International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

ELSEVIER
SSRN

IJIEMR Transactions, online available on 4th Sept 2019. Link :http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-09

Title A NOVEL MECHANISM OF DNA BY GRAPH THEORY

Volume 08, Issue 09, Pages: 570–580.

Paper Authors

MRS.M.DRGADEVI ,MS.D.PRATHIBHA, MS.P.GANGA BHAVANI
CH.S.D.ST THERESA'S COLLEGE FOR WOMEN,ELURU ELURU,ANDHRAPRADESH

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

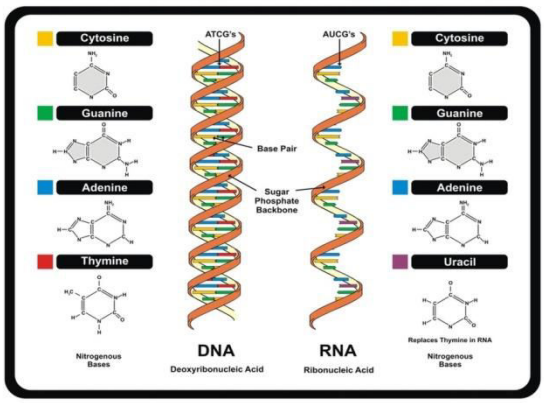
International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

A NOVEL MECHANISM OF DNA BY GRAPH THEORY

[1]MRS.M.DRGADEVI ,[2]MS.D.PRATHIBHA, [3]MS.P.GANGA BHAVANI
[1]MSC,MCA,MTECH(CSE). (LECTURER IN MATHEMATICS) CH.S.D.ST THERESA'S COLLEGE FOR WOMEN,ELURU
ELURU,ANDHRAPRADESH.
[2]MSC, CH.S.D.ST THERESA'S COLLEGE FOR WOMEN,ELURU ELURU,ANDHRAPRADESH.
[3]STUDENT CH.S.D.ST THERESA'S COLLEGE FOR WOMEN,ELURU ELURU,ANDHRAPRADESH.
m.devi.mca.06@gmail.com prathibha339@gmail.com parasagangabhavani622@gmail.com

Abstract:
DNA stands for Deoxyribonucleic Acid. It is the hereditary material found in all living organisms. It contains the genetic instructions for the development and functioning of an organism. DNA is made up of molecules called nucleotides. Each nucleotide contains a phosphate group, a sugar group and a nitrogen base. The four types of nitrogen bases are adenine (A), thymine (T), guanine (G) and cytosine (C).DNA replication is a fundamental process for cell proliferation in humans and other living organisms. It involves transfer of genetic information from the original DNA molecule into two copies. Understanding this process is of great importance towards unveiling the underlying mechanisms that control genetic information transfer between generations**.** DNA fragment assembly requirements have generated an important computational problem created by their structure and the volume of data. Therefore, it is important to develop algorithms able to produce high-quality information that use computer resources efficiently. Such an algorithm, using graph theory, is introduced in the present article. Molecular biology which aims to study DNA and protein structure and functions, has stimulated research in different scientific disciplines, discrete mathematics being one of them. One of the problems considered is that of recognition of DNA primary structure. It is known that some methods for solving this problem may be reduced (in their computational part) to graph-theoretic problems involving labeled graphs.
Keywords:DNA, Molecular biology, DNA fragment assembly, Graph theory Algorithms.

