EXTENDED DNA PLAYFAIR WITH IMAGE STEGANOGRAPHY

Varsha Tyagi [1] and Ravindra Chauhan[2]

[1] *M.Tech (CSE), Department of Computer Science & Engineering, MIET, Meerut*

[1]*Tyagivarsha17@gmail.com*

[2]*Department of Computer Science & Engineering, MIET, Meerut*

---***---

Abstract: *Text to image encryption technique is a new and promising direction in cryptographic research .it would be more effective when this technique merge with another two techniques of encryption, such as play fair cipher and DNA cryptography. cryptography referred almost exclusively to encryption, which is the process of converting ordinary information (called plaintext) into unintelligible gibberish (called cipher text). Decryption is the reverse, in other words, moving from the unintelligible cipher text back to plaintext. A cipher is a pair of algorithms to create the encryption and the reversing decryption. This paper discusses a significant modification to the old play fair cipher by introducing DNA - based ,amino acid based structure and encrypted it into an image.*

Key Words: Cryptography, DNA Cryptography, Text Encryption, Text Decryption, Image Encryption, Image Decryption, Playfair Algorithm.

1.INTRODUCTION

The concept of using DNA computing in the fields of cryptography and steganography is a possible technology that may bring forward a new hope for powerful, or even unbreakable, algorithms. The main purpose behind our work is to discover new fields of encoding the data in addition to the conventional used encryption algorithm in order to increase the concept of confusion and therefore increase security. In our work, we applied the conversion of character form or binary form of data to the DNA form and then to amino acid form and then encrypted with an image . In this study, a binary form of data, such as plaintext messages, or images are transformed into sequences of DNA nucleotides. Subsequently, these nucleotides pass through a Play fair encryption process based on amino acids structure. The fundamental idea behind this encryption technique is to enforce other conventional cryptographic algorithms which proved to be broken, and also to open the door for applying the DNA and Amino Acids concepts to more conventional cryptographic algorithms to enhance their security features by adding steganography technique. DNA cryptography is a relatively new field of cryptography which is being explored for advancement in the existing cryptography makes use the concept of DNA computing.

2. CRYPTOGRAPHY

2.1. DNA: In terms of biology, a Deoxyribonucleic acid (DNA) is the master molecule whose structure encodes all the information needed to create and direct the chemical machinery of life [2]. In 1953, the structure of DNA was correctly predicted by Watson and Francis Crick that DNA molecule consists of two long polynucleotide produces some redundancy in the code: most of the amino acids being encoded by more than one codon. The genetic code can be expressed as either RNA codons or DNA chains each of these chains is known as a DNA chain, or a DNA strand which is made from simple subunits, called nucleotides. The DNA segments that carry this genetic information are called genes, but other DNA sequences have structural purposes, or are involved in regulating the use of this genetic information.

The DNA double helix is stabilized by hydrogen bonds between the bases attached to the two strands. The four bases found in DNA are adenine (abbreviated A), cytosine (C), guanine (G) and thymine (T). These four bases are attached to the sugar/phosphate to form the complete nucleotide, as shown for adenosine monophosphate[1].

2.2 Genetic Code:

The genetic code consists of 64 triplets of nucleotides. These triplets are called codons. With three exceptions, each codon encodes for one of the 20 amino acids used in the synthesis of proteins. That codons. RNA codons occur in messenger RNA (mRNA) and are the codons that are actually "read" during the synthesis of polypeptides (the process called translation). But each mRNA molecule acquires its sequence of nucleotides by transcription from the corresponding gene. The DNA Codons is read the same as the RNA codons Except that the nucleotide thymine (T) is found in place of uridine(U). So in DNA codons we have (TCAG) and in RNA codons, we have (UCTG).

