Kereselidze et al., J Proteomics Bioinform 2018, 11:9
DOI: 10.4172/jpb.1000483

Journal of
Proteomics & Bioinformatics

Research Article Open Access

Quantum-Chemical Description of Some Physical-Chemical Properties of Proteinogenic Amino Acids

Jumber Kereselidze*, George Mikuchadze and Lia Bobokhidze

Department of Chemistry, Ivane Javakhishvili Tbilisi State University, 3 I. Chavchavedze Ave, 0179 Tbilisi, Georgia

Abstract

We describe an impact of the inductive and steric effects of R-groups of amino acids on the reaction center (carboxy and amine groups) to estimate the propensity of amino acids for the peptide bond formation. These effects were quantitatively evaluated using the orders of the C-O and N-H bonds (P_{CO} and P_{NH}), the charges on the carbon, nitrogen and oxygen atoms of the carboxy, amine and hydroxy groups (q(C3), q(N6), q(O2)) and the dipole moments of all amino acids (μ). The calculations were carried out by means of modern quantum-chemical method - Density Functional Theory (DFT).

Keywords: Amino acids; Physical chemical properties; DFT calculations

Introduction

It is well known that only twenty of amino acid are used for the synthesis of protein. The reason for this restriction is still unknown, but such choice is obviously caused by the physical-chemical properties of these molecules [1]. Electronic properties of amino acid side chains, such as inductive and field effects still have not been investigated in details. Quantum-chemical calculations of the influence of R-groups can help to evaluate their role in amino acid pairing. Dwyer [2] and Grantham [3] indicated fundamental role of substituent effects of amino acid side chains in the protein structure. Dwyer attempted, to estimate quantitatively these effects using quantum mechanics calculations [2]. Kolaskar et al. selected 6 'obligatory' amino acids (Ser, Val, Leu, Asp, Gly and Pro) based on the comparative analysis of the conformational similarity of amino acid residues [4]. Amino acid residues cover a wide range of shapes, sizes in many atomic and molecular interactions. The residues determine the broad variety of bio-physicochemical properties that are fundamental in ascertaining macromolecular structures and functional activities. Based on their electronic properties a classification of amino acid type was described - known as Taylor classification [5]. Amino acid residues are classified into three groups, depending on their polarity: polar (Arg, Lys, His, Gln, Asn, Asp and Glu), weak polar (Ala, Pro, Gly, Thr and Ser) and nonpolar (Cys, Val, Met, Ile, Leu, Phe, Tyr and Trp) [6]. Naturally occurring amino acids can be grouped based on their similarity of physical-chemical properties. A collection of physical-chemical properties of amino acids will be helpful to study macroscopic properties of proteins (such as aggregation), perform sequence comparison or understand conservation of functionally important residues in a protein family (physico-chemical signatures). Venkatarajan et al. collated 242 properties for the 20 naturally occurring amino acids and created a database named APDbase (Amino acid Physico-chemical properties Database) [7]. The huge majority of researches of electronic properties of amino acid side chains relate to conformation preference, little or nothing is known about their role in peptide synthesis. From our point of view, this process is not less important. This article is our modest attempt to shed light on this issue (Figure 1).

Material and Method

The density functional theory (DFT) is a quantum computational method used in physics, chemistry and biology for investigation of the electronic structure of atoms and molecules [8]. 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 of density functional theory comes from the use of functionals of 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 molecules. 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 [9-11] most commonly used for computational chemistry and Biology. Calculations were performed using software,"Priroda-8" in regime of the reaction coordinate [12].

