Codon Characterization Based on Electrical Response

Monalisa Dutta
Institute of Radio Physics and Electronics
University of Calcutta
92 A.P.C. Road, Kolkata-700009, India

Soma Barman
Institute of Radio Physics and Electronics
University of Calcutta
92 A.P.C. Road, Kolkata-700009, India

ABSTRACT
Electrical circuit modeling of biological system is a significant research topic in recent years. Codon is an important element in DNA sequence which is responsible for formation of protein in genes. 64 codons of genes are modeled using MOSFET from elementary level and characterized them based on purine and pyrimidine property. The model is realized in Spice domain and a new genetic code classification table is described based on simulated transient voltage and current responses of purine and pyrimidine bases.The effect of base position is investigated using frequency domain analysis of voltage and phase characteristics of codon electrical circuit.

Keywords
Codon, DNA, nucleotides electrical model, MOSFET.

1. INTRODUCTION
One of the most important scientific discoveries of the 20th century by Crick and Watson is the double-helix structure of DNA (Deoxyribonucleic acid) and the mechanism of base pairing by which genetic information is stored in living organisms [1]. DNA is blueprint of every living organism [2]. The DNA is string of genetic code (or codon). George Gamow,1953, was first proposed that the sets of three bases must be employed to encode the 20 standard amino acids used by living cells to build proteins [3]. From several decades, the researchers from different fields show their interest on study of electrical response of DNA, theoretically as well as experimentally [4]. Many researchers are studied charge migration through DNA molecules [5, 6] and the electrical conductivity of DNA molecules for its possible application in molecular electronics in different ways [7-11] and created a revolution in traditional pharmaceutical industries [4]. In order to study the electrical characteristic theoretically, different type of electrical equivalent circuit of DNA molecules were modeled by many researchers from different fields.DNA consists of a long chain of codon and each codon of DNA was modeled as a R-C ladder circuit to study its nature [11].DNA also was modeled as L-C oscillator based on chemical structure of DNA molecules [12]. The PSpice model of electrical behavior of DNA used in nanoelectronics was described by Vedrana Hodzic and also studied the nonlinear I-V characteristic [13]. Researchers also proposed diode like behavior of codons [14].H- Bond has capacitive property proposed by many researchers [15].

The equivalent electrical circuits using passive circuit components become very bulky and consume high power compared to active circuit components. The equivalentelectrical circuit design using MOSFET is a challenging research for molecular device designers [17]. Inpresent paper the authors have described equivalent electrical circuit realization of each codon using MOSFET and study their transient and frequency response characteristic. Based on transient responses, the codons are grouped and a new genetic code table is designed. The

frequency response of the equivalent electrical circuit of codon is analyzed to study the effect of position of nucleotide base in codon.

The rest of the paper is organized as background of DNA, methods of equivalent electrical circuit modeling of codon, time and frequency domain response of codon equivalent electrical circuit and conclusion.

2. BACKGROUND
Deoxyribonucleic acid, commonly known as DNA,has two strands, running in opposite directions. The backbone of strands made up by sugar phosphate.Since each strand of a DNA is represented by an alphabet of four elements, named as adenine (A), guanine (G), cytosine (C) and thymine (T)(Fig. 1a)[18].A small portion of single strands given below.

