Electrical Network Modeling of Amino Acid and its Characterization

T. Roy
Institute of Radio Physics & Electronics
University of Calcutta
Kolkata, India
tanusree.rinki@gmail.com

S. Barman
Institute of Radio Physics & Electronics
University of Calcutta
Kolkata, India
barmanmandal@gmail.com

***Abstract*—Electrical Circuit Modeling of Amino acid based on their structural and chemical properties using R, L, C passive components has been reported in this paper. The time responses of the electrical circuit systems are studied to characterize the twenty amino acids and establish correspondence between biological and electrical properties.**

Keywords- amino acid; genomics; network modeling; proper system

I. INTRODUCTION

Amino acids are the primary building blocks or chemical units for proteins. The genes of DNA are responsible for the creation of amino acids [1-3]. There are twenty amino acids joined together by peptide bonds to form the basic structure of proteins present in every living cells, lack of which can cause indigestion, depression, stunted growth, neurological problems, physical disorders, etc. Amino acids are carbon compounds, contain two functional groups: an amino group (NH2) and a carboxylic acid group (COOH). A side chain (r) attached to this compound gives each amino acid a unique set of characteristics. Only Proline has a slightly different structure as the "r" group is bent into a circle to attach itself with nitrogen in place of one of the hydrogen atoms. Fig. 1 illustrates the basic structure of an amino acid.

DNA and protein structure modeling is gaining popularity in genomics research to learn their structure and functionality. Different complex network approaches have been proposed by several authors to model DNA/RNA, amino acid and protein structure [4-9]. Marshall [4], [5] has developed resistor-capacitor ladder network for DNA/RNA strings and amino acid strings. The passive analog electrical circuits for protein structure have been modeled by Sampath [6]. The secondary structure and the secondary structure linkages have been modeled using resistor, capacitor and inductor by Marshall [7]. Alfinito et al. [8] have proposed a network model to study the structural and electrical properties of protein. Vedrana et al. [9] have developed a PSpice model to study the electrical behaviour of DNA molecules. It has been understood from various literature survey that the property and structure of amino acid is very important as twenty amino acids are coded by DNA to form proteins . The structure of protein and its property is also important for designing structure based drugs and prediction of genetic disease [10].

Realising their importance in the present society, the authors in this paper have modeled each amino acid using electrical passive component R, L, C and study their time response to characterize the amino acids based on their physicochemical properties. The paper is organised into following sections: Section 2 illustrates classification of amino acids based on their properties; section 3 describes the methodology for the electrical circuit modeling of amino acids. In section 4, discussions have been made based on the results obtained and finally, conclusions are drawn in section 5.

II. CLASSIFICATION OF AMINO ACIDS

Gene is a segment of DNA that contains genetic code, translated into twenty amino acids. These amino acids are essential to form antibodies, to combat invading bacteria and viruses; they are part of the enzyme and hormonal system; carry oxygen throughout the body and are part of all muscular activity [11].

Researchers have classified the amino acids in different ways [12-18]. Amino acids are generally categorized based on the structure and property i.e. physicochemical property of their side groups. Proteins are combination of various amino acids and the properties of the side chains determine the properties of the protein. The amino and carboxyl both chemical group are bonded to the alpha carbon (C) which is common for all amino acids except Proline and the side chain (r) functional group is distinctive for all amino acid which defines a particular amino acid. Knowledge of side chain property is important for understanding the method of analysis and identification of protein. Generally the amino acids are categorized according to their hydropathy index [14], volume, physicochemical property, pK value (define the pH-dependent characteristics of a protein) and polarity of side chain as represented in Table 1.

III. ELECTRICAL CIRCUIT MODELING OF AMINO ACIDS

The method of electrical circuit modeling and characterization of amino acids is described into following steps:

Step 1: Design twenty electrical network systems by modeling each amino acid using R, L and C based on their chemical and structural property.

ISBN: 978-81-923777-9-7

Step 2: Check whether the systems are proper or not by observing their transfer function individually.

