Fullerene Derivative as Anti-HIV Protease Inhibitor: Molecular Modeling and QSAR Approaches

M. Ibrahim*[,1], N.A. Saleh[2], W.M. Elshemey[2] and A.A. Elsayed[2]

[1]*Spectroscopy Department, National Research Centre, Dokki, Cairo, Egypt*

[2]*Biophysics Department, Faculty of Science, University of Cairo, Giza, Egypt*

Abstract: A Fullerene based system is modified in order to increase its solubility and enhance its ability to carry a protein-like structure. The modified structure, which is proposed to act as HIV-1 protease inhibitor, is [C_{60}–C_2H_4N–(2,4-$XCOCH_2OH)C_6H_4$], where the X atom is either O, S or Se. The geometry optimization, vibrational spectra and thermodynamics were performed using semiempirical quantum mechanical PM3 method in order to study the proposed compounds. Furthermore, the quantitative structure activity relationship (QSAR) properties of the compounds are calculated at the same level of theory. Results indicate a possible use of the investigated structures as HIV-1 protease inhibitors. The compounds containing oxygen is more stable as compared to the other two compounds.

Keywords: Fulleropyrrolidine, HIV-1 Protease, Hydroxymethylcarbonyl group (HMC), Molecular modeling, QSAR.

1. INTRODUCTION

The quantitative structure-activity relationship (QSAR) is a technique for correlating certain structural or property descriptors of a given structure with its activities. Early in 1863, it was observed that the toxicity of alcohols to mammals increased as the water solubility of the alcohols decreased [1]. In 1890's, it was noticed that the toxicity of organic compounds depended on their lipophilicity [1, 2].

Drug design often involves the use of QSAR to identify given chemical structures. Density functional theory (DFT) based on QSAR was used to study the mechanism of N-phenylbenzamides as potent antimicrobial agents [3]. A hybrid system of QSAR models was developed for predicting bioconcentration factors (BCF) [4]. Both DFT and QSAR were used to study the effect of substitution on the AZT drug [5]. Recently, three QSAR programs were used to predict drug-related cardiac adverse effects [6]. As a result of the importance of QSAR applications it continues to be a hot topic of research [7-12].

On the other hand, one of the most widely spread diseases in the world is AIDS. HIV which is the virus responsible for AIDS has shown resistance towards many drugs and is still considered a challenge for scientists and doctors. Two factors are crucial for understanding the underlying principles of drug resistance. First, HIV has a very high reproduction rate. Secondly, the high error rate of the reverse transcriptase when transcribing viral RNA to DNA results in a large number of mutated forms of the virus being produced. The HIV Protease enzyme (HIV-PR) posseses catalytic triad amino acids, $Asp_{25,\ 25'}$ $Thr_{26,\ 26'}$ & $Gly_{27,\ 27'}$ which are not affected by mutation and are considered as conserved residues [13]. This property of HIV-PR active site made it a favourable target in order to inhibit HIV-PR which would in turn inhibit HIV.

C_{60} has many unique physical and chemical properties which are utilized in several technological applications in physics and biology [14-20]. In biology, C_{60} based compounds have been applied as imaging probes [20], antioxidants [21-23], drug carriers [24], and enzyme inhibitors [25]. One very important feature of C_{60} is that it has a diameter comparable to the diameter of HIV-PR active site [25].

In a previous study, the authors of the present work discussed the possible use of their first family of C_{60} – based compounds [C_{60}-C_2H_4N-(2,4-$XCOCH_2OH)C_6H_4$] suggested as HIV-1 protease inhibitors, where X atom is O, S or Se respectively (Fig. **1**). These derivatives were considered to follow a competitive mode of inhibition because of their ability to form hydrogen bonds with the catalytic aspartic acids in addition to van der Waals contacts with the non-polar HIV-1 PR surface, thereby improving the binding [26]. No QSAR properties have been investigated at that time and no other fullerene based families were available for comparison of performance.

