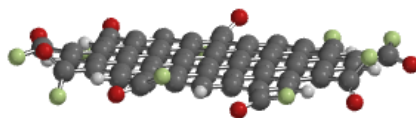
The setting for functionalized graphene – DNA base adsorption is presented in Figure 3.3 using density functional option to calculate the equilibrium adsorption energy of the interaction with B3LYP functional and 6-31G* basis set.

Figure 3.4 is an illustration of the optimized graphene and functionalized graphene – DNA base adsorption system, and the detail relative to all functional groups is available in appendices.



(c) Methyl Acid-I – Graphene +
Adenine (A).

(d) Methyl Acid-I – Graphene +
Thymine (T).



(e) Optimized structure of Methyl Acid-I – Graphene

Figure 3.4 Optimized graphene and functionalized graphene – DNA base (Guanine, Adenine, Thymine and Cytosine) adsorption system.

CHAPTER FOUR

4.0 RESULTS AND DISCUSSION

4.1 Quantum Chemical Parameter Study

Quantum chemical calculations were performed on the eighteen (18) different functional groups in order to relate their adsorption potentials to molecular structures as shown in Table 4.1 (a) and (b). According to the frontier orbital approximation, donor-acceptor interactions do occur between frontiers molecular orbitals (HOMO and LUMO) of reacting species [38]. The adsorption process of a functional group molecule onto a graphene-based material surface increases with an increase in the HOMO energy (E-HOMO) and a decrease of the LUMO energy (E-LUMO). This is because, from the HOMO orbital the functional group molecule will donate the electrons to the d-orbital of the graphene nano structure (GNS), and the LUMO orbital of the functional group will receive the electrons from the d-orbital of the graphene material, in-electron-donation, and electron back-donation process. Thus, E-HOMO is often associated with the electron donating ability of a molecule; a high value of E-HOMO indicates the tendency of the molecule to donate electrons to the acceptor graphene material. Frontier molecular orbital energies of the studied functional groups (E-HOMO and E-LUMO) are listed in Table 4.1 (a). According to the frontier molecular orbital theory, a high E-HOMO value for a molecule implies the tendency of the molecule to donate electrons to the appropriate vacant d-orbital of the metal/semi-metal atoms. The higher the HOMO energy the more the probability of a molecule to donate an electron to the DNA bases. From the results obtained, E-HOMO decreases in the order: Amine > Amide > ketone > Ether > Methyl-I > Alcohol=Aldehyde > Acid-I > Ester > Methyl-Br > Carboxylic > Acid-Br > Carboxylate > Methyl-Cl > Acid-F > Nitrile > Methyl-F. A comparison of the E-HOMO of the studied functional group shows that Amine has the highest while Methyl-F has the lowest, implying that the amine group would be the most preferred electron donor while Methyl-F would be the least preferred.

On the other hand, a low E-LUMO value for a molecule indicates the relative tendency of the molecule to accept electrons from GNR orbital during back-donation. From Table 4.1 (a), the E-LUMO value in the descending order are: Methyl-F > Ether > Amine > Nitrile > Amide > Carboxylate > Methyl-Cl > Ester > Carboxylic > Methyl-Br > ketone > Acid-F > Aldehyde > Methyl-I > Acid-Cl > Methyl-I > Acid-Br > Acid-I. This implies that Acid-I has a stronger tendency to accept electrons from GNR orbital as they adsorb on its surface. Therefore, Acid-I significantly adsorb on the GNR surface to achieve better DNA identification efficiency.

Energy gap E-gap (E-HOMO - E-LUMO) is an important parameter as efficient DNA sequencing is characterized by a small energy gap [39]. From Table 4(b), the trend of decreasing the energy gap shows that: Methyl-F > Nitrile > Carboxylate > Methyl-Cl > Ether > Acid-F > Carboxylic > Ester > Amide > Methyl-Br > Acid-Cl > Amine > Acid-Br > Aldehyde = ketone = Alcohol > Methyl-I > Acid-I. This suggests that Acid-I with the lowest energy gap of 5.68 eV is the best functional group in terms of this parameter among the series.

The Functional group with the large hardness (η) value is expected to be a weaker functionalizing agent compared to those with lower values. From the results in Table 4(a) the hardness value decreases in the order Methyl-F > Nitrile > Carboxylate > Methyl-Cl > Ether > Acid-F > Carboxylic > Ester > Amide > Methyl-Br > Acid-Cl > Amine > Acid-Br > Aldehyde = ketone = Alcohol > Methyl-I > Acid-I. This indicates that Acid-I with the lowest value of 2.84 will be the best functional group in DNA discrimination.

Additionally, adsorption usually occurs in the region of the molecule where chemical softness (σ) has the highest value. The order across structures in the σ values as reported in Table 4.1 (a) and decreases in the order, Acid-I > Methyl-I > Aldehyde = ketone = Alcohol > Acid-Br > Amine > Acid-Cl > Methyl-Br > Amide > Ester > Carboxylic > Acid-F > Ether > Methyl-Cl >

Carboxylate > Nitrile > Methyl-F. This suggests that Acid-I has the highest and therefore the most reactive functional group molecule.

Table 4.1 (a) Quantum chemical parameters of interacting species.

S/N	Species	Formula	Energy (ev)	E-HOMO (ev)	E-LUMO (ev)	E-gap (ev)	η	σ
1	Carboxylate	C ₂ H ₃ O ₂	-6216.86	-8.25	0.39	8.64	4.32	0.23
2	Carboxylic	C ₂ H ₄ O ₂	-6235.39	-7.75	0.07	7.82	3.91	0.25
3	Nitrile	C ₂ H ₃ N	-3613.38	-9.14	0.72	9.86	4.93	0.20
4	Alcohol	C ₂ H ₄ O	-4187.05	-7.14	-0.78	6.36	3.18	0.31
5	Ester	C ₃ H ₆ O ₂	-7305.14	-7.57	0.2	7.77	3.88	0.25
6	Aldehyde	C ₂ H ₄ O	-4187.05	-7.14	-0.78	6.36	3.18	0.31
7	Ketone	C ₃ H ₆ O	-5257.36	-6.89	-0.53	6.36	3.18	0.31
8	Ether	C ₂ H ₆ O	-4219.51	-7.03	1.21	8.24	4.12	0.24
9	Methyl-F	CH ₃ F	-3803.56	-9.33	1.34	10.67	5.33	0.18
10	Methyl-Cl	CH ₃ Cl	-13609.6	-8.29	0.21	8.5	4.25	0.23
11	Methyl-Br	CH ₃ Br	-71132.4	-7.66	-0.41	7.25	3.62	0.27
12	Methyl-I	CH ₃ I	-189376	-7.1	-1.24	5.86	2.93	0.31
13	Acid-F	C ₂ H ₃ FO	-6889.02	-8.63	-0.61	8.02	4.01	0.24
14	Acid-Cl	C ₂ H ₃ ClO	-16694.6	-8.4	-1.34	7.06	3.53	0.28
15	Acid-Br	C ₂ H ₃ BrO	-74217.3	-7.9	-1.39	6.51	3.25	0.30
16	Acid-I	C ₂ H ₃ IO	-192460	-7.22	-1.54	5.68	2.84	0.35
17	Amine	C ₃ H ₉ N	-4748.75	-5.84	1.07	6.91	3.45	0.28
18	Amide	C ₂ H ₅ NO	-5694.53	-6.88	0.7	7.58	3.79	0.26
19	Guanine	C ₅ H ₅ N ₅ O	-14767.2	-5.86	-0.34	5.52	2.76	0.36
20	Adenine	C ₅ H ₅ N ₅	-12719.3	-6.1	-0.65	5.45	2.72	0.36
21	Thymine	C ₅ H ₆ N ₂ O ₂	-12360.8	-6.78	-1.23	5.55	2.77	0.36
22	Cytosine	C ₄ H ₅ N ₃ O	-10749.2	-6.39	-1.01	5.38	2.69	0.37

Dipole moment (μ) can give information on the polarity of a molecule. Molecules with higher value of the dipole moment have a greater tendency to interact with other molecules through electrostatic interactions [14]. Based on this parameter presented in Table 4.1 (b), the trend of increase in the μ follow the order: Amine > Ether > Amide > Ester > ketone > Carboxylic > Carboxylate > Aldehyde > Methyl-F > Methyl-Br > Methyl-Cl > Methyl-I > Nitrile > Acid-I > Acid-F > Acid-Br > Acid-Cl. Amine has the highest and therefore the best to be adsorbed on the GNS.

Moreover, the tendency of a functional group to accept electron/s is measured by global electrophilicity index (ω), thus, the values for this parameter decreases as: Acid-I > Acid-Br > Acid-Cl > Methyl-I > Acid-F > Aldehyde = Alcohol > Methyl-Br > ketone > Methyl-Cl > Carboxylic > Nitrile > Carboxylate > Ester > Methyl-F > Amide > Ether > Amine. According to this parameter Amine with lower electrophilicity index should exhibits higher identification efficiency.