2.3 Transcription And Translation:

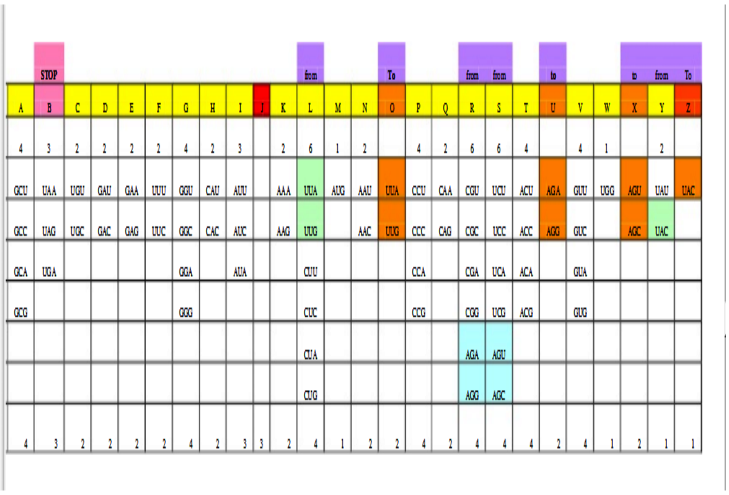
A gene is a sequence of DNA that contains genetic information and can influence the phenotype of an organism. Within a gene, the sequence of bases along a DNA strand defines a messenger RNA sequence, which then defines one or more protein sequences. The relationship between the nucleotide sequences of genes and the amino acid sequences of proteins is determined by the rules of translation, known collectively as the genetic code. The genetic code consists of three-letter 'words' called codons formed from a sequence of three nucleotides (e.g. ACT, CAG, TTT). In transcription, the codons of a gene are copied into messenger RNA by RNA polymerase. This RNA copy is then decoded by a ribosome that reads the RNA sequence by base-pairing the messenger RNA to transfer RNA, which carries amino acids. Since there are 4 bases in 3-letter combinations, there are 64 possible codons (43 combinations). These encode the twenty standard amino acids, giving most amino acids more than one possible codon. There are also three 'stop' or 'nonsense' codons signifying the end of the coding region; these are the TAA, TGA and TAG codons.

3. DNA-BASED PLAYFAIR ALGORITHM

Playfair used to be applied to English alphabet characters of plaintext. It was unable to encode any special characters or numbers which is considered a drawback that enforces the sender to write everything in the English letters. This problem appears while sending numerical data, equations or symbols. But in our algorithm, we can use any numbers, special characters or even spaces in our plaintext. The encryption process starts by the binary form of data (message or image) which is transferred to DNA form according to Table 1. Then the DNA form is transferred to the Amino acids form according to Table 2 which is a standard universal table of Amino acids and their codons representation in the form of DNA table.

Bit1	Bit2	DNA
0	0	A
0	1	C
1	0	G
1	1	U

Table-1: Standardized table for DNA representation of bits



	STOP										from			To			from	from		to			to	from	To
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
4	3	2	2	2	2	4	2	3		2	6	1	2		4	2	6	6	4		4	1		2	
GCU	UAA	UGU	GAU	GAA	UUU	GGU	CAU	AUU		AAA	UUA	AUG	AAU	UUA	CCU	CAA	CGU	UCU	ACU	AGA	GUU	UGG	AGU	UAU	UAC
GCC	UAG	UGC	GAC	GAG	UUC	GGC	CAC	AUC		AAG	UUG		AAC	UUG	CCC	CAG	CGC	UCC	ACC	AGG	GUC		AGC	UAC	
GCA	UGA					GGA		AUA			CUU				CCA		CGA	UCA	ACA		GUA				
GCG						GGG					CUC				CCG		CGG	UCG	ACG		GUG				
											CUA						AGA	AGU							
											CUG						AGG	AGC							
4	3	2	2	2	2	4	2	3	3	2	4	1	2	2	4	2	4	4	4	2	4	1	2	1	1

Table -2: Standard alphabetical table in correspondence with its codons

Amino acid codon is considered about, there are 20 amino acids in addition to 1 start and 1 stop. But, we need 25 letters for 5*5 playfair matrix formation (ignoring „J‟ due to its lowest frequency)[3]. We already have the standard alphabetical table for the amino acid codons(fig. table2). When the coupling of "three" of the generated codon is formed, after the formation of DNA representation of bits. The corresponding alphabet of its respective codon is noted down. The resultant alphabets are further taken for performing the playfair cipher action. The one thing to be noticed in this table is the number of ways of formation of codons in several alphabets. This number of terms is noted in the correspondence column of that particular alphabet so called "ambiguity".