This work is conducted in order to explore the QSAR properties and biological activity of [C_{60}-C_2H_4N-(2,4-$XCOCH_2OH)C_6H_4$] compounds (first family) in order to examine their potency to act as HIV-PR inhibitors and compare their QSAR properties with the QSAR properties (specially Log P) of the second and third families of compounds which we have previously reported [27], (Hydroxy-chalca-acetic acid-(4-pyrrolidin-1-yl-phenyl) ester [C_{60}–C_2H_4N–(4-$XCOCH_2$ $OH)C_6H_4$]; second family; and Hydroxy-chalcoacetic acid-[2-(2-hydroxy-acetyl chalcanyl)-4-pyrrolidin-1-yl-phenyl]ester [C_{60}–C_2H_4N–(3,4-$XCOCH_2$ $OH)C_6H_4$]; third family; as shown in Fig. (**2a** & **b**),

*Address correspondence to this author at the Spectroscopy Department, National Research Centre, Dokki, Cairo, Egypt; Tel: 0020122 2727636; Fax: 00202 33370931; E-mail: methat.ibrahim@gmail.com

1875-5607/12 $58.00+.00

respectively where X atom is either O, S or Se for each family.

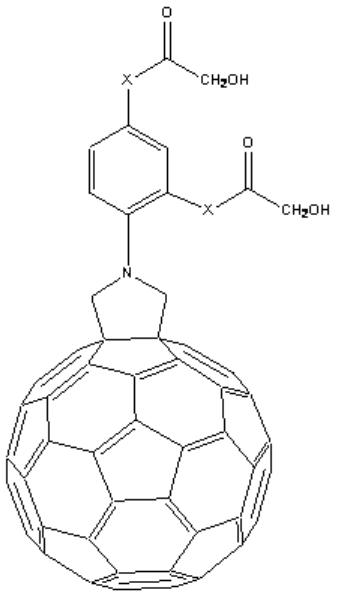


Fig. (1). The general structures of the first family, [C_{60}-C_2H_4N-(2,4-$XCOCH_2OH$) C_6H_4] where x is S, O or Se [26].

2. CALCULATIONS DETAILS

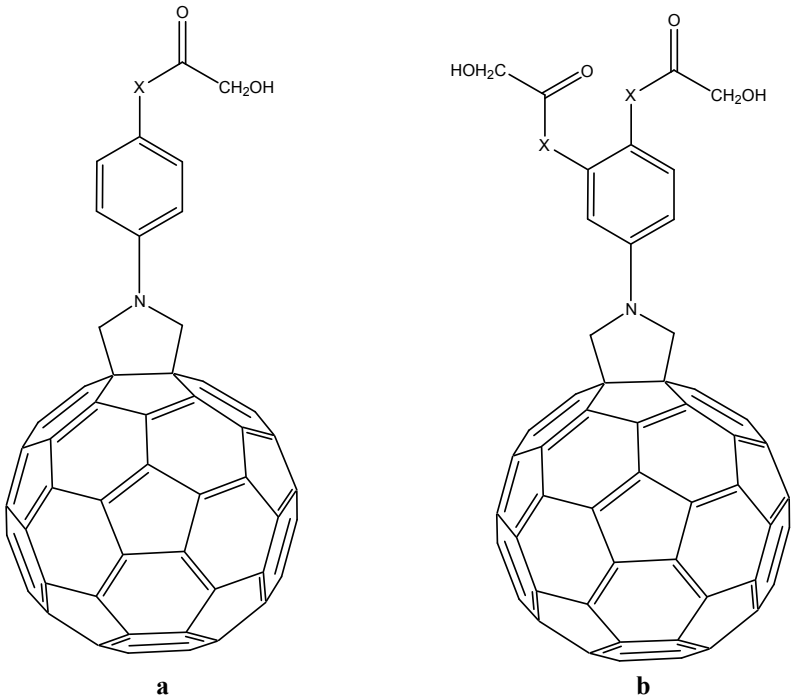
Calculations are carried out on a personal computer. The geometry of the interaction systems between studied compounds and two aspartic acids is optimized by performing PM3 semiempirical quantum mechanical calculation level using a semiempirical quantum mechanics package, MOPAC 2002 which is implemented within the CAChe Program [28].