Table 4.1(b) Quantum chemical parameters of interacting species

S/N	Species	μ	ω	Dipole (debye)	Solv (kJ/mol)	Oval	PSA (Å ²)	Log p	Pol
1	Carboxylate	-3.93	1.787	2.77	-----	1.10	29.412	-----	-----
2	Carboxylic	-3.84	1.885	1.53	-20.02	1.12	34.628	-0.31	44.5
3	Nitrile	-4.21	1.797	3.86	-16.91	1.08	15.278	0.67	43.35
4	Alcohol	-3.96	2.465	2.65	-12.63	1.09	14.803	-0.57	44.29
5	Ester	-3.685	1.747	1.78	-10.6	1.17	21.116	-0.04	46.17
6	Aldehyde	-3.96	2.465	2.65	-12.63	1.09	14.803	-0.57	44.29
7	Ketone	-3.71	2.164	2.83	-12.89	1.13	14.855	0.20	45.78
8	Ether	-2.91	1.027	1.29	-9.93	1.11	7.653	0.09	44.36
9	Methyl-F	-3.995	1.495	1.90	3.17	1.05	0.00	0.43	41.96
10	Methyl-Cl	-4.04	1.920	2.15	-3.62	1.07	0.00	0.95	43.20
11	Methyl-Br	-4.035	2.245	1.97	-5.45	1.08	0.00	1.07	43.86
12	Methyl-I	-4.17	2.967	1.89	-8.44	1.09	0.00	1.46	44.76
13	Acid-F	-4.62	2.661	2.85	-8.63	1.10	14.595	-0.01	44.26
14	Acid-Cl	-4.87	3.359	3.02	-14.60	1.13	14.559	0.27	45.26
15	Acid-Br	-4.645	3.314	2.95	-16.20	1.14	14.541	0.60	45.78
16	Acid-I	-4.38	3.377	2.97	-18.96	1.15	14.539	0.60	46.56
17	Amine	-2.385	0.823	0.50	-16.02	1.16	2.70	0.24	46.48
18	Amide	-3.09	1.2596	3.78	-33.16	1.14	39.362	-0.96	44.83
19	Guanine	-3.1	1.740	6.55	-98.82	1.24	78.924	-1.74	50.60
20	Adenine	-3.375	2.090	2.42	-80.62	1.21	59.963	-0.24	50.00
21	Thymine	-4.005	2.890	4.19	-41.96	1.22	51.40	-0.60	49.73
22	Cytosine	-3.7	2.544	6.37	-92.60	1.19	57.565	-0.99	48.53

4.2 Quantitative Structure–Property Relationship (QSPR) Studies

Genetic function approximation was used on the training set to select the significant descriptors and it was found that among all the computed descriptors, Solvation Energy, Polarizability, E-

HOMO, Polar surface area, Ovality, Polarizability, and Log P construct the best model. The selected descriptors were subjected to regression analysis with the computationally determined structural property E-gap as the dependent variable and the selected descriptors as the independent variables using Genetic function approximation (GFA) method, new GFA-MLR equation as shown in Equation (16) were developed on the basis of the training set.

QSPR model for predicting E-gap against DNA identification:

$$\mathbf{E\text{-}gap} = 41.22 + 0.562*\mathbf{dipole} - 0.0156*\mathbf{solvation\ energy} + 104.7*\mathbf{Ovality} - 0.0349*\mathbf{Polar\ surface\ area} + 0.393*\mathbf{LogP} - 3.381*\mathbf{Polarizability}$$

$$R^2_{train} = 0.79, Q^2_v = 0.47, N_{test} = 5, R^2_{test} = 0.74, N_{train} = 12$$

Where N is the number of compounds in the training and test set, R^2_{train} is the squared correlation coefficient, Q^2_{cv} is the cross-validation coefficients of the training set and R^2_{test} is the squared correlation coefficient of the prediction (test) set.

4.3 QSAR Model Validation

Furthermore, the built model from the training data set was used to evaluate the predictive ability of the model by predicting the E-gap values in the prediction set (test set). The results are given in Table 4.2.

Table 4.2 Residuals from model validation for trained and test data.

S/N	Functional Group	Simulated E-gap (eV)	Predicted E-gap (eV)	Residual (eV)
1	Carboxylic	7.82	7.8545	-0.0345
2	Nitrile	9.86	9.8929	-0.0329
3	Alcohol	6.36	6.1952	0.1648
4	Ester	7.77	8.0313	-0.2613
5	Aldehyde	6.36	6.7471	-0.3871
6	Ketone	6.36	6.0298	0.3302
7	Ether	8.24	8.104	0.136
8	Methyl-F	10.67	10.4756	0.1944
9	Methyl-Cl	8.5	8.8279	-0.3279
10	Methyl-Br	7.25	7.618	-0.368
11	Methyl-I	5.86	4.6295	1.2305
12	Acid-F	8.02	7.97	0.05
13	Acid-Cl	7.06	8.0299	-0.9699
14	Acid-Br	6.51	7.4348	-0.9248
15	Acid-I	5.68	5.8989	-0.2189
16	Amine	6.91	6.0541	0.8559
17	Amide	7.75	9.8984	-2.1484

The predicted E-gap values for the training and test sets were plotted in a bar chart against the simulated E-gap as shown in Figure 4.1. The predicted E-gap results obtained for both the training set and test set (Table 4.2) are in good agreement with the simulated E-gap obtained. The residual values obtained between predicted and experimental E-gap were very low.

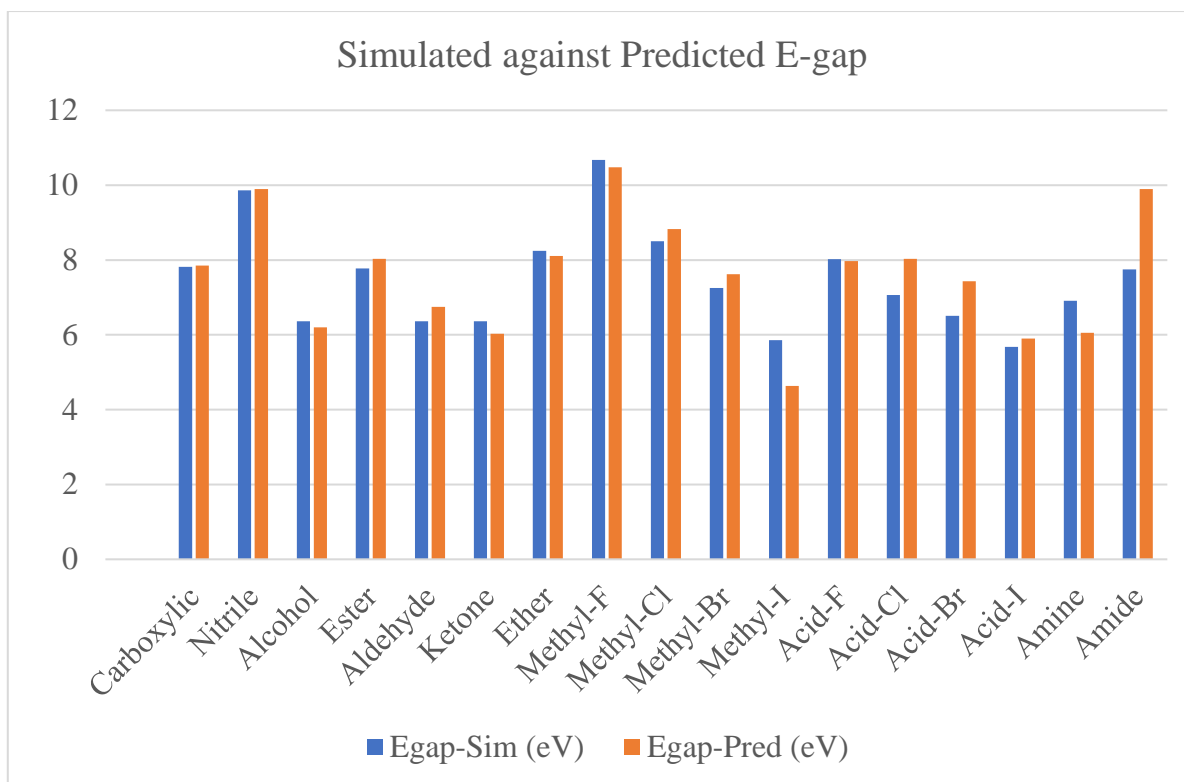


Figure 4.1 Bar chart of comparison between trained and predicted E-gap for the various functional groups.

The result of the QSPR model is in conformity with the standard shown in Table 2.1 as seen from the built model. The closeness of coefficient of determination (R^2) to its absolute value of 1.0 is an indication that the model explained a very high percentage of the response variable (descriptor) variation, high enough for a robust QSPR model. Its 0.79 value illustrates that 79 % of variation is residing in the residual meaning that the model is very good.

The high adjusted R^2 (R^2 adj) value as seen in the model and its closeness in value to the value of R^2 implies that the model has excellent explanatory power to the descriptors in it. It also illustrates the real influence of applied independent variables on the Dependent Variable. Also, the high and closeness of Q^2_{cv} to R^2 train revealed that the model was not over-fitted. The high R^2 test as seen in the model is an indication that the model is capable of providing valid predictions for new molecules.

Furthermore, the equation contains five descriptors and each descriptor has a positive or negative coefficient attached to it. These coefficients along with the value of descriptor have a significant role in deciding the overall DNA identification efficiency of the functional groups. Validation of the QSPR-model shows that the coefficients of each descriptor play an important role in deriving the DNA identification efficiency. From the point of view of identification efficiency of the functional groups in terms of E-gap values, the weight of a positive coefficient is very significant because it contributes towards the increased value of E-gap. So, the descriptors with high weight positive coefficients are most important followed by descriptors with the low weight negative coefficient and lastly the descriptors with high weight negative coefficients [40]. On the basis of values of the coefficients on the model, the associated descriptors are arranged in a sequence pertaining to their contribution towards overall efficiency of the DNA identification in the following increasing order towards sequencing of GNS:

Ovality > Dipole-moment > LogP > Solvation Energy > Polar Surface Area > Polarizability

4.4 Molecular Dynamic Simulation Studies

Electronic parameters alone are not enough to determine the trend of the DNA identification performances of these studied functional groups relative to GNS despite its success in explaining the mechanism of action of the identification. Thus, it is essential to conduct thorough modelling of the direct interactions of the studied inhibitors with the GNS surface. It is believed that the major mechanism of the DNA identification (sequencing) is through adsorption.

Therefore, adsorption of these DNA bases on the GNS surface was simulated to predict the nature of interactions between the studied DNA bases and 12.0Å GNS cluster surface in

vacuum. The first step of this simulation process was the geometry optimization of the studied eighteen (18) functional groups namely: Methyl Carboxylic Acid (acetic acid), methyl-Nitrile, Methyl-Alcohol, Methyl-Ester, Methyl-Aldehyde, Methyl-Ketone, Methyl-Ether, Methyl-F, Methyl-Cl, Methyl-I, Methyl-Br, Methyl-Acid-F, Methyl-Acid-Cl, Methyl-Acid-I, Methyl-Acid-Br, Methyl-Amine, Methyl-Amide, Methyl-Carboxylates as mentioned in Table 4.1(a & b).

This study begins with determining the energy gap of four (4) DNA bases namely Guanine (G), Adenine (A), Thymine (T) and Cytosine (C) as shown in Table 4.3.