Introduction



The discovery of DNA structure 55 years ago marked the beginning of a process that has transformed the foundations of biology and medicine, and accelerated the development of new fields, such as molecular biology or genetic engineering. Today, we know much about DNA, its properties, and function. We can determine the structure of short DNA fragments with picometer precision, find majority of the

genes encoded in DNA, and we can manipulate, stretch and twist individual DNA molecules. We can utilize our knowledge of gene regulatory apparatus encoded in DNA to produce new microorganisms with unexpected properties. Yet, there are aspects of DNA function that defy our understanding, mostly because the molecule is just one, albeit essential, component of a complex cellular machinery. The basic idea behind numerical characterization is that specific gene sequences are generally unique and therefore possess a characteristic signature in the composition and distribution of the nucleotides that make up the genes. The departure from uniqueness will come from mutations although some degree of homology will be maintained. Numericalcharacterization will seek to capture the essence of this homology so that each gene canbecharacterized by one number or a vector that identifies a gene. The same construct can be applied to significant regio-specific motifs that may be identified within the gene, corresponding to, say, particular structural aspects of the downstream protein or enzyme, or within a DNA or RNA sequence segments such as promoter sequences. In a broader perspective, numerical characterization can play an important role in the identification of coding segments in newly emerging sequences, or prediction of functions from sequences.

The primary step in creating a mathematical descriptor is to develop reliable techniques forcharacterizing DNA/RNA sequences. While algorithms can be constructed to generate mathematical representations directly from DNA primary sequences, it is intuitively more appealing to represent a long DNA sequence in the form of a graph and visually identify regions of interest or the distribution of bases along the sequence. Most methods that have been proposed in the literature to numerically characterize DNA sequences are based on one or more graphical representations of such sequences, and several applications have been made using these techniques. This is a new field of enquiry and has been gathering momentum over the last decade. In this review we focus on the different mathematical techniques for characterizing DNA sequences. We briefly enumerate the graphical representations of DNA sequences that form the foundations of these numerical techniques, and then discuss the techniques themselves. We propose a set of criteria of what the numerical descriptors are supposed to achieve, and then compare the different methods on the basis of the results they have demonstrated for a set of gene sequences measured against the corresponding amino acid sequences. We hope that this will highlight both the utilities and limitations of the current crop of numerical methods and thus lead the way towards more sophisticated analysis and improvements in techniques for better understanding of what information the DNA sequences contain and how numerical techniques can help. Mathematical descriptors of DNA sequences and their use in rationalizing biological properties of DNA follow from the structure-property similarity principle.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Biological information processing is very different from classical computing architectures.Biological systems' elementary components respond slowly compared to solid-stateswitches-but they implement much higher-level operation. A second striking feature, particularlyduring development, is biological systems' self-assembly growth, which lets themachieve high interconnection densities. A third fundamental point is that biological systemsare implemented without being planed.Molecular computing is based on the idea that data can be encoded as bio molecules such as DNA strands and molecular biology tools can be used to transform this data to perform, for example, arithmetic or logic operations.Adleman(1994) demonstrated how to solve an instance of the Directed Hamilton Path (DHP) problem by encoding it in DNA and subsequently using a biological protocol that can create and extract the solution in a small number of steps. The main attraction of this method of performing computation lies in the potential of massive parallelism resulting in a greater number of computations per second than the fastest supercomputers could perform.

Infact, a mathematical model that tackled thisquestionsuccessfullywasfirstproposedin 1948by C.E.Shannon[1].Inhiswork,Shannonintroducedan operationalframeworktoestablishawelldefinedmodelforwirelesscommunicationchannels.Thisgenericmodelstartswithaninformationsourcethatproducesadatastreamcomposedofelementspickedfromafixedalphabet. The mostsimplestreamisthebinarystreamcomposedofbitsthatbelongtothemathematical group F = {0, 1}, which isthebasisformoderndigitalcommunications ystems.

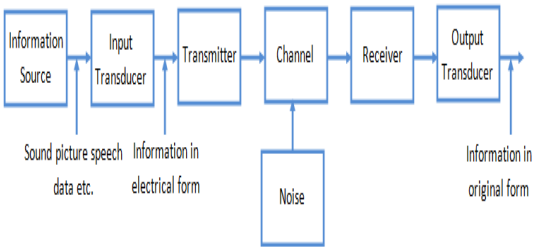


Fig1: Block diagram of a basic communication system.