The calculations of QSAR properties are performed using the Hyperchem version 7.5 program [29]. The geometry of studied compounds is optimized at the PM3 level, using the restricted Hartree–Fock (RHF) procedure. The QSAR calculated properties include partial charges, surface area (Approx & Grid), volume, hydration energy, Log P, refractivity, polarizability and mass.

3. RESULTS AND DISCUSSION

3. 1. Building the Model Molecules

In this work there are a number of modifications introduced to the fullerene based system. The first modification is the fusion of pyrrolidine ring to a 6,6 ring-junction of C_{60}. The addition to 6,6 ring-junction is expected to be better than the addition to 6,5 ring-junction. This is because the electron density across 6,6 ring junction is much higher than the 6,5 ones, due to the dominant resonance structure for C_{60} in which formal double bonds are located at 6,6 ring fusions [30] and play an important role owing to their easy preparation with many possible synthetic variations**.** Furthermore, the mono-functionalization of C_{60} does not considerably alter the basic fullerene properties [31]. Thus, retaining the fullerene properties and having the advantage of possessing higher solubility in polar solvents. Fulleropyrrolidines are among the most studied fullerene derivatives, which have been used for numerous biological applications [32, 33]. There are many polar and hydrophilic groups, which have been introduced onto the fulleropyrrolidine skeleton *via* substitution to the nitrogen or carbon atoms of pyrrolidine moiety in order to improve its



a **b**

Fig. (2). The general structures of (**a**) Second family, Hydroxy-chalca-acetic acid-(4-pyrrolidin-1-yl-phenyl) ester and (**b**) Third family, Hydroxy-chalcoacetic acid-[2-(2-hydroxy-acetylchalcanyl)-4-pyrrolidin-1-yl-phenyl] ester [27].

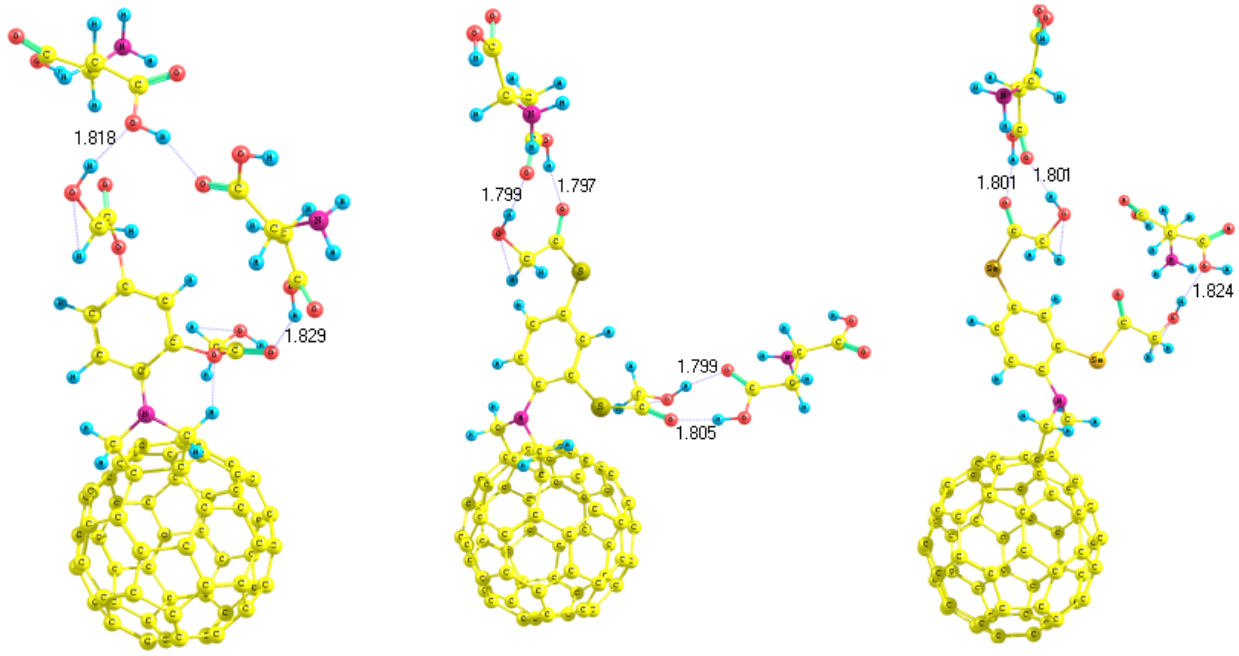


Fig. (3). Optimized interaction systems according to PM3 method, whereas each structure is interacting with two aspartic acids through hydrogen bonding [26].

