Study of Biomolecule Under UV Irradiation at CASSCF/ NEVPT2 Level Theory

Satish Kumar[1], Ashok Jangid[2]
[1]Assistat Professor, [2]Associate Professor
Department of Physics and Computer Science
Dayalbagh Educational Institute, Dayalbagh- Agra-05,India

Abstract

The photostability of biomolecules at CASSCF/NEVPT2 level theory has been investigated. Under UV irradiation excited states energies of the Gly-Gly-Gly molecule are calculated. Potential energy surfaces have been obtained as a function of hydrogen bond length. The potential energy surfaces crossing along conical intersection and avoided crossing have been found. Through these conical intersections and avoided crossing molecules return to its ground state that is photostability and fundamental phenomenon of evolution of life on the earth.

Keywords

Conical intersection-CIs, Avoided crossing-AC, Potential energy surfaces –PES, CASSCF, NEVPT2

I. INTRODUCTION

How life is evolve on the earth in early days? It is the burning topic of research in scientific community to understand the evolution of life under heavy electromagnetic irradiation on biological molecules such as adenine, thymine, guanine, glycine and cytosine and many more. High, ultra violet (UV) energy exposure can structurally deform these organic molecules. Although these biological molecules are retain in their stable state or close shell state (CS) under UV irradiations [1]. Photon irradiation can breaks the covalent bond and thus induce a high chemical deformation to molecules. In view this, Life is running still under full exposure electromagnetic (EM) spectrum. So survival of biomolecules under UV radiation is interesting aspect of life.

The organic molecules are combination of atoms of oxygen, carbon, hydrogen and nitrogen. These elements of biomolecules are affected by UV radiation. The behavior of molecule at atomic and molecular level can understand by quantum principles as dynamical behavior of the molecules or its constituent are probabilistic in nature. Theoretically investigations are very helpful to understand distribution of electrons in various spin orbitals and their wave functions. Photo stability of

glycine is investigated at theoretical level as glycine is one of the important constituent of biological proteins. Three glycine molecules which are combine form of three amino acids through peptide bond. Hydrogen bond is formed between CO…HN group of two amino acids. The photo stability of amino acids is explains on basis of proton transfer process through hydrogen bond experimentally [2] and theoretically [3]. UV light is incident to molecules, it get energy from photon and excite into the higher state. Once molecules is reached to higher state after that it try to come in lowest state of energy through potential energy surface (PES) crossing. These crossing are two type conical intersections CIs and avoided crossing ACs of two potential energy surfaces. The path it follow through various potential energy surfaces is known as minimum energy path (MEP). Energy transfer to higher state by direct photon radiation through charge is termed as charge transfer state. Some excited states which are lower in energy from charge transfer states corresponding to high oscillator strength is known as locally excited states. The CIs are key mechanistic element in photo stability [4-7] which are considered as barrier less and it allows molecules to pass through crossings. The photo stability of (7-(2'-pyridyl) indole) is shown through intramolecularly hydrogen-bonded π systems [6]. Molecules de-excited to lower state through radiation less process i.e the energy is dissipated to environment in form of vibrational energy and finally return to its initial structure [8-9]. This mechanism has proposed in DNA base pair model [10-11] and it has been proved by experimentally in two amino pyridine dimer electron and spectroscopy [12]. There is AC between the lower excited state and higher excited state this process is less efficient and it depends on the energy gap between the lower point of higher excited state and higher point of the Lower excited state. If the AC gap is very high then the system remains in higher lower state and its stability is affected under UV radiation. Lower the gap will lead to take molecules to its ground states through avoided crossings [13]. It is least efficient method to stability of the molecules. In this paper the photostability two glycine molecule is investigated on basis of completely active space of

self consistent field CASSCF which is based on quantum principle. Once the energies and wave functions of the exicted states are calculated by CASSCF method. These CASSCF calculated energies are corrected by *N* electron valace state purtabation theory NEVPT2 [14] with including DC effect at single point energy calculation.

