Kereselidze et al., J Comput Sci Syst Biol 2017, 10:1
DOI: 10.4172/jcsb.1000241

Journal of
Computer Science & Systems Biology

Research Article Open Access

Quantum-Chemical Study of the Propensity of the Amino Acid Pairs for the Peptide Bond Formation

Jumber Kereselidze[1*], George Mikuchadze[1] and Lia Bobokhidze[1]

[1]Department of Chemistry, Ivane Javkhishvili Tbilisi State University, Tbilisi 0179, Georgia

Abstract

The Modern method of quantum chemistry - the density functional theory (DFT) was used to quantitative describe the formation of the peptide bond between amino acid pairs. In particular the formula of parameter of propensity (K_p), which is a function of six variables: the length and order of the CO and NH bonds, (R_{CO}, R_{NH}, P_{CO} P_{NH}), the activation energy for the formation of the peptide bond ($\Delta E^{\#}$), and difference between charges of the carbon atom of the carbonyl group and the amino nitrogen atom (Δq) was constructed. By means of the proposed formula the K_p parameter for 400 amino acid pairs was calculated. Among them only 26 amino acid pairs are most likely to take part in the synthesis of proteins that have been selected based on the value of the parameter K_p. This approach may have important meaning for quantitative description of the amino acid sequences in proteins.

Keywords: Amino acids; Peptide bonds; Parameter of propensity; DFT calculations

Introduction

The theoretical description of biochemical processes is the main focus of modern natural science - Biophysical Chemistry. In recent years, for the quantitative description of complex biochemical processes are widely used modern quantum chemistry methods based on density functional theory (DFT). Including for research of peptide bond formation mechanism [1]. It is assumed that the inductive and field effects of amino acid side chains have an essential effect on peptide bond formation [2]. The quantum-mechanical study of different possible mechanisms of peptide synthesis in the ribosome has been carried out using density functional also [3]. Analysis of a database of protein sequences for all possible binary patterns of polar and non-polar amino acid residues revealed that alternating patterns occur significantly less often than others with similar composition [4]. To facilitate understanding of the information available for protein structures, has been constructed the structural classification of proteins (scop) database. This database provides a detailed and comprehensive description of the structural and evolutionary relationships of the proteins of known structure [5]. Analysis of extant proteomes has the potential of revealing how amino acid frequencies within proteins have evolved over biological time. Evidence presented here indicates, that cysteine, tyrosine, and phenylalanine residues have substantially increased in frequency [6]. To understand more fully how amino acid composition of proteins has changed over the course of evolution, a method has been developed for estimating the composition of proteins in an ancestral genome. The method was used to infer the amino acid composition of a large protein set in the Last Universal Ancestor (LUA) of all extant species. It is proposed that the inferred amino acid composition of proteins in the LUA probably reflects historical events in the establishment of the genetic code [7].

Protein sequences contain many local regions of low compositional complexity. These include different types of residue clusters, some of which contain homopolymers, short period repeats or aperiodic mosaics of a few residue types. Several different formal definitions of local complexity and probability are presented and are compared for their utility in algorithms for localization of such regions in amino acid sequences and sequence databases [8]. The occurrence of all di- and tripeptide segments of proteins was counted in a large data base containing about 119 000 residues [9]. Systematic conformational analysis study of the tripeptide units (Gly-X-Pro) and (Gly-Pro-X), with X=Pro, Ala, Ser, Val, Leu, Ile, and Phe it has been reported. The low-energy conformers obtained by quantum computations are discussed with respect to other theoretical investigations and experimental protein structures [10]. Pairing of Amino acids is only possible on a parallel β ribbon and involves both the polypeptide backbones and the side chains. Model building revealed that of the 210 possible amino acid pairs of the standard 20 amino acids, no more than 26 could be built to meet standard criteria for bonding. Of these 26, 14 were found to be genetically encoded when the codons are read as if they paired in a parallel manner [11].

