Chemical Physics Letters 610–611 (2014) 345–350

Contents lists available at ScienceDirect

Chemical Physics Letters

journal homepage: www.elsevier.com/locate/cplett

Stabilizing the zwitter-ionic form of amino acids in the gas phase: An ab initio study on the minimum number of solvents and ions

Arup Kumar Pathak

Theoretical Chemistry Section, Chemistry Group, Bhabha Atomic Research Centre, Mumbai 400085, India

ARTICLE INFO

Article history:
Received 22 May 2014
In final form 24 July 2014
Available online 1 August 2014

ABSTRACT

Second order Moller–Plesset perturbation theory with 6-311++G(d,p) as well as aug-cc-pvtz basis sets is applied to unearth the minimum number of solvent molecules or ions necessary to stabilize the zwitter-ionic form of amino acids by inhibiting the intra-molecular proton transfer. It is observed that the electrostatic interaction between the amino acids and solvents or ions is responsible for the stability of zwitter-ionic form. It is also observed that minimum two solvent (water and methanol) molecules and a single cation or anion are sufficient to stabilize the zwitter-ionic form of all the twenty standard amino acids.

1. Introduction

Proteins play a very important role in nearly all biological processes. Amino acids are the basic structural units of proteins. The side chains of these building blocks, differ in charge, and hydrogen-bonding capacity that varies from amino acids to the other. Theoretical studies on the interaction of solvents and ions with amino acids are very important in developing a molecular view of how different residues interact with solvent molecules and ions. Such studies provide the information of conformational stability and the nature of active sites of proteins. Knowledge of the contribution of the individual amino acids to the electrostatic field and energetics of proteins is of considerable value in designing enzymes or proteins with enhanced or altered function and stability. It is reported that all the naturally occurring amino acids prefer nonzwitter-ionic form over the zwitter-ionic form in the gas phase as an isolated species and zwitter-ionic form is stable only in the condense phase. The zwitter-ionic form can exists over a wide pH range in aqueous solution [1]. Recently, interests have been shown to study the microsolvation of few amino acids (namely, glycine, alanine, arginine, cysteine, tryptophan, tyrosine, lysine, etc.) out of twenty standard amino acids [2–40]. However, there have been limited number of studies to understand the stability of zwitter-ionic form in the presence of solvents or ions [2–14,34–40]. The minimum number of water molecules required for dissociation of acids, bases, and salts are also studied [41–46]. Solvation can be studied at both the macroscopic and the molecular level. Solvation at the microscopic level permits the quantitative measurement of the intrinsic bond strengths between ions and solvents molecules with amino acids. As physical chemists, we are interested in probing the systems at the atomic level in order to understand the basic principles of solvation. Molecular level knowledge of solute–solvent or solute–ion interactions is crucial to the understanding and prediction of biological phenomenon in vivo. Solute–solvent or solute-ion interactions may play major role on the conformational preference of amino acids. Some of the interesting questions that come up are how do solvent molecules position themselves around the solute amino acids? Are the interaction energies of all amino acids same? How ions do stabilize the zwitter-ion? How many solvent molecules or ions do require to attain the stability of the zwitter-ion in some minimum energy structure in the multidimensional potential energy surface by inhibiting the intra-molecular proton transfer? Answering these questions involves the theoretically challenging task of calculating structures and energetics of the complex of zwitter-ion with solvents and ion in the gas phase. In the work presented here, the interaction of cations, anions, water and methanol with all twenty standard amino acids is investigated at the microscopic level in details. This study will also be helpful to those who are interested in the calculation of microscopic pK_a of amino acids [47].

2. Theoretical methodology

Geometry search for all the systems consisting of twenty amino acids and solvent molecules (water and methanol) and ions (lithium cation, di-positive beryllium cation and chloride anion) have been carried out using electron correlated post Hartee–Fock