Table 4.3 Adsorption potential of DNA bases relative to E-gap

DNA Bases	Total Energy (eV)	E-HOMO (eV)	E-LUMO (eV)	E-gap (eV)
G	-14763.44	-5.61	-0.16	-5.45
A	-12716.31	-6.08	-0.68	-5.40
T	-12357.58	-6.52	-0.88	-5.64
C	-10746.41	-6.25	-0.83	-5.42

The higher the HOMO energy the more the probability of a molecule to donate an electron to the functionalized GNR. Thus, as shown in Table 4.3 E-HOMO decreases in the order: $G > A > C > T$, A comparison of the E-HOMO of the studied DNA bases shows that G has the highest while A has the lowest which indicates that, G would be the most preferred electron donor while A would be the least preferred.

On the other hand, a low E-LUMO value for a DNA molecule indicates the relative tendency of the molecule to accept electrons from functionalized GNS orbital during back-donation. From Table 4.3, the E-LUMO value in the descending order are: $G > A > C > T$, indicating that G has a stronger tendency to accept electrons from functionalized GNS orbital as they are

been adsorb on its surface and T would be the least. Therefore, G can more powerfully adsorb on the GNR surface to achieve better DNA identification efficiency.

Energy gap E-gap ($E\text{-HOMO} - E\text{-LUMO}$) is an important parameter as efficient DNA sequencing is characterized by a small energy gap [39]. From Table 4.3, the trend of decreasing the energy gap shows that: $G > A > C > T$. This suggest that T with the lowest energy gap of - 5.64 eV is the best functional group in term of this parameter among the series.

The geometry optimization of the DNA sequencing system was conducted until the total energy of the system reached a local minimum on the potential energy surface.

The molecular dynamics (MD) process was conducted with adsorption distance of 3.0 Å and the system attained equilibrium when the energy of the system was balanced. All the other functionalized GNS were studied similarly. After the system attained equilibrium, the adsorption parameters of the DNA bases adsorbed on the GNS surface was calculated according to equation (9) respectively and are presented in Table 4.4.

Table 4.4 Adsorption parameters of DNA nucleobases relative to functional groups.

FunG	Ead-G (eV)	Ead-A (eV)	Ead-T (eV)	Ead-C (eV)	Egap-G (eV)	Egap-A (eV)	Egap-T (eV)	Egap-C (eV)
Carboxylic	-1.39	-1.53	-0.5	-0.38	-2.63	-2.63	-2.7	-2.62
Nitrile	-0.56	-0.61	-3.97	-0.5	-1.96	-0.98	-0.93	-1.55
Alcohol	-0.66	-0.99	-0.07	-0.59	-2.63	-2.63	-2.7	-2.62
Ester	-1.15	-0.6	-0.76	-0.72	-2.26	-2.38	-2.3	2.33
Aldehyde	-1.13	-0.32	-0.86	-0.91	-1.97	-2.19	-2.15	-2.2
Ketone	-1.28	-1.03	-0.82	-0.87	-2.31	-2.29	-2.3	-2.29
Ether	-1.06	-0.52	-1.02	-0.69	-2.35	-2.38	-2.37	-2.37
Methyl-F	-0.76	-1.2	-0.75	-0.64	-2.29	-2.35	-2.37	-2.35
Methyl-Cl	-0.34	-0.56	-0.56	-0.33	-2.32	-2.33	-2.31	-2.32
Methyl-Br	-0.9	-0.56	-0.44	-1.11	-2.46	-2.54	-2.55	-2.48
Methyl-I	-0.66	-1.06	-0.74	-1.06	-2.3	-2.28	-2.31	-2.29
Acid-F	-0.68	-0.91	-1.3	-2.46	-1.97	-2.19	-2.2	-2.15
Acid-Cl	-0.57	-1.89	-0.35	-0.35	-2.19	-2.2	-2.32	-2.3
Acid-Br	-0.64	-0.3	-1.03	-0.57	-2.14	-2.29	-2.21	-2.27
Acid-I	-1.23	-0.6	-0.62	-0.13	-2.14	-2.26	-2.27	-2.27
Amine	-0.98	-0.8	-2.88	-0.56	-2.25	-2.24	-2.18	-2.2
Amide	-1	-0.69	-0.69	-0.61	-2.33	-2.33	-2.35	-2.33
Methyl	-0.92	-0.47	-1.28	-0.48	-2.72	-2.81	-2.85	-2.82

FunG= Functional group, Eads-G, A, T, C = Adsorption Energy of G, A, T, C respectively, E-gap- G, A, T, C = Energy gap of G, A, T, C respectively.

It can be observed from the molecular structures of the examined DNA bases that these molecules contain various electrons (lone pair) on the N and O atoms, and also π -bonds of the

aromatic systems. Therefore, these electrons (lone pair) on the heteroatoms will be donated to an empty d-orbitals of functionalized.

Among these eighteen (18) graphene functionalized molecules, nine (9) of them namely: Methyl-carboxylic acid, Methyl-ester, Methyl-ketone, Methyl-ether, Methyl-F, Methyl-I, Acid-F, Amine and Amide were found to be more promising in the DNA sequencing process and preference will be given to them.

Methyl-Carboxylic (Acetic) Acid:

The relationship between adsorption energy of methyl-carboxylic acid functionalized GNS is presented in Fig.12 below relative to all the DNA bases.

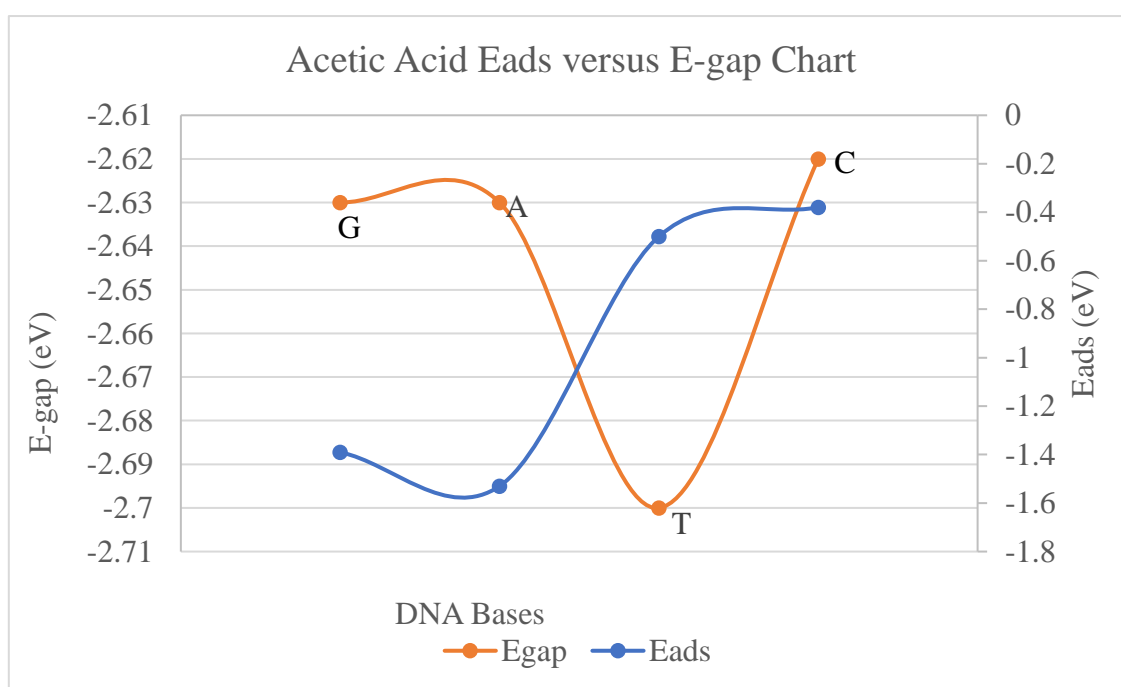


Figure 4.2 Comparing the adsorption energy and energy gap of the functionalized GNS system with each adsorbed DNA base.

The strength of DNA bases adsorption for methyl-carboxylic acid functionalized GNS is of the decreasing order: C > T > G > A, having A to be the most stable and C been the least stable. However, the variation the adsorption energies of these bases is relatively minimal. While on the other hand, E-gap decreases in this order: C > G = A > T, with C having the highest energy

gap and T with the least. This comparison reveals that, G and A are more electronically stable and will require similar and moderate energy (E-gap) between C and T in order to be identified by the methyl-carboxylic acid functionalized GNS and will therefore be easily adsorbed on its surface. Carboxylic acid functionalized graphene has been found to improve thermal properties of graphene [52], therefore it can be understood that the functionalization can also improve the surface properties of pristine graphene.

Methyl-Ketone:

Fig.13 present the relationship between adsorption energy of methyl-ketone functionalized GNR relative to all the DNA bases.

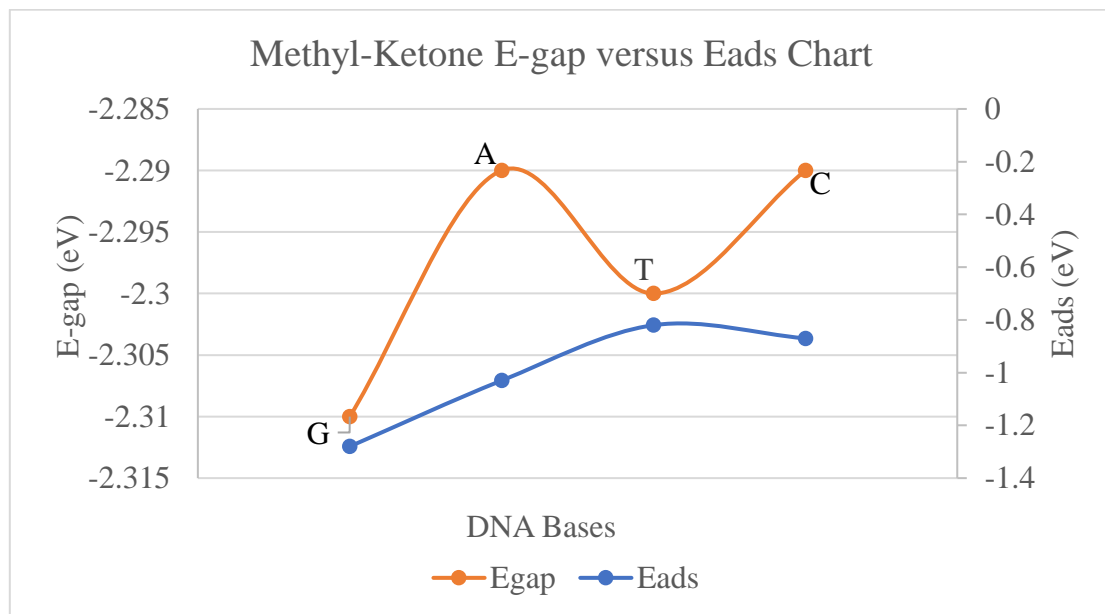


Figure 4.3 Comparing the adsorption energy and energy gap of the Methyl-ketone functionalized GNR system with each adsorbed DNA base.