Methods

Density functional theory (DFT) is a computational quantum mechanical method used in physics, chemistry and biology for investigate the electronic structure in particular atoms, and molecules [12]. The properties of a many-electron system can be determined by using functionals, which in this case is the spatially dependent electron density. Hence the name density functional theory comes from the use of functionals of the electron density. DFT is among the most popular and versatile methods available in computational biology. Unlike the wave function, which is not a physical reality, electron density is a physical characteristic of all molecules. The electron density is a function with three variables - x^-, y^-, and z^-position of the electrons. Hybrid methods, as the name suggests, attempt to incorporate some of the more useful features from ab initio methods (specifically Hartree-Fock methods) with some of the improvements of DFT mathematics. Hybrid methods, such as B3LYP [13-15], tend to be the most commonly used methods

***Corresponding author:** Jumber Kereselidze, Department of Chemistry, Ivane Javkhishvili Tbilisi State University, Tbilisi 0179, Georgia, Tel: 995322225107; E-mail: jumber.kereselidze@tsu.ge

Received January 16, 2017; **Accepted** January 30, 2017; **Published** January 31, 2017

Citation: Kereselidze J, Mikuchadze G, Bobokhidze L (2017) Quantum-Chemical Study of the Propensity of the Amino Acid Pairs for the Peptide Bond Formation. J Comput Sci Syst Biol 10: 014-016. doi:10.4172/jcsb.1000241

for computational chemistry and Biology. Calculations were performed using software, "Priroda" in regime of the reaction coordinate [16].

Results

The aim of this work is a quantitative description of the propensity of 400 pairs of 20 amino acids for the formation of peptide bond using quantum - chemical method of density functional theory (DFT). For the study of the conformation of 20 amino acids by Kolaskar and Ramabraham 400 options were selected also [17]. As know, at the stage of activation of the formation of the peptide bond between two amino acids the C-O bond of the first and N-H bond of the second one are ruptured, what leads to C-N bond formation with release of water molecule (Figure 1). With the aim of quantitative description of peptide bond formation by the quantum-chemical method DFT the bond orders of carbonyl and amino groups (P_{CO}, P_{NH}) and corresponding bond lengths (R_{CO}, R_{NH}), difference values of charge between C and N atoms (Δq), as well as the activation energy ($\Delta E^{\#}$) and the reaction energy (ΔE) have been calculated. In addition, it must satisfy the condition of exothermic formation of the peptide bond ($\Delta E<0$). The increase of bond lengths R_{CO} and R_{NH}, as well as the difference of charges between carbon atom of carbonyl group and nitrogen atom the amino group (Δq) causes an increase in the value of the parameter of the propensity of amino acids of formation of peptide bond (K_p). Hence, these values in the formula K_p placed in the numerator. On the other hand, the decrease of bond orders P_{CO} and P_{NH}, as well as the activation energy $\Delta E^{\#}$ the parameter of propensity of the peptide bond formation (K_p) also increases. Hence, these values are in denominator of the formula. Consequently, the formula, for the quantitative description of the effect of R-groups of the formation of peptide bond was built.

$$K_P = \frac{\Delta q.R_{NH}.R_{CO}}{P_{NH}.P_{CO}.\Delta E^{\#}}$$

From the analysis of the data of Table 1, one can make a general conclusion relatively of the parameter of propensity of amino acids for the formation of peptide bond (K_p). In Particular, K_p inversely proportional of the activation energy $\Delta E^{\#}$ of peptide bond formation. In this row the pairs of the amino acids are listed in order of increase of the values of K_p and we expect that, they must be presented in proteins more often (for example 21-26 amino acid pairs, from Table 1). Such conclusion, at first glance may seem a less convincing, however from the point of view of the R-group's influence on the ability of the peptide bond formation of amino acids K_p parameter can be successfully applied to the quantitative description of this process. In addition, the value of the reaction energy is always negative, which is common for exothermic reactions. In other cases, values of K_p is low, which indicates on difficulty of peptide bond formation. It is known that the pairing of amino acids in proteins are encoded by nucleotide base sequences of RNA. On the other hand, the formation of peptide bonds is much dependent on the influence of inductive and steric effects of R-groups. One can consider that, the energy ($\Delta E^{\#}$, ΔE), electronic (q_N, q_C, P_{CO} and P_{NH}) and structural (R_{CO} and R_{NH}) values contained in the Table 1, has completely reasonable values, from the viewpoint of chemical reactions. Such choice of amino acid pairs is probably caused by different direction of influence of inductive and steric effect of R-groups with respect to formation of peptide bonds. It means, that one R-group can promote, but other one prevent the formation of peptide bond. The proposed quantum-chemical formula of amino acid pairs for the peptide bond formation, in our opinion may contribute to the prediction of the direction of peptide bond formation of a quantitative description of the propensity.