II. METHODOLOGY AND COMPUTATIONAL DETAILS

To calculate the electronic excited states energies or wave functions a multiconfigurational ab initio method is widely used by computational physicist and chemists. A CASSCF/NEVPT2 method is applied to many chemical, physical and biological systems to investigate the physical and chemical properties of matter with high accuracy [15]. This method is combination of energy derivative with single point correction to NEVPT2 including dynamical correlation effect. CASSCF is extended form of Hartree Fock method and its state wave function is written as

$$|\psi_I^s\rangle_{CASSCF} = \sum \chi C_{\chi L} |\phi_\chi^S\rangle$$

$$(2.1)$$

Where $|\psi_I^s\rangle_{CASSCF}$ is the N-electron, CASSCF wave functions for the state *I* and its spin S and $|\phi_\chi^S\rangle$ is the set of all configurational states (CSF) in the form of Slater determinant and $C_{\chi L}$ is coefficient of variational parameters. The CSF is combination of all orthonormal set of molecular orbitals $\psi_i(r)$ which are expended form of the basis functions

$$\psi_i(r) = \sum_q C_{qi} \phi_q(r) \qquad (2.2)$$

Where, C_{qi} is molecular orbital constant from the 2nd set of variational parameter. By knowing CASSCF wave functions the energy of selected states are calculated by Rayleigh quotient method [16]. The expression for energy is given as

$$E = \frac{\langle \psi_I^s | \hat{H} | \psi_I^s \rangle}{\langle \psi_I^s | \psi_I^s \rangle}$$

$$(2.3)$$

The energy calculated by this method is not the total energy but it gives correct picture of the upper bound to total energy. This method is fully variational with energy that is made static with regarding to variations in both set of MO$_s$ and CI constants. At conervergence point the gradient of energy with respect to MO and CI constants vanishes and it can be written as

$$\frac{\delta E}{\delta C_{qi}} = 0 \qquad (2.4)$$

$$\frac{\delta E}{\delta C_{\chi L}} = 0 \qquad (2.5)$$

NEVPT2 is second order perturbative correction to calculated wave functions by CASSCF method. NEVPT2 is equvalent to comple active space pertubation to second order (CASPT2) method [16-17]. NEVPT2 is purturbation method that include DC effect and differs from CASPT2 mainly in the definition in the zeroth-order Hamiltonian. In NEVPT2 the Fock operator is replaced by the Dyall Hamiltonian [18].NEVPT2/CASSCF results are highly accurate as comparison to single use of CASSCF and these results are well matched with experimental results.
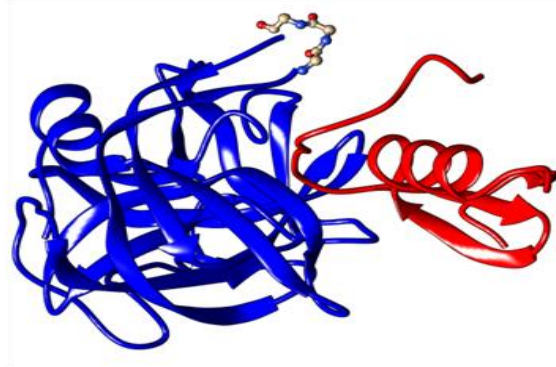

**Fig.1 Structure of Protease B (SGPB) and ball and stick diagram of 3GLY.**

Here we selected SGPB protein from protein data bank [19]. The selected molecules are combination of most crucial amino acids in *β* chain of the proteins and it is written as GLY-GLY-GLY (G-G-G). The ribbon structure of the protein is shown fig. 1 is combination of many biological molecules. In fig.1 ball and stick portion of the molecules is combination three primary glycine molecules. Incident of UV light damage G-G-G molecules. So its middle glycine is removed from three glycine combination. Remaining two glycine molecules are used to study photostability of the molecules. The Extended view of the G-G combination is shown in fig 2. All electrons are not participating in excitation process. It is necessary to find out the highest occupied molecular orbit HOMO and lowest unoccupied molecular orbit LUMO energy gap of studied molecules where the transitions of electrons are occur under UV radiation. These electrons and orbitals are known active electrons and active orbitals. In CASSCF calculations all degeneracy effect are included, consequently no large terms in perturbation expansion. Electronic spectra of G-G combination initially include for each peptide unit the π, π^* orbitals and the n lone pair orbitals on the oxygen together with six electrons. In active space 12 electrons and 8 molecular orbitals are selected

around the HOMO and LUMO gap. All NEVPT2/CASSCF calculations were carried out by using ORCA software package [20].

