Original Article   Open Access

Infrared Spectrum for the New Exobiological Nanomolecules Asi, Csi, Tsi and Gsi

Ricardo Gobato[*]
Green Land Landscaping and Gardening, Seedling Growth Laboratory, Bela Vista do Paraíso, 86130-000, Parana, Brazil
Email: ricardogobato@gardener.com

Alireza Heidari
Faculty of Chemistry, California South University, 14731 Comet St. Irvine, CA 92604, USA

Lauro Figueroa Valverde
University Autonomous of Campeche (Faculty of Chemical-Biological Sciences), Calle Avenida Agustín Melgar s/n, Buenavista, 24039 Campeche, Mexico

Abhijit Mitra
Department of Marine Science, University of Calcutta, 35 B. C Road, Kolkata, 700019, West Bengal, India

Abstract
The core of the work is based on the replacement of carbon atoms by silicon atoms, on the basis of four standard bases of DNA: A, C, G and T (adenine, cytosine, guanine, thymine). Determining with minimum computational methods via *ab initio* Hartree-Fock methods, infrared spectrum and their peak absorbance frequencies. The option for simple replacement of carbon by silicon is due to the peculiar characteristics between both. Atomic interactions under non-carbon conditions were studied, with only the Hydrogen, Silicon, Nitrogen and Oxygen atoms, in CNTP, for the four standard bases of DNA, A, C, G and T, thus obtaining by quantum chemistry four new compounds, named here as: ASi, CSi, GSi and TSi. Computational calculations admit the possibility of the formation of such molecules, their existence being possible via quantum chemistry. Calculations obtained in the *ab initio* Unrestricted and Restrict Hartree-Fock method, (UHF and RHF) in the set of basis used Effective core potential (ECP) minimal basis, UHF CEP-31G (ECP split valance) and UHF CEP-121G (ECP triple-split basis), CC-pVTZ (Correlation-consistent valence-only basis sets triple-zeta) and 6-311G**(3df, 3pd) (Gaussian functions quadruple-zeta basis sets).
Keywords: DNA; Adenine; Cytosine; Guanine; Thymine; Hartree-Fock method; Nano-molecule; Infrared spectroscopy; CEP-31G; CEP-121G; CC-pVTZ; 6-311G** (3df, 3pd).

1. Introduction

The Deoxyribonucleic acid (DNA) is a molecule composed of two polynucleotide chains that coil around each other to form a double helix carrying genetic instructions for the development, functioning, growth and reproduction of all known organisms and many viruses. The A, C, G, and T, - adenine, cytosine, guanine and thymine, respectively, representing the four nucleotide bases of a DNA strand – adenine, cytosine, guanine, thymine – covalently linked to a phosphodiester backbone [1].

The core of the work is based on the replacement of carbon atoms by silicon atoms, on the basis of four standard bases of DNA: A, C, G and T (adenine, cytosine, guanine, thymine). Determining with minimum computational methods via *ab initio* Hartree-Fock methods, infrared spectrum and their peak absorbance frequencies [2-14].

The option for simple replacement of carbon by silicon [15, 16] is due to the peculiar characteristics between both. Atomic interactions under non-carbon conditions were studied, with only the Hydrogen, Silicon, Nitrogen and Oxygen atoms, in STP (Standard Temperature and Pressure), for the four standard bases of DNA, A, C, G and T, thus obtaining by quantum chemistry four new compounds, named here as: ASi, CSi, GSi and TSi.

In comparing the carbon and silicon has: the Si lies in the same column of the periodic table of the elements, and it has been investigated as a possible alternative for building up biological molecules in exobiology [17, 18].

Silicon based chemistry, however, is by far less flexible than carbon chemistry, not able to form double covalent bonds with the same easiness as C does. Other fact is the larger volume occupied by the external electronic orbitals of silicon tend to reduce the superposition of p orbitals [17, 18].

Through the chemical abundances of biological elements in the earth crust, terrestrial life has chosen carbon instead of silicon, in spite of the larger abundance of silicon. This fact suggests that carbon is better suited to form biological molecules [17, 18].

However, this paper assumes conditions without the presence of Carbon.

Calculations obtained in the *ab initio* Unrestricted and Restrict Hartree-Fock method, (UHF and RHF). The set of basis used Effective core potential (ECP) minimal basis, UHF CEP-31G (ECP split valance) and UHF CEP-121G

*Corresponding Author

(ECP triple-split basis), CC-pVTZ (Correlation-consistent valence-only basis sets triple-zeta) and 6-311G**(3df, 3pd) (Gaussian functions quadruple-zeta basis sets) [2-14].

Studies did not reveal any works with characteristics studied here. There is an absence of a referential of the theme, finding only one work in [19-27].

2. Methods

2.1. Hartree-Fock Methods

The molecular Hartree-Fock [2-14] wave function is written as an antisymmetrized product (Slater determinant) of spin-orbitals, each spin-orbital being a product of a spatial orbital ϕ_i and a spin function (either α or β).