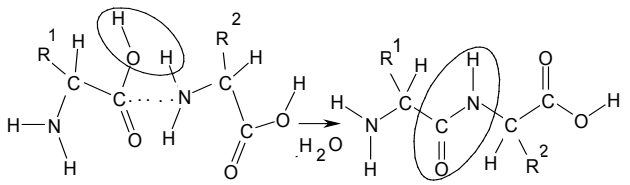

Figure 1: Scheme of the peptide bond formation resulting from the interaction of two different amino acid.

N	Amino acid pairs	$\Delta E^{\#}$, kJ/mol	Δq	P_{CO}	P_{NH}	R_{CO}, Å	R_{NH}, Å	$K_P \cdot 10^{-2}$	ΔE, kJ/mol
1	Thr-Asn	19.4	0.24	0.86	0.77	1.45	1.10	3.0	-23.1
2	Val-Met	24.1	0.22	0.81	0.67	1.55	1.20	3.1	-68.2
3	Pro-Lys	24.4	0.19	0.71	0.66	1.55	1.20	3.1	-34.1
4	Ileu- Phe	26.5	0.25	0.81	0.65	1.55	1.20	3.3	-23.6
5	Gln-Trp	21.0	0.26	0.88	0.76	1.55	1.20	3.4	-30.4
6	Gln-tyr	26.2	0.29	0.81	0.66	1.55	1.20	3.8	-11.8
7	Leu-Ser	17.1	0.24	0.83	0.74	1.50	1.15	3.9	-27.8
8	Tyr-His	16.2	0.25	0.90	0.73	1.50	1.15	4.1	-41.9
9	Thr-Trp	18.4	0.24	0.79	0.63	1.55	1.20	4.8	-39.4
10	Pro-Arg	19.7	0.28	0.79	0.66	1.55	1.20	5.1	-32.8
11	Arg-Trp	15.0	0.22	0.78	0.67	1.55	1.20	5.2	-5.5
12	Thr-Tyr	14.1	0.22	0.79	0.68	1.55	1.20	5.4	-60.4
13	Trp-Gln	10.5	0.24	0.86	0.73	1.50	1.15	6.2	-23.5
14	His-Ala	11.6	0.22	0.77	0.68	1.55	1.20	6.7	-39.4
15	Tyr-Asp	10.6	0.25	0.82	0.72	1.50	1.15	6.9	-49.9
16	Thr-Phe	7.6	0.25	0.86	0.78	1.45	1.10	7.8	-78.7
17	Thr-His	6.3	0.23	0.80	0.69	1.55	1.20	12.3	-57.2
18	Gln-Arg	6.7	0.25	0.74	0.67	1.55	1.20	14.0	-4.5
19	Glu-Lys	5.7	0.25	0.81	0.66	1.55	1.20	15.5	-41.5
20	Tyr-Glu	3.7	0.18	0.79	0.66	1.55	1.20	17.3	-57.7
21	Thr-Glu	6.3	0.24	0.75	0.58	1.60	1.25	17.4	-50.0
22	Ala-Ser	6.3	0.29	0.77	0.59	1.60	1.25	20.3	-44.6

Table 1: The values of activation energy ($\Delta E^{\#}$), difference of charge between carbon atom of carbonyl group and nitrogen atom of amino group (Δq), bond orders and bond lengths of CO and NH bonds (P_{CO}, P_{NH}, R_{CO}, (Å) R_{NH}, (Å) and the parameter of propensity of peptide bond formation (K_P).