E-mail address: akpathak@barc.gov.in

http://dx.doi.org/10.1016/j.cplett.2014.07.058
0009-2614/

second order Moller–Plesset (MP2) perturbation theory with triple split valence, 6-311++G(d,p) as well as aug-cc-pvtz basis sets. A few initial structures (10–15) have been considered for geometry search for the zwitter-ionic form of the twenty amino acids with one or two solvent (water and methanol) molecules and ions. Pseudo Newton–Raphson based algorithm has been employed for the geometry search. The major drawback of this minimum energy structure search method is that it cannot be guaranteed to locate the global minimum in the multidimensional potential energy surface. Hence, global minimum structure search have also been performed adopting Monte-Carlo based simulated annealing procedures. The obtained structure is then re-optimized at MP2 level of theory. Hessian calculations have also been carried out on all the optimized geometry to ensure that there is no imaginary frequency. Morkuma energy decomposition method is applied to partition the total interaction energies into various physically meaningful components namely, electrostatic, exchange, polarization and charge transfer energy [48]. It is based on restricted Hartee–Fock level of theory to decompose the total interaction energy. The electrostatic part corresponds to the interaction energy for the process of bringing individual monomers into the final configuration of the system. On the other hand the exchange-repulsion term is a quantum mechanical phenomenon and is always positive. The interaction between individual constituent monomers in the system will result in distortion of their respective electron densities. The induction effect gives rise to the polarization stabilization of the system. The complex is also stabilized by charge transfer or charge delocalization between the monomers. All electronic structure calculations have been carried out applying the GAMESS suite of program [49]. All visualizations of molecular geometries are carried out through MOLDEN program [50].

3. Results and discussion

The number of solvent molecules and ions, necessary to form the zwitter-ionic conformation of all the amino acids at MP2/6-311++G(d,p) level of theory are provided in Table 1. The absolute energies at the same level of theory are also provided in Table 1. The most stable zwitter-ionic forms of all the amino acids are shown in Figures 1–4, respectively, for water, methanol, lithium cation and chloride anion. Present investigation shows that eight amino acids, namely, arginine, asparagines, aspartic acid, cysteine, phenylalanine, serine, threonine and tyrosine form stable zwitter-ionic form with only one water molecule. It is interesting to mention that all the above amino acids are either hydrophilic or polar except phenylalanine. However, minimum two water molecules are needed to stabilize the zwitter-ionic form of other twelve amino acids, namely, alanine, glutamic acid, glutamine, glycine, histidine,

Table 1
Number of solvent molecules and ions required (bold numeral) to stabilize the zwitter ionic form of the amino acids and absolute energies of the complexes at MP2/6-311++G(d,p) level of theory.