The expression for the Hartree-Fock molecular electronic energy E_{HF} is given by the variation theorem as $E_{HF} = \langle D|\hat{H}_{el} + V_{NN}|D\rangle$ where D is the Slater-determinant Hartree-Fock wave function and $\hat{H}_{el}$ and V_{NN} are given by

$$\hat{H}_{el} = -\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 - \sum_\alpha \sum_i \frac{Z_\alpha e'^2}{r_{i\alpha}} + \sum_j \sum_{i>j} \frac{e'^2}{r_{ij}}$$

$$V_{NN} = \sum_\alpha \sum_{\beta>\alpha} \frac{Z_\alpha Z_\beta e'^2}{r_{\alpha\beta}} \tag{1}$$

Since V_{NN} does not involve electronic coordinates and D is normalized, we have $\langle D|V_{NN}|D\rangle = V_{NN}\langle D|D\rangle = V_{NN}$. The operator $\hat{H}_{el}$ is the sum of one-electron operators $\hat{f}_i$ and two-electron operators $\hat{g}_{ij}$; we have $\hat{H}_{el} = \sum_i \hat{f}_i + \sum_j \sum_{i>j} \hat{g}_{ij}$, where $\hat{f}_i = -\frac{1}{2}\nabla_i^2 \sum_\alpha \sum_\alpha /r_{i\alpha}$ and $\hat{g}_{ij} = 1/r_{ij}$. The Hamiltonian $\hat{H}_{el}$ is the same as the Hamiltonian $\hat{H}$ for an atom except that $\sum_\alpha \sum_\alpha /r_{i\alpha}$ replaces Z/r_i in $\hat{f}_i$. Hence

$$E = \langle D|\hat{H}|D\rangle = 2\sum_i^{n/2}\langle \phi_i(1)|\hat{f}_i|\phi_i(2)\rangle + \sum_{j=1}^{n/2}\sum_{i=1}^{n/2}(2J_{ij} - K_{ij}) \tag{2}$$

where

$$J_{ij} = \langle \phi_i(1)\phi_j(2)| e'^2/r_{12} |\phi_i(1)\phi_j(2)\rangle$$

and

$$K_{ij} = \langle \phi_i(1)\phi_j(2)| e'^2/r_{12} |\phi_j(1)\phi_i(2)\rangle$$

$$\hat{f}_i = -(\hbar^2/2m_e)\nabla_i^2 - Z e'^2/r_1 \tag{3}$$

can be used to give $\langle D|\hat{H}_{el}|D\rangle$.

Therefore, the Hartree-Fock energy of a diatomic or polyatomic molecule with only closed shells is

$$E_{HF} = 2\sum_{i=1}^{n/2} H_i^{core} + \sum_{j=1}^{n/2}\sum_{i=1}^{n/2}(2J_{ij} - K_{ij}) + V_{NN} \tag{4}$$

$$H_i^{core} \equiv \langle \phi_i(1)|\hat{H}^{core}(1)|\phi_i(1)\rangle \equiv \left\langle \phi_i(1)\left|-\frac{1}{2}\nabla_i^2 \sum_\alpha Z_\alpha/r_{1\alpha}\right|\phi_i(1)\right\rangle \tag{5}$$

$$J_{ij} \equiv \langle \phi_i(1)\phi_j(2)| 1/r_{12} |\phi_i(1)\phi_j(2)\rangle$$

and

$$K_{ij} \equiv \langle \phi_i(1)\phi_j(2)| 1/r_{12} |\phi_j(1)\phi_i(2)\rangle \tag{6}$$

where the one-electron-operator symbol was changed from $\hat{f}_i$ to $\hat{H}^{core}(1)$ [6].

The vast literature associated with these methods suggests that the following is a plausible hierarchy:

$$HF << MP2 < CISD < CCSD < CCSD (T) < FCI \tag{7}$$

The extremes of 'best', FCI, and 'worst', HF, are irrefutable, but the intermediate methods are less clear and depend on the type of chemical problem being addressed. The use of HF [2-14] in the case of FCI was due to the computational cost.

2.2. Hardware and Software

For calculations the computer used for MD (molecular Dynamics) was a Desktop with SUSE Linux Enterprise Desktop, AMD Ryzen 7 1800X processor, ASUS Prime A320M-K motherboard, 16GB of RAM, with 500GB SSD [28].

The *ab initio* calculations have been performed to study the equilibrium configuration of ASi, CSi, GSi and TSi molecules. The set of programs GaussView 5.0.8 [29], Mercury 3.8 [30], HyperChem 8.0.6 Evaluation [31]. are the advanced semantic chemical editor, visualization, and analysis platform and GAMESS [15, 32] is a computational chemistry software program and stands for General Atomic and Molecular Electronic Structure System [15], Evaluation Gordon [33], BIOVIA Draw [34], CHARMM22 [35] set of programs were used.