Conclusion

It is unknown to what extent the condensation of mixtures of amino acids occurs preferably randomly or between certain amino acids. In this case, the terminal amino acid residue of the polypeptide could influence the choice of the next amino acid. Such a mechanism could the determine the organization of the first polypeptides. The proposed quantum-chemical formula, which includes the calculated values of energetic, electronic and structural characteristics of amino acids can be applied for quantitative description of the propensity of amino acids for the peptide bond formation as these values quantitatively describe inductive and steric influence of R-groups of amino acids on the reaction center during the peptide bond formation. These factors together encoding nucleotide bases can have an important role in the study of the amino acid sequences in proteins. According to Atkins the orderliness can turn into chaos, and in certain conditions, from chaos it is possible the formation of the orderliness also [18]. On the other hand, F. Chapeville [19] wrote that in nature, there are many different amino acids in the free state, but only 20 of them are used for protein synthesis. The reason for this limitation is not known, but it is apparently related to the fact that during the evolution of selection showed severe requirements for the physicochemical properties of these molecules, resulting in all the other amino acids were discarded. Based on the above we can assume that our results quantitatively describe prediction F. Chapeville. The proposed theoretical formula for the probability of the sequence of amino acids in the polypeptides can be used for building an adequate correlation with data of the morphological study of the structure of proteins.

Acknowledgement

This work was supported by the Shota Rustaveli National Science Foundation Grant No: 217732.

References

1. Kastner J, Sherwood P (2010) The ribosome catalyzes peptide bond formation by providing high ionic strength. Mol Phys 108: 293-306.
2. Dwyer DS (2005) Electron properties of amino acid side chains: quantum mechanics calculation of substituent effects. BMC Chem Biol 5: 2.
3. Acosta-Silva C, Bertran J, Branchadell V, Oliva A (2012) Quantum-mechanical study on the mechanism of Peptide Bond Formation in the Ribosome. J Am Chem Soc 134: 5817-5831.
4. Broome BM, Hecht MH (2000) Nature disfavors sequences of alternating polar and non-polar amino acids: implications for amyloid genesis. J Mol Biol 296: 961-968.
5. Urzin AG, Brenner SE, Hubbrad T, Chothia (1995) C SCOP: Structural Classification of proteins database for the investigation of sequences and structures. J Mol Biol 247: 536-540.
6. Brooks DJ, Fresco JR (2002) Increased Frequency of Cysteine, Tyrosine, and Phenylalanine Residues Since the Last Universal Ancestor. Mol Cell Prot 1: 125-131.
7. Brooks DJ, Fresco JR, Lesk AM, Singh M (2002) Evolution of Amino Acid Frequencies in Proteins Over Deep Time: Inferred Order of Introduction of Amino acids into Genetic Code. Mol Biol Evol 19: 1645-1655.
8. Wootton JC, Federhen S (1993) Statistics of local complexity in amino acid sequences and sequence database. Computer & Chemistry 17: 149-163.
9. Vonderviszt F, Matrai GY, Simon I (1986) Characteristic sequential residue environment of amino acids in protein. Int J Pept Prot Res 27: 483-492.
10. Vasilescu D, Cabrol D, Broch H (1983) Conformons in amino acid sequences typical of collagen: A quantum mechanical study. Int J Quant Chem 24:109-122.
11. Root-Bernstein RS (1982) Amino acid pairing. J Theory Biol 94: 885-894.
12. Kohn W, Becke A, Parr R (1996) Density Functional Theory of Electronics Structure. J Phys Chem 100: 12974-12980.
13. Becke A (1988) Density Functional Exchange: Energy approximation with correct asymptotic behavior. Phys Rev. A 38: 3098-3100.
14. Lee C, Yang W, Parr R (1988) Development of the Colle-Sallvetti correlation - energy formula into a functional of the electron density. Phys Rev B 37:785-789.
15. Perdew J, Wang Y (1992) Accurate and simple analytic representation of the electron-gas correlation energy. Phys Rev B 45: 13244-13249.
16. Laikov D, Ustynyuk Yu (2005) PRIRODA 04: A Quantum-Chemical Program Suite. New Possibilities in the Study of Molecular Systems with the Application Parallel Computing. Russ Chem Bull Int Ed. 540:820-826.
17. Kolaskar A, Ramabrahmam V (1983) Conformational properties of pairs of amino acids. Int J Pept and Prot Res 22:83-91.
18. Atkins PW (1984) The Second Law, Scientific American Books, An Imprint of W.H. Freeman and Company, New York. Chapeville F, Haenni A-L Biosynthesis des proteins. Hermann Collection Methods. Paris