ATCCAGCTAGGGCAAGTAGGCAAATATCATGATAGG

The nucleotide bases C and T are known as pyrimidine (aromatic heterocyclic organic compound) and A and G are known as purine(heterocyclic aromatic organic compound and It consists of two pyrimidine ring).The nucleotide A, G, C and T of a single strand connected with T, C, G and A respectively by hydrogen bonds.There are two and three hydrogen bonds between the base pair A, T and pair C, Gshown in Figure1b.Three adjacent nucleotides form a codon (Figure 2).This provides genetic code information for a particular amino acid. Different combination of four bases form 64 codons.The 64 codons and their corresponding amino acids shown in table I[19].
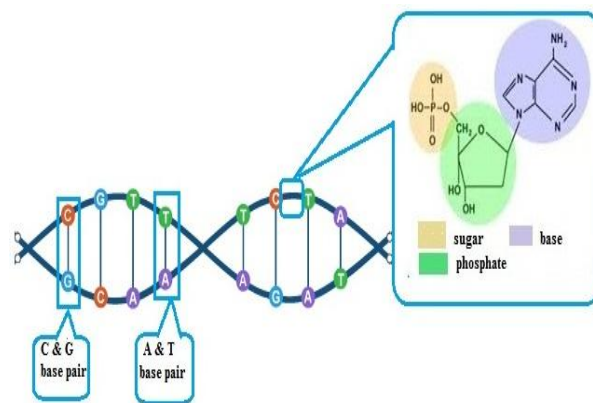
Fig1: (a)Double helix structure of DNA. A small portion of sugar phosphate bond and A-T and G-C base pair are highlighted in box
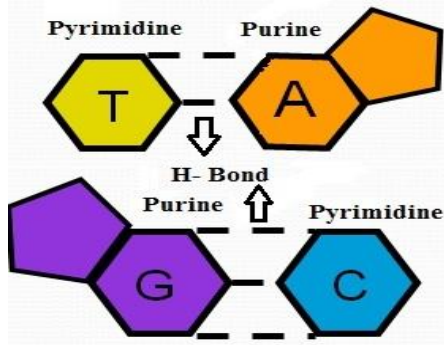
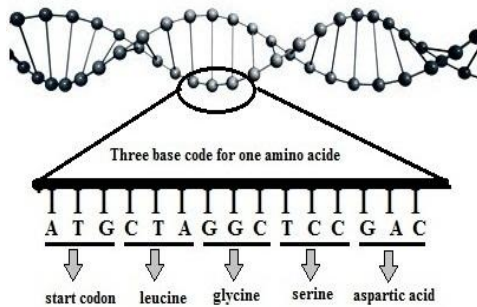**Fig1: (b) The detailed structure of base pair of A, T, C, G.**

Fig2: A model of mapping of three adjacent nucleotides intoan amino acid.

At first, the authors are realized a model of each nucleotide A, G, C, T of DNA using MOSFET and cascading them to form an equivalent electrical model of codons. Designingthe electrical model of codon, consideration is given on the number of benzene rings present in pyrimidine and purine structure, physiochemical property of hydrogen bond and sugar phosphate bond of each nucleotide base.The main objectives of the paper are:

1. Realized equivalent electrical model of 64 codons.
2. Study of the transient and frequency characteristic of 64 codon electrical circuit in Spice domain.
3. A new genetic code table is designed based on transient behavior of codon electrical model.
4. Analyze the effect of base position in codon based on frequency characteristic.

3. METHODS

For designing 64 codons electrical equivalent circuitusing MOSFET, pyrimidine and purine are considered as a resistor because of their constant ring characteristic[11]. The net atomic charge in H bond produces a screening effect or atomic shielding which is viewed as a capacitor[15]. Intermolecular distance between two consecutive base pair is much shorter, so hopping mechanism is considered for charge transfer from one molecule to another. This hopping phenomenon is equivalently representedby diode [14]. The schematic is realized using Tanner S-Edit tool (version 16.0).

3.1 Design specification for equivalentresistors, capacitor and diode

For resistor modeling,values are taken 100 ohms and 200 ohms for pyramidine (one ring) and purine (two ring) base respectively. In triode region the drain current (I_D) is linear function of drain to source voltage (V_{DS}), and in this region

MOSFET behaves as a resistor [20,21], the equation for drain current is shown below:

$$I_D = K_p \frac{W_{eff}}{L_{eff}}[(V_{GS} - V_{TN})V_{DS} - \frac{V_{DS}^2}{2}](1+\lambda V_{DS}) \qquad (1)$$

For triode region $V_{DS} < (V_{GS}-V_{TN})$ and $V_{GS} \geq V_{TN}$.

The threshold voltage V_{TN} is calculated as

$$V_{TN} = V_{TO} + \gamma(\sqrt{\left|2\phi_F\right| + V_{SB}} - \sqrt{\left|2\phi_F\right|}) \qquad (2)$$

Where K_p, V_{TO}, γ, Φ, λ, t_{ox}, W_{eff}, L_{eff}, V_{TN}, $(V_{GS}-V_{TN})$are represent transconductance parameter, zero bias threshold voltage, body effect parameter, surface inversion potential($2\Phi_F$),channel length modulation,oxide thickness, effective channel width,effective channel length,thresholdvoltage,overdrive voltage respectively.

- $W/L=15\mu/0.60\mu$, $V_{GS}=2.67$volt for $R_{ON} = 100$ ohms
- $W/L=15\mu/0.60\mu$, $V_{GS}=1.68$volt for $R_{ON} = 200$ ohms.