The data stream then passes through a transmitter that subsequently modulates it and encodes it to form the transmitted signal. The transmitted signal, in turn, propagates through a communication channel until it is delivered to the target receiver. The modulation and coding schemes are to be chosen carefully to secure the transmitted signal's propagation capability and overcome the channel distortion effects. Many noise sources can impact the transmitted signal including additive white Gaussian noise (AWGN) and multipath fading. Upon receiving the transmitted signal, mixed with noise, the receiver is responsible for several procedures including noise cancellation, decoding, and demodulation. Basically, the receiver's function is to reverse every process the original data stream had to go through, either intentionally by the transmitter, or unintentionally due to the channel conditions. Finally, the receiver delivers the recovered stream to the information destination which is the intended target by the data transaction process. Figure 1 summarizes the aforementioned steps.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Structure-property similarity principle

The development of descriptors, including descriptors for the characterization of DNA sequences, follows from the structure-property similarity principle which states that similar structures usually have similar properties. This can be formally represented in terms of mapping of the Set C (chemicals or DNA sequences) to the real number line R.As opposed to the direct mapping of C to R by experimental means, the composition of mappings C→D→R, based on the base sequences of DNA, will give us power of associating functions (properties) to sequences based on their composition only. Such method can also compare hypothetical or hazardous sequences with existing data sets in the growing genomesequence bank and make predictions about their biological activities, hazardous nature etc. It has to be emphasized that neither of these two mappings described in Fig.1 is unique, i.e., both the experimental (C→R) and theoretical (C→D→R) approaches can assign the same magnitude of certain properties to mutually different sequences. This is also true for descriptors of small molecules1. This is not a big handicap for property prediction because even a degenerate descriptor may quantify important structural aspects of DNA or chemical species. Of course, the less degenerate the descriptor, the better it is as a tool for documentation.

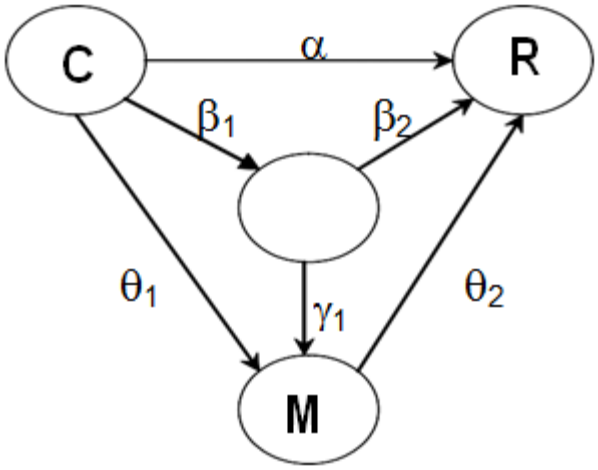


Figure 2. Composition functions for structure-property similarity principle1, where C = A set of chemicals, R = The set of real numbers, D = A set of structural descriptors and M = A set of molecular properties Recent literature on the topic shows that the structure-property similarity principle is a general paradigm where C might represent a small organic molecule or a macromolecule likeDNA. The Set D might be topological, geometrical, or quantum chemical descriptors, and Mmight represent experimental or calculated molecular properties. In some instances, elements of the Set C might be proteomics patterns which are represented by matrices or matrix invariants2, 3, 4, 5, 6, 7, 8.

Graphical representation of Sequencing DNA Chains

As it is known DNA is a double helix in which the two coiled strands (chains) are composed each of only four different molecule types -nucleotides. Every nucleotide consists of phosphate, sugar and one of the following bases: adenine (abbreviated A), guanine (G), cytosine (C) and thymine (T). The two chains are held together by hydrogen bonds which exist only between pairs of complementary

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

bases, which are A-T and C-G. It follows that knowing one chain, the other (complementary) can beeasily reconstructed.