Amino acid	Number of water molecule(s) and absolute energy[a] (a.u.)	Number of methanol molecule(s) and absolute energy[a] (a.u.)	Number of lithium cation and absolute energy[a] (a.u.)	Number of beryllium cation and absolute energy[a] (a.u.)	Number of chloride anion and absolute energy[a] (a.u.)
Alanine	**2**; −475.5794518 (−21.4; −38.4)	**2**; −553.9232429 (−22.9; −43.6)	**1**; −330.3130205 (−83.6; −79.7)	**1**; −337.0321891 (−325.4; −218.5)	**1**; −782.7331125 (−41.2; −55.3)
Arginine	**1**; −681.9127742 (−17.1; −34.5)	**1**; −720.5520239 (−16.9; −35.5)	**1**; −612.4409603 (−97.8; −94.7)	**1**; −619.2462992 (−400.1; −232.3)	**1**;−1064.8388904 (−35.0; −50.8)
Asparagines	**1**; −567.6267487 (−14.2; −34.7)	**1**; −606.7735434 (−14.3; −37.4)	**1**; −498.6632494 (−86.7; −82.8)	**1**; −505.4010835 (−354.5; −243.6)	**1**; −951.0829461 (−51.1; −69.6)
Aspartic acid	**1**; −587.4747527 (−14.8; −34.1)	**1**; −626.6197828 (−14.7; −36.6)	**1**; −518.4998237 (−81.8; −77.7)	**1**; −525.2191695 (−323.3; −214.0)	**1**; −970.9343100 (−55.3; −76.8)
Cysteine	**1**; −796.9464986 (−15.1; −38.5)	**1**;−736.968277 (−15.4; −42.6)	**1**; −727.9844851 (−81.4; −76.8)	**1**; −734.7019973 (−321.4; −210.9)	**1**; −1180.4114471 (−43.3; −57.8)
Glutamic acid	**2**; −702.9682165 (−39.0; −84.2)	**2**; −781.3165693 (−39.7; −90.2)	**1**; −557.7070336 (−87.6; −83.5)	**1**; −564.434372 (−327.3; −216.2)	**1**; −1010.1405561 (−56.2; 78.7)
Glutamine	**2**; −683.1334918 (−19.7; −36.7)	**2**; −761.4555507 (−19.4; −39.1)	**1**; −537.8690550 (−92.3; −88.5)	**1**;−544.603647 (−371.1; −258.9)	**1**; −990.2693064 (−43.8; 58.1)
Glycine	**2**; −436.3737441 (−32.1; −71.7)	**2**; −514.7198014 (−32.4; −75.4)	**1**; −291.1086136 (−81.5; −78.0)	**1**; −297.8202768 (−319.7; −216.5)	**1**; −743.5310251 (−42.1; −56.3)
Histidine	**2**; −700.0600340 (−30.7; −69.8)	**1**; −662.9276496 (−11.5; −23.5)	**1**; −554.7838508 (−84.1; −79.1)	**1**; −561.5104887 (−331.4; −216.9)	**1**; −1007.2230229 (−52.0; 71.7)
Isoleucine	**2**; −593.1700556 (−30.0; −68.7)	**1**; −556.0389625 (−14.6; −40.9)	**1**; −447.9045880 (−85.5; −80.5)	**1**; −454.6344699 (−334.1; −219.7)	**1**; −900.3241158 (−38.54; −53.5)
Leucine	**2**; −593.1731290 (−35.3; −81.2)	**2**; −671.5195422 (−35.2; −84.9)	**1**; −447.9072503 (−85.9; −81.2)	**1**; −454.6371082 (−335.1; −223.0)	**1**; −900.3265509 (−37.6; −52.6)
Lycine	**2**; −648.3927048 (−33.6; −64.7)	**2**; −726.7392757 (−32.8; −67.6)	**1**; −503.1229388 (−88.0; 82.6)	**1**; −509.8502558 (−332.4; −220.4)	**1**; −955.5685989 (−39.6; −55.0)
Methionine	**2**; −951.6421504 (−29.2; −66.9)	**1**; −914.5115958 (−13.0; −30.1)	**1**; −806.3806675 (−88.8; −84.4)	**1**; −813.1157602 (−342.1; −229.1)	**1**; −1258.7984731 (−41.1; −57.4)
Phenylalanine	**1**; −629.6987989 (−14.9; −36.9)	**1**; −668.8443538 (−15.1; −40.7)	**1**; −560.7341068 (−87.4; −82.4)	**1**; −567.4687010 (−339.1; −223.4)	**1**; −1013.1487988 (−36.8; −50.5)
Proline	**2**; −552.7776558 (−30.5; −68.9)	**2**; −631.1239449 (−30.3; −73.0)	**1**; −407.5187162 (−87.2; −82.8)	**1**; −414.2476191 (−335.2; −223.3)	**1**; −859.9351731 (−38.2; −53.6)
Serine	**1**; −474.3585137 (−16.0; −41.4)	**1**; −513.5089973 (−16.6; −46.3)	**1**; −405.3947073 (−86.1; −82.2)	**1**; −412.1187682 (−333.1; −224.6)	**1**; −857.8202344 (−43.7; −58.4)
Threonine	**1**; −513.5598815 (−14.6; −34.4)	**1**; −552.7093120 (−15.0; −39.1)	**1**; −444.5875090 (−79.4; −74.7)	**1**; −451.3154551 (−344.3; −230.9)	**1**; −897.0230097 (−43.8; −58.1)
Tryptophan	**2**; −837.2561748 (−34.1; −65.9)	**1**; −800.1226541 (−15.6; −38.1)	**1**; −691.9843244 (−92.1; −85.7)	**1**; −698.7275704 (−352.9; −241.8)	**1**; −1144.3971363 (−39.7; −52.1)
Tyrosine	**1**; −704.7828724 (−13.1; −23.9)	**1**; −743.9248768 (−11.6; −22.7)	**1**; −635.8165556 (−87.8; −82.8)	**1**; −642.5535125 (−341.3; −225.1)	**1**; −1088.2331221 (−43.9; −61.1)
Valine	**2**; −553.9737648 (−30.2; −69.0)	**2**; −632.3204399 (−30.0; −73.4)	**1**; −408.7087804 (−85.0; −80.3)	**1**; −415.4362397 (−332.0; −220.1)	**1**; −861.1273327 (−39.1; −53.9)

[a] Values in the parentheses are the basis set superposition error corrected total interaction energy and electrostatic energy in kcal/mol using Morokuma scheme.