dissolvability, biocompatibility profile [33] and increase the ability of blocking HIV-PR [26,27]. Nitrogen-containing compounds are used as structural components of pharmaceuticals and agrochemicals due to their high biological activities [34]. In fact, the nitrogen in pyrrolidine and proline derivatives of fullerene C_{60} is much less basic than in the corresponding, non-fullerenic, amines [35].

The second modification is the addition of chalcogen atoms (O, S & Se) and hydroxymethylcarbonyl group (HMC). The NMR, X-ray crystallography and molecular modeling studies [26, 27, 36] showed that the HMC group interacted excellently with the aspartic acid carboxyl groups of the HIV protease active site.

Fig. (**3**) presents the optimized structure of the interaction between the aspartic acids of HIV-PR active site and C_{60}-C_2H_4N-(2,4-$XCOCH_2OH$)C_6H_4 compounds [26]. This figure also shows the distance in angstroms (Å) of the hydrogen bonds in the interaction between studied molecules and aspartic acids.

The optimization energy of the interaction between studied compounds and two aspartic acids at PM3 level are listed in Table **1**. It also shows the number of hydrogen bonds which is formed in this interaction.

3. 2. QSAR Study

QSAR is one of the most important calculations that are useful for optimizing the groups that modulate the potency of a molecule and express the biological activities of compounds.

According to Table **1**, the optimization energy of the interaction systems for O, S and Se molecules are 0.3066 au, 0.4762 au and 0.3893 au respectively. The interaction with oxygen molecule has the lowest optimization energy which is expected to make the interaction more stable. The interaction with sulfur molecule has the highest optimization energy. This will potentially affect the stability of the interaction making it less stable and consequently less suitable for achieving the goal. Comparing these results with the previous work on the interaction systems of the second and third families [27], it can be concluded that the oxygen molecules of all families have the lowest optimization energy of the interaction systems.

As presented in Table **1**, all compounds formed hydrogen bonds with the two aspartic acids. Generally, the hydrogen bonds are formed mainly between aspartic acids and the oxygen and hydroxyl groups of hydroxymethylcarbonyl (Fig. **3**).

Table 1. Optimization Energy Values of Interaction Systems Between Studied Compounds and Two Aspartic Acids in (au) and Number of Hydrogen Bonds According to PM3 Method

Properties	O	S	Se
Optimization energy (au)	0.3066	0.4762	0.3893
No. of H-bond	2.0000	4.0000	3.0000

Table 2. QSAR Properties of the Studied Molecules Calculated for the PM3 Optimized Structures

Properties	O	S	Se
Partial charges e	0.00	0.00	0.00
Surface area (Approx) $Å^2$	361.37	382.33	397.30
Surface area (Grid) $Å^2$	853.53	864.02	869.25
Volume $Å^3$	1807.82	1848.79	1858.66
Hydration energy Kcal/mol	-17.94	-15.36	-12.76
Log p	7.47	8.16	6.93
Refractivity $Å^3$	294.82	307.70	292.12
Polarizability $Å^3$	105.28	110.00	108.51
Mass amu	987.90	1020.02	1113.82

Table **2** shows some molecular descriptors of QSAR properties of the investigated molecules. The presented values include partial charges, surface area (Approx & Grid), volume, hydration energy, Log P, refractivity, polarizability and mass. All three compounds have zero partial charge which indicates that all compounds are in ground state.

