

**COMPUTER APPLICATIONS IN NUCLEIC ACID CHEMISTRY****\*Okore S.A.<sup>1</sup>, Njoku V.O.<sup>2</sup> and Onwusonye J.C.<sup>3</sup>**

<sup>1</sup>Department of Chemistry, Federal Polytechnic Nekede, Owerri, P.M.B 1036 Owerri, Imo State, Nigeria.

<sup>2</sup>Department of Chemistry, Imo State University, P.M.B 2000, Owerri Imo State, Nigeria.

<sup>3</sup>Department of Biochemistry, Federal Polytechnic Nekede, Owerri, P.M.B 1036 Owerri, Imo State, Nigeria.

Article Received on  
07 April 2017,

Revised on 28 April 2017,  
Accepted on 18 May 2017

DOI: 10.20959/wjpr20176-8590

**\*Corresponding Author\***

**Dr. Okore S.A.**

Department of Chemistry,

Federal Polytechnic

Nekede, Owerri, P.M.B

1036 Owerri, Imo State,

Nigeria.

**ABSTRACT**

This study focused on how computer applications could be used for the study of deoxyribonucleic acid (DNA) chemistry and its applications as a technological tool for sequencing the genes, computing the base pairing of DNA and in modeling biomolecules. The works involves mining a public database for related genes with available software programs. With this idea, computational chemistry becomes a bridge between computer science and biochemistry.

**KEYWORDS:** This study focused science and biochemistry.

**INTRODUCTION**

There are many recent advances in deoxyribonucleic acid (DNA) technology which have helped in the understanding of its biological role and chemical structure. The rate of developments in recent years have been dramatic not only on the amount of information available about DNA sequencing, but also the way which that knowledge has been applied to the benefit of humanity.

DNA and ribonucleic acid (RNA) are molecular polymers which carry genetic information. These two nucleic acids are long sequences of much smaller molecular structure called nucleotides each of which is built from phosphate group, a sugar and a ring structure called a base. In DNA the four possible bases are Adenine (A), Guanine (G), Thiamine (T) and Cytosine (C) whereas in RNA the thiamine is replaced by Uracil (U). Thus one can use the order of the base pairing letters to specify the order of the nucleotides. As a cell requires

protein, another type of nucleic acid RNA translates the genetic information in DNA and carries that information to the ribosome where the synthesis of protein takes place.<sup>[1-3]</sup>

By means of Computer simulations a lot of work has been done in the area of DNA supercoiling, determination of differential geometric chain folding, protein interaction and fighting viral infections and in the building of biomolecules.<sup>[4]</sup> One recent tool of the use of computer in DNA analysis is for the production of therapeutic agents by genetic engineers. Genetic engineering is not restricted to organisms. Recently there have been great advances in the ability to genetically modify plants.<sup>[5]</sup>

Computerized finger print search systems are now widely applied in forensic technology. The computerized systems being developed independently have incorporated automatic scanning devices that can code and read genetic fingerprint characteristics from standard fingerprint cards.<sup>[6]</sup> As computational power increased over the years so too is the size of the system which the computational chemist has the potential to investigate. This has opened up new field of application in the field of biological molecules. It is in one of these fields that this paper is written to show chemists and biological scientists the application of computer as a sequence tool and how to compute base pairing of DNA and in the modeling of biomolecules. We shall indicate how some existing software packages in the market can be used to achieve this goal from a public database.

## COMPUTATIONAL METHODS AND DISCUSSIONS

Modern DNA chemistry would not be possible without the use of computational techniques. Chemists can use computers as a technological tool to collect, store and analyze the output from automated sequencing experiments. DNA computing is a new vista of computation that bridges between computer science and biological sciences. In DNA base computation information is carried out through DNA molecules and its processing is accompanied through series of biochemical reactions and the high information density of the DNA molecules. DNA computing has become amore attractive research field. Our bodies store something like a gigabyte of information per cubic micrometer. Compare their storage capacity with the information density of ten megabytes per square centimeter for a modern solid state electronic device.<sup>[7-10]</sup> Let us examine the following programming methods.

**(A) Dynamic Programming Method(DPM)**

There are many databases to choose from, but one of the most useful methods applied in the study is the one at the National Center for Biotechnology information in the U.S.A.<sup>[11-12]</sup> There is a vast amount of information in this database and the only way to get appreciation of it is to go on and around mining it from internet. MATHLAB is a useful software tool to sequence the information from this database.<sup>[13]</sup> A vast number of sequence alignment tools have been developed to carry out a kind of screening efficiently. The general principles behind the general methods of dynamic programming are considered.