Step 3: If system is improper, using pole addition make it as proper system.

Step 4: Study the time response of the system to characterize the amino acids.

Step 5: Finally, study the correlation between the electrical and the biological properties.

As in the basic structure of amino acids (Fig. 2a) contain a central alpha-carbon to which the carboxyl group (COOH), amino group (NH_2) and the variable side chain (r) are attached, an attempt has been made to realize the electrical circuit model of amino acid based on its structure using passive component R, L and C. The generalized electrical circuit model of an amino acid is consist of three distinct circuits, two parallel circuits one of which represent the carboxyl and other represent the amino group (backbone structure) , common for all amino acids except Proline and the third circuit which is connected in series with the backbone structure represents the side chain.

The impedance of the carboxyl (COOH) and amino (NH_2) circuit model is Z_{CE} and Z_{AE} respectively and the side chain (r) impedance is $Z_{SC.}$ The alpha-carbon and hydrogen atom is taken as a node shown in Fig. 2b.

In this circuit model, the carboxyl group (COOH) is represented by a resistor (1Ω) and its impedance is given by,

$$Z_C = 1. \quad (1)$$

Similarly, the amino group is represented by an inductor L (2mH as NH_2 group has two hydrogen atom) and its impedance is given by,

$$Z_A = 2e\text{-}003s. \quad (2)$$

The electrical circuits of COOH and NH2 groups are connected in parallel. The reduced equivalent impedance for the backbone circuit is given by,

$$Z = [Z_C \parallel Z_A] = (2e\text{-}003s) / (2e\text{-}003s + 1) \quad (3)$$

But the backbone of Proline is composed of carboxyl group and the side chain bent into a cycle which attaches itself to the nitrogen in place of one of the hydrogen atoms. The impedance of COOH for Proline is same as other amino acids, but NH group is represented by an inductor L of 1mH as NH group has one hydrogen atom and its impedance is given by,

$$Z_{PA} = 1e\text{-}003s \quad (4)$$

So for Proline, the electrical circuits of COOH and NH groups are connected in series as it is of heterocyclic structure and the reduced equivalent impedance for the backbone circuit is given by,

$$Z_P = [Z_C + Z_{PA}] = (1e\text{-}003s + 1) \quad (5)$$

As the amino acid property varies with the side chain therefore instead of considering it as a group, individual atom is considered here for designing the circuit models using the passive components R, L, C based on their atomic numbers. The carbon atom ($_6$C) is modeled by six resistor R (each of 1Ω) connected in parallel, nitrogen ($_7$N) by seven inductor L (each of 1mH) connected in parallel, sulphur ($_{16}$S) by sixteen inductor L (each of 1mH) connected in parallel, oxygen ($_8$O) by eight capacitor C (each of 1μF) connected in series and hydrogen ($_1$H) by one resistor R (0.2Ω). The side chain consists of more number of hydrogen atoms compared to others atoms. In order to reduce circuit complexity less value of resistor is chosen for hydrogen atoms to model the amino acids. The electrical circuits for the atoms are shown in Fig. 3a-3e. The equivalent impedance Z_{eq} of 20 amino acids is different and computed using (3) and except for Proline (5) is used. Due to space constraint some of the electrical circuit realization based on amino acid structure is depicted in Table 2.

Each amino acid model is considered as individual system and checked whether the system is proper or not by observing the transfer function (TF). The systems are changed into proper system by pole addition method. Hence, a first order system is cascaded with the each system to make the system proper. The summarized details of amino acid transfer function model are shown in Table 3.

After converting the improper system to proper system, study the correlation between electrical and biological properties by observing their time response plots.