As we mentioned, one of the methods of recognition of the primary structure of DNA(i.e. a sequence of nucleotides) is sequencing by hybridization. Its biochemical phase isbased on the property of single-stranded acids to form a complex with a complementarystrand of nucleic acid. All short fragments of nucleic acids (oligonucleotides) of length l(a library composed of 4^lsub chains) are used in the hybridization experiment and thus,the formation of the complex indicates the occurrence of a sequence complementaryto the(oligonucleotide) in the DNA chain. It is detected by a nuclear or spectroscopicdetector. As a result of the experiment one gets a set (called Spectrum) of all l-longoligonucleotides which are known to hybridize with the investigated DNA sequence Nof length n (i.e. they are substrings of string N). In case of ideal data (when no l-longoligonucleotide appears more than once in the sequence) we have thus Spectrum=n − l + 1 (We will not consider here experiments with errors).

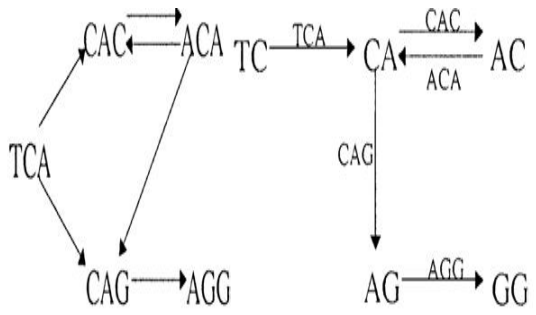


Fig 3:The graph H for the example TCACAGG. Fig 4:The graph G for the example TCACAGG.

Now comes the computational phase, where for a given Spectrum one should re-construct an unknown sequence N. The first approach to this problem based on graph theory, has been described by Lysov et al. [2]. They have proposed to formulate the problem of finding original sequence N as the problem of looking for a Hamiltonian path in a special graph. A directed graph H is built from Spectrum as follows: each oligonucleotide from Spectrum becomes a vertex, two vertices are connected by an arc if the l−1 rightmost nucleotides of the first vertex overlap with the l−1 leftmostnucleotides of the second one. A Hamiltonian path found in this graph corresponds to a proper sequence of elements of the Spectrum, i.e. a possible solution.

To illustrate this procedure let us consider the original sequence TCACAGG oflength n = 7. After the hybridization with oligonucleotides of length l = 3 we get fullSpectrum{TCA, CAC, CAG, ACA,AGG}. Graph H constructed by this method is as shown in Fig. 3. The only Hamiltonian path in this graph is TCA → CAC →ACA →CAG →AGG from which the original sequence can be read.

The above approach, however, leads to an exponential-time algorithm since looking for a Hamiltonian path is in general strongly NP-complete . Fortunately, Pevzner has observed that in this particular case one can treat graph H as a directed line graph of a certain original graph G. Now, graph H can be transformed into graph G, each vertex of H corresponding to an arc of G (the set of arcs in the new graph corresponds, in fact, to the Spectrum). The

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

arc connects vertices labeled as l − 1 left and l −1right nucleotides of the oligonucleotide corresponding to this arc. As a result of the transformation one gets the new graph in which aEulerian path is looked for. This reduces the complexity of the algorithm solving the DNA sequencing problem since finding aEulerian path can be done in polynomial time. Coming back to our example we have the graph G given in Fig. 4.

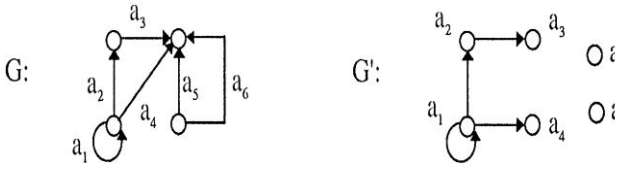


Fig5: A graph G and its adjoint G'.

The Eulerian path is TC →CA→ AC → CA→ AG →GG from which the same original sequence can be obtained.