The starting point in all dynamic programming is to construct a two dimensional matrix where the elements represent the similarity between any residue in sequence A with each of the residue B. Next, we must find the highest scoring path through which the comparison matrix to obtain the optimum alignment. This is known as generating the maximum match matrix. Now consider what a matrix element ( $M_{ij}$ ) in the maximum match matrix represents. It is the highest scoring path which can be obtained starting at that point. Clearly since we are restricted to obtaining the sequence order, the matches which follow positions  $i$  and  $j$  can arise only from higher values of the subscripts. In practice we start from the bottom right hand corner and systematically go through the comparison matrix converting it to a maximum match matrix by finding the maximum value in the row and column  $i+1, j+1$  of the maximum match matrix and adding this to the value of comparison matrix as shown in the example (fig1).

	A	C	F	C	G	S	T
A	1	0	0	0	0	0	0
C	0	1	0	1	0	0	0
Q	0	0	0	0	0	0	0
C	0	1	0	1	0	0	0
A	1	0	0	0	0	0	0
S	0	0	0	0	0	1	0
T	0	0	0	0	0	0	1

Fig.1.

In the second example as depicted (fig2)

	A	C	F	C	G	S	T
A							
C							
Q			1				
C							
A					2	1	0
S		1	2	0			
T					0	0	1

**Fig. 2.**

Having completed the assembly of the maximum match at the top left hand corner and find the optimum alignment by choosing the highest value in each row and column pair such that  $I$  and  $j$  are both higher than those of the previous pair as shown in (fig3).

	A	C	F	C	G	S	T
A	5	3	3	2	2	1	0
C	3	4	3	3	2	1	0
Q	3	3	3	2	2	1	0
C	2	3	2	3	2	1	0
A	3	2	2	2	2	1	0
S	1	1	1	1	1	2	0
T	0	0	0	0	0	0	1

**Fig: 3.**

For the two similar sequences in the example the highest value commonly fall on the diagonal but occasionally a letter is skipped leaving it unpaired. Hence gap can be created in the alignment, in evolutionary terms this is justifiable as additional base can be inserted into DNA as two sequences diverge from a common point. Therefore it is necessary to obtain an exact match. However the creation of gaps must be controlled otherwise if dissimilar stretches of residues crop up in matched sequences, the algorithm will react by increasing the length of the gaps to overcome this problem and gap penalties are introduced. These are subtracted from the maximum match score to disfavor moving of the diagonal when different residue types are matched.

The second issue in sequence alignment is the value used in the comparison matrix. In the example we used the identity matrix. Many different relationships have been constructed. In homologue proteins, take for example the hydrophobic residue leucine, if this were matched against isoleucine or valine, the property of the side at that site in protein would be conserved.

Therefore, it will not be an unwise alignment. Thus some account must be taken of conservative substitutions which would be accounted as zero by the identity matrix. Having screened a data base using pairwise alignments to obtain a family of sequence which shows evidence of being related, the next task is to try and optimally align all the sequences. This provides new information than a single alignment. Clearly the dynamic programming method could be extended into many dimensions for simultaneous alignment but this would soon run into problems with memory requirement. An alternative approach is to generate pairwise alignments between all the members of the family. The most similar pair is then taken as the core of the alignment to which successive sequences are matched with average values being placed in the comparison matrix where there is not identity across the growing set of alignments.

This method can be run iteratively to find a self-consistent solution. The main consequence of a multiple alignment is that position in the sequence which shows absolute conservation, or where conservative changes are made can be spotted. Comparisons of known related three dimensional structures have shown that the rate of successful mutation is much greater on the surface of a protein core. Thus by monitoring the degree of conservation across a family of proteins we can start to make a prediction about tertiary structure without knowledge of the three dimensional structure.<sup>[14-17]</sup>

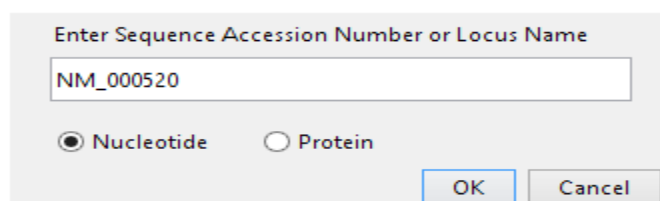
### **(B) Sequence Analysis Method (SAM)**

This is another computational method for obtaining information about a nucleotide or amino acid sequence. The common task are identifying genes, determine the protein coded by gene, determining the function of a gene by finding a similar gene in another organism. Other examples include sequence statistics, multiple sequence alignment and we can explore some basic information from World Wide Web public databases for sequencing nucleotides. The MATHLAB command window is one of the mathematical software packages that provide an integrated environment for bringing sequence information from the web. The following procedure explains how to sequence the gene from a public database.

### **(C) Importing a Sequence using MATHLAB Software.**

The following procedure illustrates how to retrieve sequence information from the NCBI Web database. The first step when analyzing a nucleotide or amino acid sequence is to get sequence information into MATLAB. The seqtool, using functions in Bioinformatics Toolbox, can connect to Web databases and read information into MATLAB.