III.RESULT AND DISCUSSION

A. Ground *state structure*

The combination of G-G-G molecules is selected from the protease B (SGBP) and optimize the molecules by applying MP2/6-31-G (d) [21-22] level theory. The remaining two glycines are optimize with same level theory by imposing some constrained to remove the effect of damage glycine molecule by UV light. The hydrogen bond length between NH---CO group is 1.970 A^0 when it is optimized in free space and bond length with between N--H is found $1.015 A^0$. The optimize structure of the ground state of two glycine molecules is shown in fig2. in free space.
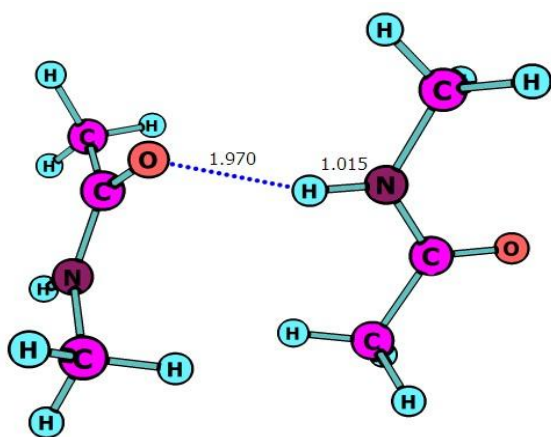
Fig.2 Optimize ground state (GS) of G-G molecule in free space.

The calculated energy for six excited states is tabulated in the table 1. The calculated energies are more close to experimental data.

Table1:Nevpt2/Casscf Calculated Absorption Spectra Of Two G-G Molecules With And Without Water.

States	NEVPT2	f_{osc}	CASSCF	f_{osc}	Exp. results [24] ΔE
	G-G Absorption				
E1	6.065393	0.0006	5.754639	0.0007	5.8
E2	5.839308	0.0008	6.140776	0.0006	6.0
E3	7.254191	0.4036	8.616132	0.4794	6.9
E4	8.308602	0.0093	9.013551	0.0101	7.0
E5	7.385589	0.4160	9.170156	0.5166	8.2
E6	8.723279	0.0700	9.605799	0.0771	10.0

[#] All energies are in eV units.

Hydrogen bond is active with light as it transfers charge (electron or protons) through hydrogen bond to take whole molecules to higher states and finally return to it initially state. The two charge transfers E_5 and E_6 states are found in the molecules. The two orbits are corresponding to weak band are E_1 and E_2. The two orbitals E3 and E4 are found in between the weakly excited and two charge transfer states. are known as locally excited states. The experimental data and calculated data is written in table is calculate in gases environment. The potential energy surfaces of first six excited states of G-G molecule are investigated. The first six singlets excited states energies are calculated named as E1, E2, E3, E4, E5 and E6. The two singlet states E5 and E6 are transfer states. The NEVPT2 energy profile is shown in the fig.3 of two glycine molecules with hydrogen bond length of molecules G1CO…HNG2. It has seen that CIs and AC are found with OH distance between computed energies of the excited sates of the system. These calculations are being done in gases medium. The first crossing between two charge transfer states E5 and E6 is found. CI crossing of E6 and E5 with hydrogen bond length is shown in fig. 3.

It has been seen the charge transfer from glycine G1 to G2 or G2 to G1 as an oscillatory in nature with hydrogen bond length. If the charge transfers from G1 to G2 molecule then the energy of E6 state is high and if charge transfers from G2 molecule to G1 molecule the energy of E5 states is high as it is clear by the energy profile of PES. There are some crossings of energy states which are found between two excited states. CI crossing is being seen between two charge transfer states E5 and E6 with hydrogen bond length. CI crossing is formed because of the coupling of two states wave functions of the molecules. It transfers the energy from higher state to lower state through steepest path along GD vector at coupling point. This junction point provides a funnel to the lower states of the energy. The crossing between E5 and E6 is also found around bond length 1.4 angstrom but molecule will follow the path along gradient differences vector and derivative coupling vector. The hydrogen bond length of the optimized structure of the molecule is 1.97 angstrom. During stabilization process the molecules passes through various crossing. So the next CI crossing is found between E4 and E5 states. The every CI crossing will provide a funnel to transfer the energy to next lower excited states. The locally excited states E3 and weakly excited states E1 will not cross each other through conical intersection as shown in fig. 3. A avoided crossing AC between E1 and E3 states is found. The energy difference is 2 kcal-mol[-] between E1 and E3 is found which is consistent with result in ref [23]. It is small difference of energy which can be consider for AC crossing that will provide a path to lower energy state. Molecule will reach to lower sate E1 from the state E3. E1 and E2 states are cross each other around 1.15 angstrom OH