3. Results

The molecular structure of ASi, CSi, GSi and TSi molecules, were obtained through computationally calculated molecular dynamics, using the *ab initio* Hartree-Fock (HF) method.

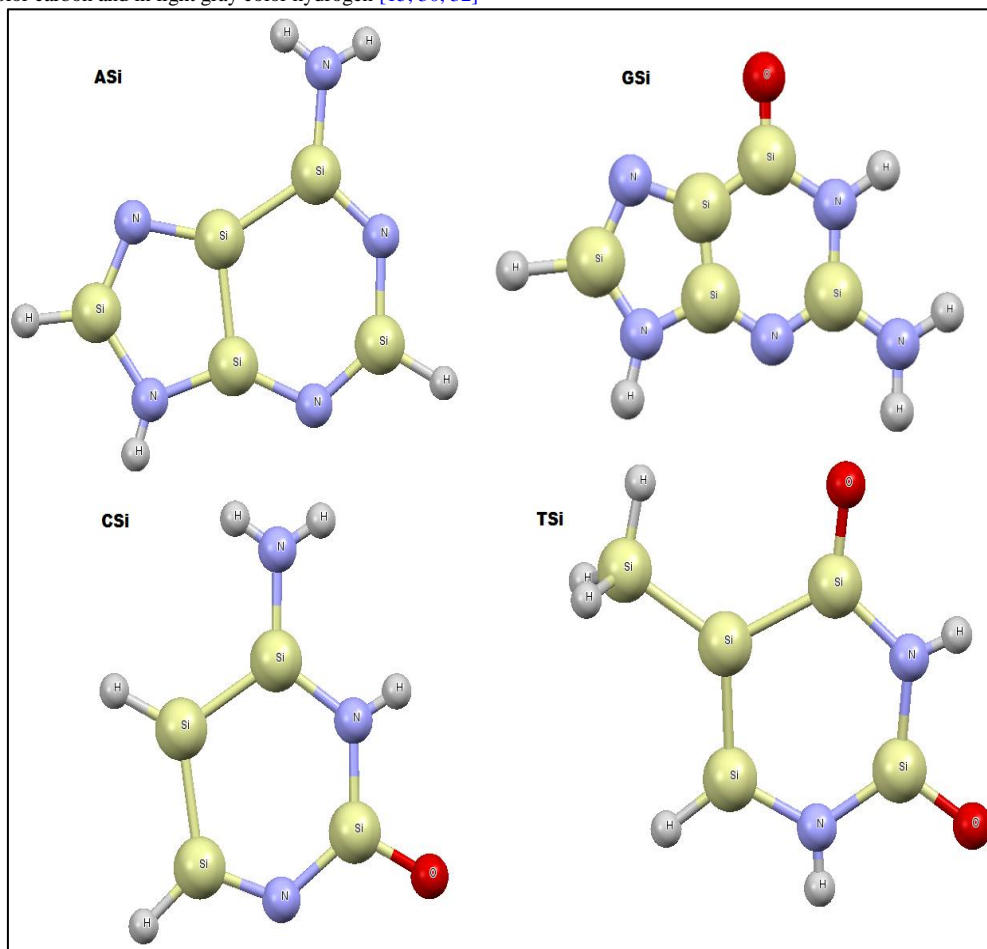The names in Figure (1) of the new molecules obtained ASi, CSi, GSi and TSi, are:

2,3,4,5,6,7,8,9-octahydro-1H-[1,3,2,4,5]diazatrisilolo[4,5-d][1,3,2,4,5,6]diazatetrasilin-8-amine;

2-hydroxy-1,3,2,4,5,6-diazatetrasilinan-4-amine;

8-oxo-3,7-dihydro-[1,3,2,4,5]diazatrisilolo[4,5-d][1,3,2,4,5,6]diazatetrasilin-6-amine and

(2,4-dihydroxy-1,3,2,4,5,6-diazatetrasilinan-5-yl)silane, respectively.

Figure-1. Representation of the molecular structure of ASi, CSi, GSi and TSi molecules, obtained through computer via *ab initio* calculation method RHF/UHF [2-14] sets basis obtained using computer programs GAMESS[15, 32]. Images obtained in the software Mercury 3.8 [30]. Represented in bluish gray color the atom of silicon, in the purple color lithium, in the lemon yellow color beryllium, in the orange the selenium, in dark gray color carbon and in light gray color hydrogen [15, 30, 32]

3.1. ASi

Properties of molecule Asi:
IUPAC name:
2,3,4,5,6,7,8,9-octahydro-1H-[1,3,2,4,5]diazatrisilolo[4,5-d][1,3,2,4,5,6]diazatetrasilin-8-amine;
PSA(Polar Surface Area): 74.14;
ALogP: 7.1232;
Stereo Center Count: 3;
Hydrogen Acceptor Count: 5;
Hydrogen Donor Count: 5;
Composition: H: 5.9% N: 31.3% Si: 62.8%;
Formula Weight: 223.56421;
Exact Mass: 223.001728116;
Molecular Formula: $H_{13}N_5Si_5$.
[34]