1. Open the sequence viewer. In the MATLAB Command Window, type seqtool. The Sequence Viewer window opens without a sequence loaded. Notice that the panes to the right and bottom are blank.
2. To get a sequence from the NCBI database, select File> Download Sequence from NCBI to open the Download Sequence From NCBI dialog box as depicted in (fig.4)



(fig.4)

3. In the Enter Sequence box, type an accession number for an NCBI database entry. For example, enter NM\_000520 and select the Nucleotide option button. This is the human gene HEXA that is associated with Tay-Sachs disease. MATLAB goes to the Web, loads information for the accession number you entered, and calculates some basic statistics.

### Viewing Nucleotide Sequence Information

After you import a sequence into seqtool, you can read information stored with the sequence, or you can view graphic representations for ORFs and CDSsn.

1. In the left pane tree, click Comments. The right pane displays general information about the sequence.
2. Now click Features. The right pane displays NCBI feature information, including index numbers for a gene and any CDS sequences
3. Click ORF to show the search results for ORFs in the six reading frames.

### In clicking comments gives

LOCUS NM\_000520 2437 bp mRNA linear PRI 15-MAR-2015  
 DEFINITION Homo sapiens hexosaminidase A (alpha polypeptide) (HEXA), mRNA.  
 ACCESSION NM\_000520  
 VERSION NM\_000520.4 GI:189181665

In summary this gene encodes the alpha subunit of the lysosomal enzyme beta-hexosaminidase that, together with the cofactor GM2 activator protein, catalyzes the degradation of the ganglioside GM2, and other molecules containing terminal N-acetyl hexosamines. Beta-hexosaminidase is composed of two subunits, alpha and beta, which are encoded by separate genes. Both beta-hexosaminidase alpha and beta subunits are members of family 20 of glycosyl hydrolases. Mutations in the alpha or beta subunit genes lead to an accumulation of GM2 ganglioside in neurons and neuron degenerative disorders termed the GM2 gangliosidoses. Alpha subunit gene mutations lead to Tay-Sachs disease (GM2-gangliosidosis type I) [provided by RefSeq, Jul 2009]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data because no single transcript was available for the full length of the gene. The extent of this transcript is supported by transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications.

Sequence View

NM\_000520: Homo sapiens hexosaminidase A (alpha polypeptide) (HEXA)

- Sequence
  - ORF
  - Full Translation
  - Annotated CDS
  - CDS with Translation
  - Complement Sequence
  - Reverse Complement Sequence
  - Features
  - Comments

Base Count

A:	526	21.6%
C:	653	26.8%
G:	644	26.4%
T:	614	25.2%

Features

```

source      1..2437
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="15"
            /map="15q24.1"
gene        1..2437
            /gene="HEXA"
            /gene_synonym="TSD"
            /note="hexosaminidase A (alpha polypeptide)"
            /db_xref="GeneID:3073"
            /db_xref="HGNC:HGNC:4878"
            /db_xref="MIM:606869"
exon        1..460
            /gene="HEXA"
            /gene_synonym="TSD"
            /inference="alignment:Splice:1.39.8"
misc_feature 121..123
            /gene="HEXA"
            /gene_synonym="TSD"
            /note="upstream in-frame stop codon"
CDS         208..1767
            /gene="HEXA"
  
```

Map View

Sequence

CDS

### In clicking features we obtain

source 1..2437

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606" ...

Sequence View

NM\_000520: Homo sapiens hexosaminidase A (alpha polypeptide) (HEXA)

- Sequence
  - ORF
  - Full Translation
  - Annotated CDS
  - CDS with Translation
  - Complement Sequence
  - Reverse Complement Sequence
  - Features
  - Comments

Base Count

A:	526	21.6%
C:	653	26.8%
G:	644	26.4%
T:	614	25.2%

Position: 2437 bp

```

1 agttgccgac gccccggcaaca atccctctgca cgtagcagga gccctcaggtc cagggccggaa
+1
+2
+3
-1
-2
-3
61 gTgaagggc aggggtgtggg tccctctggg gtcgcaggcg cagagccgcc tctgttcaag
+1
+2
+3
-1
-2
-3
121 Tgattcgccy ataagtcacg gggggcccg ctaactgacc agggctctaac gtggccagcc
+1
+2
+3
-1
-2
-3
181 cccctccgaga ggggagacaa gccggccatg acaagctcca ggccttgggt ttcgctgctg
+1
+2
+3
-1
-2
-3
241 cTaccaccag ctttcaccag acaccaccac cccctctaac cctaccctca aaacttccaa
  