distance of hydrogen bond length of the molecule. CI crossing between the two states will provide path along the GD vector. Finally molecule will reach to the ground state GS through MEP. The MEP is

shown in figure through arrows. The photo stability will check in two glycine molecule system in free space.
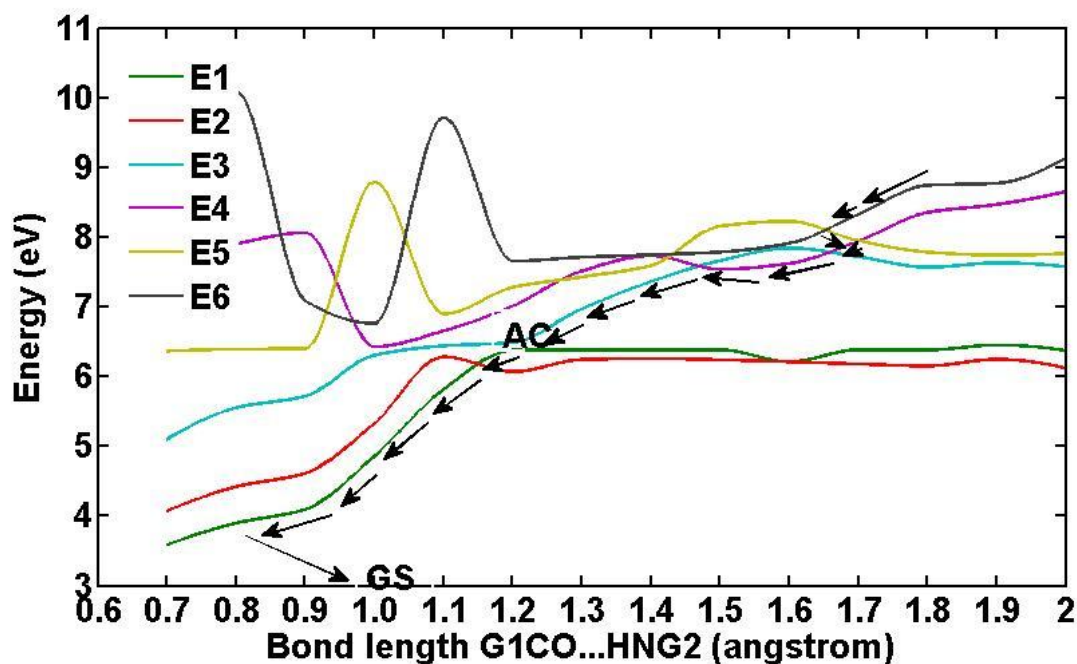

Fig 1: Potential energy surfaces G-G molecule with RCO…HNR in water.

The state E1 and E2 are stable states with hydrogen bond length as oscillator strength is approximate to zero.

IV.CONCLUSIONS

The photostability of GLY-GLY-GLY (G-G-G) molecule has been studied using first principal. Phototability in the molecule has been obtained using CASSCF and the results are validated by experimental data. Good agreement between calculated results and experimental data proves the goodness and effectiveness of approach used in the paper. It is proved that calculation of minimum energy path and full description of all electronic states crossing is necessary to validate the photo induced process in biological proteins through the MEPs. It is strongly supported to CI crossing for excited states deactivation funnel to obtain it ground states which minimum energy states of the molecules. The approach used in paper is of fundamental importance and can be extended to other biomolecules to characterize them.

REFERENCES

[1] C. E. Crespo-Hernandez, B. Cohen, P.M. Hare, B. Khole, "Ultrafast Excited State Dynamics in Nucleic Acids,". Chem. Rev.vol.104, pp. 1977-2019, 2004.

[2] A.L. Sobolewski, W. Domcke, "Computational Studies of the Photophysics of Hydrogen-Bonded Molecular Systems," J. Phys. Chem. A, vol. 111, pp. 11725, 2007.

[3] W. Kohn, L. J. Sham, "Self consistent Field Equations Including Exchange Correlation Effect," vol. 140,Num. 4A, 15 November 1965.

[4] J.P. Roth, S. Lovel, J.M Mayer, "Intrinsic Barriers for Electron and Hydrogen Atom Transfer Reactions of Biomimetic Iron Complexes," J.Phys.Chem. Soc., vol.122, pp. 5486, 2000

[5] J. Von Neumann, E. Wigner, "Conical intersection theory and Computational," Physik. Z, vol. 30, 467, 1929.