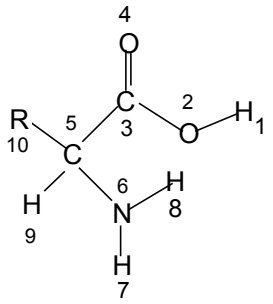

Figure 1: Amino acids: R = H **(Gly)**, CH_3 **(Ala)**, CH_2OH **(Ser)**, CH_2COOH **(Asp)**, $(CH_2)_3NHC=NHNH_2$ **(Arg)**, $CH(CH_3)_2$ **(Val)**, $CHOHCH_3$ **(Thr)**, $(CH_2)_2COOH$ **(Glu)**, $CH_2CH(CH_3)_2$ **(Leu)**, CH_2SH **(Cys)**,$CH_2C_3H_3N_2$ **(His)**, $(CH)_4NH_2$ **(Lys)**, $CHCH_3CH_2CH_3$ **(Ile),** $CH_2C_6H_4OH$ **(Tyr)**, $CH_2C=ONH_2$ **(Asn)**, $(CH_2)_2SCH_3$ **(Met)**, $(CH_2)C_8H_6N$ **(Trp),** $CH_2C_6H_5$ **(Phe),** $(CH_2)_2C=ONH_2$**(Gln),** C_4H_8N **(Pro)**.

***Corresponding author:** Jumber Kereselidze, Department of Chemistry, Ivane Javakhishvili Tbilisi State University, 3 I. Chavchavedze Ave, 0179 Tbilisi, Georgia, Tel: + 995 32 25 30 48; E-mail: jumber.kereselidze@tsu.ge

Received June 05, 2018; **Accepted** October 01, 2018; **Published** October 08, 2018

Citation: Kereselidze J, Mikuchadze G, Bobokhidze L (2018) Quantum-Chemical Description of Some Physical-Chemical Properties of Proteinogenic Amino Acids. J Proteomics Bioinform 11: 169-172. doi: 10.4172/jpb.1000483

Results and Discussion

It is known that the electronic structure of the amino acids is one of the main factors, which promote or hinders the formation of a peptide bond. Concerning with this, the values of charges on the carbon atom of the carbonyl group - $q(C_3)$, on the oxygen atom of the hydroxyl group $q(O_2)$ and on the nitrogen atom of the amine group $q(N_6)$, and also the orders of the CO and NH bonds (Pco, P_{NH}) and the dipole moments (μ) were calculated using DFT. The results of the calculations are given in Table 1. The search for a quantitative correlation between the calculated physicochemical characteristics of amino acids and their percentage content in proteins (%) (Table 1) is not always successful. However, in some cases the qualitative dependence is observed. In particular, comparative analysis shows a symbate correlation between percentage of content (%) and the dipole moment.

From the Tables 2 and 3 is clear that the values of the content of amino acids (%) and their dipole moments (μ) are grouped in three ranges: high, average and low. Consequently, depending on the calculated values of the amino acid dipole moments (μ), their percentage in proteins (%) can be estimated.

Table 4 shows the values of the charges on the carbon atoms of the carbonyl group - q (C_3), the nitrogen of the amino group - q (N_6) and oxygen - the q (-O-) hydroxyl group, and also the order of the C-O bond – P_{CO}. Using the numerical values of these characteristics, a number of amino acids were constructed in the direction of their decrease in content:

$q(C_3)$: Arg > Thr, Asn, Lys > Ala, Glu, Gln, Met, Thr; (a)

$q(N_6)$: Arg > Thr, Asn > Lys > Ala, Glu, Gly; (b)

q(-O-): Pro> Lys > Thr > Arg > His > Trp > Val > Asp > Gln > Phe > Gly, Glu > Leu, Ala, Asn; (c)

P(C-O): Cys > Gly > Thr > Leu > His > Ala, Trp > Met, Ile, Phe, Tyr, Pro, Val, Asn, Asp > Arg, Glu, Gln, Lys. (d)

The increase in the positive charge C_3 of the carbon atom of the carbonyl group promotes the nucleophilic attack of aminoacyladenyl acid on the carboxy group to form an ester and the subsequent break of the C-O bond. As can be seen from the series (a), arginine has the highest nucleophilic property of the C_3 atom, and threonine has the lowest one. A similar distribution is observed for the nitrogen atom-N_6 (row b). The charge of hydroxyl oxygen (-O-) is given in the row (c). The highest values are observed for proline and lysine (-0.163 and -0.160), and the lowest values for asparagine (-0.089). Consequently, asparagine must have a high acidity, since the hydroxyl oxygen atom weakly retains the acid proton due to the low value of the electronic charge. The row (d) shows that Lysine has the lowest value (1.07) of the order of the C-O (Pco) bond, and Cysteine has the highest one (1.17). Consequently, Lysine promotes the breakdown of the C-O bond and therefore easily forms the peptide bond, but cysteine - has the difficulties for the C-O bond breakage.