```

Map View

Sequence

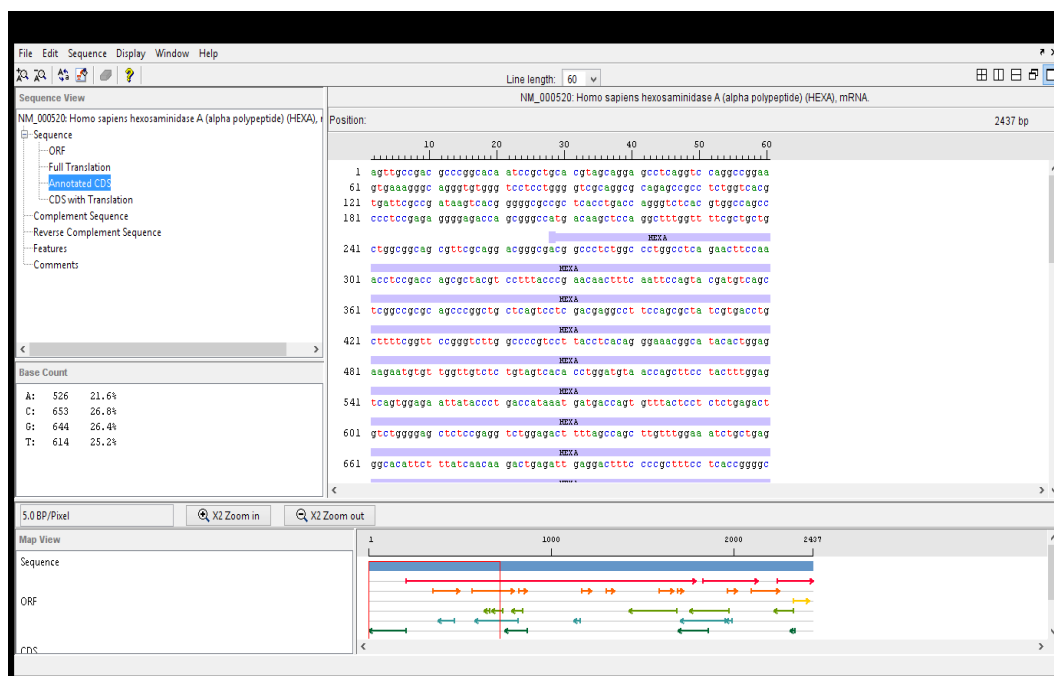
ORF

CDS



**Window with a click on ORF showing six reading frames**

agttgccgacgcccggcacaatccgctgcacgtagcaggagcctcaggctcaggccggaagtgaaagggcagggtgtgggtcctc  
ctggggctgcaggcgcagagccgcctctggcacgtgattgccgataagtcacggggcgccgctcacctgaccagggtctcacg  
tggccagccccctccgagaggggagaccagcgggcatgacaagctccaggctttggtttcgtgctgctggcggcagcgttcgc  
aggacggggcagggccctctggccctggcctcagaactccaacctccgaccagcgtacgtcctttaccgaacaactttcaattc  
cagtacgatgtcagctcggccgcagccccggctgctcagtcctcgacgaggcctccagcgtatcgtgacctgtttcggttccg  
ggctttggccccgtccttacctcacagggaacggcatacactggagaagaatgtgttggttctctgtagtcacacctggatgtaacc  
agcttctactttggagtcagtgaggagaattataacctgaccataaatgatgaccagtgtttactcctctctgagactgtctggggagctctc  
cgaggtctggagacttttagccagcttgttgaaatctgctgagggcacattctttatcaacaagactgagattgaggactttccccgctt  
tctcaccggggcttctgttgatgatactctgccattacctgccactctctagcatcctggacactctggatgcatggcgtacaataaat  
tgaacgtgttccactggcatctggtagatgatccttctccatagagagcttactttccagagctcatgagaaggggtcctacaac  
cctgtcaccacatctacacagcacaggatgtgaaggaggtcattgaatacgcacggctccggggatccgtgtgcttcagagtttga  
cactctggccacactttgtctctggggaccagggtatccctggattactgactcctgtactctgggtctgagccctctggcacctttgga  
ccagtgaatcccagctcaataataacctatgagttcatgagcacattctttagaagtcagctctgtcttccagatttttatcttcatctgg  
aggagatgagggtgattcacctgctggaagtccaacccagagatccaggactttatgaggaagaaaggcttcggtgaggacttcaag  
cagctggagctcttctacatccagacgctgctggacatcgtcttcttatggcaagggtatgtggtgtggcaggaggtgtttgataata  
aagtaaagattcagccagacacaatcatacagggtgtggcgagaggatattccagtgaactatatgaaggagctggaactggtcaccaa  
ggccggcttccggcccttctctgccccctggtacctgaaccgtatcctatggccctgactggaaggatttctacatagtggaacc  
cctggcatttgaaggfaccctgagcagaaggctctggtgattggtggagaggcttgtatgtggggagaatgtggacaacacaaaac  
ctggtccccaggctctggcccagagcaggggctgttccgaaaggctgtggagcaacaagttgacatctgacctgacatttgcctatg  
aacgtttgtcacacttccgctgtgaattgctgagggcaggtgtccaggcccaacctcaatgtaggcttctgtgagcaggagtttgaac  
agacctgagccccaggcaccgaggagggtgctggctgtaggtgaatggtagtgagccaggcttccactgcatcctggccagggg  
acggagccccttgccttctgccccttgcctgctgcccctgtgcttggagagaaaggggcccgggtgctggcgtcgcattcaataaag  
agtaatgtggcattttctataataaacatggattacctgtgttaaaaaaaaaagtgtgaatggcgtagggtaagggcacagccaggct  
ggagtcagtgctgcccctgaggtctttaagttgagggtgggaatgaaacctatagcctttgtgctgttctgccttgcctgtgagctatg  
tactccccctccactctgaccatattccagacacctgccctaatcctcagcctgctcaacttctgacttatctccaaggcgttgg  
tatatgaaaaagatgtaggggcttgagggttctggacagtggggagggtccagaccaacctggtcacagaagagcctctccc  
ccatgcatactcatccacctcccctccctagagctattctccttgggttcttctgcttcaattttatacaaccattttaaatattttaaac  
acatattgttctta.

