

# A Novel Peptidomimetic Compounds as HCV-NS3 Protease Inhibitors: Spectroscopic Analysis

Medhat Ibrahim<sup>a</sup>, Noha A. Saleh<sup>\*b</sup>, Wael M. Elshemey<sup>b</sup>, and Anwar A. Elsayed<sup>b</sup>

<sup>a</sup>Spectroscopy Department, National Research Centre, Dokki, Cairo, Egypt

<sup>b</sup>Biophysics Department, Faculty of Science, University of Cairo, Giza, Egypt

**Abstract:** The presented molecular modeling is utilized to innovate new peptidomimetic compounds. These proposed compounds are designed to act as HCV NS3 protease antiviral. The suggested antivirals are divided into two groups. The first group has hexapeptide (Glu-Asp-Val-Val-Cys-Cys) binding to cellulose monomer at positions 2, 3 or 6 while the second group has hexapeptide (Glu-Asp-Val-Val-Cys-Cys) binding to cellulose dimer at positions 2, 3, 6, 2', 3' or 6'. Semi-empirical PM3 quantum mechanical method is first utilized for optimization, then to calculate the vibrational spectra of these novel compounds. It is found that higher dipole moment (11.907 Debye) corresponds to the hexapeptide (Glu-Asp-Val-Val-Cys-Cys) binding to cellulose dimer at position 2' compound. Accordingly, calculation is repeated at HF/3-21g\*\* and B3LYP/3-21g\*\* for such compound for verification.

**Keywords:** FTIR, HCV, Molecular modeling, NS3 protease, NS5A/NS5B junction, PM3 and subtype 4a.

## INTRODUCTION

The hepatitis C virus (HCV) is one of the most widely spread epidemic diseases in the world. HCV belongs to the Flaviviridae family and is the only member of the Hepacivirus genus [1]. HCV was discovered in 1989 as the etiological agent for non-A, non-B hepatitis. Since its discovery, HCV is becoming a hot topic of research due to the large number of reported worldwide HCV infections (about 200 million cases) [2]. Of those infected, over 85% will develop chronic hepatitis, and 20% of the chronic infections progress to liver cirrhosis and hepatocellular carcinoma (HCC) [3]. HCV has six major genotypes and more than 70 subtypes. The genotype 4 is prevalent in the Middle East and different parts of Africa. Egypt is considered as one of the highest countries infected with HCV in the world [4]. HCV genotype 4 (HCV-4); especially subtype 4a; is common in Egypt [5-9]. Although HCV is a major health problem especially in Egypt, there is only a single available treatment (a combination of interferon- $\alpha$  and ribavirin) which is expensive and effective in only 50–60% of treated patients. Hence, there is a pressing need for new and broadly active antiviral agents to combat HCV infections [10].