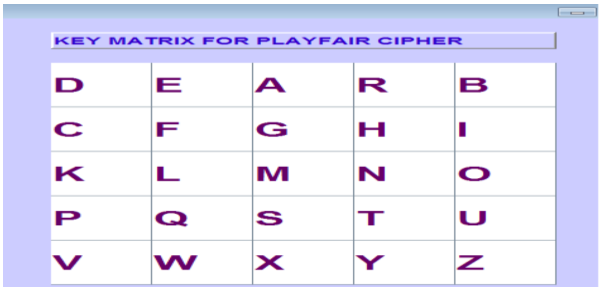
4.STEGANOGRAPHY

Steganography is the art and science of writing hidden messages within another seemingly innocuous message. Modern Steganography techniques using digital information offering wonderful opportunities not only to hide information, but also to develop a general theoretical framework for hiding different kinds of data such as sound tracks, images, videos, and even 3D objects [4].

5.PROPOSED WORK

Encryption Module:

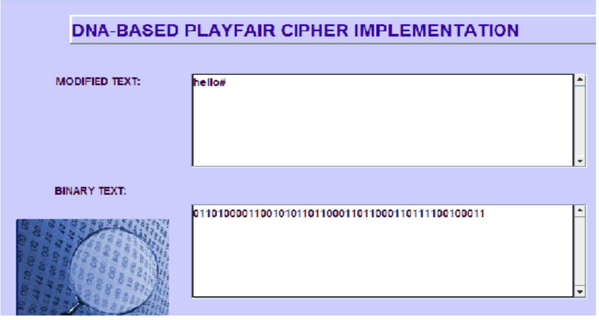
1.Browse Module – The file that has to be encrypted has to be fetched from the system and its text has to be read.



KEY MATRIX FOR PLAYFAIR CIPHER

D	E	A	R	B
C	F	G	H	I
K	L	M	N	O
P	Q	S	T	U
V	W	X	Y	Z

Fig -1



DNA-BASED PLAYFAIR CIPHER IMPLEMENTATION

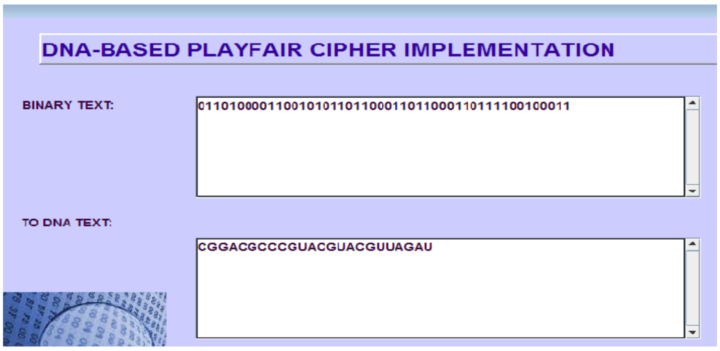
MODIFIED TEXT: hello#

BINARY TEXT: 011010000110010101101100011011000110111100100011

Fig -2

2.Encrypt Text Module – This module is used for the encryption process. It has the following **sub-modules**

3.DNA Converter: The binary form is converted into the DNA bases form in this module.



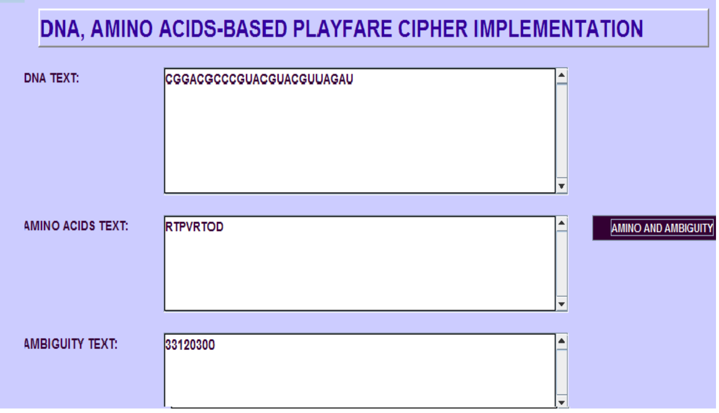
DNA-BASED PLAYFAIR CIPHER IMPLEMENTATION

BINARY TEXT: 011010000110010101101100011011000110111100100011

TO DNA TEXT: CGGACGCCCGUACGUACGUUAGAU

Fig -3

a. **4.Amino Acids Converter:** The DNA form is further encrypted in the Amino Acids form in this module.