Adenine (A) and Thymine (T) have two hydrogen bonds and Guanine (G) and Cytosine (C) have three hydrogen bonds,so 20fF and 30fF capacitor is considered respectively.To model bias-independent overlap capacitor we consider equations (3).

$$C_{mos} = WL_{eff}C_{OX} + 2WL_{OV}C_{OX} \qquad (3)$$

Where, $V_{SG} > V_{TP}$ and L_{ov}, $L_{eff}$are represent overlapping length, effective channel length respectively.

- $W/L=5.1364\mu/0.60\mu$ for $C_{MOS} = 20$fF.
- $W/L=7.704\mu/0.60\mu$ for $C_{MOS} = 30$fF.

In order to represent the charge hoppingphenomena between two adjacent nucleotides,the authors short the gate and drain terminal whichbehaves like a diode. Based on the specifications, the equivalentelectrical circuit of 64 codons are designed in the paper. Out of the 64 electrical equivalent circuits of codonsonlyGTA is shown in figure 3.

4. RESULT AND DISCUSSION

Here a new genetic table is designed based on electrical responses of codon circuit. The electrical codon circuits are excited by an input voltage 900 mV and frequency 100 kHz. In table II, the authors grouped 64 codons based on transient voltage and current response which correlates the ring structure of nucleotide. The effect of base position in codon is studied by observing frequency response of all the codon electrical circuit.

The authors observedall the codons belong to the group '2**', show a stable nature in voltage as well as stable nature in current.Butthe codons belong to one group differin voltage and currentvalues from other groups. For example the triplets which are from the group '211'(such as GCC, GCT, ACC, ACT, GTC, GTT, ATC, ATT) has uniformly same current and voltage values, but they are different from other groups i.e. '212'(such as GCA, GCG etc), '221'(such as GGC, GGT etc), '222'(such as GGG,GGA etc)(Figure 3a, b).
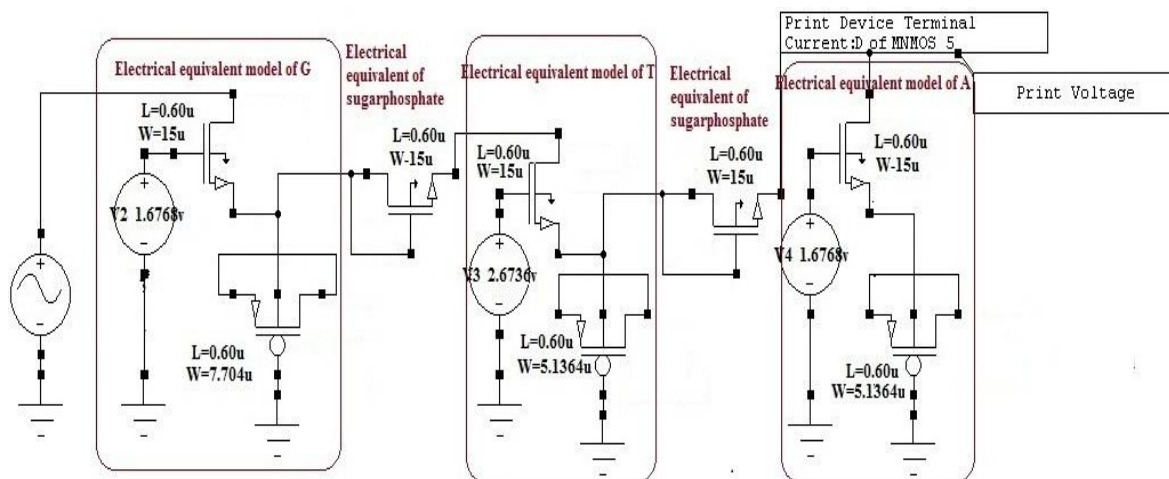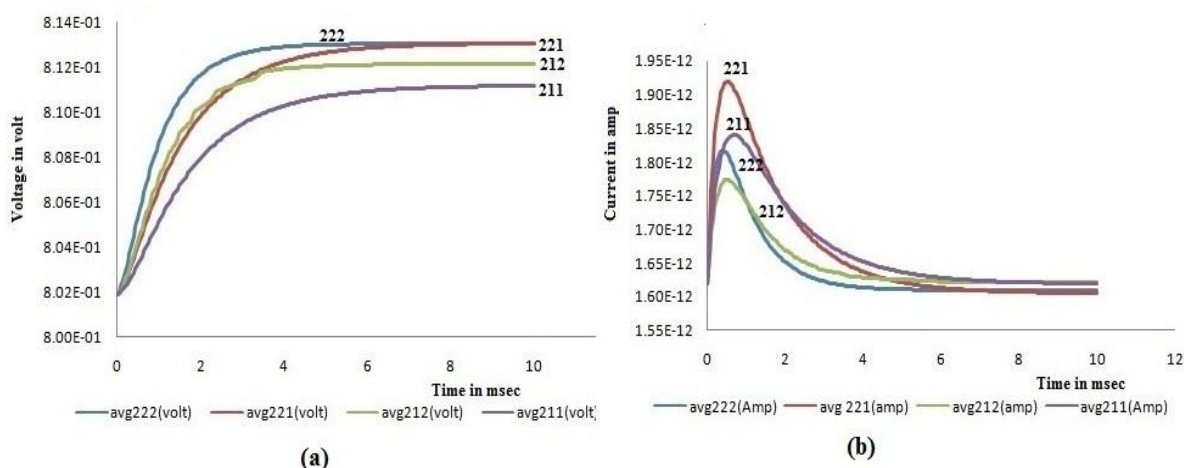
Fig 3: The electrical equivalent circuit of GTA.