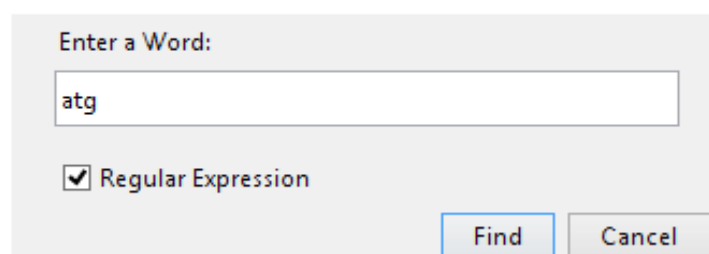


Click Annotated CDS to show the protein coding part of a nucleotide sequence.

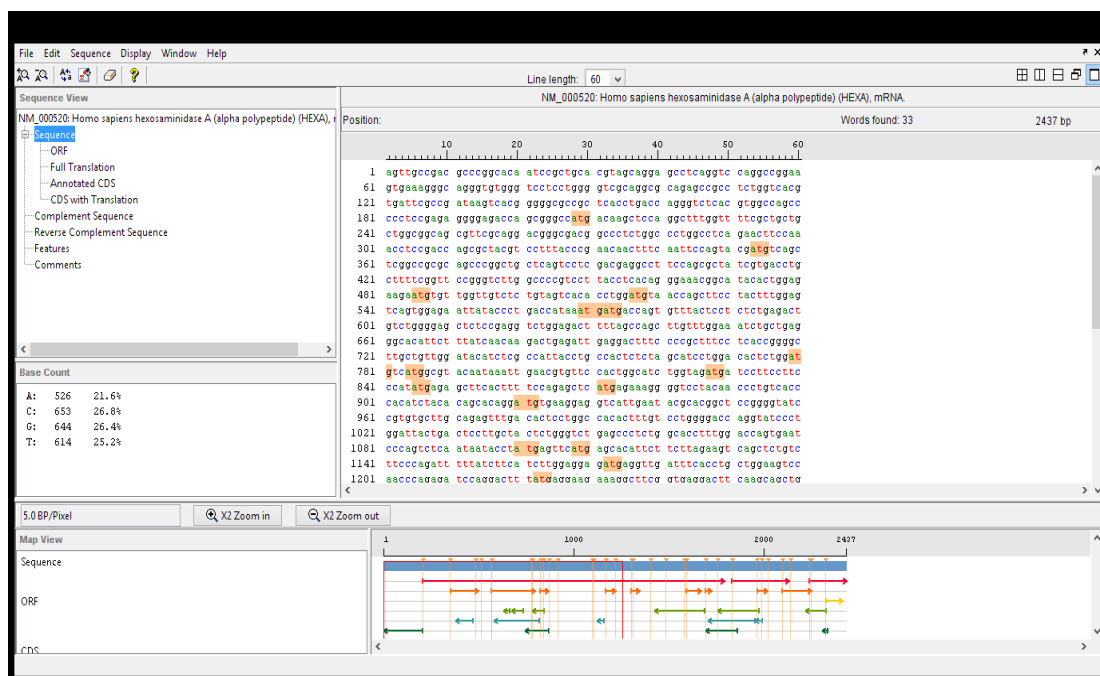
### Searching for Words

The following procedure illustrates how to search for characteristic words and sequence patterns. You will search for sequence patterns like the TATAA box and patterns for specific restriction enzymes.

1. From the Sequence menu, click Find Word.
2. In the Enter a Word box, type a sequence word or pattern. For example, enter atg.



seqtool searches and displays the location of the selected word.