DNA, AMINO ACIDS-BASED PLAYFAIR CIPHER IMPLEMENTATION

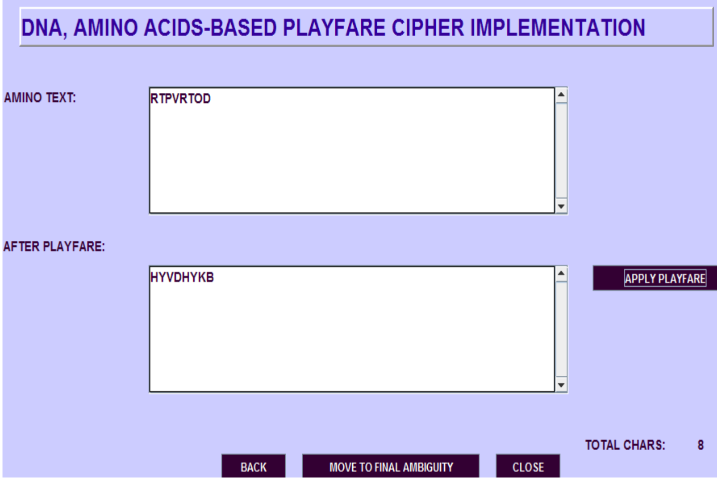
DNA TEXT: CGGACGCCCGUACGUACGUUAGAU

AMINO ACIDS TEXT: RTPVRTOD

AMBIGUITY TEXT: 33120300

Fig -4

b. **5,Playfair Implement:** This module will implement the Playfair algorithm with the secret key that the user provides.



DNA, AMINO ACIDS-BASED PLAYFARE CIPHER IMPLEMENTATION

AMINO TEXT: RTPVRTOD

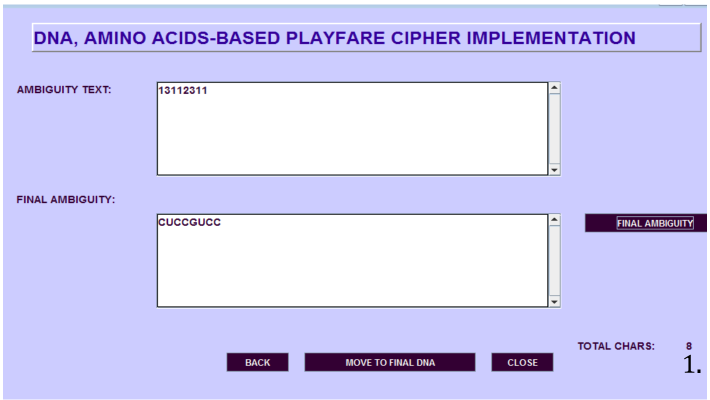
AFTER PLAYFARE: HYVDHYKB

TOTAL CHARS: 8

c.

Fig -5

6.Final Append: This module will further convert the encrypted message into the DNA form and add the ambiguities to further convert into final encrypted form.