(a)

(b)

Same phenomena is observed for the codon under the group '1**' theyexhibit same nature in electrical behavior. For example all the codons from the group '111'(such as CCC, CCT, TCC, TTT, CTC,CTT, TTC, TTT) has uniformly same current or voltage values with respect to time, but they are different in value from other groups such as '112'(such as CCG,CCA etc.), '121'(such as CGG,CGT etc), '122'(such as CGG,CGA etc). All '1**'groups show stable nature in voltage but oscillatory nature in current. (Figure 4c, 4d).

The transient voltage response show exponentially increasing in nature for codon whose first base is purine. Whereas the codon start with pyrimidine base exhibitsexponentially decreasing voltage response.

Transient currentresponse oscillatory in nature for codon begins with Pyrimidine base, whereas the codon start with purine base is stable. Therefore, transient characteristics of codon equivalent circuit can easily classify the total 64 codons into two groups based on transient voltage and current

response. For voltage response either exponentially increasing or decreasing in nature and for current response either stable or unstable in nature.

Based on transient response of codon equivalent circuit, a new genetic table is proposed by authors which is depicted in Table II and III. The authors broadly divided genetic table into two group, one ring and tworings, each group divided into four subgroups i.e. 11, 12, 21, 22 shown in table II.The dimension of original genetic code table is 4×4 and based on electrical response which is transformed into 2×4 genetic code table.

The frequency domain response is analyzed to examine the effect of base position in codon. The codon which starts with pyrimidine are very weak in nature but purines are very strong base, it is also observed that the effect of third base position is negligible in codon. Figure 5 represents a comparisons between the group 111 and112(change in third base position)
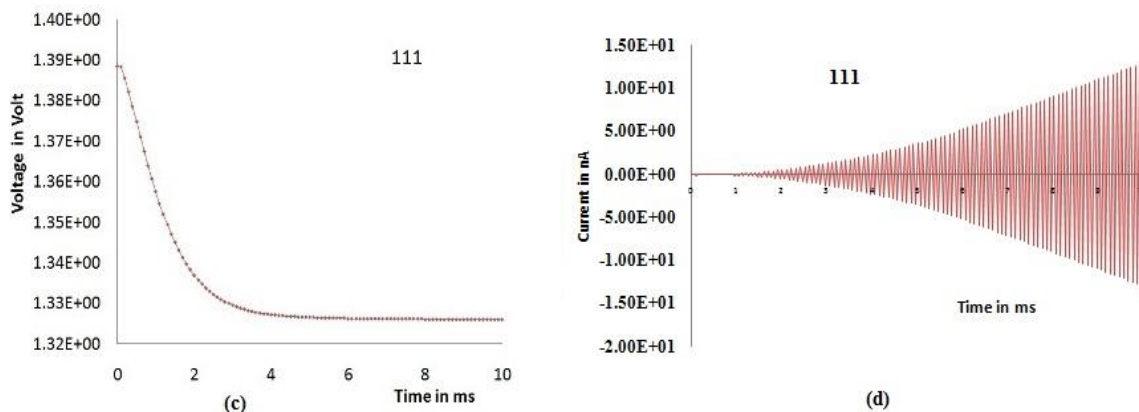
Fig 4: (c), (d) Shows transient characteristic of 111groups