In the presented calculations, the total surface areas and volumes increased from oxygen molecule to selenium molecule. This was due to increasing the atomic number and size of the chalcogen atoms from O atom to Se atom. The surface area (Approx) and (Grid) varied from 361.37 $Å^2$ to 397.3 $Å^2$ and from 853.53 $Å^2$ to 869.25 $Å^2$ respectively. The total volume of the studied compounds varied from 1807.82 $Å^3$ to 1858.66 $Å^3$ respectively.

The effect of increasing the atomic number of chalcogen atoms is quite clear in the values of the mass of presented compounds. The mass values are 987.90 amu, 1020.02 amu and 1113.82 amu for the O, S and Se molecules respectively.

Table **2** also shows the values of hydration energies of compounds. The hydration energy is the amount of energy released when a mole of the ion dissolves in a large amount of water forming an infinite dilute solution. It varies for the studied compounds from -17.94 Kcal/mol to -12.76 Kcal/mol.

The polarizability of the investigated compounds is also presented in Table **2**. As the polarizability increases, the magnitude of the instantaneous dipole will also increase resulting in a stronger attraction to other systems. This table also shows that the variation in polarizability values of studied molecules is very small. The oxygen molecule has the lowest polarizability (105.28 $Å^3$), while the sulfur molecule has the highest polarizability (110.00 $Å^3$).

Log P is the most important parameter of the QSAR presented descriptors. A high value of log P indicates that the molecule is hydrophobic while a low value of log P indicates that the molecule is hydrophilic. The calculated log P values of the investigated compounds are 7.47, 8.16 and 6.93 for O, S and Se molecules respectively. This indicates that, the O and Se molecules are more hydrophilic and soluble than the S molecule.

The log P values of oxygen molecules of the second and third families are -12.46 and -13.52 respectively [27]. This indicates that the oxygen molecules of the second and third families are more hydrophilic and soluble than that of the first family and that the compound with the oxygen molecule of the second family is the most soluble of all compounds (-13.52).

3. CONCLUSION AND RECOMMENDATIONS

Based on a related previous study [26, 27] and the present work involving the calculation and comparison of the QSAR properties of three suggested C_{60} families of compounds, it can be concluded that the oxygen compounds of the three families offer the most stable interaction with the two aspartic acids of HIV-PR active site. This is based on the fact that they have the lowest optimization energy of interaction among the investigated compounds with aspartic acid and considerably low log P values. Therefore, the oxygen compounds of the three families are probably the best of all of the investigated compounds for HIV-1 protease inhibition.

CONFLICT OF INTEREST

None declared.

REFERENCES

[1] Borman, S. New QSAR Techniques Eyed for Environmental Assessments. *Chem. Eng. News*, **1990**, *68*, 20-23.

[2] Lipnick, R. L. Charles Ernest Overton: Narcosis Studies and a Contribution to General Pharmacology. *Trends Pharmacol. Sci.*, **1986**, *7*, 161-164.

[3] Pasha, F.A.; Muddassar, M.; Lee, C.; Cho, S. J. Mechanism based QSAR studies of N-phenylbenzamides as antimicrobial agents. *Enviro. Toxicol. Pharmacol.*, **2008**, *26*, 128-135.

[4] Zhao, C.; Boriani, E.; Chana, A.; Roncaglioni, A.; Benfenati, E. A new hybrid system of QSAR models for predicting bioconcentration factors (BCF). *Chemosphere*, **2008**, *73*, 1701-1707.

[5] Essa, A. H.; Ibrahim, M.; Hameed, A. J.; and Al-Masoudi, N. A. Theoretical investigation of 3'-subtituted-2'-3'-dideoxythymidines related to AZT. QSAR infrared and substituent electronic effect studies. *ARKIVOC*, **2008**, *xiii*, 255-265.

[6] Frid, A. A.; Matthews, E. J. Prediction of drug-related cardiac adverse effects in humans-B: Use of QSAR programs for early detection of drug-induced cardiac toxicities. *Regul. Toxicol. Pharmacol.,* **2010**, *56*, 276-289.