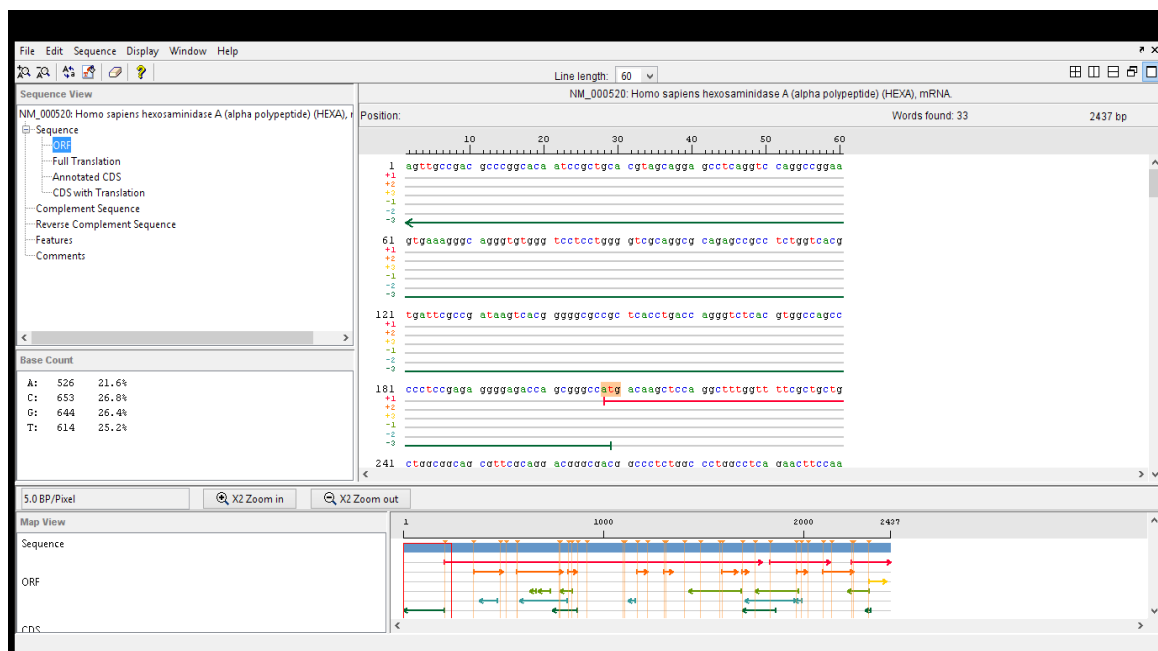


To clear the display, on the toolbar, click the Clear Word Selection button.

### Exploring Open Reading Frames

The following procedure illustrates how to identify the protein coding part of a nucleotide sequence and copy it into a new view. Identifying coding sections of a nucleotide sequence is a common bioinformatics task. After locating the coding part of a sequence, you can copy it to a new view, translate it to an amino acid sequence, and continue with your analysis.

1. In the left pane, click ORF. seqtool displays the ORFs for the six reading frames in the right and lower window.



2. Click the longest ORF on reading frame 3. The ORF is highlighted to indicate the part of the sequence that is selected
3. Click the longest ORF on reading frame 3. The ORF is highlighted to indicate the part of the sequence that is selected.
4. Right-click the selected ORF and then select Export to Workspace. Enter the name of a variable. For example, enter NM\_000520\_ORF
5. From the File menu, click Import from Workspace. Enter the name of a variable with an exported ORF. For example, enter NM\_000520\_ORF.seqtool adds a tab at the bottom for the new sequence while leaving the original sequence open.
6. In the left pane, click Full Translation. From the Display menu, point to Amino Acid Residue Display and click One Letter Code .seqtool displays the amino acid sequence below the nucleotide sequence.

The Bioinformatics Toolbox offers computational molecular biologists, chemists and other research scientists an open and extensible environment in which to explore ideas, prototype new algorithms and build applications in drug research, genetic engineering and other genomics and proteomics projects. The toolbox provides access to genomic and proteomic data formats, analysis techniques and specialized visualizations for genomic and proteomic sequence and microarray analysis. Most functions are implemented in the open MATLAB language, enabling scientists to customize the algorithms or develop their own.

#### **(D) Electronic Spread Sheet Model of Base Pairing in DNA**

Two helical, polynucleotide chains that constitute a DNA fiber are held together by a hydrogen bonds between pairs of bases. Adenine is always paired with thymine; Guanine is always paired with cytosine. Thus if the adenine content of one of the polynucleotide chains is known, the thymine content of the complimentary chain is also known. If the guanine content of one chain is known, the cytosine content of the complementary chain is also known.<sup>[18]</sup> See the table below as shown in fig.5.

<b>Strand 1</b>	<b>Strand 2</b>
A	T
A	T
G	C
G	C
C	G
T	A
T	A

C	G
G	C

fig.5.