DNA, AMINO ACIDS-BASED PLAYFARE CIPHER IMPLEMENTATION

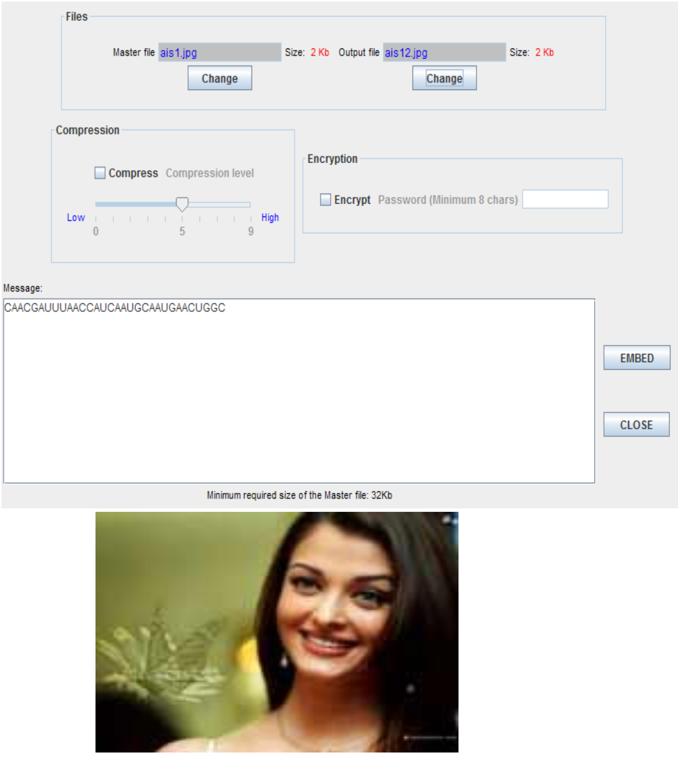
AMBIGUITY TEXT: 13112311

FINAL AMBIGUITY: CUCCGUCC

TOTAL CHARS: 8

Fig -6

7. Then cipher text will go to image Steganography module.



Files

Master file ais1.jpg Size: 2 Kb Output file ais12.jpg Size: 2 Kb

Compression

□ Compress Compression level

Encryption

□ Encrypt Password (Minimum 8 chars)

Message:

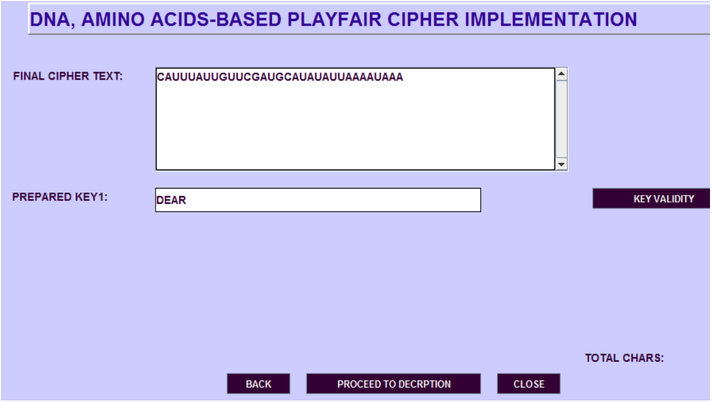
CAACGAUUUAACCAUCAAUGCAAUGAACUGGC

Minimum required size of the Master file: 32Kb

Fig -7

Decryption Module:

1.Extract cipher text from image Steganography module.



DNA, AMINO ACIDS-BASED PLAYFAIR CIPHER IMPLEMENTATION

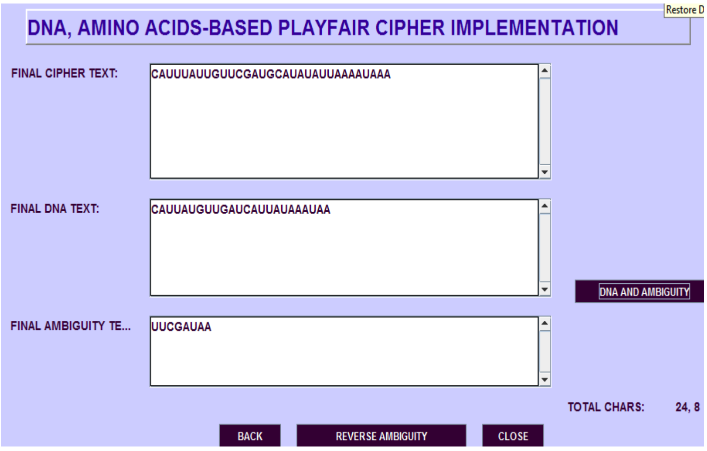
FINAL CIPHER TEXT: CAUUUAUUGUUCGAUGCAUAUAUUAAAAUAAA

PREPARED KEY1: DEAR

TOTAL CHARS:

Fig -8

1. **2.DNA I:** The module will convert the Initial DNA to the Amino Acid form with the help of the ambiguity numbers.