The strength of DNA bases adsorption for methyl-ketone functionalized GNS is of the decreasing order: C > T > A > G, having G to be the most stable and C been the least stable. However, the variation the adsorption energies of these bases is relatively minimal. While on the other hand, E-gap decreases in this order: C = A > T > G, with C having the highest energy gap and T with the least. This comparison reveals that, G and A are more electronically stable

and will require similar and more energy (E-gap) relative to G and T in order to be identified by the methyl-ketone functionalized GNS and will therefore be easily adsorbed on its surface.

Acid-F:

Fig14. present the relationship between adsorption energy of methyl-ketone functionalized GNR relative to all the DNA bases.

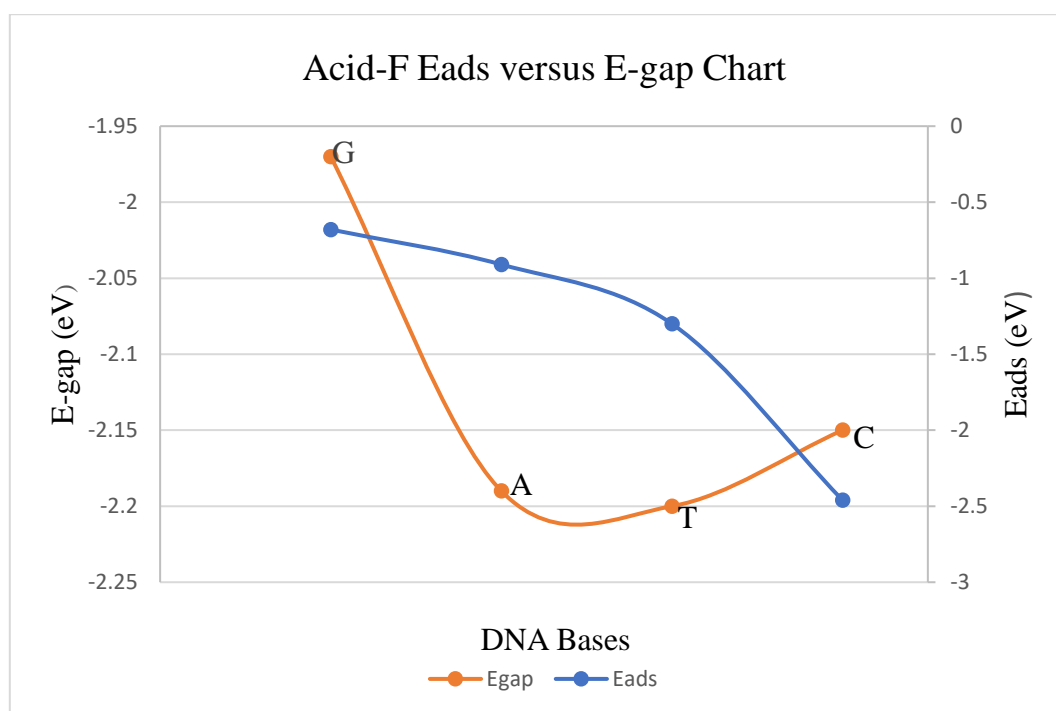


Figure 4.4 Comparing the adsorption energy and energy gap of the Acid-F functionalized GNS system with each adsorbed DNA base.

The strength of DNA bases adsorption for Acid-F functionalized GNS is of the decreasing order: $G > A > T > C$, having C to be the most stable and G been the least stable. However, the variation the adsorption energies of these bases is relatively minimal. While on the other hand, E-gap decreases in this order: $G > C > A > T$, with C having the highest energy gap and T with the least. This comparison shows that, T and A are more electronically stable and will require similar and more energy (E-gap) relative to G and C in order to be identified by the methyl-ketone functionalized GNS and will therefore be easily adsorbed on its surface.

Methyl-F:

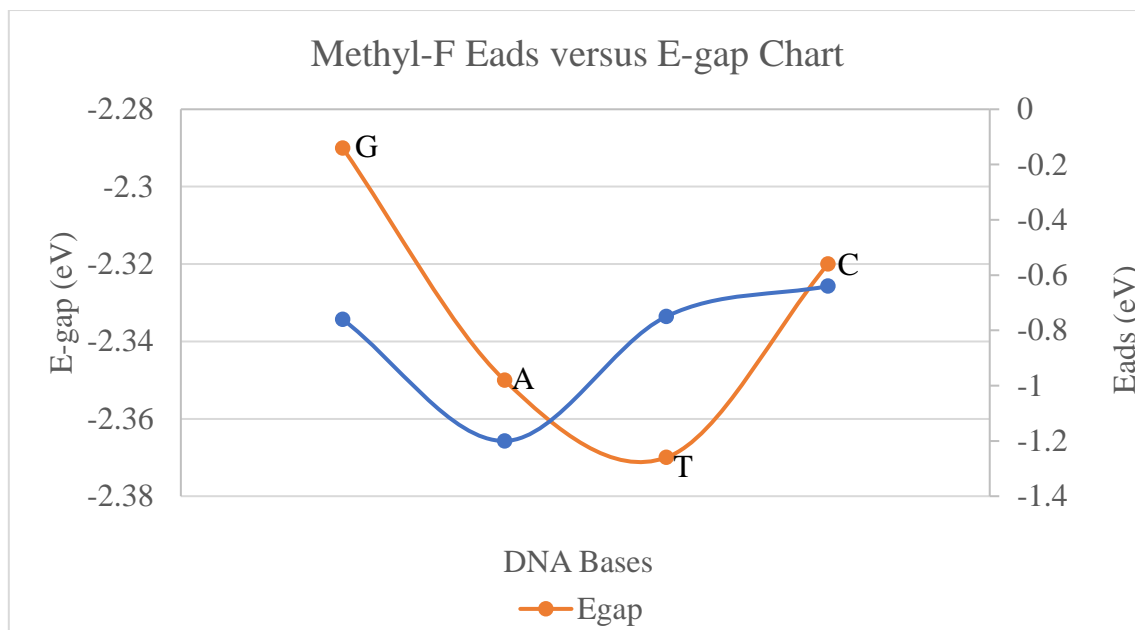


Figure 4.5 Comparing the adsorption energy and energy gap of the Methyl-F functionalized GNS system with each adsorbed DNA base.

The strength of DNA bases adsorption for Acid-F functionalized GNS is of the decreasing order: $G > T > C > A$, having A to be the most stable and G been the least stable. However, the variation the adsorption energies of these bases is relatively minimal. on the other hand, E-gap decreases in this order: $G > C > A > T$, with G having the highest energy gap and T with the least. This comparison shows that, T and A are more electronically stable and will require similar and more energy (E-gap) relative to G and C in order to be identified by the methyl-F functionalized GNS and will therefore be easily adsorbed on its surface.

Amide:

Fig.16 present the relationship between adsorption energy of amide functionalized GNS relative to all the DNA bases.

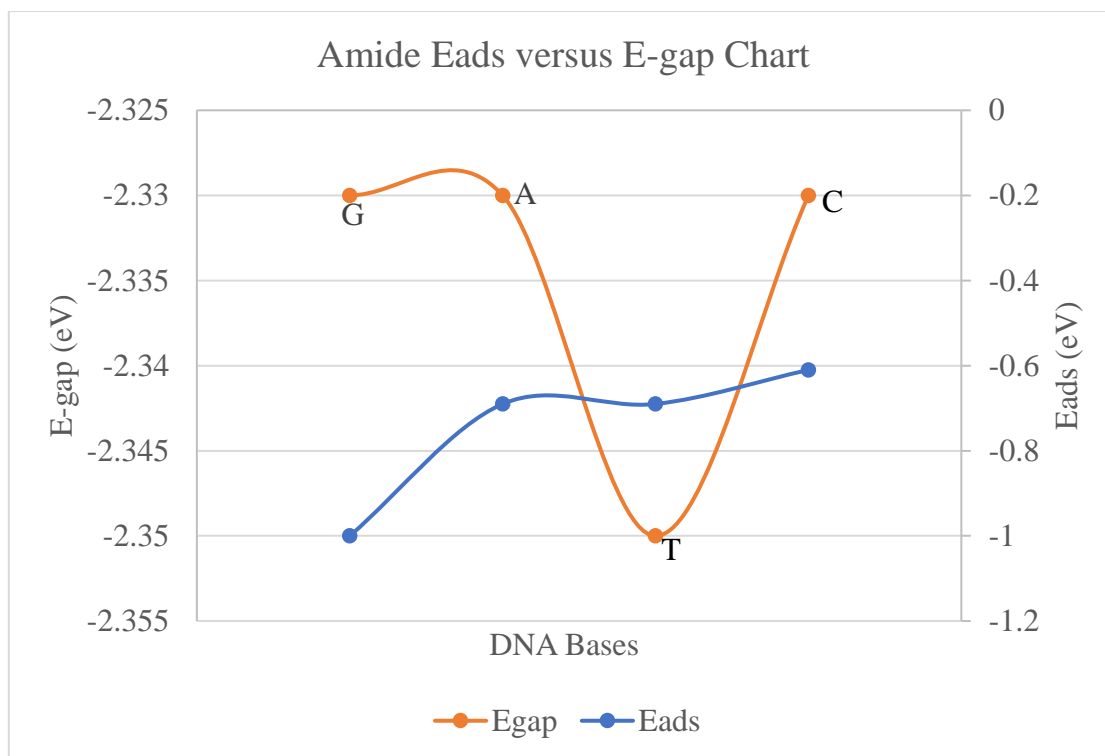


Figure 4.6 Comparing the adsorption energy and energy gap of the Amide functionalized GNS system with each adsorbed DNA base.