IV. Results and Discussion

Each amino acid circuit is treated as an individual system and the step response of the system for time range (0-12) sec has been observed to study the circuit model behavior. Here the amino acids are characterized in terms of hydropathy index, volume and pK value. The system response plots shown in Figs. 4a–4e characterize the amino acids as hydrophobic, hydrophilic, acidic, basic, polar, nonpolar, large, small etc. The aromatic (ring type structure), hydrophobic (water fearing amino acid), hydrophilic (water loving amino acid), polar (form hydrogen bond with water) and nonpolar (cannot form hydrogen bond with water) amino acids are classified based on their hydropathy index value [14]. Whereas the acidic (negatively charged) and basic (positively charged) amino acids are classified with respect to their pK (define the pH-dependent characteristics of a protein) value. The large, medium and small amino acids are distinguished according to their volume.

In Fig. 4a, the step response of the nonpolar amino acid is larger than the polar amino acid which correlates that the nonpolar amino acids have greater hydropathy index values than the polar amino acids [14]. Fig. 4b exhibits the time responses of the hydrophobic amino acid with larger amplitude values and the hydrophilic amino acid with smaller amplitude values which is analogous to the hydropathy index values of the hydrophobic and hydrophilic amino acids. Fig. 4c differentiates the aromatic amino acids (W, F, Y and H)

ISBN: 978-81-923777-9-7

according to their hydrophobicity of side chains [18] which also correlates that the aromatic amino acid tryptophan and Phenylalanine have very hydrophobic aromatic side chains, whereas tyrosine and histidine have less hydropathy index values. The step responses in Fig. 4d shows the basic (Histidine, Lysine and Arginine) amino acids have higher amplitude values whereas the acidic (Aspartic acid and Glutamic acid) amino acids have smaller values for time span (0-12) sec which satisfy the chemical property i.e. pK values of the basic amino acid is greater than the pK values of acidic amino acid. According to the volume, very large, large, medium, small and very small amino acids are distinguished as shown in Fig. 4e. Thus the electrical responses of circuit models are truly correlate the biological properties.

V. CONCLUSION

The electrical circuit model realization of amino acids and investigation of system responses for their characterization play a significant role in the present era of genomic research. The time response of the systems truly reflects the biological properties. This circuit model concept may help in structure based drug designing and also be useful in VLSI technology to develop electronic and bio-sensing devices. It is increasingly important for engineers and scientist to process the biological data which is available in public domain [19] in a way that will be helpful for humankind.

ACKNOWLEDGMENT

The first author wishes to thank the University Grant Commission (UGC), New Delhi for providing fellowship.

REFERENCES

[1] P. P. Vaidyanathan, "Genomics and proteomics: A signal processor's tour," IEEE Circuits Syst. Mag., vol. 4, no. 4, pp. 6–29, April 2004.

[2] J. Koolman, K. H. Rohm, Colour Atlas of Biochemistry, 1st ed, Thieme, Stuttgart, 1996.

[3] The AP john cancer institute website. [Online]. Available: https://www.apjohncancerinstitute.org/caat-protocol

[4] R. Marshall, "Modeling DNA/RNA Strings Using Resistor–Capacitor (RC) Ladder Networks," The Computer Journal, vol. 53 no. 6, pp. 644-660, 2010.

[5] R. Marshall, "Modeling amino acid strings using electrical ladder circuits," Nature & Biologically Inspired Computing., IEEE, pp. 1580-1583, Dec. 2009.

[6] G. Sampath, "RLC(M) circuit models of protein structure: analysis, visualization, shape synthesis, and pattern matching," Proceedings of the 40th Conf. on Information Science and Systems (CISS 06), Princeton University, Princeton, NJ, USA, pp. 1623-1628, March 2006.

[7] R. Marshall, "Secondary structures and secondary structure linkages of protein sequences," Systems Man and Cybernetics (SMC), IEEE, pp. 2995-3000, Oct. 2010.

[8] E. Alfinito, C. Pennetta, and L. Reggiani, "A network model to investigate structural and electrical properties of proteins," Mar. 2007.

[9] V. Hodzic, V. Hodzic, and R. W. Newcomb, "Modeling of the Electrical Conductivity of DNA," IEEE Transactions on Circuits and Systems, vol. 54, no. 11, Nov. 2007.