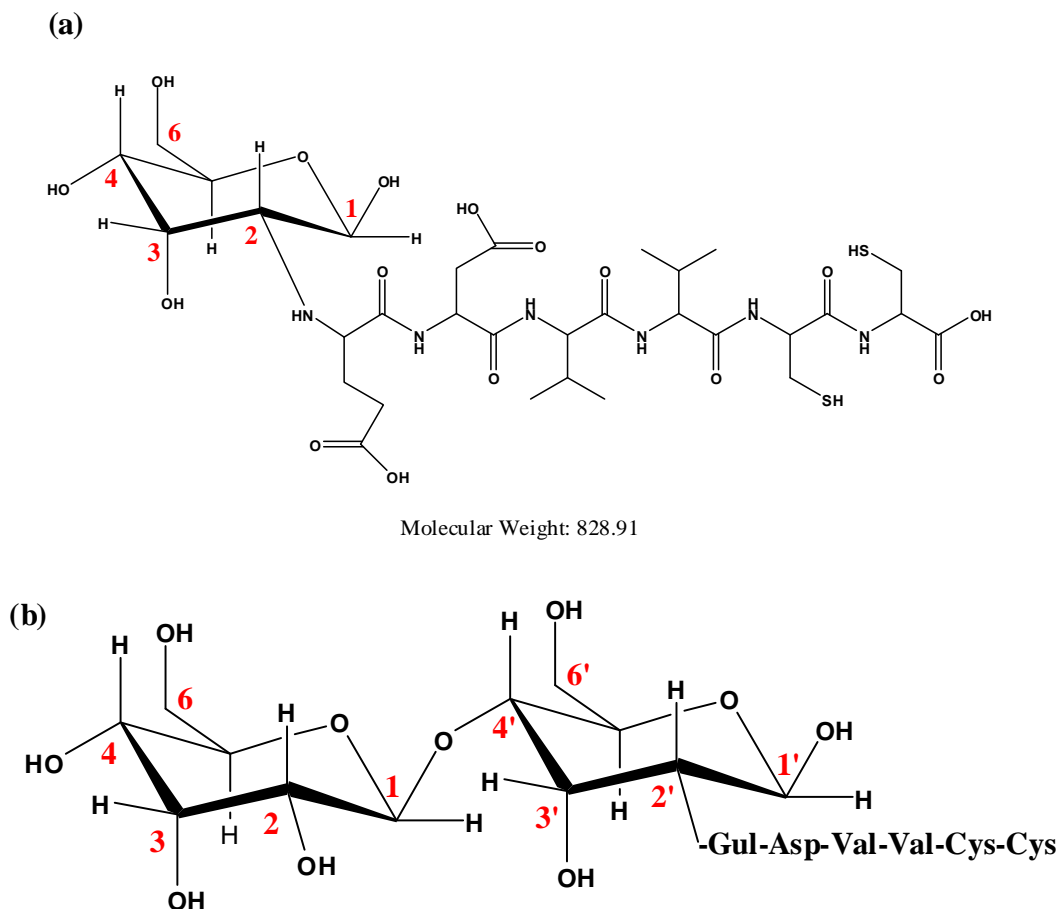
The progress of research in HCV is very difficult due to the inability to cultivate HCV *in vitro* and the absence of a small animal model susceptible to HCV infection. The chimpanzee is the only animal known to be susceptible to HCV infection. Due to these difficulties, the molecular modeling simulation can overcome the problems in cultivating HCV and enhance the progress of HCV research [2, 3].

Molecular modeling can be simply considered as a range of computerized techniques based on the basic laws of physics and experimental data that can be used either to analyze molecules (number and types of atoms, bond, bond lengths, angles and dihedral angles also the basic characteristics of the molecule such as molecular energy, geometry optimization energy, enthalpy and vibrational frequency) and molecular systems (nucleophilicity, electrophilicity and electrostatic potentials) or to predict molecular and biological properties (which are primarily important in the understanding of structure-activity relationships and in rational drug design); [11-18]. Therefore, molecular modeling helps in studying the biological activity of the new inhibitors and the physical aspect of the interaction with the virus and selecting the possibly best compound to synthesize, thus, saving time, effort and money.

The NS3 protease of HCV is considered as one of essential enzymes for HCV life cycle. So this enzyme is considered as major antiviral target. NS3 protease is responsible for the proteolytic cleavage at four junctions of the HCV polyprotein precursor: NS3/NS4A (self cleavage), NS4A/NS4B, NS4B/NS5A, and NS5A/NS5B [19-22]. In our previous work, the hexapeptide compounds were suggested and built as HCV NS3 protease inhibitors especially for Egyptian genotype 4a. The electronic properties of these suggested inhibitors were studied theoretically by molecular modeling programs at semi-empirical method (Fig. 1).

The Fourier transform infrared spectroscopy (FTIR) is one of the classical methods for structure determination of small molecules. The Fourier transform infrared spectroscopy is very sensitive to the presence of chemical functional groups and architecture of the compounds. The high information content in an infrared spectrum also covers biological systems. The most powerful aspect of Fourier transform in-

\*Address correspondence to this author at the Science faculty, Cairo University Egypt; Tel: +202-01144605000; Fax: +202 -35728843; E-mail 1: noha\_saleh\_kh@yahoo.com



**Fig. (1).** Scheme 1, the general structure of the suggested compounds (a) First group; cellulose monomer binding with hexapeptide at position 2, 3 or 6 with molecular weight equal to 828.91. (b) Second group, cellulose dimer binding with hexapeptide at position 2, 3, 6, 2', 3' or 6'.

frared spectroscopy is the identification of unknown compound and comparing the spectra of two compounds with each other to determine whether the compounds have the same composition [23, 24]. In order to help in the process of production of novel structures it is important to predict theoretically their vibrational characteristics. So in this work, the vibrational analyses and assignment of the proposed inhibitors are supported by modeling programs at semi-empirical PM3 level of theory.