**TABLE I. The standard genetic code and corresponding amino acids**

		2nd Base									
		T		C		A		G			
1st Base	T	TTT	Phe	TCT	Ser	TAT	Tyr	TGT	Cys	T	3rd Base
		TTC		TCC		TAC		TGC		C	
		TTA	Leu	TCA		TAA	Stop	TGA	Stop	A	
		TTG		TCG		TAG		TGG	Trp	G	
	C	CTT	Leu	CCT	Pro	CAT	His	CGT	Arg	T	
		CTC		CCC		CAC		CGC		C	
		CTA		CCA		CAA	Gln	CGA		A	
		CTG		CCG		CAG		CGG		G	
	A	ATT	Ile	ACT	Thr	AAT	Asn	AGT	Ser	T	
		ATC		ACC		AAC		AGC		C	
		ATA		ACA		AAA	Lys	AGA	Arg	A	
		ATG	Met	ACG		AAG		AGG		G	
	G	GTT	Val	GCT	Ala	GAT	Asp	GGT	Gly	T	
		GTC		GCC		GAC		GGC		C	
		GTA		GCA		GAA	Glu	GGA		A	
		GTG		GCG		GAG		GGG		G	

**Table 2. New genetic code table based on transient response of electrical equivalent circuits of codons**

	The number of ring present in 2nd and 3rd base position			
The Number of ring present in 1st base position	11	12	21	22
1	111	112	121	122
2	211	212	221	222