Based of tabular data of dipole moments, it is possible to construct a series of amino acid polarities:

Amino Acids	q (C3)	q (N6)	q (O2)	P2,3 (CO)	P6,7 (NH)	μ, D	% [13,14]	% [15]
Ala	0.195	-0.225	-0.151	1.1	0.91	2.38	7.4	7.6
Gly	0.192	-0.224	-0.155	1.15	0.91	0.94	7.4	6.8
Leu	0.194	-0.213	-0.151	1.12	0.91	1.66	7.6	9.4
Ser	0.189	-0.227	-0.148	1.07	0.9	3.7	8.1	7.1
Lys	0.197	-0.223	-0.16	1.07	0.91	2.1	7.2	5.9
Val	0.193	-0.223	-0.156	1.09	0.91	2.77	6.8	6.6
Glu	0.195	-0.212	-0.152	1.08	0.91	2.64	5.8	6.4
Thr	0.197	-0.227	-0.159	1.14	0.91	2.46	6.2	5.7
Asp	0.194	-0.221	-0.155	1.09	0.9	4	5.9	5.3
Arg	0.198	-0.223	-0.158	1.08	0.93	2.37	4.2	5.2
Pro	0.194	-0.164	-0.163	1.09	0.9	2.17	5	4.9
Ile	0.194	-0.213	-0.149	1.09	0.9	4.61	3.8	5.8
Asn	0.197	-0.213	-0.089	1.09	0.9	3.27	4.4	4.4
Gln	0.195	-0.213	-0.154	1.08	0.91	2.84	3.7	4
Phe	0.195	-0.224	-0.154	1.09	0.91	2.87	4	4.1
Cys	0.196	-0.224	-0.148	1.17	0.9	6.14	3.3	1.7
Trp	0.191	-0.219	-0.157	1.1	0.91	1.29	1.3	1.2
Tyr	0.192	-0.214	-0.148	1.09	0.92	3.5	3.3	3.2
His	0.193	-232	-0.157	1.11	0.9	3.13	2.9	2.2
Met	0.195	-0.215	-0.152	1.09	0.9	1.99	2.4	2.4

Table 1: Electronic (q_i - charges on atoms, P_{ij}-bond orders, μ-dipole moments, content of amino acids in proteins (%)

R = H **(Gly)**, CH_3 **(Ala)**, CH_2OH **(Ser)**, CH_2COOH **(Asp)**, $(CH_2)_3NHC{=}NHNH_2$ **(Arg)**, $CH(CH_3)_2$ **(Val)**, $CHOHCH_3$ **(Thr)**, $(CH_2)_2COOH$ **(Glu)**, $CH_2CH(CH_3)_2$ **(Leu)**, CH_2SH **(Cys)**, $CH_2C_3H_3N_2$ **(His)**, $(CH)_4NH_2$ **(Lys)**, $CHCH_3CH_2CH_3$ **(Ile)**, $CH_2C_6H_4OH$ **(Tyr)**, $CH_2C{=}ONH_2$ **(Asn)**, $(CH_2)_2SCH_3$ **(Met),** $(CH_2)C_8H_6N$ **(Trp)**, $CH_2C_6H_5$ **(Phe),** $(CH_2)_2C{=}ONH_2$**(Gln)**.