The above approach raised some interesting questions in graph theory itself. They are concerned with the above class of labeled graphs which will be referred to asDNA graphs in the following. Specifically, one is interested in the characterizationand recognition of these labeled graphs as well as in finding conditions for whichtheabove transformation is possible. In the paper, these issues will be studied forunbounded and bounded alphabets used for graph labeling. Before doing this, we willset up the subject more formally in terms of graph theory. The definitions not givenhere can be found in . Note that by graph, we mean directed graph.

Graphical representations in 2D,3D,4D view:

Representations based on two dimensional Cartesian coordinates remain the staple form of graphical methods for their simplicity and intuitive feel. The idea was to read a DNA sequence base by base and plot succeeding points on the graph. According to the Nandyprescription[3], a point was plotted by moving one step in the negative x- direction if the base was an adenine (A) and in the opposite direction if it was a guanine (G) and a walk of one step in the positive y-direction if the base was a cytosine (C) and in the opposite direction if it was a thymine (T). The Gates method prescribed the bases GTCA and the Leong Morgenthalermethod[4] prescribed CTAG reading clockwise starting from the negative x-axis for the walks. Thus a sequence like ATGGTGCACC will display in the three systems plots as shown in Fig.6. It is interesting to note that these three co-ordinate systems exhaust all possibilities of representation of the four bases in a 2D system and thus together form a complete set of descriptionsforagivensequence.Thistechniquehasbeenusedby,Nandy[5][6],Raychaudhury and Nandy[7], Nandy and Basak[8], Nandy, Nandy, and Basak[9], Wu, Liew, Yan, and Yang[10], Yao, Nan, and Wang[11] and Ghosh, Roy, Adhya and Nandy[12] for variousapplications.

All of these prescriptions of the rectangular walk had the inherent limitation that sequences of bases that alternated between two types along one axis will cause overlapping paths in one or the other of these representations. Thus a repetitive sequence like GAGAGAGAG will show up in the Nandy plot as only one step along the positive x-direction. Such degeneracies lead to loss of information, and while it was recognised that the chances of two

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

sequences leading to identical plots were minuscule, several authors proposed alternative systems where such degeneracies would not arise. Li, Tang, and Wang[13] used a directed graph method to circumvent this problem, and Randic[14] proposed a condensed representation of DNA sequences that would bypass the degeneracies of graphical representationsaltogether.

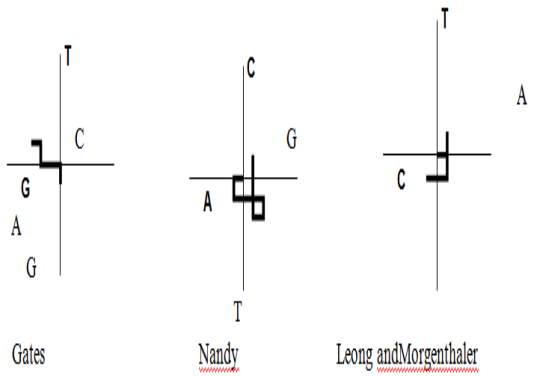


Fig 6:Sequence segment ATGGTGCACC plotted in the axes systems of the 2D graphical representation schemes of Gates[15], Nandy[16] and Leong and Morgenthaler[17].

Graphical representation can also be done by using a binary method. The four bases are split into their three classifications, amino(M)/keto(K), purine(R)/pyrimidine(Y), and weak(W)/strong(S). Then, a value of 1 is ascribed to a R, M, or W type of base in the sequence, and value of 0 is ascribed to a Y, K, or S type of base in the sequence. The graphing is done by placing two horizontal lines, each labelled with a 1 or a 0, one unit distance apart. The binary sequence is then placed along the bottom of the horizontal lines with each number being separated by one unit distance. For each number in the sequence, a dot is place on the corresponding horizontal line, and the dots are connected. There will be three of these characteristic graphs for each DNA sequence at hand. Thus, a sequence such as ATGGTGCACC will have 3 graphs such as the ones shown in Fig. 7. Among those to use this method were Li and Wang[18], Liao and Wang[19], Liao and Ding[20], and Wang andZhang[21].