For example, suppose that strand 1 contains 25% Adenine and 35% guanine. Strand 2 must contain 25% Thymine and 35% cytosine it also followed that the remaining 40% of strand 2 must be composed of adenine and guanine. That is if the percent adenine content of strand 1 equals  $x$  and the percent guanine content of strand 1 equals  $y$ , then:

T content (strand 2) =  $x$

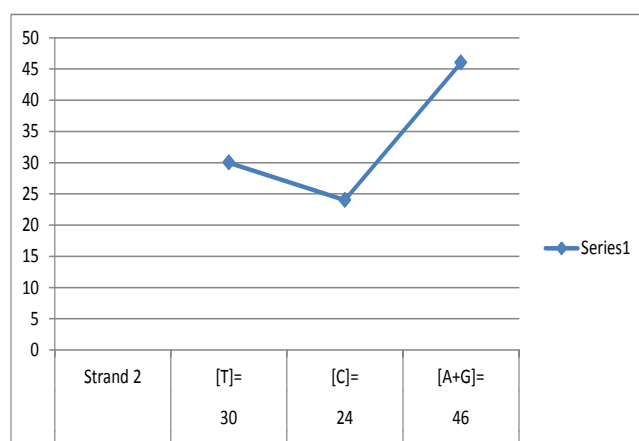
C content (strand 2) =  $y$

T + C content (strand1) =  $100-x-y$

A + G content (strand 2) =  $100-x-y$

This simple relationship can be used to form the basis of an electronic spreadsheet model that defines the base pairing content of a fiber of DNA if the adenine and guanine content of one of the strands is known.

Strand 1	Strand 2	
[A]=	30 [T]=	30
[G]=	24 [C]=	24
[T+C]=	46 [A+G]=	46

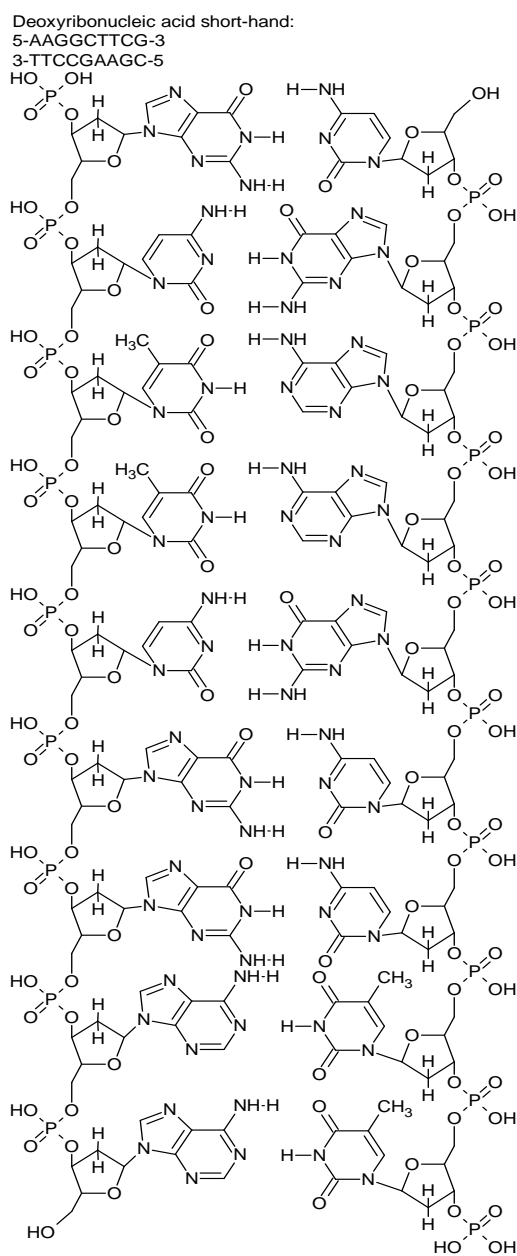


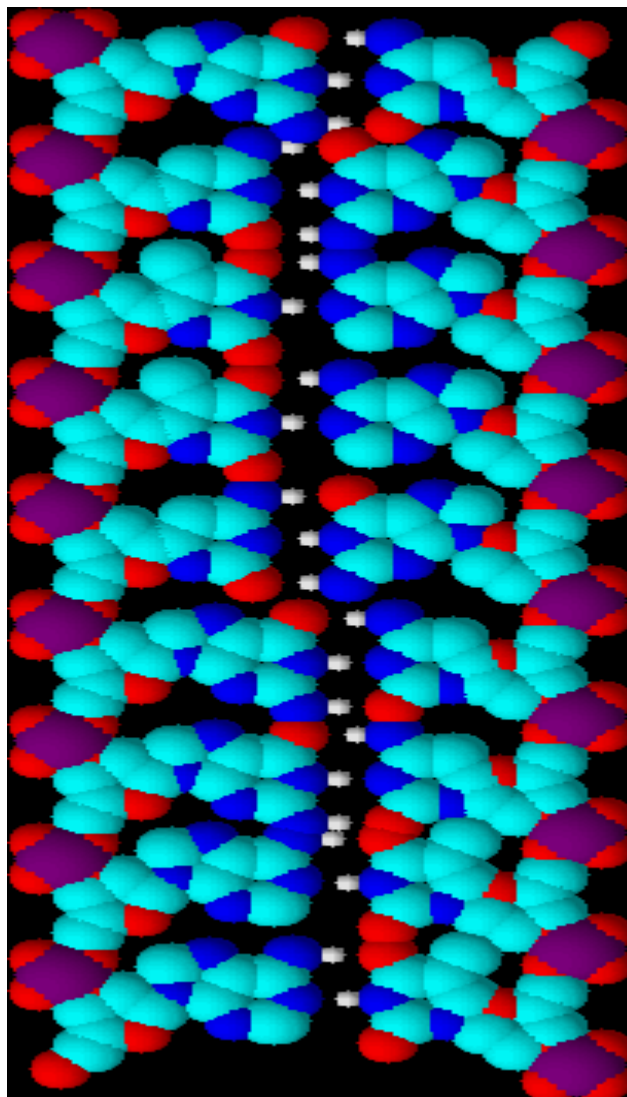
The graph is an original trend line model of DNA Strand.fig.6.

### (E) Modeling the DNA Structure 5-AAGGCTTCG-3

Applying computer simulation to model biomolecules of interest is very exciting, because such molecules are mostly essential elements in life process. But extracting some useful information is very challenging because the systems are so inherently complex. The impact of

computer graphics has had the biggest influence in modeling biomolecules. Chemists have always used molecular models to understand their subject. In this study we employ computer graphics from two different programs to model the DNA Strand 5-AAGGCTTCG-3. The first methods involve modeling the strand structure based on the percent of the DNA content using electronic spreadsheet program as shown in fig6. The second modeling does not take into account of the DNA content but take into account the chemical structures of the genetic bases. Various modeling tools for modeling DNA from their chemical structures can be sourced from the internet.<sup>[19]</sup> In general computer graphics can enable us make accurate prediction of structure of a protein from its sequence as shown from fig.7 and 8 below.

**Fig.7**



**Fig 8:** space field model of the sequence structure.

5-AAGGCTTCG-3.

## CONCLUSION

The role of computational chemistry in diverse range of chemical discipline is a clear sign of its maturity as a field. Chemistry should be concerned with molecular explanations and it is pictures of this type which computational chemistry provides in this study. A computer application to nucleic acid has been used to form a bridge between computer science and biochemistry by applying computer in sequencing the genes through dynamic programming methods. Computer graphics have helped us to build a better picture of DNA strands and their molecules by simulation methods. Computers are essential in nucleic acid chemistry. Their applications include analysis of sequence data, storage and analysis of completed sequences and assignment of functions by similarity search and in building a complex model.

**REFERENCES**

1. Williams, J. K. and Larry, I.S. (1993): Supercomputing and the Transformation of Science, Printed in U.S.A, 118-128.
2. Timberlake, K.C (2006): An Introduction to General, Organic, and Biological Chemistry, Published in San Francisco C.A 94111, U.S.A, 574-582.
3. Mathews, A. A-C, Mastering chemistry for biochemistry. Online textbook. 10 August 2015 <[www.aw-bc.com/chemistry](http://www.aw-bc.com/chemistry)>.
4. Beveridge, D.L. and Jorgensen W.L, (1986): Computer Simulations of Chemical and Bimolecular System, The New York Academy of Science, New York, 69-80.
5. Olson, S (2002): Mapping History, Discovering Our Past through Our Genes, Bloomsbury London.
6. Suferstern R, (1987): An introduction to forensic science, Prentice–Hall Inc., New Jersey, 357-358.
7. Alderman L. (1994): Molecular Computation to Combinatorial Problems Science, 266: 1021-1024.
8. Lui, W. Gao, L., Xiangrong, L., Wang, S. and Xung. J (2003): Solving the 3-SAT problem Based on DNA Computing. J. Chem. Inf. Sc, 43: 1872-1875.
9. Wu, H. (2001): An Improved Surface Based methods for DNA Computation, 59: 1-5.
10. Doonan S (2004): Nucleic Acids, Royal Society of Chemistry, 116-157.
11. Swiss Institute of Bioinformatics (SIB), 12 August 2015 <<http://ca.expasy.org>>.
12. Nat. Center for Biotech. Inform. (NCBI), 12 August 2015 <<http://www.ncbi.nlm.nih.gov>>.
13. Packson, R.: Computational Biology (R2025a), MATHLAB, [Computer Software] Retrieved 13 August 2015 <<http://www.mathworks.com/products/Mathlab>>.
14. Grant G.H and Richard W.G (1994): Computational Chemistry, Oxford University Press Inc. New York, 66-78.
15. Branden .C and Tooze. J(1991): Introduction to protein Structure, Garland New York.
16. Lukas, C.M. et-al (2000): Nat. Struct. Biology, 7: P134.
17. Barton, G.J (1990). Methods in Enzymology, 183: 403-128.
18. Atkinson, D.E., Clark, S.G., Rees D.C. (1987): Dynamic Models In Biochemistry, Benjamine/ Cumming Publishing Company, Inc U.S.A.
19. Wheeler, P. and Pautler, B (ChemSketch Software) 12 August 2015 <[www.acdlabs.com](http://www.acdlabs.com)>.