The strength of DNA bases adsorption for Acid-F functionalized GNS is of the decreasing order: $C > T > A > G$, having G to be the most stable and C been the least stable. However, the variation the adsorption energies of these bases is relatively minimal. on the other hand, E-gap decreases in this order: $G = C = A > T$, with G, C, and A having the highest and similar energy gap and T with the least. This comparison shows that, T will be more electronically stable and will require similar and more energy (E-gap) relative to G, A and C in order to be identified by the methyl-F functionalized GNS and will therefore be easily adsorbed on its surface.

Amine:

Figure 4.7 present the relationship between adsorption energy of amine functionalized GNS relative to all the DNA bases.

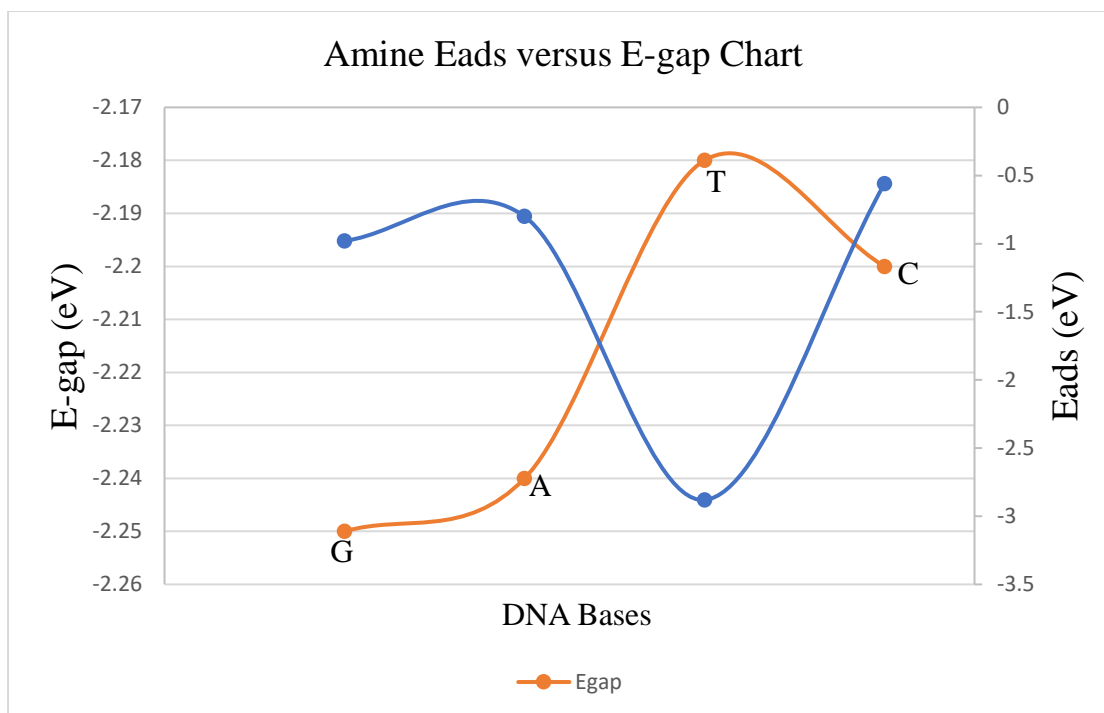


Figure 4.7 Comparing the adsorption energy and energy gap of the Amine functionalized GNS system with each adsorbed DNA base.

The strength of DNA bases adsorption for Acid-F functionalized GNS is of the decreasing order: $C > A > G > T$ having T to be the most stable and C been the least stable. However, the variation the adsorption energies of these bases is relatively minimal. on the other hand, E-gap decreases in this order: $T > C > A > G$, with T having the highest and similar energy gap and G with the least. This comparison shows that, T will be more electronically stable and will require similar and more energy (E-gap) relative to G, A and C in order to be identified by the amine functionalized GNS and will therefore be easily adsorbed on its surface. Amine functionalized graphene has been found to improve thermal properties of graphene [52], therefore it can be understood that the functionalization can also improve the surface properties of pristine graphene.

Methyl-I:

Fig.18 present the relationship between adsorption energy of methyl-I functionalized GNS relative to all the DNA bases. The strength of DNA bases adsorption for Acid-F functionalized

GNS is of the decreasing order: $G > T > C > A$, having A to be the most stable and G been the least stable.

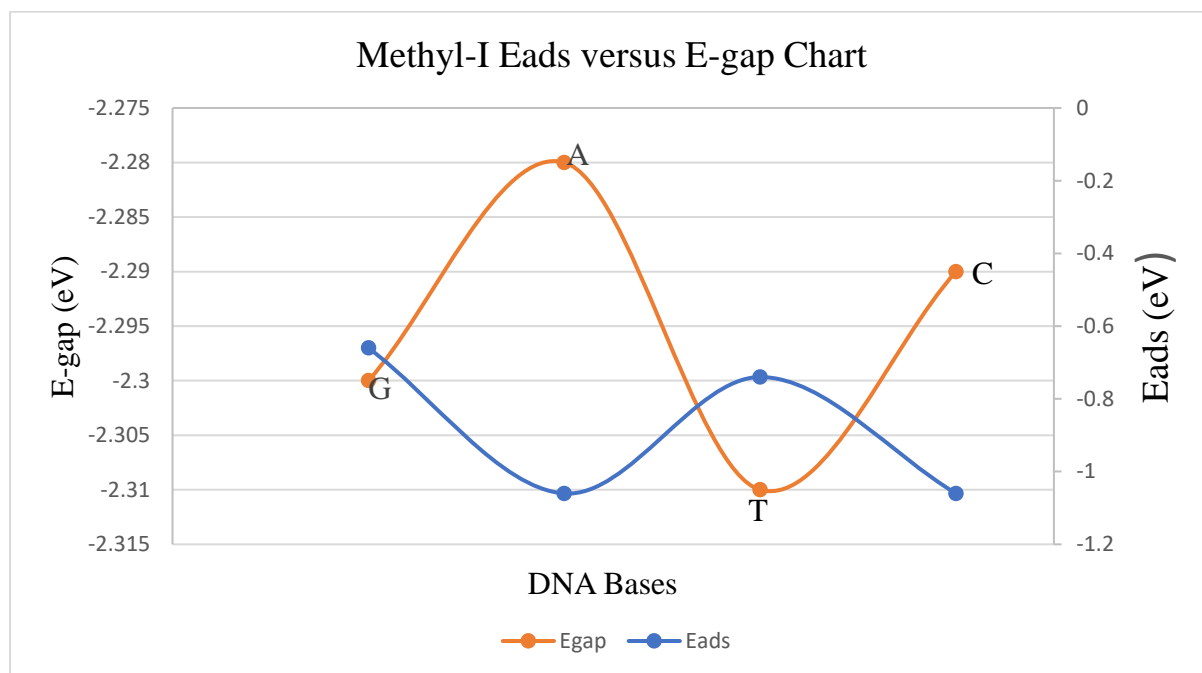


Figure 4.8 Comparing the adsorption energy and energy gap of the Methyl-I functionalized GNS system with each adsorbed DNA base.

However, the variation the adsorption energies of these bases is relatively minimal. on the other hand, E-gap decreases in this order: $A > C > G > T$, with T having the highest and similar energy gap and A with the least. This comparison shows that, A will be more electronically stable and will require similar and more energy (E-gap) relative to G, T and C in order to be identified by the Methyl-I functionalized GNS and will therefore be easily adsorbed on its surface.

Ester:

Fig.19 present the relationship between adsorption energy of ester functionalized GNS relative to all the DNA bases. The strength of DNA bases adsorption for Acid-F functionalized GNS is of the decreasing order: $A > C > T > G$, having G to be the most stable and A been the least stable.

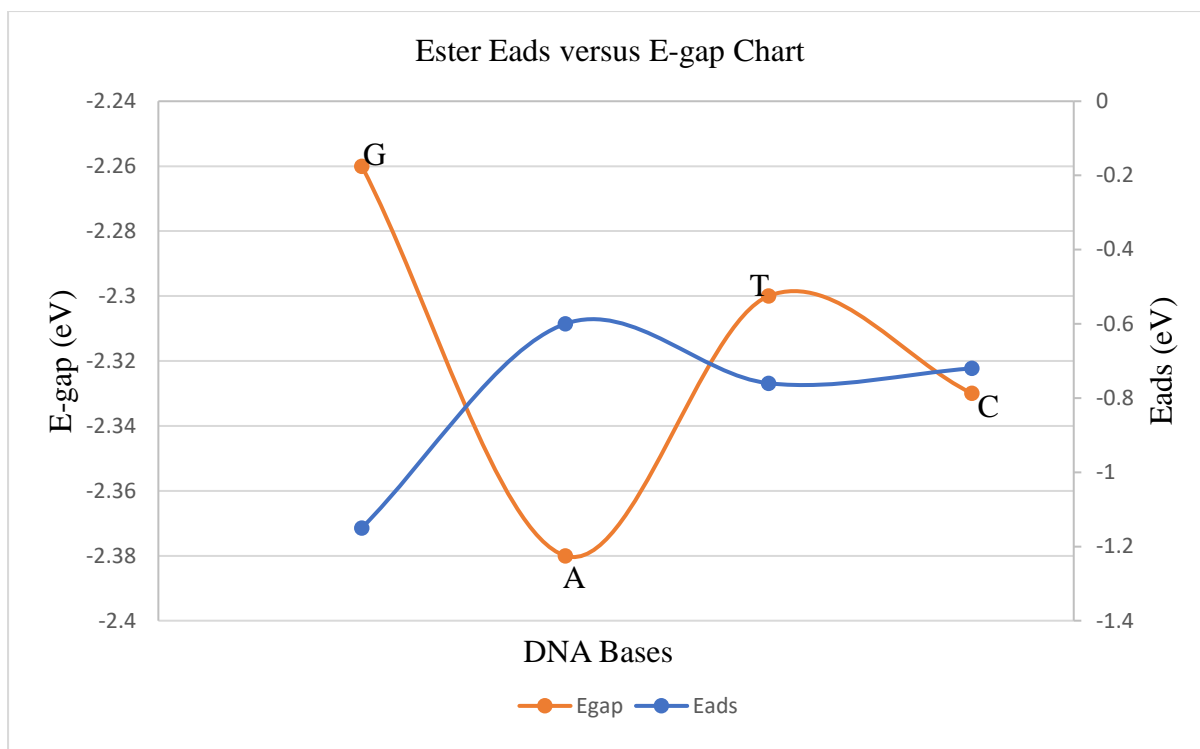


Figure 4.9 Comparing the adsorption energy and energy gap of the ester functionalized GNS system with each adsorbed DNA base.