[10] S. Mitra, S. Das, P. Pal Choudhuri, and S. Nandi . "Archtecture of a VLSI chip for modelling Amino acid sequence in Proteins", 9th International Conference on VLSI Design, Jan. 1996.

[11] S. Barman, S. Saha, A. Mondal, and M. Roy, "Signal Processing Techniques for the Analysis of Human Genome Associated with Cancer Cells," 2nd Annual international Conference IEMCON, Jan. 2011.

[12] K. C. Timberlake, Chemistry, 5th ed, Haper-Collins Publishers Inc, NY, 1992.

[13] T. M. Devlin, The Textbook of Biochemistry, 3rd ed, Wiley-Liss Inc, NY, 1992.

[14] J. Kyte, and R. F. Doolittle, "A simple method for displaying the hydropathic character of a protein," J. Mol. Biol. 157, pp. 105–132, 1982.

[15] A. A. Zamyatnin, "Protein volume in solution," Prog. Biophys. Mol. Biol., 24, pp. 107-123, 1972.

[16] The iitkgp website. [Online]. Available: http://www.chem.iitkgp.ernet.in/faculty/SDG/ProteinStructure.pdf.

[17] C. Pommie, et al., "IMGT standardized criteria for statistical analysis of immunoglobulin V-REGION amino acid properties", J. Mol. Recognit., 17, pp. 17-32, 2004.

[18] M. R. Barnes, Bioinformatics for Geneticists, 2nd ed, 2007, pp. 311–342.

[19] The NCBI website. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/15032603.