3.2. CSi

Properties of molecule CSi:
IUPAC name:
2-hydroxy-1,3,2,4,5,6-diazatetrasilinan-4-amine;
PSA: 70.31;
ALogP: 6.022;
Stereo Center Count: 2;
Hydrogen Acceptor Count: 4;
Hydrogen Donor Count: 4;
Composition: H: 6.1% N: 23.2% O: 8.8% Si: 61.9%;
Formula Weight: 181.44883;
Exact Mass: 180.997918122;
Molecular Formula: $H_{11}N_3OSi_4$.
[34]

3.3. GSi

Properties of molecule GSi:
IUPAC name:
8-oxo-3,7-dihydro-[1,3,2,4,5]diazatrisilolo[4,5-d][1,3,2,4,5,6]diazatetrasilin-6-amine;
PSA: 96.16;
ALogP: 5.1999;
Stereo Center Count: 0;
Hydrogen Acceptor Count: 4;
Hydrogen Donor Count: 3;
Composition: H: 2.2% N: 30.3% O: 6.9% Si: 60.7%;
Formula Weight: 231.50009;
Exact Mass: 230.93404248;
Molecular Formula: $H_5N_5OSi_5$.
[34]

3.4. TSi

Properties of molecule TSi:
IUPAC name:
(2,4-dihydroxy-1,3,2,4,5,6-diazatetrasilinan-5-yl)silane;
PSA: 64.52;
ALogP: 8.6476;
Stereo Center Count: 3;
Hydrogen Acceptor Count: 4;
Hydrogen Donor Count: 4;
Composition: H: 5.7% N: 13.2% O: 15.1% Si: 66.1%;
Formula Weight: 212.53497;
Exact Mass: 211.974510294;
Molecular Formula: $H_{12}N_2O_2Si_5$.
[34]