High content in percentage, %	8.1(Ser); 7.6 (Leu); 7.4 (Ala); 7.4 (Gly); 7.2 (Lys).	High value of dipole moment μ, D	6.1(Cys);4.6(Ile);4.0(Asp); 3.7(Ser) 3.5(Tyr); 3.3(Asn); 3.1(His).
Average content in percentage, %	6.8(Val); 6.2(Thr); 5.9(Asp); 5.8(Glu); 5.0(Pro); 4.4(Asn); 4.2(Arg); 4.0(Phe).	Average value of dipole moment μ, D	2.9(Phe); 2.8(Gln); 2.8(Val); 2.6(Glu); 2.5(Thr); 2.4(Ala); 2.4(Arg); 2.2(Pro); 2.1(Lys).
Low content in percentage, %	3.8(Ile); 3.7 (Gln); 3.3(Cys); 3.3(Tyr). 2.9(His);1.8 (Met); 1.3(Trp)	Low value of dipole moment μ, D	2.0 (Met); 1.7(Leu); 1.3(Trp); 0.9(Gly).

Table 2: The values of the percentage content in proteins and dipole moments of proteinogenic amino acids.

High content in percentage, %	9.4(Leu); 7.6(Ala); 7.1(Ser); 6.8(Gly); 6.6(Val); 6.4(Glu).	High value of dipole moment μ, D	6.1(Cys);4.6(Ile);4.0(Asp); 3.7(Ser) 3.5(Tyr); 3.3(Asn); 3.1(His).
Average content in percentage, %	5.9(Lys); 5.7(Thr); 5.8(Ile); 5.3(Asp); 5.2(Arg); 4.9(Pro); 4.4(Asn); 4.1(Phe); 4.0(Gln).	Average value of dipole moment μ, D	2.9(Phe); 2.8(Gln); 2.8(Val); 2.6(Glu); 2.5(Thr); 2.4(Ala); 2.4(Arg); 2.2(Pro); 2.1(Lys).
Low content in percentage, %	3.2(Tyr); 2.4 (Met); 2.2(His); 1.7(Cys); 1.2(Trp).	Low value of dipole moment μ, D	2.0(Met); 1.7(Leu); 1.3(Trp); 0.9(Gly).

Table 3: The values of the percentage content in proteins and dipole moments of proteinogenic amino acids.

Am. Ac	Gly	Ala	Ser	Cys	Thr	Val	Asn	Asp	Glu	Gln
q (C_3)	0.192	0.195	0.189	0.186	0.197	0.193	0.197	0.194	0.195	0.195
q (N_6)	-0.213	-0.224	-0.213	-0.212	-0.227	-0.214	-0.227	-0.221	-0.224	-0.223
q (-O-)	-0.155	-0.151	-0.148	-0.148	-0.159	-0.156	-0.089	-0.155	-0.152	-0.154
P(C-O)	1.15	1.10	1.07	1.17	1.14	1.09	1.09	1.09	1.08	1.08
Am. Ac	Met	Leu	Ile	Lys	Arg	Phe	Tyr	His	Trp	Pro
q (C_3)	0.195	0.194	0.194	0.197	0.198	0.195	0.192	0.193	0.191	0.184
q (N_6)	-0.224	-0.219	-0.223	-0.225	-0.232	-0.223	-0.213	-0.215	-0.213	-0.164
q (-O-)	-0.152	-0.151	-0.149	-0.160	-0.158	-0.154	-0.148	-0.157	-0.157	-0.163
P(C-O)	1.09	1.12	1.09	1.07	1.08	1.09	1.09	1.11	1.10	1.09

Table 4: Charges on C_3, N_6, -O- atoms and bond order P_{CO} of proteogenic amino acids.