Table 3. Illustrated view of proposed table

| 111	CC(C/T)	TC(C/T)	CT(C/T)	TT(C/T)
112	CC(G/A)	TC(G/A)	CT(G/A)	TT(G/A)
121	CG(C/T)	TG(C/T)	CA(C/T)	TA(C/T)
122	CG(G/A)	TG(G/A)	CA(G/A)	TA(G/A)
211	GC(C/T)	AC(C/T)	GT(C/T)	AT(C/T)
212	GC(A/G)	AC(G/A)	GT(G/A)	AT(G/A)
221	GG(C/T)	AG(C/T)	GA(C/T)	AA(C/T)
222	GG(G/A)	AG(G/A)	GA(G/A)	AA(G/A)
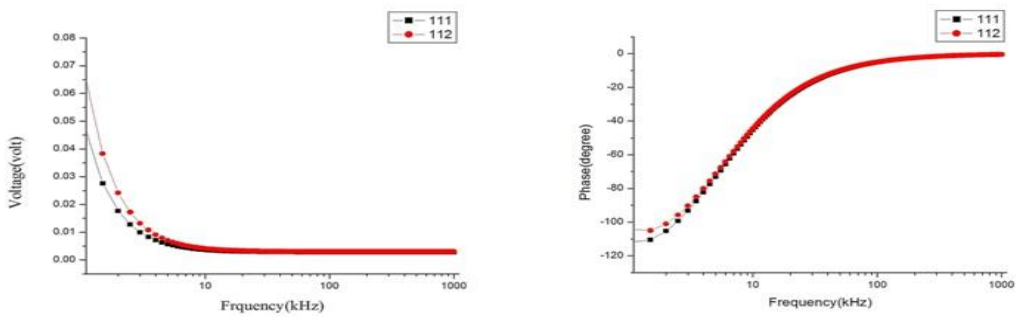
Fig 5: The average amplitude and phase response for the codons of the group 111 and 112
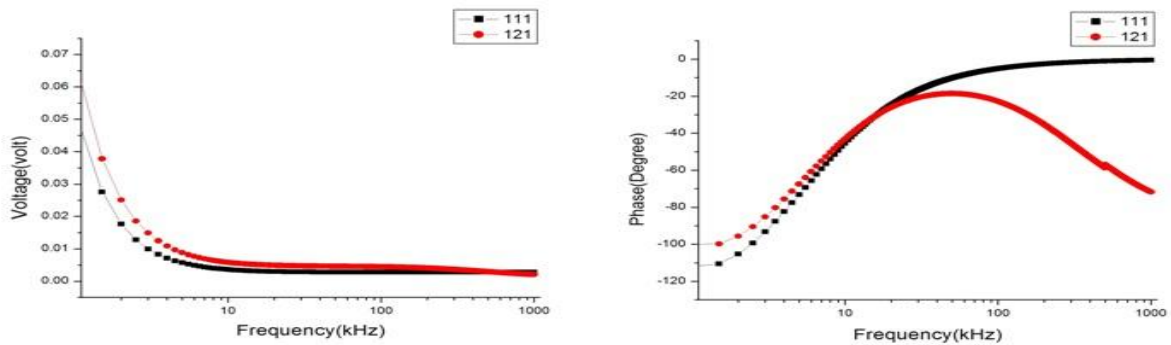
Fig 6: The average amplitude and phase response for the codons of the group 121 and 121.
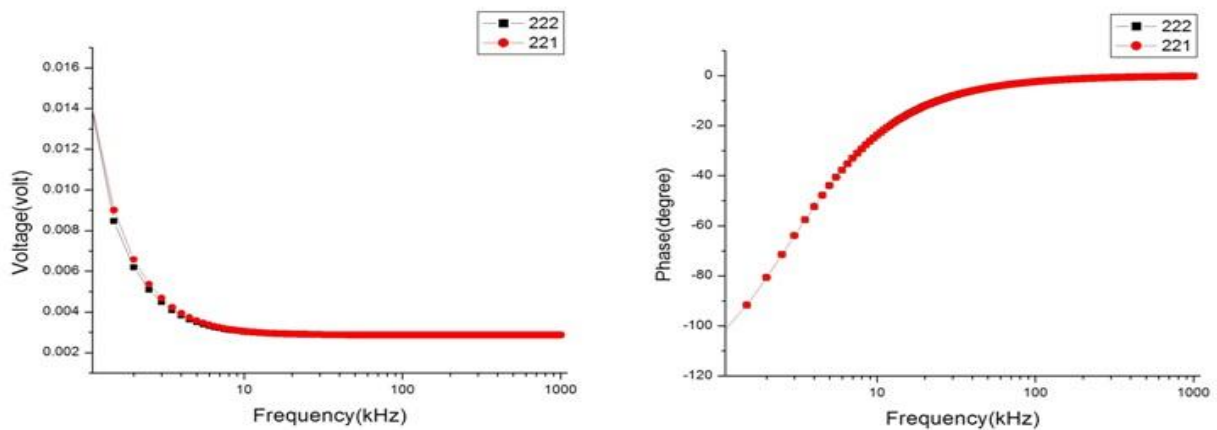
Fig 7: The average amplitude and phase response for the codons of the group 222 and 221.
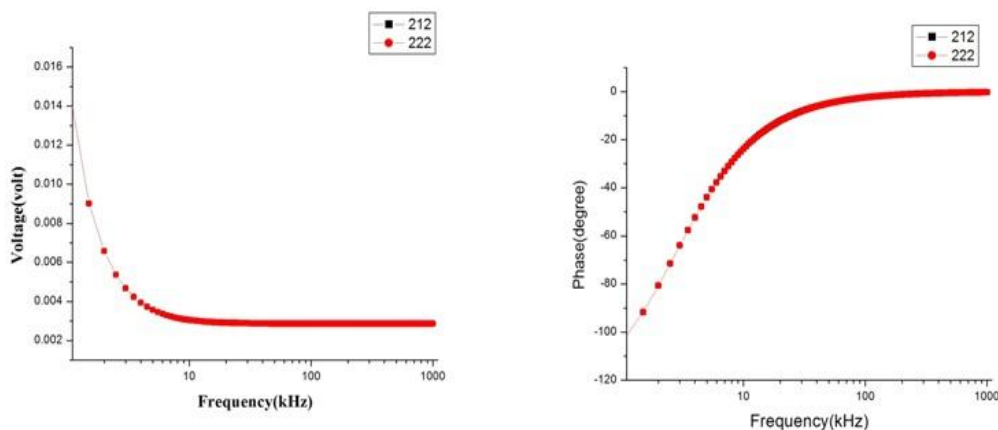
Fig 8: The average amplitude and phase response for the codons of the group 222 and 212.

where the change in amplitude and phase is very negligible. But the change is very significant for the groups 111 and 121, shown in figure 6, because the second base position is varying.