### COMPUTATIONAL DETAILS

Calculations are carried out on a personal computer. Semiempirical PM3 quantum mechanical calculations are performed using Gaussian 03 program system with total charge equal 0.0 and spin multiplicity 1 [25]. The vibrational spectra calculations are performed upon the optimized compounds at the same level of theory method. Based on calculated dipole moment one of the studied compounds is reoptimized using HF/3-21g\*\* and B3LYP/3-21g\*\* [26-28].

### RESULTS AND DISCUSSION

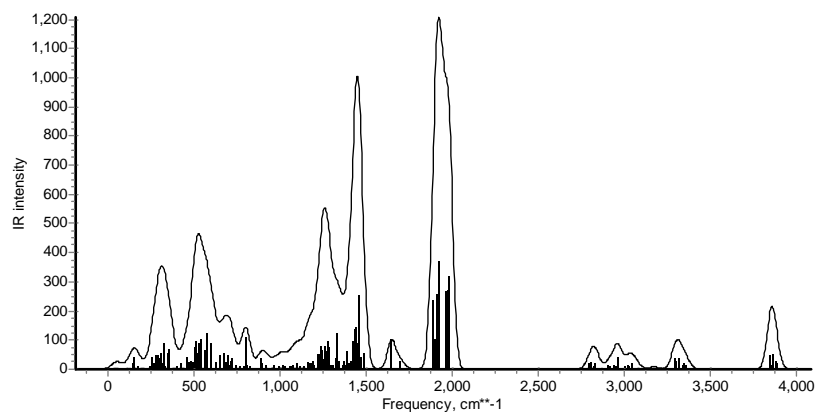
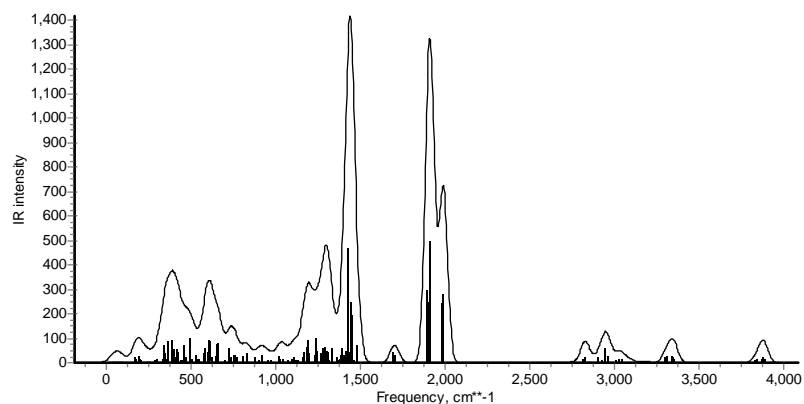
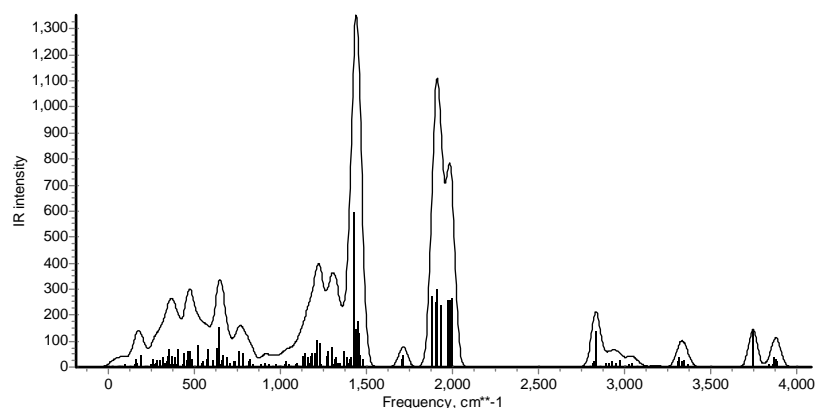
The NS5A/NS5B junction is one of the four junctions with proteolytic cleavage by NS3 protease. The sequence of

NS5A/NS5B junction for Egyptian genotype 4 sequence is Glu-Asp-Val-Val-Cys-Cys which is related to hexapeptide (P6-P5-P4-P3-P2-P1) [7]. This sequence may be used as a good target to propose new antiviral for HCV-4a genotype. These new inhibitors are considered as peptidomimetic inhibitors and acting as competitive mode of inhibition.

One of the most common organic compounds on the earth is cellulose; which composes about 33% of all plant matter [29-30]. Cellulose and its derivatives are becoming an emerging materials in biology as well as many other fields of research [31, 32].