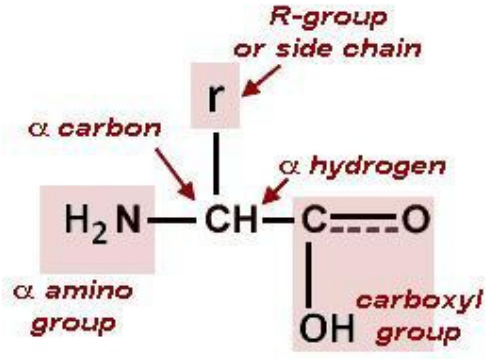


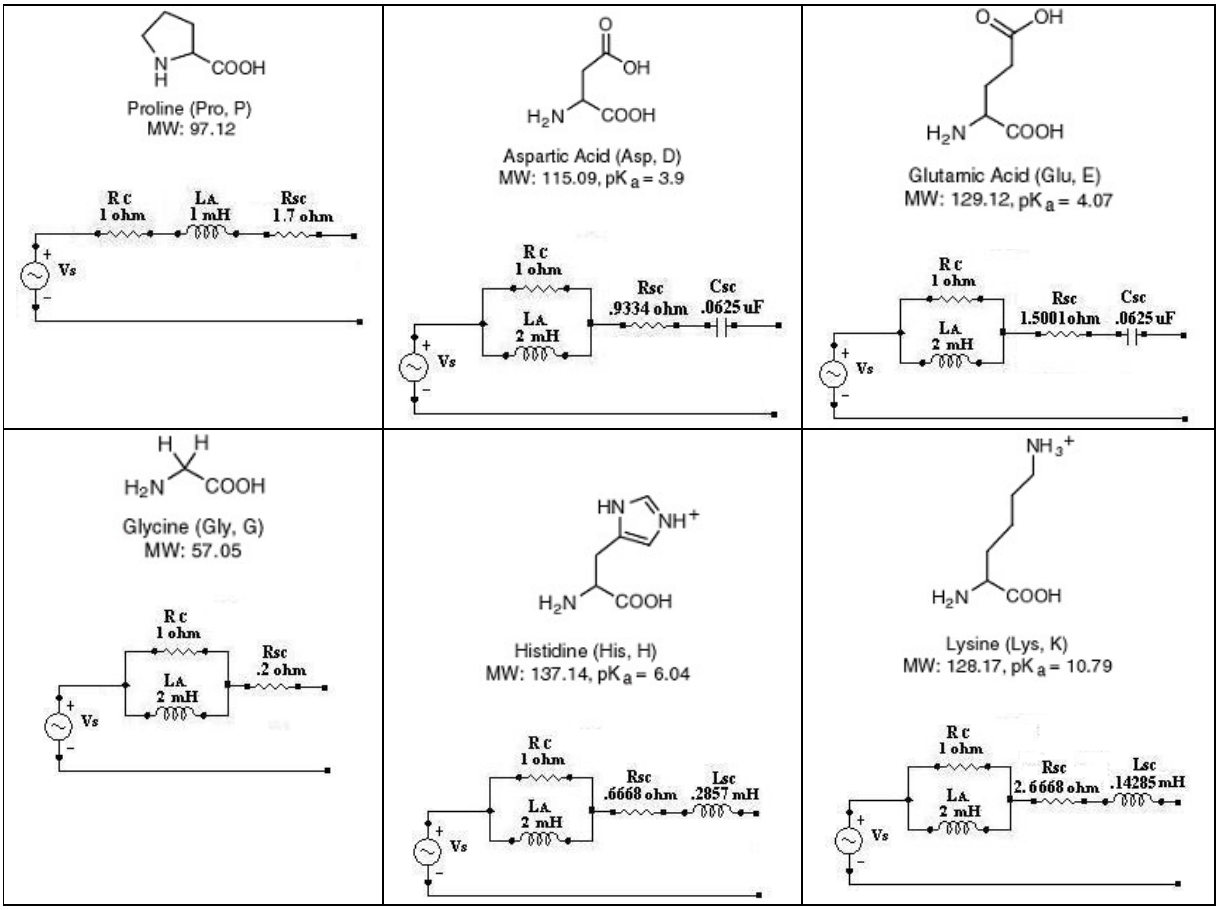
Figure 1. Basic structure of amino acid

ISBN: 978-81-923777-9-7

TABLE I. AMINO ACIDS AND THEIR CLASSIFICATIONS

Amino Acid	Abbreviations		Hydropathy	Volume	Chemical	Charge and pK of side chain (r)	Polarity
Alanine	Ala	A	hydrophobic	very small	aliphatic	neutral	nonpolar
Arginine	Arg	R	hydrophilic	large	basic	positive (12.5)	polar
Asparagine	Asn	N	hydrophilic	small	amide	neutral	polar
Aspartic acid	Asp	D	hydrophilic	small	acidic	negative (3.9)	polar
Cysteine	Cys	C	hydrophobic	small	sulfur	neutral	nonpolar
Glutamine	Gln	Q	hydrophilic	medium	amide	neutral	polar
Glutamic acid	Glu	E	hydrophilic	medium	acidic	negative (4.2)	polar
Glycine	Gly	G	hydrophilic	very small	aliphatic	neutral	polar
Histidine	His	H	hydrophilic	medium	basic	positive (6.0)	polar
Isoleucine	Ile	I	hydrophobic	large	aliphatic	neutral	nonpolar
Leucine	Leu	L	hydrophobic	large	aliphatic	neutral	nonpolar
Lysine	Lys	K	hydrophilic	large	basic	positive (10.5)	polar
Methionine	Met	M	hydrophobic	large	sulfur	neutral	nonpolar
Phenylalanine	Phe	F	hydrophobic	very large	aromatic	neutral	nonpolar
Proline	Pro	P	hydrophobic	small	aliphatic	neutral	nonpolar
Serine	Ser	S	hydrophilic	very small	hydroxyl	neutral	polar
Threonine	Thr	T	hydrophilic	small	hydroxyl	neutral	polar
Tryptophan	Trp	W	hydrophobic	very large	aromatic	neutral	nonpolar
Tyrosine	Tyr	Y	hydrophilic	very large	aromatic	neutral	polar
Valine	Val	V	hydrophobic	Medium	aliphatic	neutral	nonpolar