However, the variation the adsorption energies of these bases is relatively minimal. on the other hand, E-gap decreases in this order: $G > T > C > A$, with G having the highest and similar energy gap and A with the least. This comparison shows that, G will be more electronically stable and will require similar and more energy (E-gap) relative to A, T and C in order to be identified by the ester functionalized GNS and will therefore be easily adsorbed on its surface.

Ether:

Fig.20 present the relationship between adsorption energy of ether functionalized GNS relative to all the DNA bases. The strength of DNA bases adsorption for ether functionalized GNS is of the decreasing order: $A > C > T > G$, having G to be the most stable and A been the least stable.

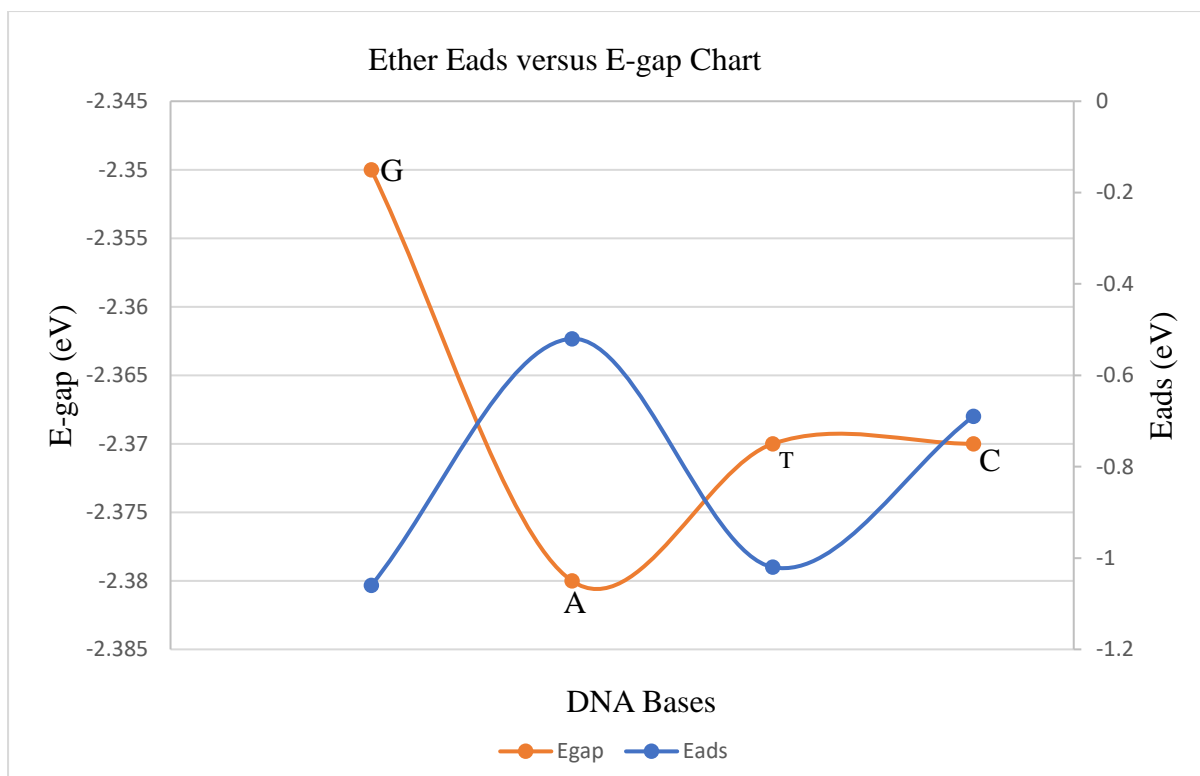


Figure 4.10 Comparing the adsorption energy and energy gap of the ether functionalized GNS system with each adsorbed DNA base.

However, the variation the adsorption energies of these bases is relatively minimal. on the other hand, E-gap decreases in this order: $G > T > C > A$, with G having the highest and similar energy gap and A with the least. This comparison shows that, G will be more electronically stable and will require similar and more energy (E-gap) relative to A, T and C in order to be identified by the ether functionalized GNS and will therefore be easily adsorbed on its surface.

4.5 Comparing the most stable functionalized GNS and pristine.

To evaluate the central idea of this study, the adsorption energies of high potential functionalized GNS must be compared with pristine which can be seen in Fig.21

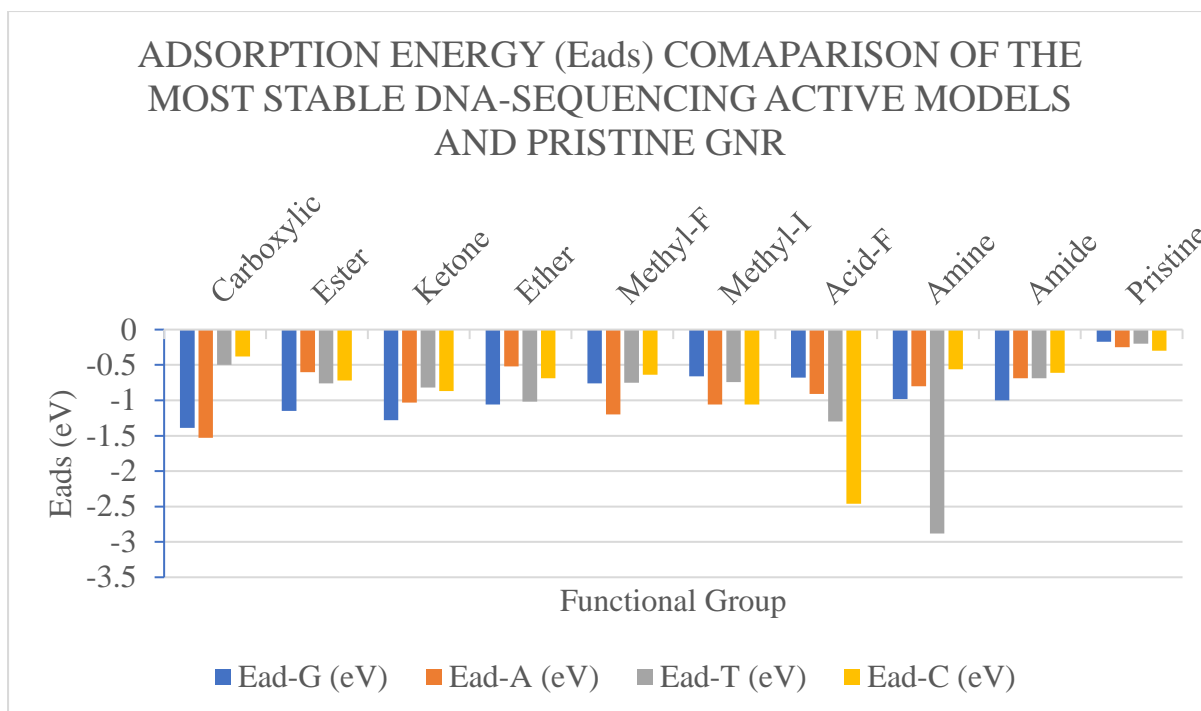


Figure 4.11 Comparison of the most stable DNA-sequencing active models and pristine GNS

It can be observed that, the most underscoring DNA molecule (G) in a pristine GNS has an appreciable potential of being detected when GNS is functionalized due to the presence of hetero atoms that are highly energetic. Despite observing functional groups to have higher affinity of some specific DNA bases, the adsorption energy of all functional the nine (9) functional groups underscore the pristine graphene.

CHAPTER FIVE

5.0 CONCLUSION AND RECOMMENDATION

5.1 Conclusion

The role of functional groups in enhancing the adsorption capacity of graphene of graphene nano structure (GNS) was studied using eighteen (18) different functional groups, among which nine (9) of them namely: Methyl-Carboxylic Acid, Methyl-Ester, Methyl-Ketone, Methyl-Ether, Methyl-F, Methyl-I, Acid-F, Amine and Amide were found to be more promising in the DNA sequencing process base on their appreciable quantum chemical parameters such as highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO), chemical hardness, chemical softness, electrophilicity index etc. DNA bases adsorption was carried out using these functionalized GNS of 12.0Å diameter and adsorption distance of 3.0Å. The most underscoring DNA molecule (G) in a pristine GNS has an appreciable potential of being detected when GNS is functionalized due to the presence of hetero atoms that are highly energetic. Despite observing functional groups to have higher affinity of some specific DNA bases, the adsorption energy of all functional the nine (9) functional groups underscore the pristine graphene. This is an indicator that, the common practice of doping heteroatoms or semiconductor atoms is not the only approach enhancing materials for DNA sequencing application, since functionalized graphene-based materials has the potential of competing with doped and nano-pore graphene.

5.2 Recommendation

Further research can be focused on the electronic properties such as density of state (DOS), partial density of states (pDOS) and band gap of DNA sequencing using these functionalized graphene nano structure (GNS). A comparison should be made using a functionalized GNS unitcell (supercell) and a cluster.