[6] D.R. Yarkony, "Diabolical intersections," Rev. Mod.Phys. vol. 68, pp.985, 1996.

[7] H. Norbert, "Photochemical Reactions as Key Steps in Organic Synthesis," Chemical reviews, vol. 108. pp. 1052-103, 2008.

[8] W. Domcke, G. Stock, "Theory of ultrafast Non adiabatic excited-state processes and their spectroscopic detection in real time," Adv. Chem. Phys., vol.100, pp.1, 1997.

[9] S. Perun, A. L. Sobolewski, W. Domcke, "Photophysics of Organic Photostabilizers. Ab Initio Study of the Excited-State Deactivation Mechanisms of 2-(2'-Hydroxyphenyl) benzotriazole" J. Phys. Chem. A, vol.110, pp., 9031, 2006.

[10] Z. Lan, L. M. Frutos, W. Domcke, "Photochemistry of hydrogen-bonded aromatic pairs: Quantum dynamical calculations for the pyrrole–pyridine complex," Proc. Natl. Acad. Sci. U. S. A. vol. 105, pp. 12707, 2008.

[11] W. Domcke, D.R. Yarkony, H. Koppel, "Conical Intersections: Electronic Structure, Dynamics and Spectroscopy," Eds., World Scientific Singapore, 2004.

[12] L. Sobolewski, W. Domcke, "Computational Studies of the Photophysics of Hydrogen-Bonded Molecular Systems." *J. Phys. Chem. A*, vol. *111* (46), pp 11725–11735, 2007.

[13] C.P. Husu, "The Electronic Couplings in Electron Transfer and Excitation Energy Transfer," *Acc. Chem. Res.*, Vol. *42* (4), pp. 509–5, 2009.

[14] M. Roemelt, F. Neese, "Excited States of Large Open-Shell Molecules: An Efficient, General, and Spin-Adapted Approach Based on a Restricted Open-Shell Ground State Wave function," J. Phys.Chem. A, vol.117, pp. 3069, 2013.

[15] M. K.G. Kruse, J. M. Conroy, H. G. Miller, "The Rayleigh Quotient," Physics Department, UOA, Arizona, USA, 2011.

[16] Schapiro, K. Sivalingam, F. Neese, "Assessment of n-Electron Valence St Infrared Depletion Spectroscopy and Structure of the 2-Aminopyridine Dimer.'' *J. Phys. Chem. A*, vol. *108* (45), pp 9715–9720, 2004

[17] C. Wiebeler, V. Borin, A. Vasconcelos, S. de Araújo, I.Schapiro , A. C. Borin**,** "Excitation Energies of Canonical Nucleobases Computed by Multiconfigurational Perturbation Theories," Photochemistry and Photobiology, vol. 93., pp.3, 2017.

[18] K. G. Dyall, "Relativistic and correlation effects in CuH, AgH, and AuH: Comparison of various relativistic methods," J. Chem. Phys. Vol.102, pp.4909, 1995.

[19] Potein structure data is avalabile https://www.rcsb.org/

[20] F. Neese, "The ORCA program system," WIREs Comput Mol Sci., vol.2, pp.73–78,2012,doi: 10.1002/wcms.81

[21] V. Rassolov, J.A. Pople, M. Ratner and T.L. Windus, "6-31G* basis set for atoms K through Zn," J. Chem. Phys. Vol.109, pp. 1223, 1998.

[22] M.M. Francl, W.J. Pietro, W.J. Hehre, J.S. Binkley, M.S. Gordon, D.J. De Frees, J.A. Pople, "Self consistent molecular orbital methods. XXIII. A polarization type basis set for second row elements," J. Chem. Phys. Vol. 7, pp. 3654, 1982.

[23] M. Marazzi, U. Sancho, O. Castano, L. M. Frutos, " First principles study of photostability within hydrogen-Bonded amino acids, " Phys. Chem. Chem. Phys.,vol.13,pp. 7805-7811, 2011

[24] L. Serrono-Andres, M. P., Fülscher, "Theoretical Study of the Electronic Spectroscopy of Peptides 2. Glycine and *N*-Acetylglycine," J.Phys. Chem. B, Vol. 105, pp. 9323-9330,20