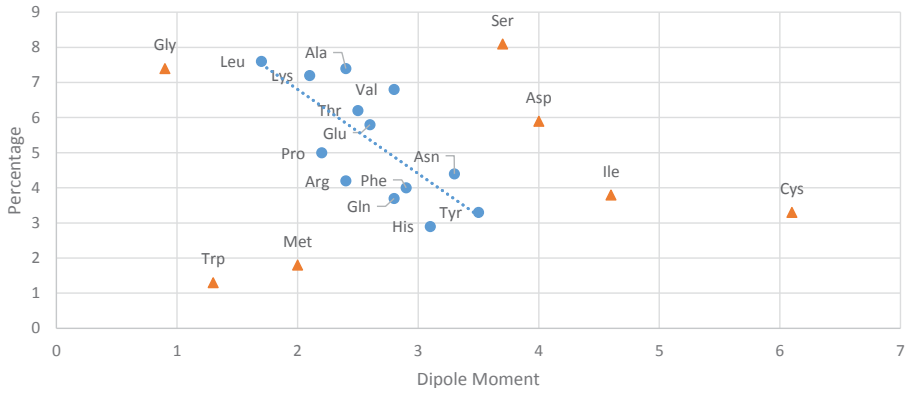

Figure 2: Percentage of amino acids appearance in proteins on its dipole moment.

Gly<Trp<Leu<Met<Lys<Pro<Arg<Ala<Thr<Glu<Val<Gln<Phe<His<Asn<Tyr<Ser<Asp<Ile<Cys.

For the thirteen proteinogenic amino acids: alanine, arginine, asparagine, glutamine, glutamic acid, histidine, leucin, lysine, phenylalanine, proline, threonine, tyrosine and valine, are observed the antibatic dependence of their percentage [13] content from the dipole moment. The remaining amino acids (red triangles on the chart) are eliminated from the observed correlation, which may be caused by additional steric effects (Figure 2).

Conclusion

Electronic characteristics of proteogenic amino acids: charges on carbonyl carbon atom qC_3, amine nitrogen atom qN_6 and hydroxyl oxygen atom qO_2, as well as the orders of CO and NH bonds P_{CO} and P_{NH} and dipole moments μ are calculated using techniques of DFT. The qualitative correlation between the polarity (dipole moment) and the percentage content of amino acids in proteins was found. Using the numerical values of these characteristics, the row of the amino acids was built in the direction of their decrease. Among them, we can single out (a) and (b) series that are similar in composition and structure, which can be explained by the main contribution of C_3 and N_6 charges in the formation of the peptide bond [14,15].

Acknowledgement

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

Conflict of Interest

The authors declare that they have no conflict of interest.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent

Informed consent was obtained from all individual participants included in the study.

References

1. Chapeville F, Heanni AL (1974) Biosynthsis of Proteins, Hermann Collection Paris Methodes.
2. Dwyer D (2005) Electronic properties of amino acid side chains: quantum mechanics calculation of substituent effects. BMC Chem Biol 5: 2.

3. Grantham R (1974) Amino acid difference formula to help explain protein evolution. Science 185: 862-864.

4. Kolaskar AS, Ramabrahmam V (1982) Obligatory amino acids in primitive proteins. Biosystems 15: 105-109.

5. Taylor WR (1986) The classification of amino acid conservation. J Theor Biol 119: 205-218.

6. Mitiko Go, Sanzo Miyazawa (2009) Relationship between mutability, polarity and exteriority of amino acid residues in protein evolution. Int J Peptide Res 15: 211-224.

7. Mathura VS, Kolippkkam D (2005) APDbase: Amino acid Physico-chemical properties Database. Bioinformation 1: 2-4.

8. Kohn W, Becke AD, Parr RG (1996) Density Functional Theory of electronic structure. J Phys Chem 100: 12974-12980.

9. Becke AD (1988) Density functional exchange-energy approximation with correct asymptotic behavior. Phys. Rev A, Gen Phys 38: 3098-3100.

10. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation energy formula into of the electron density a functional. Phys Rev B Condensed Matter 37: 785-789.

11. Perdew JP, Wang Y (1992) Accurate and simple analytic representation of the electron-gas correlation energy. Phys Rev B Condence Matter 45: 13244-13249.

12. Laikov DN, Ustynyuk Yu A (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.

13. Dyer KF (1971) The quit revolution: A new synthesis of biological knowledge. J Biol Edu 5: 15-24.

14. King JL, Jukes TH (1969) Non-Darwinian Evolution. Science 164: 788-798.

15. Katti MV, Sami-Subbu R, Ranjekar PK, Gupta VS (2000) Amino acid repeat patterns in protein sequences: Their diversity and structural-functional implications. Protein Sci 9: 1203-1209.