REFERENCES

- [1] L. Liang, J. Shen, Z. Zhang, and Q. Wang, “Biosensors and Bioelectronics DNA sequencing by two-dimensional materials : As theoretical modeling meets experiments,” *Biosens. Bioelectron.*, vol. 89, pp. 280–292, 2017.
- [2] M. K. Jena, R. L. Kumawat, and B. Pathak, “First-Principles Density Functional Theory Study on Graphene and Borophene Nanopores for Individual Identification of DNA Nucleotides,” 2021.
- [3] S. J. Heerema and C. Dekker, “Graphene nanodevices for DNA sequencing,” *Nat. Publ. Gr.*, vol. 11, no. February, pp. 127–136, 2016.
- [4] P. Yadav, Z. Cao, and A. B. Farimani, “DNA Detection with Single-Layer Ti₃C₂ MXene Nanopore,” 2021.
- [5] M. B. Henry, “Identification of DNA bases using nanopores created in finite-size nanoribbons from graphene , phosphorene , and silicene Identification of DNA bases using nanopores created in finite-size nanoribbons from graphene , phosphorene , and silicene,” vol. 035324, 2021.
- [6] S. Banerjee, J. Wilson, J. Shim, M. Shankla, and E. A. Corbin, “Slowing DNA Transport Using Graphene – DNA Interactions,” pp. 936–946, 2015.
- [7] J. Lee, S. Park, and J. Choi, “Electrical Property of Graphene and Its Application to Electrochemical Biosensing,” 2019.
- [8] H. Mohammed, A. Kumar, E. Bekyarova, and Y. Al-hadeethi, “Antimicrobial Mechanisms and Effectiveness of Graphene and Biomaterials . A Scope Review,” vol. 8, no. May, 2020.
- [9] P. K. Dubey and I. Oh, “Review on Functionalized Graphenes and Their Applications Review on Functionalized Graphenes and Their Applications,” no. August, 2012.
- [10] D. Saini, “Synthesis and functionalization of graphene and application in electrochemical biosensing,” vol. 5, no. 4, pp. 393–416, 2016.
- [11] M. B. Henry and B. O. Tayo, “Evaluation of the Feasibility of Phosphorene for Electronic DNA Sequencing Using Density Functional Theory Calculations,” vol. 6, no. 2021, pp. 1–6.
- [12] M. M. Mohammadi, O. Bavi, and Y. Jamali, “Journal of Molecular Graphics and Modelling DNA sequencing via molecular dynamics simulation with functionalized graphene nanopore,” *J. Mol. Graph. Model.*, vol. 122, no. January, p. 108467, 2023.
- [13] H. Yang, Y. Liu, C. Gao, L. Meng, Y. Liu, X. Tang, and H. Ye, “Adsorption Behavior

- of Nucleobases on Doped MoS₂ Monolayer: A DFT Study,” *J. Phys. Chem. C*, vol. 123, no. 51, pp. 30949–30957, 2019.
- [14] K. Eren, N. Taktakoğlu, and I. Pirim, “DNA Sequencing Methods: From Past to Present,” *Eurasian J. Med.*, vol. 54, no. February, pp. S47–S56, 2022.
- [15] N. Sciences and H. Ce, “Busra Gheith , Mehmet Can,” vol. 2, no. 2, 2013.
- [16] I. Duarte, “DNA sequencing: Techniques and Applications,” no. December, 2022.
- [17] Y. M. Yuguda, “Application of Next Generation Sequencing (NGS) technology in forensic science : A review,” vol. 23, no. 02, pp. 155–159, 2023.
- [18] G. Lear, I. Dickie, J. Banks, S. Boyer, H. L. Buckley, T. R. Buckley, R. Cruickshank, A. Dopheide, K. M. Handley, S. Hermans, J. Kamke, C. K. Lee, R. Macdiarmid, S. E. Morales, and D. A. Orlovich, “Methods for the extraction , storage , amplification and sequencing of DNA from environmental samples,” vol. 42, 2018.
- [19] S. Aman and A. Yusuf, “Review on Different Types of Sequencing Technologies And Their Application ; Merit and Demerit,” vol. 22, no. 1, pp. 9–18, 2020.
- [20] W. J. Ansorge, “Next Generation DNA Sequencing (II): Techniques , Applications Journal of Next Generation Sequencing & Applications,” 2016.
- [21] J. Shendure, S. Balasubramanian, G. M. Church, W. Gilbert, J. Rogers, J. A. Schloss, and R. H. Waterston, “and future,” *Nat. Publ. Gr.*, 2017.
- [22] S. Moorthie, C. J. Mattocks, and C. F. Wright, “Review of massively parallel DNA sequencing technologies,” pp. 1–12, 2011.
- [23] L. Liu, Y. Li, S. Li, N. Hu, Y. He, R. Pong, D. Lin, L. Lu, and M. Law, “Comparison of Next-Generation Sequencing Systems,” vol. 2012, 2012.
- [24] I. Print, I. Online, S. Siddiraju, and N. Thota, “World Journal of Pharmaceutical Sciences,” vol. 8, no. 6, pp. 127–130, 2020.
- [25] C. S. Pareek and R. Smoczynski, “Sequencing technologies and genome sequencing,” no. May 2014, 2011.
- [26] J. A. Shendure, G. J. Porreca, and G. M. Church, “Overview of DNA Sequencing Strategies,” no. January, pp. 1–11, 2008.
- [27] S. Z. Butler, S. M. Hollen, L. Cao, Y. Cui, J. A. Gupta, H. R. Gutiérrez, T. F. Heinz, S. S. Hong, J. Huang, A. F. Ismach, E. Johnston-Halperin, M. Kuno, V. V. Plashnitsa, R. D. Robinson, R. S. Ruoff, S. Salahuddin, J. Shan, L. Shi, M. G. Spencer, M. Terrones, W. Windl, and J. E. Goldberger, “Progress, challenges, and opportunities in two-dimensional materials beyond graphene,” *ACS Nano*, vol. 7, no. 4, pp. 2898–2926, 2013.

- [28] Y. Shao, J. Wang, H. Wu, J. Liu, I. A. Aksay, and Y. Lin, "Graphene based electrochemical sensors and biosensors: A review," *Electroanalysis*, vol. 22, no. 10, pp. 1027–1036, 2010.
- [29] Y. Zhu, S. Murali, W. Cai, X. Li, J. W. Suk, J. R. Potts, and R. S. Ruoff, "Graphene and graphene oxide: Synthesis, properties, and applications," *Adv. Mater.*, vol. 22, no. 35, pp. 3906–3924, 2010.
- [30] H. Wang, T. Maiyalagan, and X. Wang, "Review on recent progress in nitrogen-doped graphene: Synthesis, characterization, and its potential applications," *ACS Catal.*, vol. 2, no. 5, pp. 781–794, 2012.
- [31] O. Parlak, A. Tiwari, A. P. F. Turner, and A. Tiwari, "Biosensors and Bioelectronics Template-directed hierarchical self-assembly of graphene based hybrid structure for electrochemical biosensing," *Biosens. Bioelectron.*, vol. 49, pp. 53–62, 2013.
- [32] O. Paul, H. Rapheal, J. Patrick " L -Cysteine-Assisted Synthesis of Layered MoS₂ / Graphene Composites with Excellent Electrochemical Performances for Lithium Ion Batteries," no. 6, pp. 4720–4728, 2011.
- [33] H. Li, Q. Zhang, C. Chong, R. Yap, K. Tay, T. Hang, T. Edwin, A. Olivier, and D. Baillargeat, "From Bulk to Monolayer MoS₂ : Evolution of Raman Scattering," no. 1 L, pp. 1385–1390, 2012.
- [34] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, "Atomically Thin MoS₂ : A New Direct-Gap Semiconductor," vol. 136805, no. September, pp. 2–5, 2010.
- [35] H. Li, J. Wu, Z. Yin, and H. Zhang, "Preparation and Applications of Mechanically Exfoliated Single-Layer and Multilayer MoS₂ and WSe₂ Nanosheets," 2014.
- [36] Y. Lee, X. Zhang, W. Zhang, M. Chang, C. Lin, K. Chang, Y. Yu, J. T. Wang, C. Chang, L. Li, and T. Lin, "Synthesis of Large-Area MoS₂ Atomic Layers with Chemical Vapor Deposition," pp. 2320–2325, 2012.
- [37] Y. Huang, J. Lin, C. Tang, and Y. Bando, "Bulk synthesis , growth mechanism and properties of highly pure ultrafine boron nitride nanotubes with diameters of sub-10 nm," vol. 145602.
- [38] S. Bhowmick, A. K. Singh, and B. I. Yakobson, "Quantum Dots and Nanoroads of Graphene Embedded in," pp. 9889–9893, 2011.
- [39] L. Song, L. Ci, H. Lu, P. B. Sorokin, C. Jin, J. Ni, A. G. Kvashnin, D. G. Kvashnin, J. Lou, B. I. Yakobson, and P. M. Ajayan, "Large Scale Growth and Characterization of Atomic Hexagonal Boron Nitride Layers," pp. 3209–3215, 2010.
- [40] B. C. Zhi, Y. Bando, C. Tang, H. Kuwahara, and D. Golberg, "Large-Scale Fabrication of Boron Nitride Nanosheets and Their Utilization in Polymeric Composites with