[7] Yangali-Quintanilla, V.; Sadmani, A.; McConville, M.; Kennedy, M.; Amy, G. A QSAR model for predicting rejection of emerging contaminants (pharmaceuticals, endocrine disruptors) by nanofiltration membranes. *Water Res.,* **2010**, *44*, 373-384.

[8] Soltani, S.; Abolhasani, H.; Zarghi, A.; Jouyban, A. QSAR analysis of diaryl COX-2 inhibitors: Comparison of feature selection and train-test data selection methods. *Eur. J. Med. Chem.*, **2010**, *45*, 2753-2760.

[9] Gozalbes, R.; Pineda-Lucena, A. QSAR-based solubility model for drug-like compounds. *Bioorg. Med. Chem.,* **2010**, *18*, 7078-7084.

[10] Zeng, H.; Zhang, H. Combined 3D-QSAR modeling and molecular docking study on 1,4-dihydroindeno[1,2-c] pyrazoles as VEGFR-2 kinase inhibitors. *J. Mol. Graphics Model.*, **2010**, *29*, 54-71.

[11] Abu Khalaf, R.; Abu Sheikha, G.; Bustanji, Y.; Taha, M. O. Discovery of new cholesteryl ester transfer protein inhibitors *via* ligand-based pharmacophore modeling and QSAR analysis followed by synthetic exploration. *Eur. J. Med. Chem.*, **2010**, *45*, 1598-1617.

[12] Srivastava, A. K.; Pandey, A.; Srivastava, A.; and Shukla, N. QSAR Based Modeling Of Hepatitis C Virus NS5B Inhibitors. *JSCS.*, **2011**,*15*, 25-28.

[13] Wang, W.; Kollman, P. A. Computational study of protein specificity: The Molecular basis of HIV-1protease drug resistance. *PNAS.*, **2001**, *98*, 14937-14942.

[14] Hebgen, P.; Goel, A.; Howard, J. B.; Rainey, L. C.; Sande, J. B. V., synthesis of fullerenes and fullerenic nanostructures in a Low-pressure benzene/oxygen diffusion flame. *Proc. Combust. Inst.,* **2000**, *28*, 1397-1404.

[15] Ibrahim, M.; El-Haes, H.; Hameed, A. J.; Jalbout, A. F. Structural and Electronic Properties of $C_{60}X_6$ (X= F, Cl, Br and I). A Theoretical Study. *J. Comput. Theor. Nanosci.,* **2008**, *5*, 2247-2251.

[16] Ibrahim, M.; El-Haes, H.; Hameed, A. J.; Essa, A. H. Spectroscopic Analysis of C_{80} Doping with Group III and Group V Elements Using Semiempirical PM3 Molecular Modelling Technique. *Ecl. Quim. Sao Paulo*., **2008,** *30*, 21-27.

[17] Jalbout, A. F.; Hameed, A. J.; Jimenez-Fabian, I.; Ibrahim, M.; de Leon, A. Chalcanthrene-Fullerene Complexes. A Theoretical Study. *J. Organomet. Chem.,* **2008,** *693*, 216-220.

[18] Ibrahim, M.; Hameed, A. J.; Essa, A. H. Analysis of the Vibrational Spectra and Thermal Parameters of C_{60} Chacognide Dimers. *J. Comput. Theor. Nanosci*., **2009**, *6*, 574-577.

[19] Elhaes, H.; Babaeer, A.; and Ibrahim, M. Effect of Metal Substitution on the Electronic Properties of Fullerene and Fulleropyrrolidine. *J. Comput. Theor. Nanosci*., **2010**, *7*, 536-541

[20] Hameed, A. J.; Ibrahim, M.; El-Haes, H. Computational notes on structural and electronic properties of Fulleropyrrolidine-1-carbodithioic acid 2; 3 and 4-substituted-benzyl esters. *J. Mol. Struct- THEOCHEM.,* **2007**, *809*, 131-136.

[21] Dugan, L. L.; Gabrielsen, J. K.; Yu, S. P.; Lin, T. S.; Choi, D. W. Buckminsterfullerenol free radical scavengers reduce excitotoxic and apoptotic death of cultured cortical neurons. *Neurobiol. Dis.,* **1996**, *3*, 129-135.