Figure-2. Representation of the normalized infrared spectrum of ASi - Frequency (cm⁻¹) for Absorbance (%). Using computer programs GAMESS, the calculations obtained in the *ab initio* HF, in a set of CEP-31G, CEP-121G, CC-pVTZ and 6-311G**(3df, 3pd). Graphic edited in Origin software, for comparison of the spectra obtained in the set of bases used
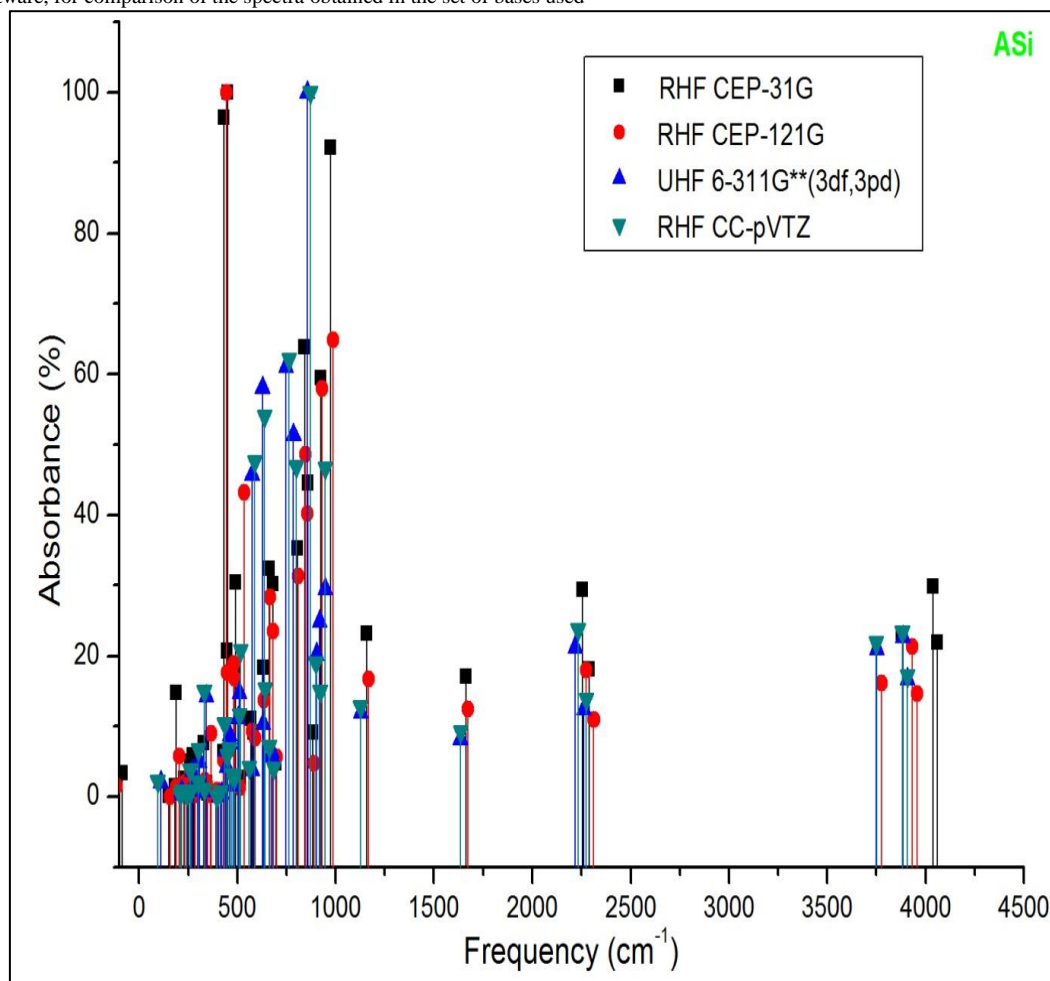
Figure-3. Representation of the normalized infrared spectrum of CSi - Frequency (cm^{-1}) for Absorbance (%). Using computer programs GAMESS, the calculations obtained in the *ab initio* HF, in a set of CEP-31G, CEP-121G, CC-pVTZ and 6-311G**(3df, 3pd). Graphic edited in origin software [33], for comparison of the spectra obtained in the set of bases used
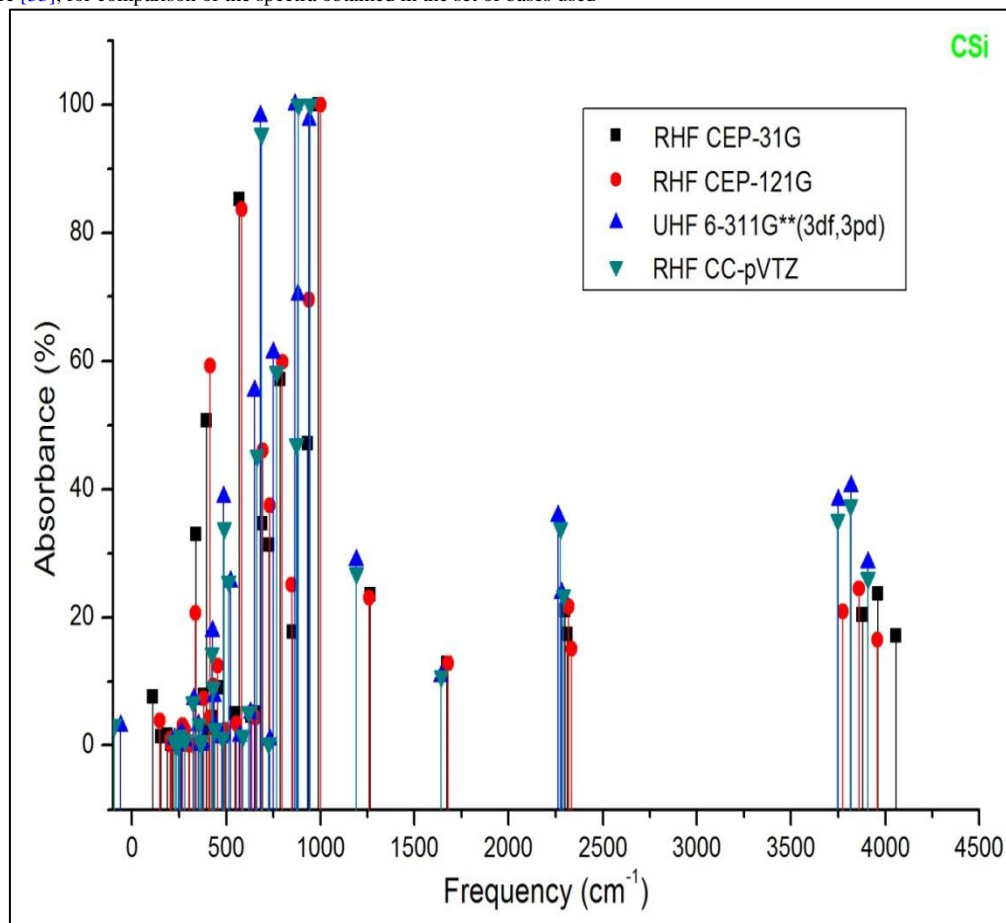

Figure-4. Representation of the normalized infrared spectrum of GSi - Frequency (cm^{-1}) for Absorbance (%). Using computer programs GAMESS, the calculations obtained in the *ab initio* HF, in a set of CEP-31G, CEP-121G, CC-pVTZ and 6-311G**(3df, 3pd). Graphic edited in origin software [33], for comparison of the spectra obtained in the set of bases used

Figure-5. Representation of the normalized infrared spectrum of TSi - Frequency (cm⁻¹) for Absorbance (%). Using computer programs GAMESS, the calculations obtained in the *ab initio* HF, in a set of CEP-31G, CEP-121G, CC-pVTZ and 6-311G**(3df, 3pd). Graphic edited in origin software [33], for comparison of the spectra obtained in the set of bases used
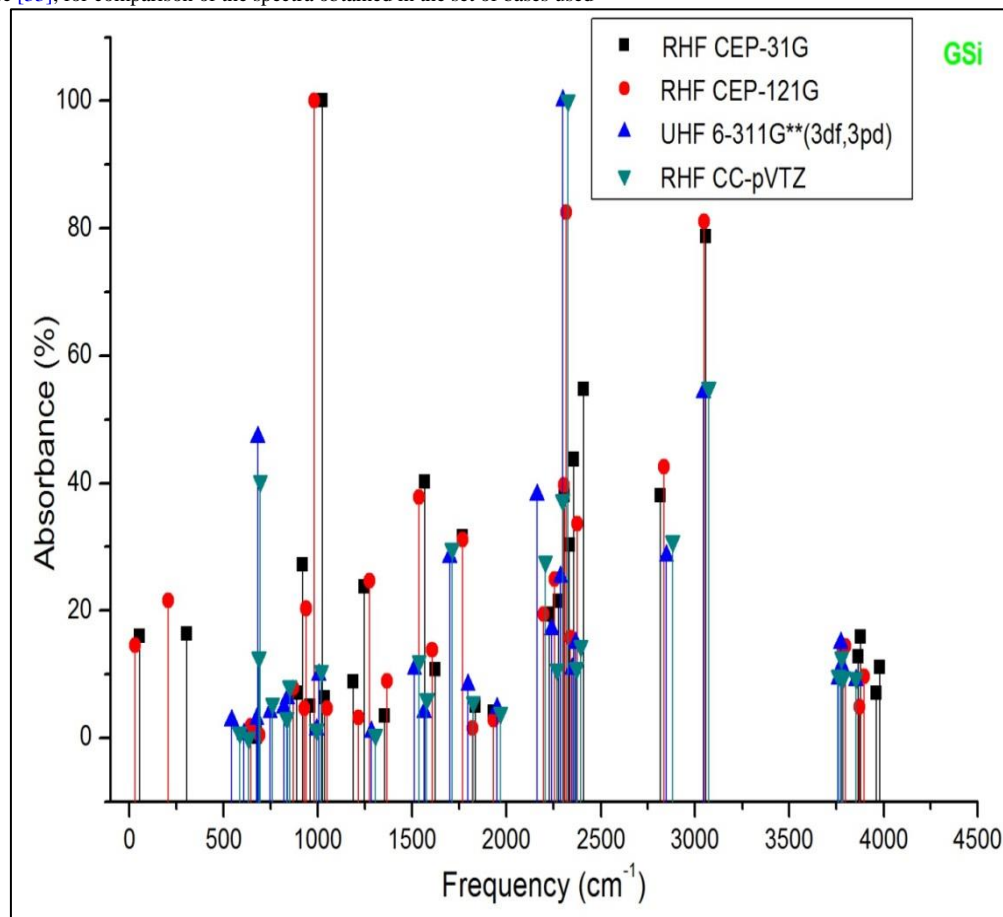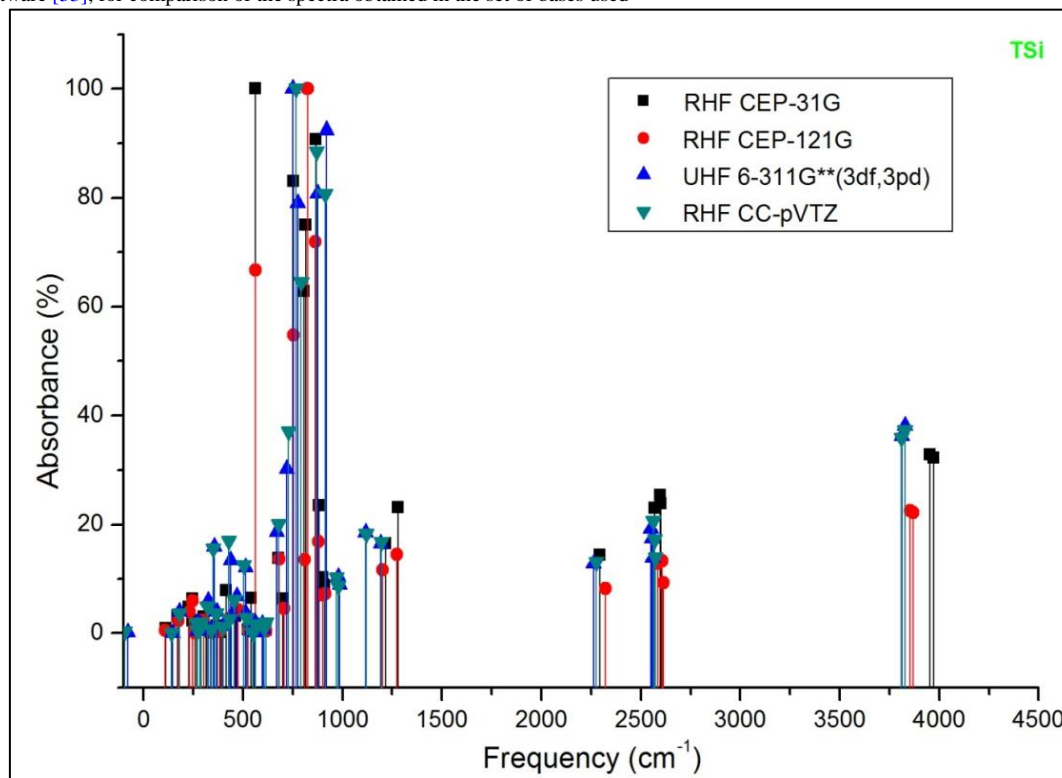

4. Conclusions and Challenges

The absorbance peaks for Asi in the set of bases then used between 400 to 100 cm⁻¹, 2250 cm⁻¹ and 3750 to 4100 cm⁻¹, Figure (2).

The absorbance peaks for Csi in the set of bases then used between 400 to 1000 cm⁻¹, 2250 to 2300 cm⁻¹ and 3750 to 4100 cm⁻¹, Figure (3).

The absorbance peaks for Gsi in the set of bases used then between the range of 1000 cm⁻¹, 2300 cm⁻¹ and 3000 cm⁻¹, Figure (4).

The absorbance peaks for Tsi in the set of bases then used between 660 to 950 cm⁻¹, 2250 to 2600 cm⁻¹ and 3500 to 4000 cm⁻¹, Figure (5) .

The fingerprint of the molecules ASi, CSi, GSi and TSI, are demonstrated in the Figures (2-5).

The infrared spectrum was calculated, indicating the characteristic of the nano-molecule genesis. Characterized its infrared spectrum, quantically calculated, accepted by quantum chemistry parameters, with *ab initio* methods, in the bases CC-pVTZ and 6-311G**(3df, 3pd). An experimental challenge to chemists.

Going beyond imagination, now the challenge is to build the basic structure of is the simulation for building a "new DNA helix", based now on the ASi, CSi, GSi and TSi molecules.

Limitations our study has so far been limited to computational simulation via quantum mechanics e molecular mechanics (QM/MM), an applied theory. Our results and calculations are compatible and with the theory of QM/MM, but their physical experimental verification depend on advanced techniques for their synthesis, obtaining laboratory for experimental biochemical.

References

[1]   Creative Commons. CC-BY 4.0, 2021. "Wikipedia, The Free Encyclopedia, Nucleic acid sequence." Available: https://en.wikipedia.org/wiki/Nucleic_acid_sequence

[2]   Dunning Jr, T. H., 1989. "Gaussian basis sets for use in correlated molecular calculations, The atoms boron through neon and hydrogen." *J. Chem. Phys.*, pp. 1007-23.

[3]   Dunning Jr, T. H. and Hay, P. J., 1977. *Modern theoretical chemistry* vol. 3. New York: Plenum.

[4]   Gordon, M. S. and Schmidt, M. W., 2005. *Advances in electronic structure theory: GAMESS a decade later. Theory and Applications of Computational Chemistry: the first forty years*. Elsevier: C. E. Dykstra, G. Frenking, K. S. Kim and G. E. Scuseria (editors). pp. 1167-1189.

[5]   Hehre, W. J., 2003. *A guide to molecular mechanics and quantum chemical calculations, wavefunction*. CA: Inc., Irvine.

[6]   Kohn, W. and Sham, L. J., 1965. "Self-consistent equations including exchange and correlation effects." *Phys. Rev.*, p. A1133.

[7]   Levine, I. N., 2003. *Quantum Chemistry. Pearson Education (Singapore) Pte.* . 5th ed. Delhi: Ltd., Indian Branch, 482 F. I. E. Patparganj.

[8]   Polak, E., 1971. "Computational methods in optimization." *Elsevier,* vol. 77, p. 10003.

[9]   Szabo, N. S. A., 1989. *Ostlund, modern quantum chemistry*. New York: Dover Publications.

[10]   Thijssen, J. M., 2001. *Computational physics*. Cambridge: Cambridge University Press.

[11]   Weigend, F. and Ahlrichs, R., 2005. "Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy." *Phys. Chem. Chem. Phys.,* pp. 3297-305.

[12]   Wilson, A. K., van Mourik, T., and Dunning Jr, T. H., 1996. "Gaussian basis sets for use in Correlated Molecular Calculations. Sextuple zeta correlation consistent basis sets for boron through neon." *J. Mol. Struct. (Theochem),* pp. 339-49.

[13]   Woon, D. E. and Dunning Jr, T. H., 1993. "Gaussian-basis sets for use in correlated molecular calculations. The atoms aluminum through argon." *J. Chem. Phys.,* pp. 1358-71.

[14]   Gobato, R., Heidari, A., Valverde, L. F., and Mitra, A., 2021. ""Applying "Ab Initio" Hartree-Fock Methods to Exobiological Nanomolecules." Available: https://en.wikipedia.org/wiki/Nucleic_acid_sequence

[15]   McDouall, J. J. W., 2013. *Computational quantum chemistry. Molecular structure and properties in silico*. Milton Road, Cambridge CB4 0WF, UK: The Royal Society of Chemistry, Thomas Graham House, Science Park.

[16]   Patai, S. and Rappoport, Z., 1989. *The chemistry of organic silicon compounds, series: The chemistry of functional groups*. Wiley.

[17]   Horneck, G. and Rettberg, P., 2007. *Complete course in astrobiology*. Series: Physics Textbook, Wiley-VCH.

[18]   Shaw, A. M., 2006. *Astrochemistry: From astronomy to astrobiology*. Wiley.

[19]   Gobato, R., Gobato, M. R. R., and Heidari, A., 2019. "Raman Spectroscopy Study of the Nano Molecule C13H20BeLi2SeSi Using ab initio and Hartree–Fock Methods in the Basis Set CC–pVTZ and 6–311G** (3df, 3pd)." *International Journal of Advanced Engineering and Science,* vol. 7, pp. 14-35.

[20]   Gobato, R., Gobato, M. R. R., and Heidari, A., 2019. "Attenuated total reflection–fourier transform infrared (atr–ftir) spectroscopy study of the nano molecule c13h20beli2sesi using ab initio and hartree–fock methods in the basis set rhf/cc– pvtz and rhf/6–311g** (3df, 3pd): An experimental challenge to chemists." *Chemistry Reports,* vol. 2, pp. 1-26.

[21]   Gobato, R., Gobato, M. R. R., Heidari, A., and Mitra, A., 2018. "Spectroscopy and Dipole Moment of the Molecule C13H20BeLi2SeSi via Quantum Chemistry Using Ab Initio, Hartree–Fock Method in the Base Set CC–pVTZ and 6–311G**(3df, 3pd)." *American Journal of Quantum Chemistry and Molecular Spectroscopy,* vol. 2, pp. 9-17.

[22]   Gobato, R. and Heidari, A., 2017. "Calculations using quantum chemistry for inorganic molecule simulation beli2sesi." *Science Journal of Analytical Chemistry,* vol. 5, pp. 76-85.

[23]   Gobato, R. and Heidari, A., 2017. "Calculations using quantum chemistry for inorganic molecule simulation beli2sesi." *Science Journal of Analytical Chemistry,* vol. 5, pp. 76-85.

[24]   Gobato, R. and Heidari, A., 2018. "Using the Quantum Chemistry for Genesis of a Nano Biomembrane with a Combination of the Elements Be, Li, Se, Si, C and H." *J. Nanomed Res.,* vol. 7, pp. 241-252.

[25]   Gobato, R. and Heidari, A., 2018. "Infrared spectrum and sites of action of sanguinarine by molecular mechanics and ab initio methods." *International Journal of Atmospheric and Oceanic Sciences,* vol. 2, pp. 1-9.

[26]   Gobato, R., Heidari, A., and Mitra, A., 2018. "The creation of c13h20beli2sesi. The proposal of a bio–inorganic molecule, using ab initio methods for the genesis of a nano membrane." *Arc. Org. Inorg. Chem. Sci.,* vol. 3,

[27]   Gobato, R., Heidari, A., and Mitra, A., 2021. "Bioinorganic membrane using a new Kurumi liquid crystal." Available: http://Doi:10.13140/RG.2.2.24868.30087

[28]   Creative Commons CC-BY 4.0, 2021. "Wikipedia, The Free Encyclopedia, March 2021." Available: https://creativecommons.org/licenses/by/4.0/

[29]   Dennington, R., Keith, T., and Millam, J., 2009. "Gaussview, Version 5."

[30]   The Cambridge Crystallographic Data Centre CCDC, 2012. *Mercury - crystal structure visualisation, exploration and analysis made easy", May 2012. Mercury 3.1 Development (Build RC5)*. The Cambridge Crystallographic Data Centre.

[31]   Hypercube. Molecular Modeling System. HyperChemTM 7.5 evaluation, 2003. "Hypercube."

[32]   OriginLab, 2018. "Evaluation licence, graphing and analysis, ©originlab corporation." Available: https://www.originlab.com/index.aspx?go=Products/Origin/2018b&pid=3289

[33]   Gordon, M. S., 1993. "General atomic and molecular electronic structure system (GAMESS)." *J. Comput. Chem.,* vol. 14, pp. 1347-1363.

[34]   BIOVIA Draw, 2017. "Enterprise, Computational results obtained using software programs from Dassault Systèmes BIOVIA." In *The ab initio calculations were performed with the DMol3 program, and graphical displays generated with Draw.*

[35]   Brooks, R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M., 1983. "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations." *J. Comp. Chem.,* vol. 4, pp. 187-217.