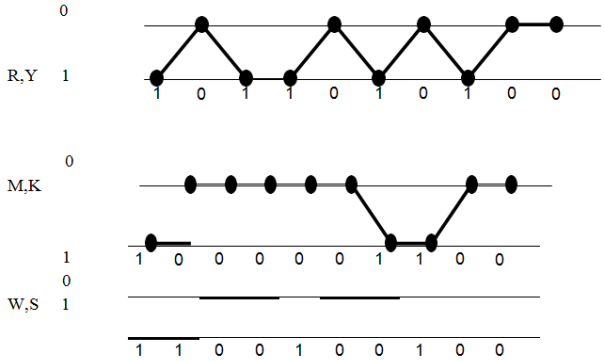


Figure 6. The 2D "two horizontal line" curves of the first 10 bases (ATGGTGCACC) in the human beta globin gene.

Another graphical method proposes the novel idea of utilizing square units called cells. The novel cell design involves a unit square in which the four points in the corners are designated as the four bases A, T, C, and G (Fig.8a). The x-coordinate of the base in the unitcell is obtained by finding which column the individual base is in. By labelling the first column as zero, the even columns are found by the formula (2(i-1)) and the odd columns are found by ((2(i- 1))+1) where i is the base number. Then the y-coordinate is found by whether the base is in the first row or the second row of the cell. In summary, the following designations are given to each base: (2(i-1), 0) = G, (2(i-1), 1) = A, (2(i-1)+1, 0) = C, and (2(i-1)+1, 1) = T where i is the

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

position of the base in the sequence. Then a sequence such as ATGGTGCACC will have a graph such as the one in Fig.7b. This methodology was used by Yao andWang[22].

Figure 8a. Representation of a cell

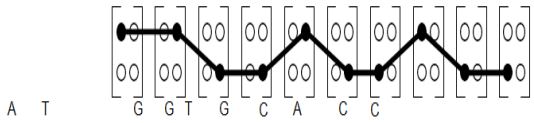


Figure7b: The 2D cell method of Yao and Wang[22] showing the plot of the first 10 bases (ATGGTGCACC) in the human beta globingene.

In another form of graphical representation[23] a square is drawn with the four corners labelled with the four nucleic acid bases. The first base in the sequence at hand is assigned to the location half way between the center of the square and the corner of the square to which the base belongs. The next base in the sequence will be placed half way between the location of the first base and the corner of the square to which it belongs. In summary, each base in the sequence will be placed half way between the position of the preceding base and the corner of the square to which it belongs. This type of representation was done originally by Jeffrey[24] and later by Randic and Zupan[23] in connection with expansion of the scope of visualrepresentations.

3D Graphical representation

A 3D graphical representation for DNA sequences was originally proposed by Hamori and his group (see, e.g. Ref 25), with the aim of facilitating numerical characterization of DNA sequences. A different 3D representation was devised by Randic, Vracko, Nandy, and Basak[26], extending the 2D methods to a 3D graph involved assigning each of the four bases to the corners of a regular tetrahedron. The bases are assigned as follows; A(+1, -1, -1), G(-1, +1, -1), C(-1, -1,+1), and T(+1, +1, +1). The graph is then plotted by placing the first base in the sequence at its correct position; say the first base was an A so its position would be (+1, -1, -1). Then if the next base is a T, it would be placed at (+2, 0, 0). The placement of any base in the sequence will depend on the position of the preceding base in the sequence. This method and its variations wereusedbyRandic,Vracko,Nandy,andBasak[26],LiandWang[27],andYao,Nan,andWang[28].A widely used 3D method of graphical representation was done by first assigning the x and y axis values to the four bases: A to the negative x-axis, G to the positive x-axis, T to the negative y-axis, and C to the positive y-axis. The z-axis value was the number of time that particular base was repeated in the DNA sequence at hand. Thus, the z values for the sequence ATGGTGCACC will be as follows: 1, 1, 1, 2, 2, 3, 1, 2, 2, 3. The points of each base in the sequence are placed in 3D space and a line connects the points. This method and its variations were used by Yuan, Liao, and Wang[29], Liao and Wang[30], Liao, Zhang, Ding, and Wang[31], Zu, Liao, and Ding[32], and Bai, Zhu, andWang[33].