[22] Dugan, L. L.; Turetsky, D. M.; Du, C.; Lobner, D.; Wheeler, M.; Almli, C. R.; Shen, C. K.; Luh, T. Y.; Choi, D. W.; Lin, T. S. Carboxyfullerenes as neuroprotective agents. *PNAS*., **1997,** *94*, 9434-9439.

[23] Zakharian, T. Y.; Seryshev, A.; Sitharaman, B.; Gilbert, B. E.; Knight, V.; Wilson, L. J. A Fullerene (C_{60})-Paclitaxel Chemotherapeutic: Synthesis, Characterization and Study of Biological Activity in Tissue Culture. *J. Am. Chem. Soc.,* **2005**, *127*, 12508-12509.

[24] Bosi, S.; Da Ros, T.; Spalluto, G.; Prato, M. Fullerene derivatives: an attractive tool for biological applications. *Eur. J. Med. Chem.,* **2003,** *38*, 913-923.

[25] Friedman, S.H.; Decamp, D.L.; Sijbesma, R.P.; Srdanov, G.; Wudl, F.; Kenyon, G. L. Inhibition of the HIV-1 protease by fullerene derivatives: model building studies and experimental verification. *Am. Chem. SOC.,* **1993**, *115*, 6506-6509.

[26] Ibrahim, M.; Saleh, N. A.; Elshemey, W. M.; Elsayed, A. A. Computational notes on Fullerene Based System as HIV-1 Protease inhibitors. *Comput. Theor. Nanosci.,* **2010,** *7*, 224-227.

[27] Ibrahim, M.; Saleh, N. A.; Hameed, A. J.; Elshemey, W. M.; Elsayed, A. A. Structural and electronic properties of new fullerene derivatives and their possible application as HIV-1 protease inhibitors. *Spectrochim. Acta A*., **2010**, *75*, 702-709.

[28] MOPAC 2002 Version 2.5.3, J. J. P. Stewart, Fujitsu Limited, Tokyo, Japan, **2002**.

[29] Hyperchem Release 7.5, Windows Molecular Modeling System; Hypercube/Autodesk, Hypercube: Gainesville, FL, **2000.**

[30] Gogotsi, Y.; Carbon Nanomaterials, Taylor and Francis group,6000 Broken Parkway NW, Suite 300, Boca Raton, FL, **2006**, 33487-2742.

[31] Echegoyen, L.; Echegoyen, L.E. Electrochemistry of Fullerenes and Their Derivatives. *Acc. Chem. Res.,* **1998**, *31*, 593-601.

[32] Benne, D.; Maccallini, E.; Rudolf, P.; Sooambar, C.; Prato, M. X-Ray Photoemission Spectroscopy Study on The Effects of Functionalization in Fulleropyrrolidine and Pyrrolidine Derivatives. *Carbon*, **2006**, *44*, 2896-2903.

[33] Jalbout, A. F.; Hameed, A. J.; Trzaskowski, B. Study of The Structural and Electronic Properties of 1-(4, 5 And 6-Selenenyl Derivatives-3-Formyl-Phenyl) Pyrrolidinofullerenes. *J. Organomet. Chem.*, **2007**, *692*, 1039-1047.

[34] Higashio, Y.; Shoji, T. Heterocyclic Compounds Such as Pyrrole, Pyridines, Pyrrolidine, Piperidine, Indole, Imidazol and Pyrazines. *Appl. Catal. A.*, **2004**, *260*, 251-259.

[35] Ros, T. D.; Prato, M. Medicinal Chemistry with Fullerenes and Fullerene Derivatives. *Chem. Commun.*, **1999**, *8*, 663-669.

[36] Kiso, Y.; Matsumoto, H.; Yamaguchi, S.; Kimura, T. Design of Small Peptidomimetic HIV-1 Protease Inhibitors and Prodrug Forms. *Lett. Pept. Sci.,* **1999**, *6*, 275-281.

Received: October 11, 2010 Revised: June 11, 2011 Accepted: June 12, 2011