Based upon the above considerations two groups of novel competitive inhibitors are introduced and built. The first group has cellulose monomer which acts as substrate to carry hexapeptide (Glu-Asp-Val-Val-Cys-Cys) at position 2, 3 or 6 as seen in scheme 1-a. While the second group consists of cellulose dimer at which the hexapeptide (Glu-Asp-Val-Val-Cys-Cys) binds at position 2, 3, 6, 2', 3' or 6' as seen in scheme 1-b.

All of the mentioned structures in Table 1 show their optimized structure then IR is calculated upon the optimized structures (Figs. 2 & 3). Table 1 represents the calculated vibrational frequencies and their intensities as well as band

**Position 2****Position 3****Position 6**

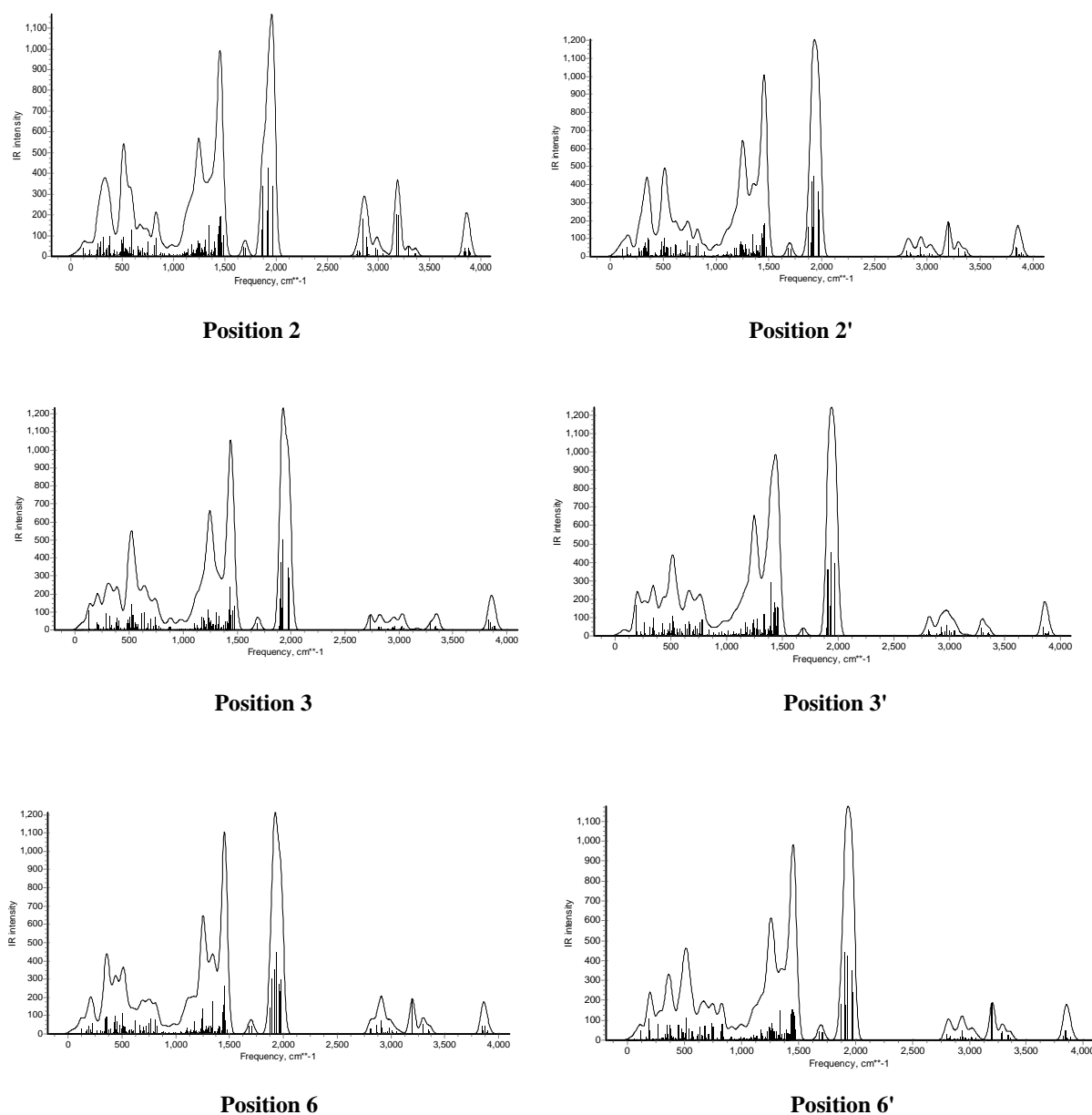
**Fig. (2).** calculated IR spectrum of First group; cellulose monomer binding with hexapeptide at position 2, 3 or 6 at PM3 semi-empirical method.

assignment at PM3 method for first and second groups of proposed compounds.

Although the normal modes for group 1 and for group 2, only 9 vibrations are mentioned in the table. These vibrations are corresponding to higher band intensities while other bands are not considered because they are not expected to happen experimentally in terms of their low or even zero intensities.

Regarding Table 1 one can characterize the vibrations as those belonging to cellulose and those belonging to amino acids. OH and CH could be assigned for cellulose, while amides are the main features of amino acids and protein structures.

The stretching vibration of OH is located for the first group from  $3881\text{ cm}^{-1}$  to  $3742\text{ cm}^{-1}$ . While for second group it varies from  $3870\text{ cm}^{-1}$  to  $3835\text{ cm}^{-1}$ . As a result of the peptide bond a stretching band of NH exists and is also named



**Fig. (3).** calculated IR spectrum of Second group, cellulose dimer binding with hexapeptide at position 2, 3, 6, 2', 3' or 6' at PM3 semi-empirical method.

as amide A. The NH vibration occurs from  $3341\text{ cm}^{-1}$  to  $3293\text{ cm}^{-1}$  for first group and occurs from  $3294\text{ cm}^{-1}$  to  $3175\text{ cm}^{-1}$  for second group. CH produces so many vibrations in its characteristic region. These bands are coming as symmetric and asymmetric vibrations of  $\text{CH}_2$  and simply for CH alone.

For the first and second groups, the bands corresponding to  $\text{CH}_2$  asymmetry show vibrations from  $2970\text{ cm}^{-1}$  to  $2945\text{ cm}^{-1}$  and from  $3047\text{ cm}^{-1}$  to  $3028\text{ cm}^{-1}$  respectively. For the same groups  $\text{CH}_2$  symmetry shows vibrations from  $2964\text{ cm}^{-1}$  to  $2887\text{ cm}^{-1}$  from  $3031\text{ cm}^{-1}$  to  $2981\text{ cm}^{-1}$  respectively. Finally CH shows vibrations at  $2888\text{ cm}^{-1}$  to  $2813\text{ cm}^{-1}$  and from  $2869\text{ cm}^{-1}$  to  $2736\text{ cm}^{-1}$ , respectively.

Amides are coming in two respective bands known as amid I and amid II, they come as a result of the existence of amino acids. The position of amide group is shifted towards

higher wave number as it is affected from the interactions with cellulose monomer and dimmers respectively.

Amide I which is mainly  $\text{C}=\text{O}$  vibration occurs from  $1926\text{ cm}^{-1}$  to  $1908\text{ cm}^{-1}$  for first group and from  $1936\text{ cm}^{-1}$  to  $1911\text{ cm}^{-1}$  for second group. Another amid attributed to aspartic and glutamic acids shows vibrations from  $1993\text{ cm}^{-1}$  to  $1981\text{ cm}^{-1}$  for first group and varies from  $1981\text{ cm}^{-1}$  to  $1968\text{ cm}^{-1}$  for second group.

The SH vibration of thiol group of cysteine amino acid appears from  $1713\text{ cm}^{-1}$  to  $1648\text{ cm}^{-1}$  for first group and from  $1690$  to  $1686$  for second group. The N-H vibration has another band which is termed also as amid II. This band is produced from  $1445\text{ cm}^{-1}$  to  $1460\text{ cm}^{-1}$  for first group compounds and from  $1441\text{ cm}^{-1}$  to  $1465\text{ cm}^{-1}$  for second group. Other CH vibrations are observed at lower frequencies but their band intensities are very low and these bands are not

mentioned because we concentrate only on the main vibrational characteristics.

The total calculated dipole moment is a good indication for the reactivity of a given structure. Higher dipole moment reflects the ability of a given structure for interaction with the surrounding molecules [33]. Accordingly, we calculate the total dipole moment for the studied compounds at PM3 level of theory. Calculated total dipole moments are listed in Table 2. Higher dipole moment values correspond to the second group compounds at positions 2' and 6' as 11.907 Debye and 10.432 Debye, respectively. Accordingly, the compound corresponding to higher dipole moment is re-

calculated at higher level of theory such as HF/3-21g\*\* and B3LYP/3-21g\*\*.

**HIGHER LEVEL OF THEORY**

Based upon the calculated dipole moment the compound 2' (cellulose dimer binded with genotype sequence Glu-Asp-Val-Val-Cys-Cys) from the second group are optimized at HF/3-21g\*\* and B3LYP/3-21g\*\* as seen in Fig. (4). Total dipole moment as well as geometrical parameters are summarized in Table 3. Total dipole moment is initially 11.907 debye and is 10.739 debye corresponding to HF/3-21g\*\*, while it is 11.246 debye corresponding to B3LYP/3-21g\*\*.

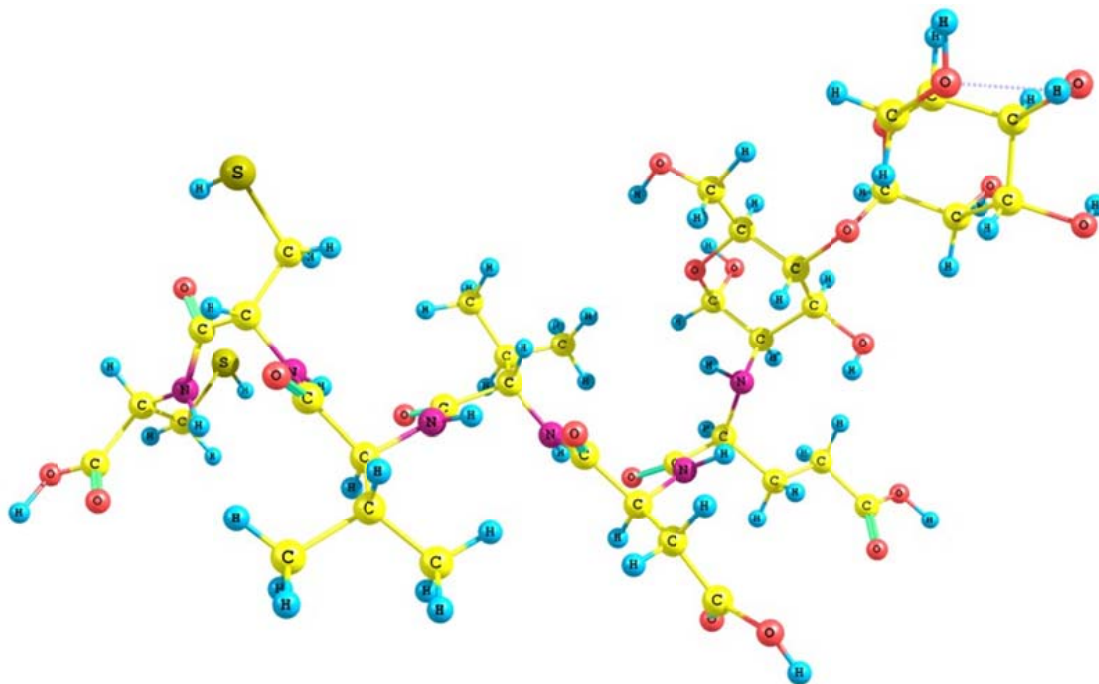


Fig. (4). HF/3-21g\*\* optimized structure for cellulose dimer bound with sequence Glu-Asp-Val-Val-Cys-Cys at position 2'.

Table 1. Calculated Band Positions and Assignment of Suggested Compounds via PM3 Semi-empirical Method

	O-H		N-H		CH <sub>2</sub> symmetry		CH <sub>2</sub> asymmetry		C-H		COOH		C=O		S-H		NH	
	Freq.	Inten.	Freq.	Inten.	Freq.	Inten.	Freq.	Inten.	Freq.	Inten.	Freq.	Inten.	Freq.	Inten.	Freq.	Inten.	Freq.	Inten.
(1 <sup>st</sup> . Group) Monomer																		
2	3864	48	3293	38	3047	16	2964	42	2813	22	1981	316	1926	367	1648	99	1460	252
3	3881	19	3341	29	3028	17	2945	58	2888	19	1993	279	1913	495	1693	43	1445	220
6	3742	144	3315	35	3046	13	2970	25	2833	141	1992	261	1908	301	1713	48	1448	177
(2 <sup>nd</sup> . Group) Dimer																		
2	3850	42	3175	203	2997	29	2887	93	2853	182	1972	340	1935	426	1688	43	1456	189
3	3835	55	3293	36	3031	22	2964	21	2736	83	1971	350	1929	501	1690	35	1441	242
6	3870	40	3198	189	2987	34	2908	72	2869	45	1981	297	1932	448	1688	42	1455	263
2'	3847	48	3200	190	3022	20	2943	51	2806	30	1970	259	1932	446	1686	45	1465	183
3'	3847	46	3294	41	2981	59	2929	45	2819	31	1971	392	1936	453	1688	43	1440	178
6'	3846	48	3203	184	3022	19	2942	50	2810	27	1968	351	1911	444	1686	46	1450	153

**Table 2. The Calculated Dipole Moment as Debye of the Suggested Compounds at PM3 Semi-empirical Method**

	(1 <sup>st</sup> . Group) Monomer			(2 <sup>nd</sup> .Group) Dimmer					
	2	3	6	2	3	6	2'	3'	6'
Total dipole moment	3.402	8.071	6.906	8.383	5.937	3.420	11.907	3.953	10.432

**Table 3. Calculated Total Dipole Moment (TDM) as Debye ; Bond Lengths C-O; C=O and Bond Angle OCO for Cellulose Dimmer Binded with Genotype Sequence Glu-Asp-Val-Val-Cys-Cys which is Related to Hexapeptid Position 2' (P2') Compound Calculated at PM3; HF/3-21g\*\* and B3LYP/3-21g\*\***

	TDM	C-O	C=O	OCO
PM3	11.907	1.350	1.219	116.182
HF/3-21g**	10.739	1.352	1.202	123.132
B3LYP/3-21g**	11.246	1.376	1.225	123.707

These results indicate that PM3 gives comparable results to what was obtained with higher level of theory for this class of compounds. The C-O; C=O and OCO angles are the geometrical parameters of carboxyl group of aspartic acid. Bond lengths are comparable to each other while the bond angle is slightly smaller than those calculated with higher level of theory.

## CONCLUSION

We present novel two families of compounds based on one and two cellulose units to act as HCV-NS3 protease inhibitors. The vibrational assignment is aided by molecular modeling work in order to describe their vibrational characteristics. The vibrational frequencies are shifted towards higher frequencies because of the interaction that was verified by theoretical method. Higher dipole moment which is 11.907 Debye indicating the existence of the compound which has cellulose dimmer bind with hexapeptide sequence Glu-Asp-Val-Val-Cys-Cys at position 2'. Comparison with higher level of theory such as HF and DFT at B3LYP indicate the suitability of PM3 for studying such compounds.

## ACKNOWLEDGEMENT

Declared none.

## CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

## REFERENCES

- Brass V, Moradpour D, Blum HE. Molecular virology of hepatitis C virus (HCV): 2006 update. *Int J Med Sci* 2006; 3: 29-34.
- Krekulova L, Rehak V, Riley LW. Structure and functions of hepatitis C virus proteins: 15 years after. *Folia Microbiol (praha)* 2006; 51: 665-80.
- Ismail NSM, El Dine RS, Hattori M, Takahashi K, Ihara M. Computer based design, synthesis and biological evaluation of novel indole derivatives as HCV NS3-4A serine protease inhibitors. *Bioorg Med Chem* 2008; 16: 7877-87.
- Deuffic-Burban S, Mohamed MK, Larouze B, Carrat F, Valleron A. Expected increase in hepatitis C-related mortality in Egypt due to Pre-2000 infections. *Hepatology* 2006; 44: 455-61.
- Bartenschlager R, Lohmann V. Replication of hepatitis C virus. *Gen Virol* 2000; 81: 1631-48.
- Kamal SM, Nasser IA. Hepatitis C genotype 4: what we know and what we don't yet know. *Hepatology* 2008; 47: 1371-83.
- Timm J, Neukamm M, Kuntzen T, et al. Characterization of full-length hepatitis C virus genotype 4 sequences. *Viral Hepat* 2007; 14: 330-7.
- Tibbs CJ. Tropical aspects of viral hepatitis. *Hepatitis C. Trans R Soc Trop Med Hyg* 1997; 91: 121-4.
- Rapicetta M, Argentini C, Dettori S, Spada E, Pellizzer G, Gandin C. Molecular heterogeneity and new subtypes of HCV genotype 4. *Res Virol* 1998; 149: 293-7.
- Lehman LS, Paulson M, Shih IH, et al. Substituted imidazopyridines as potent inhibitors of HCV replication. *Hepatology* 2009; 50: 999-1009.
- Cohen NC. *Guidebook on molecular modeling in drug design*. San Diego, USA: Academic Press Inc. 1996.
- Foresman JB, Frisch A. *Exploring chemistry with electronic structure methods*. 2<sup>nd</sup> ed. Pittsburg PA: Gaussian Inc. 1996.
- Leach AR. *Molecular modelling principle and applications*. Edinburgh Gate, Harlow, Essex CM20 2JE. England: Addison Wesley Longman Limited 1996.
- Ibrahim M, Saleh NA, Elshemey WM, Elsayed AA. Computational notes on fullerene based system as HIV-1 protease inhibitors. *J Comput Theor Nanosci* 2010; 7: 224-7.
- Ibrahim M, Saleh NA, Hameed AJ, Elshemey WM, Elsayed AA. Structural and Electronic Properties of new Fullerene Derivatives and their Possible Application as HIV-1 Protease Inhibitors. *Spectrochimica Acta Part A* 2010; 75: 702-9.
- Ibrahim M, Saleh NA, Elshemey WM, Elsayed AA. Fullerene Derivative as anti-HIV Protease Inhibitor: Molecular Modeling and QSAR Approaches. *Mini Rev Med Chem* 2012; 12: 447-51.
- Hameed AJ, Ibrahim M, ElHaes H. Computational notes on structural, electronic and QSAR properties of Fulleropyrrolidine-1-carbodithioic acid 2; 3 and 4-substituted-benzyl esters. *J Mol Struct* 2007; 809: 131-6.
- El-Sayed ME, Omar A, Ibrahim M, Abdel-Fattah WI. On the structural analysis and electronic properties of chitosan /hydroxyapatite interaction. *J Comput Theor Nanosci* 2009; 6: 1663-9.
- Kim JL, Morgenstern KA, Griffith JP, et al. Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure* 1998; 6: 89-100.
- Ortqvist P. On the design and synthesis of Hepatitis C virus NS3 protease inhibitors. Tripeptides to achiral compounds. Digital Comprehensive summaries of Uppsala. From Dissertations from the faculty of pharmacy 117. Uppsala University Uppsala, Sweden, ISBN 978-91-554-7692-2. 2010.pp. 85.

- [21] Yao N, Reichert P, Taremi SS, Prorise WW, Weber PC. Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease–helicase. *Structure* 1999; 7: 1353-63.
- [22] Dahl G. Kinetic studies of NS3 and NS5B from Hepatitis C virus. Implications and applications for drug discovery, digital comprehensive summaries of Uppsala. Dissertations from the faculty of science and technology 615. 65pp. Uppsala University. Uppsala, Sweden, ISBN 978-91-554-7445-4. 2009.
- [23] Barth A. Infrared spectroscopy of proteins. *Biochimica et Biophysica Acta* 2007; 1767: 1073-101.
- [24] Smith B. infrared spectra interpretation: a systematic approach. Boca Raton: CRC press 1999.
- [25] Frisch MJ, Trucks GW, Schlegel HB, *et al.* GAUSSIAN 03, Revision B.05. Pittsburgh, PA: Gaussian Inc. 2003.
- [26] Becke A-D. Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys* 1993; 98: 5648-52.
- [27] Lee C, Yang W, Parr R-G. Development of the Colle-Salvetti correlation energy formula into a functional of the electron density. *Phys Rev B* 1988; 37: 785-9.
- [28] Perdew J-P, Wang Y. Accurate and simple density functional for the electronic exchange energy: generalized gradient approximation. *Phys Rev B* 1986; 33: 8800-2.
- [29] Crawford RL. Lignin biodegradation and transformation. New York: John Wiley and Sons 1981.
- [30] Updegraff DM. Semimicro determination of cellulose in biological materials. *Anal Biochem* 1969; 32: 420-4.
- [31] Ibrahim M, Osman O. Spectroscopic analyses of cellulose: fourier transforms infrared and molecular modelling study. *J Comput Theor Nanosci* 2009; 6: 1054-8.
- [32] Ibrahim M, Osman O, Mahmoud A-A. Spectroscopic analyses of cellulose and chitosan: fir and modeling approach. *J Comput Theor Nanosci* 2011; 8: 117-23.
- [33] Ibrahim M, Elhaes H. computational spectroscopic study of copper, cadmium, lead and zinc interactions in the environment. *Int J Environ Pollut* 2005; 23(4): 417-24.

---

Received: October 01, 2011

Revised: December 15, 2011

Accepted: December 16, 2011

© Ibrahim *et al.*; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.