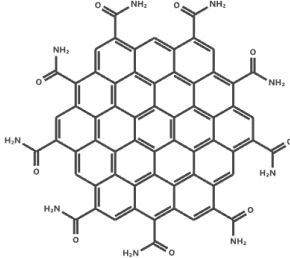
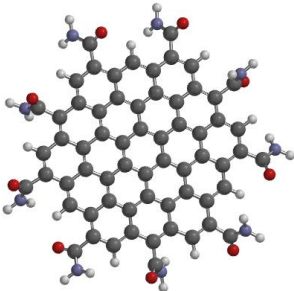
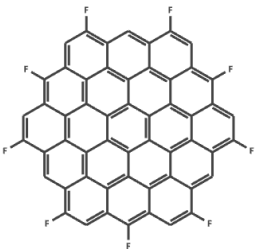
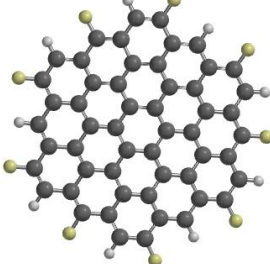
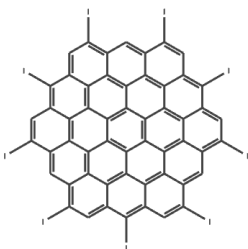
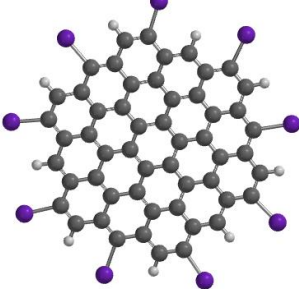
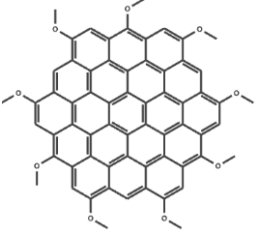
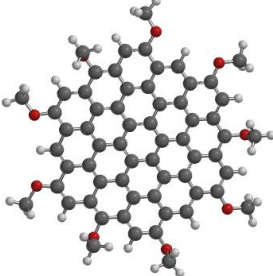
- Improved Thermal and Mechanical Properties,” pp. 2889–2893, 2009.
- [41] I. B. Obot and Z. M. Gasem, “Theoretical evaluation of corrosion inhibition performance of some pyrazine derivatives,” vol. 83, pp. 359–366, 2014.
- [42] I. B. Obot, D. D. Macdonald, Z. M. Gasem, D. D. Macdonald, and Z. M. Gasem, “Density Functional Theory (DFT) as a Powerful Tool for Designing New Organic Corrosion Inhibitors . Part 1 : An Overview,” 2015.
- [43] B. O. Tayo, “Dna base detection using two-dimensional materials beyond graphene.”
- [44] D. H. Cho, S. K. Lee, B. T. Kim, and K. T. No, “Quantitative Structure-Activity Relationship (QSAR) study of new fluorovinyloxyacetamides Quantitative Structure-Activity Relationship (QSAR) Study of New Fluorovinyloxyacetamides,” no. April 2001, 2021.
- [45] E. Sayed, H. El Ashry, A. El Nemr, and S. Ragab, “Quantitative structure activity relationships of some pyridine derivatives as corrosion inhibitors of steel in acidic medium,” no. April 2019, 2011.
- [46] O. Article, H. Sciences, C. District, and T. Nadu, “Journal of Medicinal and Chemical Sciences QSAR Modeling , Molecular Docking , and ADME Studies of Novel 5- Oxo-Imidazole Derivatives as Anti-Breast Cancer Drug Compounds against MCF-7 Cell Line,” vol. 6, pp. 3087–3112, 2023.
- [47] A. U. Bello, A. Uzairu, and G. A. Shallangwa, “Prediction of inhibition performance of some benzimidazole derivatives against steel corrosion through QSAR and molecular dynamic simulation,” vol. 2508, no. 1, pp. 1–14, 2019.
- [48] D. Ebuka, “RESEARCH ARTICLE Data Set collection,” *J. statistics* vol. 5. 3, pp. 1387–1398, 2018.
- [49] B. A. Umar, A. Uzairu, G. A. Shallangwa, and U. Sani, “cell line and ligand-based design of potent compounds using in silico virtual QSAR modeling for the prediction of - pGI 50 activity of compounds on LOX IMVI cell line and ligand - based design of potent compounds using in silico virtual screening,” *Netw. Model. Anal. Heal. Informatics Bioinforma.*, no. October 2020, 2019.
- [50] A. B. Umar, A. Uzairu, G. A. Shallangwa, and S. Uba, “Heliyon QSAR modelling and molecular docking studies for anti-cancer compounds against melanoma cell line SK-MEL-2,” *Heliyon*, vol. 6, no. 3, p. e03640, 2020.
- [51] A. U. Bello, A. Uzairu, and G. A. Shallangwa, “Prediction of inhibition performance of some benzimidazole derivatives against steel corrosion through QSAR and molecular dynamic simulation Prediction of inhibition performance of some benzimidazole derivatives against steel corrosion through QSAR and molecular dynamic simulation,”

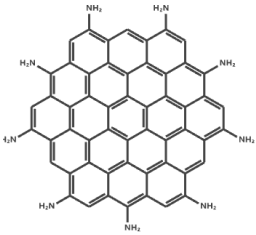
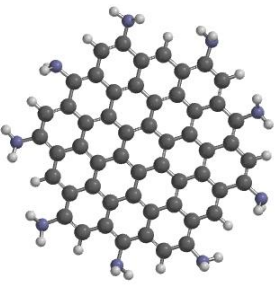
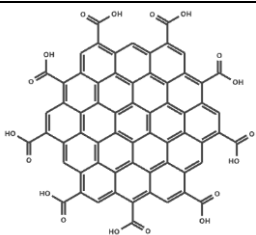
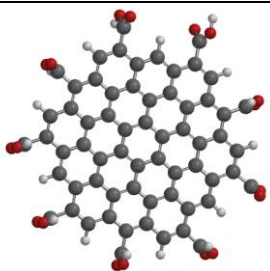
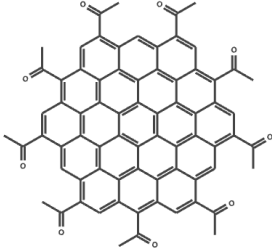
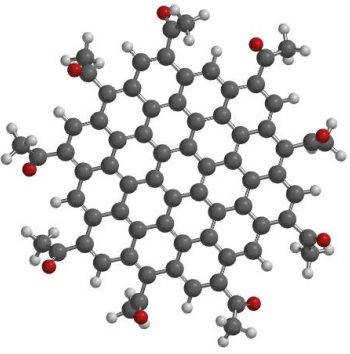
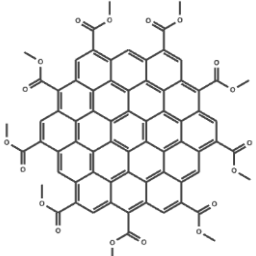
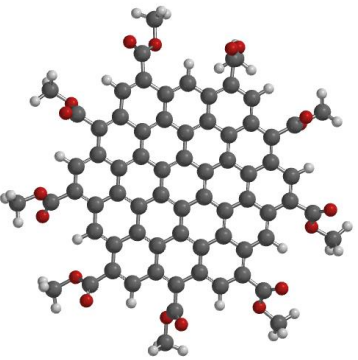
vol. 208, 1, pp. 10–19. January, 2019.

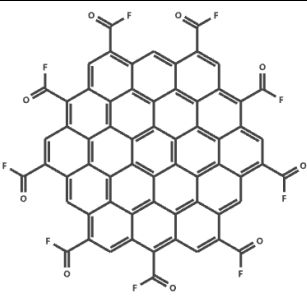
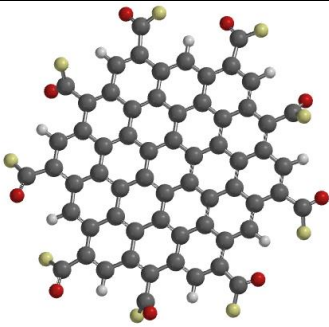
- [52] S. Nagar, “Characterizations and melting analysis of (amine) group - functionalized graphene mixed / doped with paraffin wax as phase change material for thermal management and thermal energy storage applications,” *J. Brazilian Soc. Mech. Sci. Eng.*, no. March, 2024.

APPENDICES

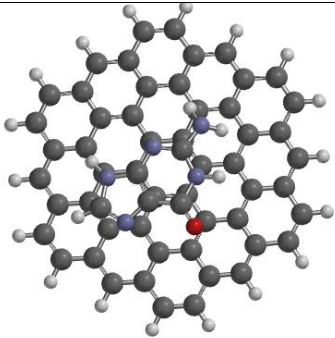
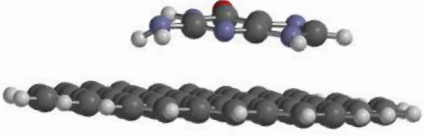
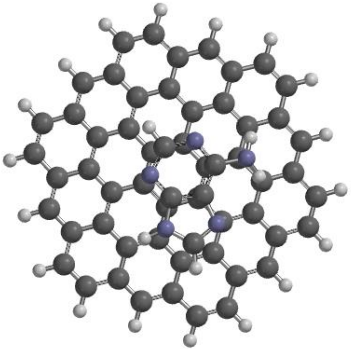
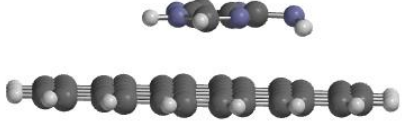
Appendix A 2D and 3D structures of functionalized graphene structure with high sequencing potential.

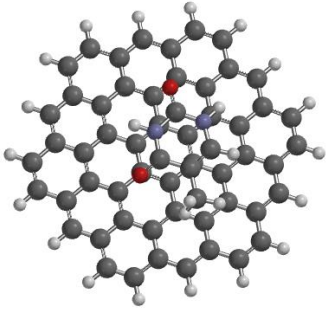
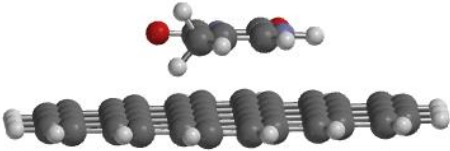
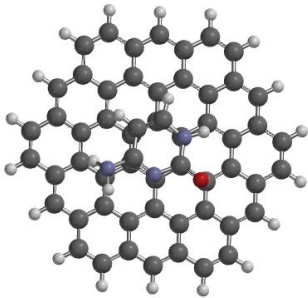
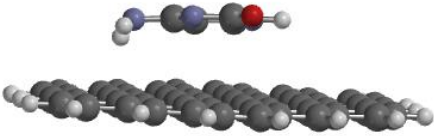
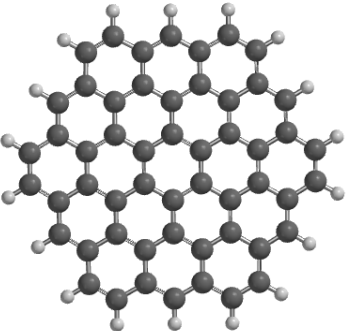
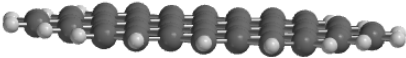
S/N	2D Functionalized Graphene	3D Functionalized Graphene
1.		
	Methyl-Amide	
2.		
	Methyl-Fluoride	
3.		
	Methyl-Iodide	
4.		

	Methyl-Ether	
5.		
	Methyl-Amine	
6.		
	Methyl-Carboxylic Acid	
7.		
	Methyl-Ketone	
8.		
	Methyl-Ester	

9.		
	Methyl-Acid Fluoride	

Appendix B 2D and 3D adsorption system of non-functionalized graphene.

S/N	Top-View	Side-View
1.		
	GNS + G	
2.		
	GNS + A	

3.		
	GNS + T	
4.		
	GNS + C	
5.	 <p style="text-align: center;">GNS</p>	

6.



2D GNS