TABLE II. CHEMICAL STRUCTURE AND ELECTRICAL CIRCUIT OF AMINO ACID



ISBN: 978-81-923777-9-7

TABLE III. SYSTEM TRANSFER FUNCTION (TF) OF AMINO ACIDS WITHOUT/WITH POLE ADDITION

Amino Acid	Transfer function model (TF) without pole addition	Transfer function model (TF_p) with pole s=-1 addition
A	$\frac{0.003533\ s + 0.7667}{0.002\ s + 1}$	$\frac{0.003533\ s + 0.7667}{0.002\ s^2 + 1.002\ s + 1}$
R	$\frac{1.224e-010\ s^3 + 2.156e-006\ s^2 + 0.004602\ s + 1.32}{5.714e-007\ s^2 + 0.001486\ s + 0.61}$	$\frac{1.224e-010\ s^3 + 2.156e-006\ s^2 + 0.004602\ s + 1.32}{5.714e-007\ s^3 + 0.001487\ s^2 + 0.6015\ s + 0.6}$
N	$\frac{2.857e-007\ s^3 + 0.00441\ s^2 + 1.133\ s + 8e-00.32}{0.002\ s^2 + s}$	$\frac{2.857e-007\ s^3 + 0.00441\ s^2 + 1.133\ s + 8e-00.32}{0.002\ s^3 + 1.002\ s^2 + s}$
D	$\frac{0.003867\ s^2 + 0.9334\ s + 1.6e-002}{0.002\ s^2 + s}$	$\frac{0.003867\ s^2 + 0.9334\ s + 1.6e-002}{0.002\ s^3 + 1.002\ s^2 + s}$
C	$\frac{1.25e-007\ s^2 + 0.003596\ s + 0.76}{0.002\ s + 1}$	$\frac{1.25e-007\ s^2 + 0.003596\ s + 0.76}{0.002\ s^2 + 1.002\ s + 1}$
Q	$\frac{2.857e-007\ s^3 + 0.005543\ s^2 + 1.7\ s + 8e-006}{0.002\ s^2 + s}$	$\frac{2.857e-007\ s^3 + 0.005543\ s^2 + 1.7\ s + 8e-006}{0.002\ s^3 + 1.002\ s^2 + s}$
E	$\frac{0.005\ s^2 + 1.5\ s + 1.6e-005}{0.002\ s^2 + s}$	$\frac{0.005\ s^2 + 1.5\ s + 1.6e-005}{0.002\ s^3 + 1.002\ s^2 + s}$
G	$\frac{0.0024\ s + 0.22}{0.002\ s + 1}$	$\frac{0.0024\ s + 0.22}{0.002\ s^2 + 1.002\ s + 1}$
H	$\frac{5.714e-007\ s^2 + 0.005619\ s + 1.667}{0.002\ s + 1}$	$\frac{5.714e-007\ s^2 + 0.005619\ s + 1.667}{0.002\ s^2 + 1.002\ s + 1}$
I	$\frac{0.003687\ s + 0.8435}{0.002\ s + 1}$	$\frac{0.003687\ s + 0.8435}{0.002\ s^2 + 1.002\ s + 1}$
L	$\frac{0.004634\ s + 1.317}{0.002\ s + 1}$	$\frac{0.004634\ s + 1.317}{0.002\ s^2 + 1.002\ s + 1}$
K	$\frac{2.857e-007\ s^2 + 0.007479\ s + 2.668}{0.002\ s + 1}$	$\frac{2.857e-007\ s^2 + 0.007476\ s + 2.667}{0.002\ s^2 + 1.002\ s + 1}$
M	$\frac{1.25e-007\ s^2 + 0.005863\ s + 1.9}{0.002\ s + 1}$	$\frac{1.25e-007\ s^2 + 0.005863\ s + 1.9}{0.002\ s^2 + 1.002\ s + 1}$
F	$\frac{0.007134\ s + 2.567}{0.002\ s + 1}$	$\frac{0.007134\ s + 2.567}{0.002\ s^2 + 1.002\ s + 1}$
P	$0.001\ s + 2.7$	$\frac{0.001\ s + 2.7}{s + 1}$
S	$\frac{0.003533\ s^2 + 0.7667\ s + 8e-006}{0.002\ s^2 + s}$	$\frac{0.003533\ s^2 + 0.7667\ s + 8e-006}{0.002\ s^3 + 1.002\ s^2 + s}$
T	$\frac{0.003042\ s^2 + 0.5545\ s + 9.067e-006}{0.001933\ s^2 + 0.9667\ s + 8e-006}$	$\frac{0.003042\ s^2 + 0.5545\ s + 9.067e-006}{0.001933\ s^3 + 0.9686\ s^2 + 0.9667\ s + 8e-006}$
W	$\frac{2.857e-007\ s^2 + 0.008343\ s + 3.1}{0.002\ s + 1}$	$\frac{2.857e-007\ s^2 + 0.008343\ s + 3.1}{0.002\ s^2 + 1.002\ s + 1}$
Y	$\frac{0.007134\ s^2 + 2.567\ s + 8e-006}{0.002\ s^2 + s}$	$\frac{0.007134\ s^2 + 2.567\ s + 8e-006}{0.002\ s^3 + 1.002\ s^2 + s}$
V	$\frac{0.0035\ s + 0.75}{0.002\ s + 1}$	$\frac{0.0035\ s + 0.75}{0.002\ s^2 + 1.002\ s + 1}$

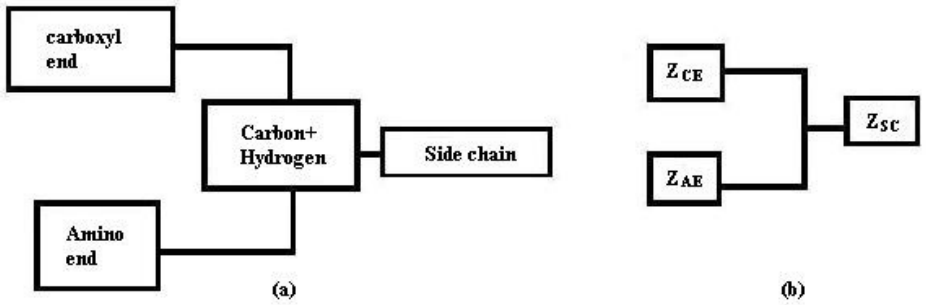


Figure 2. Basic structure and electrical equivalent model of an amino acid

ISBN: 978-81-923777-9-7

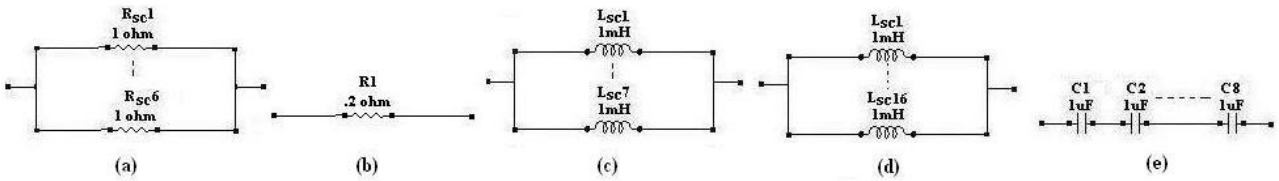


Figure 3. Circuit model for atoms (a) carbon, (b) hydrogen, (c) nitrogen, (d) sulphur and (e) oxygen

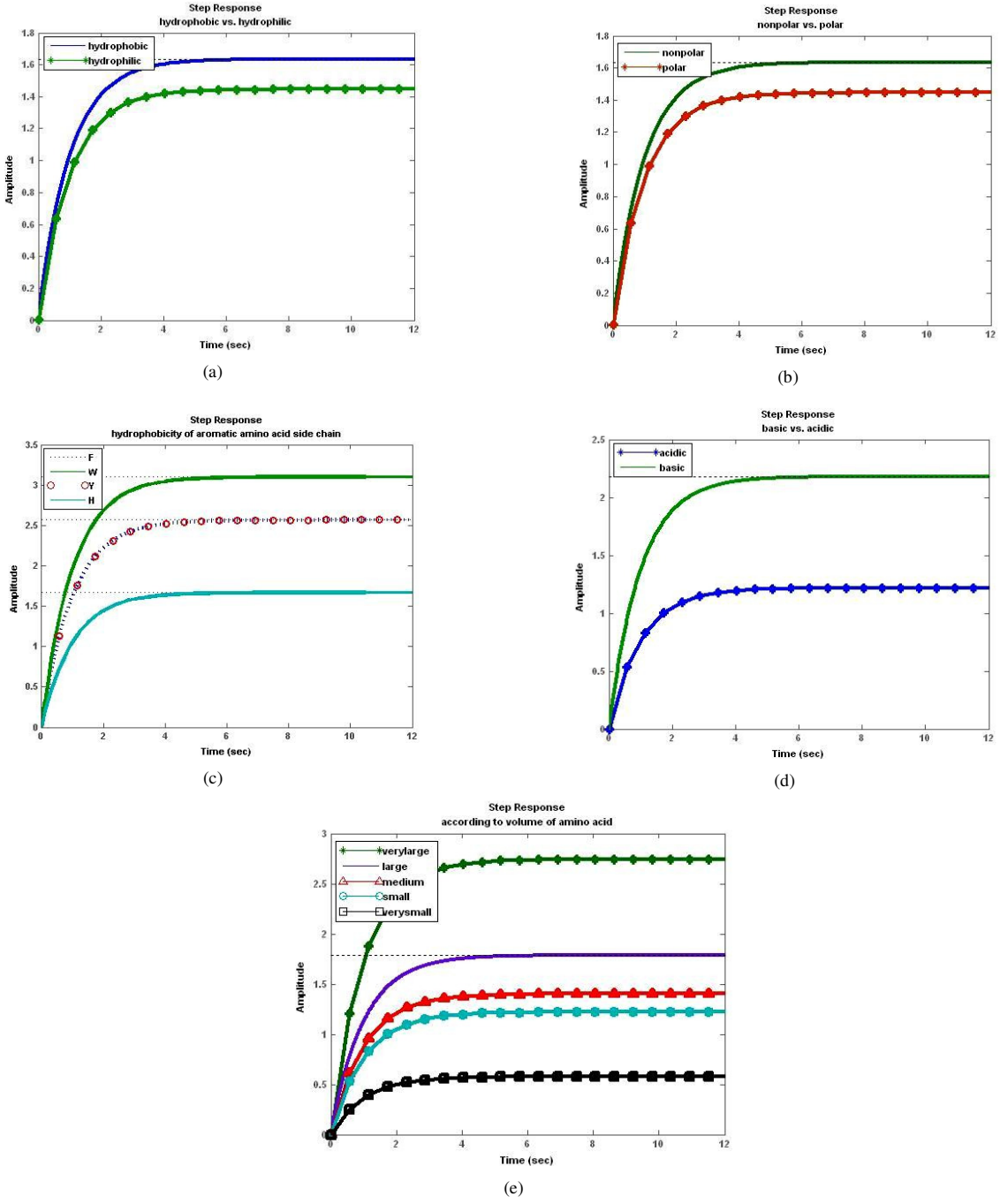


Figure 4. Step response of (a) hydrophobic vs. hydrophilic amino acid, (b) nonpolar vs. polar amino acid, (c) aromatic amino acid according to hydrophobicity of side chain (d) acidic vs. basic amino acid, and (e) all amino acid according to volume

ISBN: 978-81-923777-9-7