DNA, AMINO ACIDS-BASED PLAYFAIR CIPHER IMPLEMENTATION

FINAL CIPHER TEXT: CAUUUAUUGUUCGAUGCAUAUAUUAAAAUAAA

FINAL DNA TEXT: CAUUAUGUUGAUCAUUAUAAAUAA

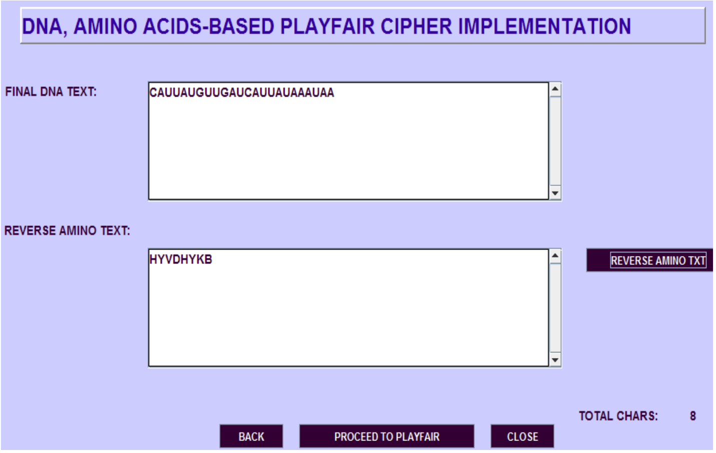
FINAL AMBIGUITY TE... UUCGAUAA

TOTAL CHARS: 24, 8

Fig -9

2.

3. **3.Amino Acid I:** The DNA form is then converted into the Amino Acid form in this module.



DNA, AMINO ACIDS-BASED PLAYFAIR CIPHER IMPLEMENTATION

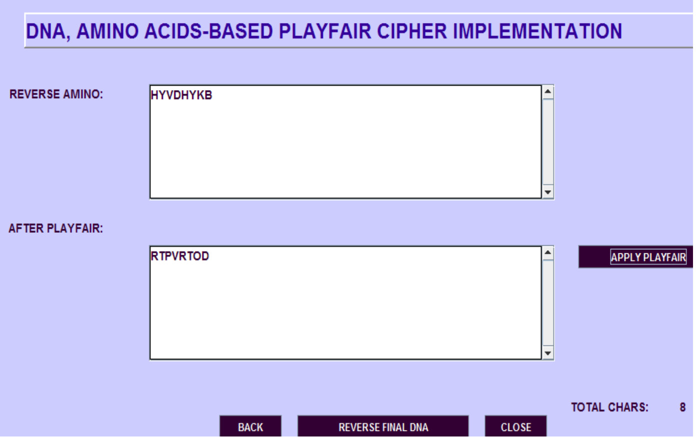
FINAL DNA TEXT: CAUUAUGUUGAUCAUUAUAAAUAA

REVERSE AMINO TEXT: HYVDHYKB

TOTAL CHARS: 8

Fig -10

4.Playfair Implement: The Playfair Cipher is implemented with the same secret key to reverse the encryption process.



DNA, AMINO ACIDS-BASED PLAYFAIR CIPHER IMPLEMENTATION

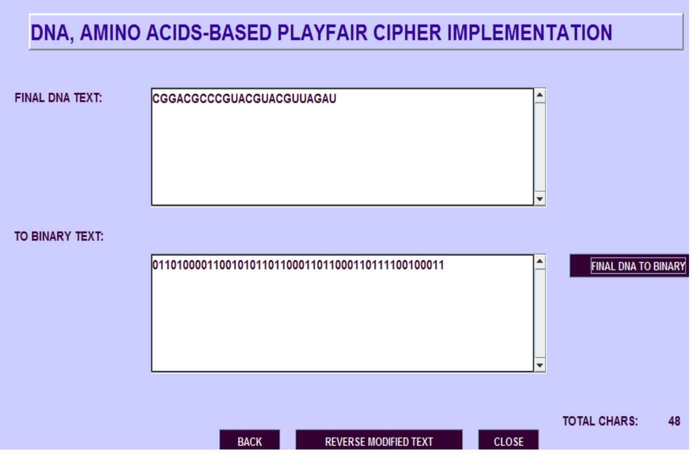
REVERSE AMINO: HYVDHYKB

AFTER PLAYFAIR: RTPVRTOD

TOTAL CHARS: 8

Fig -11

5.Final Blow: This module converts the DNA module into the binary and then the original message is obtained.



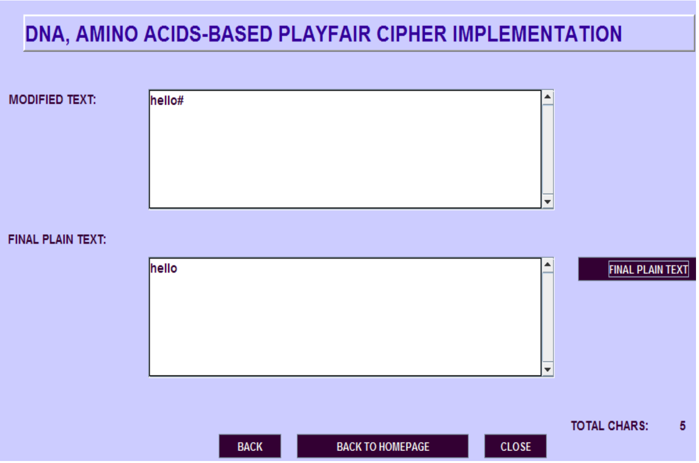
DNA, AMINO ACIDS-BASED PLAYFAIR CIPHER IMPLEMENTATION

FINAL DNA TEXT: CGGACGCCCGUACGUACGUUAGAU

TO BINARY TEXT: 011010000110010101101100011011000110111100100011

TOTAL CHARS: 48

Fig -12



DNA, AMINO ACIDS-BASED PLAYFAIR CIPHER IMPLEMENTATION

MODIFIED TEXT: hello#

FINAL PLAIN TEXT: hello

TOTAL CHARS: 5

Fig -13

6.RELATED WORKS

DNA based cryptography helps in bringing uniqueness to the concept of conventional cryptography for each user. In 1994 Adleman [5] was the first to use the concept of DNA computing to solve directed Hamiltonian path problem. This revolutionized how DNA was used in the field of communication. In 1995 Lipton [6] showed a method which could use DNA computing to solve NP-complete problem. In 1995, D. Boneh, C. Dunworth and R. Lipton [7] showed it was possible to crack the Data Encryption Standard (DES) by using DNA computing methods. In 2012 A. Atito, A. Khalifa ,S. Z. Rida gives the DNA-Based Data Encryption and Hiding Using Playfair and Insertion Techniques[8],and recently in 2016 Ayushi kansal, Shruti sneha, Manish kumar Patel show the method of data encryption by using DNA cryptography with playfair cipher technique[3].

CONCLUSIONS

The objective that a new field of encoding the data has to be explored has been achieved. The requirement of a simple algorithm implementation for DNA cryptography is achieved. The output cipher is not only an intermediate for DNA cryptography but is in itself hard to break cipher that falls in the category of image with DNA Cryptography, The project is very versatile and reliable as many amendments will be possible at any time of computing because of the support for a number of intermediary processes during encryption. The project has good current market value as it is important from the view of research and has an extremely bright future as regards the importance of DNA Cryptography with text to image conversion in an era where the other mostly used algorithm eg: DES and MD5 have been broken. All this encryption procedure increase the complexity of the project that makes it hardly unbreakable.

REFERENCES

[1]http://en.wikipedia.org.

[2] Bruce Alberts et al. Molecular Biology of The Cell Fifth Edition, 2008.

[3]Ayushi kansal et al. "Modifying Playfair Cipher by Using DNA and Amino Acids" International Journal of Education and Science Research Review, Volume-3,Issue-2,Science Research Review E-ISSN 2348-6457.

[4]K. RAMA et al. ,SURVEY AND ANALYSIS OF 3D STEGANOGRAPHY,International Journal of Engineering Science and Technology (IJEST),Vol.3,2011.

[5]L.M. Adleman,"Molecular computation of solutions to combinational problems",Science, Vol.266, pp. 1021–1024,1994.

[6]R.J.Lipton,"Using DNA to solve NP-complete problems",Science, Vol. 268, pp. 542–545, 1995.

[7]D. Boneh, C. Dunworth, R. Lipton,"Breaking DES using a molecular computer",American Mathematical Society, pp.37–65, 1995.

[8] A. Atito, A. Khalifa ,S. Z. Rida "DNA-Based Data Encryption and Hiding Using Playfair and Insertion Techniques", J. of Commun. & Comput. Eng. ISSN 2090- 6234,Volume 2, Issue 3, 2012, Pages 44: 49.