4D Graphical representation

Instead of using a 2D or 3D method, Chi and Ding[34] used a technique involving a novel 4D numerical representation of a DNA sequence. The advantage of a 4D representation is the avoidance of

overlapping and intersecting of the DNA curve with itself. The disadvantage of this method is that the graphical visualization and the ability to directly compare two DNA sequences is lost, which are the advantages of 2D and 3D methods. The idea behind this approach is to obtain the 4D coordinates of the DNA sequence based on the three classifications of DNA bases. It is known that the four nucleic acids A, T, G, and C can be separated on the basis of the distributions of purine-pyrimidine (R/Y), amino-keto (M/K), and weak-strong (W/S) bonds. The classifications are as follows: R=(A, G) and Y=(C, T), M=(A, C) and K=(G, T), W=(A, T) and S=(C, G). A binary technique assigned the value of 1 to Y, K, and S and 0 to R, M, and W. Letting R/Y, M/K, and W/S represent the first three coordinates respectively, the fourth coordinate (i) is represented by the position of the base in the DNA sequence. Therefore, the following assignments were made for the four bases: A(0,0,0,i), G(0,1,1,i), C(1,0,1,i), and T(1,1,0,i). There are 2^3=8 different arrangements of R/Y, M/K, and W/S with {0, 1}, and the 8 arrangements are as follows: I{R,M,W}, II{R,M,S}, III{R,K,W}, IV{R,K,S}, V{Y,M,W}, VI{Y,M,S}, VII{Y,K,W}, VIII{Y,K,S}. Symmetry exists among the arrangements I and VIII, II and VII, III and VI, IV and V. The four vertices of a regular tetrahedron are obtained when the four coordinates are projected along the fourth coordinate to 3D space. This 4D representation is unique since symmetry and rotation do not change thecurve.

Other graphical representations

Several other techniques of representations of DNA sequences have been proposed by different authors. Liao and Wang[35] proposed a 6-dimensional representation, while Randic, Lers, Plavsic, Basak, and Balaban[36] proposed a novel four-color map representation. In this latter method, a sequence of spiralling unit squares is drawn and the first base in the sequence is placed in the centre of the spiral. The rest of the bases in the sequence then spiral clockwise around this first base. After the last base has been placed, the map is sectioned off according to the four bases and each base is given one color. By graphing in this manner, it is possible to see regions in the map belonging to one particular base and thus get an idea of basedistribution.

Conclusion:

This survey outlines the main mathematical results and models used by researchers to discuss DNA deformability and structure at the macroscopic level. There are few important topics in DNA research that do not naturally fit under the headings above, and one of them is the connection between DNA denaturation and supercoiling. It is known that although DNA molecule is stable under the conditions mimicking the intracellular environment, the base-pairing interaction can be disrupted, in a process called denaturation, The energy required for DNA denaturation depends on base-pair composition and have been determined very accurately in calorimetric experiments.

REFERENCES

[1]C.E.Shannon,“Amathematicaltheoryofcommunication,” BellSystemTechnicalJournal, vol.27,pp.379–423,1948.

[2] Yu.P. Lysov, V.L. Florentiev, A.A. Khorlyn, K.R. Khrapko, V.V. Shick, A.D. Mirzabekov,Determination of the nucleotide sequence of DNA using hybridization with oligonucleotides. A newmethod, Dokl. Acad. Sci. USSR 303 (1988)1508{151
J. Blazewicz et al. / Discrete Applied Mathematics 98 (1999).