But intheFigure 7 and 8,no significant change occursbecause the codon start with strong base purine, therefore change in 2nd and 3rd base have a negligible effect on codon characteristic.

5. CONCLUSION

A novel approach is considered in the present paper to design a new genetic code table based on the electrical response of codon. The effect of base position in codon isinvestigated and analyzed using frequency response of codon electrical circuit.It is a new concept which may further be extended in amino acids classification and protein characterization. The work is significantfor bioscientists who are strugglingto produce proteins for drugs, biofuel and more. The work may help the biotechnologist to design devices that could be implanted in human body and programmed with genetic code. The results of this study are important in forensics, disease diagnosis, protein identification, etc.

6. ACKNOWLEDGMENTS

7. REFERENCE

[1] J. D. Watson and F. H. C. Crick, "A structure for DNA", Nature, 1953.

[2] Gullu, O. et al. "DNA-based organic-on-inorganic devices: Barrier enhancement and temperature issues, " Microelectronic Engineering Vol.85, Issue 11, 2250-2255, 2008.

[3] Nanjundiah, Vidyanand, "George Gamow and the genetic code," Resonance-Journal of Science Education , vol. 9, issue 7, 44-49, 2004Tavel, 2007 Modeling and Simulation Design. AK Peters Ltd.

[4] Collins, S.Francis. et al. "A vision for the future of genomics research", Nature, vol. 422, 835-847.

[5] Danny Porath, Gianaurelio Cuniberti, and Rosa Di Felice. "Charge transport in DNA-based devices", Long-Range Charge Transfer in DNA II, Springer Berlin Heidelberg, vol. 237, 183-228, 2004.

[6] Heath, James R., and Mark A. "Molecular electronics", Physics Today, vol. 56, issue 5, 43-49, 2003.

[7] Lu. Shun, "Conductivity measurements of single DNA molecules using conductive-atomic force microscopy", Diss. Department of Physics-Simon Fraser University, 2006.

[8] Jiaxiong Wang, "Electrical conductivity of double stranded DNA measured with ac impedance spectroscopy", Physical Review, B 78, 2008.

[9] Zdravko Kutnjak, et al. "Electrical conduction in macroscopically oriented deoxyribonucleic and hyaluronic acid samples", Physical Review, vol. 71. Issue 4, 2005.

[10] Johan Lagerqvist, Michael Zwolak, and Massimiliano Di Ventra. "Fast DNA sequencing via transverse electronic transport", Nano Letters , vol. 6, Issue 4 ,779-782.

[11] R. Marshall "Modeling DNA/RNA strings using resistor—capacitor (rc) ladder networks", The Computer Journal , vol. 53, issue 6, 644-660, 2010.

[12] Kunming Xu "DNA Circuit System and Charge Transfer Mechanism,"Engineering Vol.5,No.10,DOI: 10.4236/eng.2013.510B077, 2003.

[13] Vedrana. Hodzic, Vildana Hodzic and Robart W. Newcomb, " Modelling of electrical conductivity of DNA", vol. 54,No. 11, pp.2360-2364, 2007.

[14] N. V. Grib, J. A. Berashevich, and V. E. Borisenko,"Equivalent electrical network of the DNA molecule", Russian Microelectronics, Vol.35, issue 6,398-404, 2006.

[15] Jacob Ben, E., Z. Hermon, and S. Caspi. "DNA transistor and quantum bit element: realization of nano-biomolecular logical devices," Physics Letters A, Vol. 263, Issue 3,199-202.

[16] Muniza Zahid, et al. "DNA nanotechnology: a future perspective", Nanoscale research letters, Vol. 8, No.1, 1-13, 2003.

[17] P.P.Vaidyanathan, "Genomics and proteomics: a signal processor's tour", IEEE Circuits Syst. Mag. 4,6–29, 2004.

[18] Thomas M. Devlin, "Text Book of Biochemistry with Clinical Correlations", Wiley, 24-63, ISBN: 978-0-471-67808-3,2013.

[19] R. Jacob Baker," CMOS circuit Design and layout and simulation", Wiley, third edition, ISBN: 978-1-118-03823-9,2011.

[20] Tony Chan Carusone, David A. Johns, Kenneth W. Martin, " Analog integrated circuit design", Wiley, Second Edition, ISBN:978-0-470-77010-8.