[3]Nandy, A. Current Science **1994**, 66, 309.-12

[4] Leong and Morgenthaler, Comput. Appl. Biosci. **1995**, 11, 503.

[5] Nandy, A. Comput. Appl. Biosci. **1996**, 12, 55.
[6] Nandy, A. Internet Electron. J. Mol. Des. **2002**, 1, 545.

[7] Raychaudhury, C.; Nandy, A. J. Chem. Inf. Comput. Sci. **1999**, 39, 243.
[8] Nandy, A.; Basak, S.C. J. Chem. Inf. Comput. Sci. **2000**, 40, 915.

[9]Nandy, A.; Nandy, P.; Basak, S. C. Internet Electron. J. Mol. Des. **2002**, 1, 367.

[10]Wu, Y.; Liew, A. W.; Yan, H.; Yang, M. Chem. Phys. Lett.**2003**, 367, 170.

[11] Yao, Y.; Nan, X.; Wang, T. J. Mol. Struct. (Theochem) **2006**, 764, 101.
[12] Ghosh, S.; Roy, A.; Adhya, S.; Nandy, A. Current Science **2003**, 84, 1534.

[13]Li, C.;Tang, N.;Wang, J. J.Theo.Biol. in press doi:10.1016/j.jtbi.2005.11.023.

[14] Randic, M. J. Chem. Inf. Comput. Sci. **2000**, 40, 50.

[15] Gates, M. A. J. Theor. Biol. **1986**,119, 319.
[16] Nandy, A. Current Science **1994**, 66, 309.
[17] Leong and Morgenthaler, Comput. Appl. Biosci. **1995**, 11, 503.

[18]Li, C.; Wang, J. Combinatorial Chem. & High Throughput Screening **2003**, 6, 795.

[19] Liao, B.; Wang, T. J. Comput. Chem. **2004**, 25, 1364.

[20]Liao, B.; Ding, K. J. Comput.Chem. **2005**, 26, 1519, 1523.
[21]Wang, J.; Zhang, Y. Chem. Phys. Lett.**2006**, 423, 50.

[21] Yao, Y.; Wang, T. Chem. Phys. Lett. **2004**, 398, 318.

[22]Yao, Y.; Wang, T. Chem. Phys. Lett.**2004**, 398, 318.

[23] Randic, M.; Zupan, J. SAR and QSAR Environ. Res. **2004**, 15, 191.

[24]Jeffrey, H. J., Nucleic Acids Res. **1990**, 18, 2163.

[25] Hamori, E.; Ruskin, J. J.Biol.Chem. **1983**, 258, 1318.

[26]Randic, M.; Vracko, M.; Nandy, A.; Basak, S. C. J. Chem. Inf. Comput.Sci. **2000**, 40, 1235.
[27] Li, C.; Wang, J. Combinatorial Chemistry & High Throughput Screening **2004**, 7, 23.
[28]Yao, Y.; Nan, X.; Wang, T. Chem. Phys. Lett.**2005**, 411, 248.

[29]Yuan, C.; Liao, B.; Wang, T. Chem. Phys. Lett.**2003**, 379, 412.

[30] Liao, B.; Wang, T. J. Mol. Struct. (Theochem) **2004**, 681, 209.

[31] Liao, B.; Zhang, Y.; Ding, K.; Wang, T. J. Mol. Struct. **2005**, 717, 199.

[32] Zhu, W.; Liao, B.; Ding, K. J. Mol. Struct. **2005**, 757, 193.

[33]Bai, F.; Zhu, W.; Wang, T. Chem. Phys. Lett.**2005**, 408, 258.

[34]Chi, R.; Ding, K. Chem. Phys. Lett.**2005,** 407, 63.

[35] Liao, B.; Wang, T. J. Chem. Inf. Comput. Sci**. 2004**, 44, 1666.

[36]Randic, M.; Lers, N.; Plavsic, D.; Basak, S. C.; Balaban, A. T. Chem. Phys. Lett.**2005**, 407,
205.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr