Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

# A Mechanism for Combining Hill Cipher and Steganography Techniques in DNA Sequences to Enhance Information Security

Siham oleiwi Tuama

Ministry of Education / Babylon Education Directorate, Babylon, Iraq -E-mail: Mscsiham1994@gmail.com

*Corresponding author E-mail: Mscsiham1994@gmail.com

https://doi.org/10.46649/fjiece.v4.2.17a.26.9.2025

**Abstract.** *This research integrates the Hill Cipher encryption algorithm with DNA-based steganography as a form of dual protection. The plugins and modules mask the cryptex as DNA sequences. The proposed methods presented the full accuracy of the data recovery with low time and full recovery. The encryption and decryption steps are of low time and full recovery. The effective masking and absence of the cryptex within the cover DNA, as nominal 50%, and the absence of the cryptex confirm the absence of cryptex and detection. The proposed model gives an extremely high potential for secure communication, in high confidentiality environments, like in the exchange of medical and governmental data.*

***Keywords :*** Information Security, Hill Cipher, Steganography, DNA Sequences .

## 1. INTRODUCTION

The significant developments in communication and information technology over the last few decades have led to an increase in information security concerns. Ensuring the security of data during transmission is essential. Because the network is open and sensitive data must be exchanged, many assaults are more likely to take place. Data security is therefore essential in fields involving data communications [1].Steganography and cryptography are the two most popular and significant techniques for preserving confidentiality. These processes, in spite of their variations, both ensure the confidentiality, security, and integrity of data [2]. Cryptography is a collection of methods based on mathematical concepts. It converts understandable data into a format that unauthorized people cannot understand. In contrast, steganography conceals data under a cover item to allow for untraceable transmission [2]. Cover items, such as text files, images, videos, and DNA sequences, are used to hide the hidden data. Steganography's main objective is to conceal private information to avoid suspicion [3]. Keeping individuals in the dark about the existence of secret data is essential. Cryptography, on the other hand, offers the capacity to transmit information in a way that prevents unauthorized parties from accessing it. Combining steganography and cryptography results in a two-layer of security. After cryptography jumbles the secret information, steganography hides the encrypted data in a cover item [4].

## 2- RELATED WORK

Various new research works emphasize information security through the study of over encryption and steganography techniques. Ameen et al. [5] suggested a hybrid method integrating the Hill Cipher with modified DNA sequences yielding high security due to a double-layer approach. Huang et al. [6] explored the next stages of DNA encryption and its use with expanded universal base pairs which theoretically supports and advances DNA based encryption techniques. Li et al. [7] formulated a hyperchaotic Lorenz

and DNA-encoded image encryption algorithm which improved security of digital images. Nasr et al. [8] implemented strong audio steganography based on images which were enciphered demonstrating the hidden potential of steganography with multimedia files. Mfungo et al. [9] proposed elevated image encryption techniques where the Hill Cipher is used together with the Kronecker XOR product and the Sigmoid Logistic Map which is a development as compared to using the Hill Cipher solo.

The value of these works is in their approaches, but a majority pay attention to one encryption technique, whether Hill Cipher or DNA based encryption, or to multimedia steganography. The proposed method sets itself apart through the use of Hill Cipher in conjunction with DNA based steganography to achieve high security, large data capacity, and attack resistance. The following table 1 compares the proposed method with these existing techniques.

**Table 1: Comparison of the Proposed Method with Recent Related Work**

| References / Remarks | Resistance to Attack | Speed | Capacity | Security level | Method |
|---|---|---|---|---|---|
| Traditional encryption; see Mfungo et al., 2023 | Low | High | Medium | Medium | Hill Cipher |
| Standard modern encryption, no Steganography; Huang et al., 2025 | Very High | Medium | Medium | Very High | AES |
| Uses DNA encoding for data; Huang et al., 2025 Li et al., 2024 | Medium | Medium | High | High | DNA-based Encryption |
| Hides encrypted data in images; Nasr et al., 2024 | Medium | Medium | High | Medium-High | Image Steganography (with encryption) |
| Hides encrypted data in audio; Nasr et al., 2024 | Medium | Medium | Medium | Medium-High | Audio Steganography (with encryption) |
| Combines Hill Cipher+ DNA-based Steganography; Siham Oleiwi Tuama, 2025 | High | Medium | High | Very High | Proposed Method: Hill Cipher + DNA Steganography |

## 3. STEGANOGRAPHY

Steganography is characterized as a collection of methods aimed at hiding confidential information in a cover item so that people cannot see it [10]. The Greek word "covered writing" is where the name "steganography" originates. Steganography cFan conceals any kind of data, including text, pictures, videos, and anything else that can be represented as a bit-stream, behind a cover file. Additionally, there are other kinds of cover objects, including DNA sequences, text files, photos, and movies. The primary goal of every steganography technique is to enable absolutely undetected, effective, and safe communication.Any object containing a significant amount of redundant data is regarded as ideal, as it becomes challenging to detect any alterations made to those redundant bits [11]. In this context, redundancy refers to the excess data within a cover object that offers a level of detail or precision beyond what is necessary for its intended function or visual representation. Therefore, a steganographic method is considered secure if the modified (stego) object remains indistinguishable from the original cover object, whether by human perception or automated systems .

 

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

## 4. DNA

### BIOLOGICAL OVERVIEW

molecule known as deoxyribonucleic acid, or DNA, contains the genetic information necessary for all known living creatures, including viruses, to develop and function. The information contained in DNA is encoded by a group of four chemical bases called nucleotides: adenine (A), guanine (G), cytosine (C), and thymine (T) [12].

### 4-1 DNA BASED STEGANOGRAPHY

DNA computing is a subfield of computing that uses deoxyribose nucleic acid (DNA) to carry out computer activities. Numerous biological characteristics of DNA sequences can be used to address unsolvable issues [14]. Thus, secure and reliable cryptography and steganography approaches are obtained by taking advantage of the biological characteristics of DNA sequences. DNA steganography is a very new and advanced scientific field that was founded in 1999. The secret information was effectively transported by the DNA sequences for several reasons. DNA has a huge storage capacity; one gram of DNA can store over 108 terabytes [13].

Furthermore, DNA sequences' intricacy and unpredictability offer reliable security, resilience, and protection [15]. Additionally, the ease of converting data into a DNA sequence and vice versa. Original DNA eventually becomes indistinguishable from modified DNA [10]. Converting each DNA base into two binary bits is a popular and simple method of converting data into a DNA format and vice versa. The binary representation of DNA is seen in Table 2. Nevertheless, a few of the techniques created their own conversion guidelines.

**Table 2 :DNA-binary representation**

| DNA base | Binary representation |
|---|---|
| A | 00 |
| C | 11 |
| G | 10 |
| T | 01 |

## 5. HILL CIPHER

The famous substitution cipher known as the Hill cipher was developed by Lester S. Hill [16][17]. To encrypt a block of the plaintext of size n, the Hill Cipher approach needs a key matrix (k) with entries between 0 and (b-1), but the determinant needs to be approximately prime to b. The base in this case is b. Then, in order to generate the Base in this instance, b, we encrypt the plaintext vector, p. The cipher text vector, c, is then created by encrypting the plaintext vector, p, using the linear algebra equation below:

$$c = (k*p) \bmod b$$

We must first determine the inverse of the key matrix, k-1, which must be invertible, in order to decrypt the cipher text vector, c. The plaintext vector, p, can then be calculated using the mathematical model:

$$p= (k\text{-}1*c) \bmod b$$

Since the inverse of the matrix must be calculated during the decryption process, the key matrix should be invertible. As a result, while selecting key matrices, fewer possibilities are available. A temporary solution

to this problem would be to use involutory matrices, in which the matrix's inverse equals the matrix itself. The long-term solution is described in the section that follows.

## 6.METHODOLOGY

In this research, a secure encryption and concealment system combining the Hill Cipher algorithm for text encryption and DNA steganography is proposed and implemented to enhance the security of sensitive data transmission. The system goes through the following stages :

1. **Text Encryption Using The Hill Cipher**

A traditional cipher based on linear algebra is the Hill Cipher . In this system, the input text (plaintext) is converted to numbers according to the letter encoding (A=0 to Z=25), then divided into three blocks. Each block is encrypted by multiplying it by a 3×3 key matrix, followed by modulo 26. If the text length is not divisible by 3, padding characters of the type 'X' are added to complete the length.

2. **Converting the ciphertext to DNA representation**

After obtaining the ciphertext, each character is converted to its ASCII value and then to an 8-bit binary representation. The resulting binary string is divided into two-bit pairs, and each pair is converted to a DNA base using the following rules:

- 00 → A
- 01 → T
- 10 → G
- 11 → C

This step produces a DNA string consisting of four letters (A, T, G, C) representing the encrypted data.

3. **Generating a Cover DNA Strand**

A cover DNA string is generated with a length approximately twice that of the encrypted string. The string consists of random symbols (A, T, G, C) and serves as a medium for hiding the secret data.

4. **Hiding the Ciphertext Within the Cover Strand**

The encrypted DNA string is inserted within the cover string at specific locations (in this model, inserted at the beginning of the string), producing a hidden DNA string containing the secret data without any apparent differences, making it more difficult to decipher the message, for this demonstration, the ciphertext DNA is placed at the start of the cover strand for ease of demonstration and decryption. That said, as described, the system is adaptable and can be adjusted to accommodate random or spread embedding of segments of ciphertext throughout the cover DNA sequence. This improvement would help increase the strength of the system and help reduce the predictability of the system to steganalysis attacks.

5. **Extracting and Decrypting the Ciphertext**

To retrieve the data, a specific number of bases are extracted from the beginning of the hidden DNA string (the same length as the ciphertext), then converted to binary representation, and from there to ciphertext. The inverse of the matrix (mod 26) is used to decrypt the text using the Hill Cipher, and recover the original text. .Figure (1) show the proposal structures

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

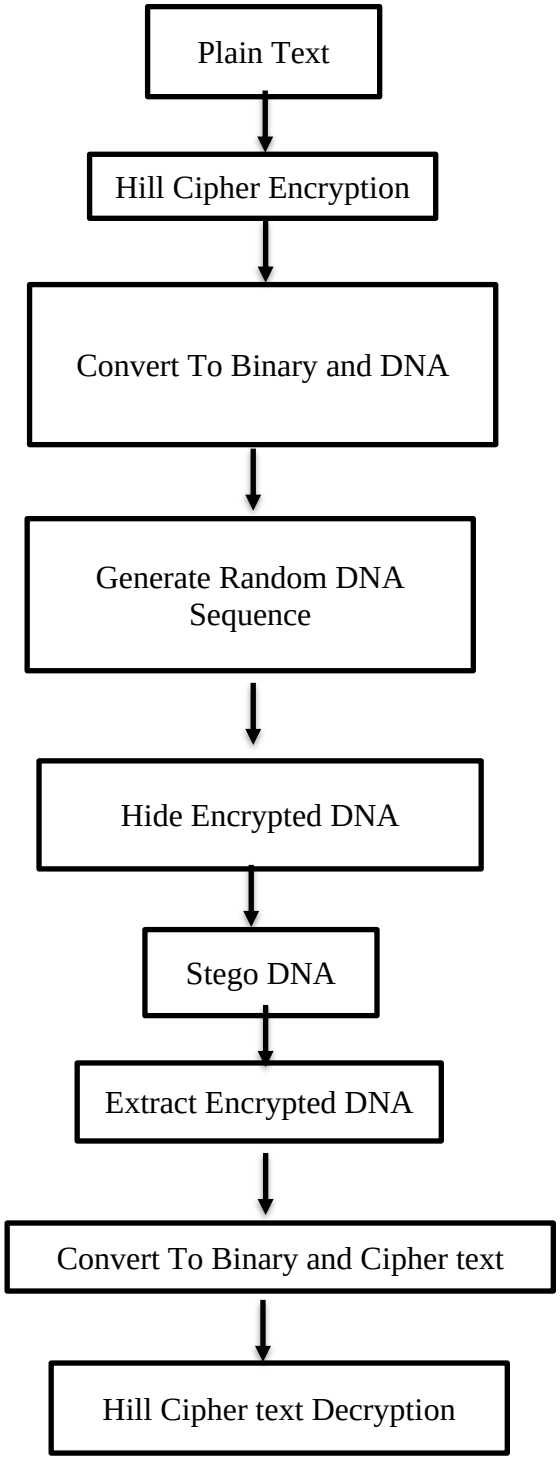


Fig 1: the proposal structures

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

## 7. RESULTS AND DISCUSSION

### *7.1 TIME ENCRYPTION AND DECRYPTION*

The time performance of the proposed encryption system was measured by recording the encryption time using the Hill Cypher algorithm, and the full decryption time after going through the process of concealment within the DNA chain and extraction from it. The results showed that the encryption time was 0.01900 seconds, a very short time that indicates the efficiency of the Hill algorithm in processing words quickly. The decryption time, which also includes converting text to and from DNA form, reached 0.01063 seconds, which is also low time and indicates the usability of the system in environments that require a near-real-time response. These findings indicate that the proposed system achieves excellent performance in terms of speed and is suitable for real-time or near real-time secure communication applications.Figure2 show the result in matlab

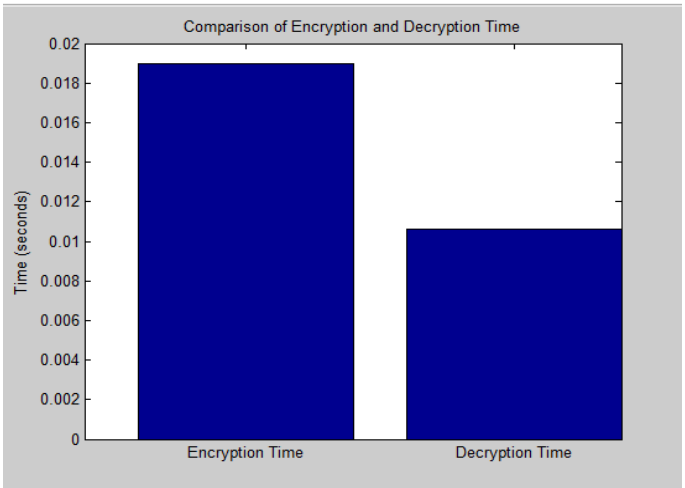


Fig2: Show Result Time Encryption and Decryption

### *7.2 RECOVERY ACCURACY*

The recovery accuracy of the proposed hybrid encryption and DNA-based steganography system was evaluated by comparing the decrypted text with the original plaintext. The accuracy was calculated as the percentage of correctly recovered characters relative to the original message length. Experimental results demonstrated a 100.00% recovery accuracy, indicating that the system is capable of preserving the full integrity of the original data through all processing stages—encryption, DNA transformation, embedding, extraction, and decryption. This high accuracy confirms the robustness and reliability of the implemented method, making it suitable for secure applications where data fidelity is critical. The figure 3 shows the results of the *recovery* accuracy scale through a two-column vertical drawing. The first column (in blue) represents the number of characters that were retrieved correctly after going through the stages of encryption, conversion to DNA, concealment, extraction and decryption, and this number reached 13 characters. The second column, which represents the number of characters that were not recovered correctly or lost, was valued zero (0), indicating that the system managed to retrieve the entire content of the original text without any loss of data. These results confirm that the proposed system has high credibility in restoring original data. The absence of errors or loss in the data (0 misleagus) means that the system maintains the

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

integrity of the entire data. This indicator can be seen as strong evidence that the combination of Hill Cipher and maskin techniques within DNA strings provides effective protection without affecting text retrieval

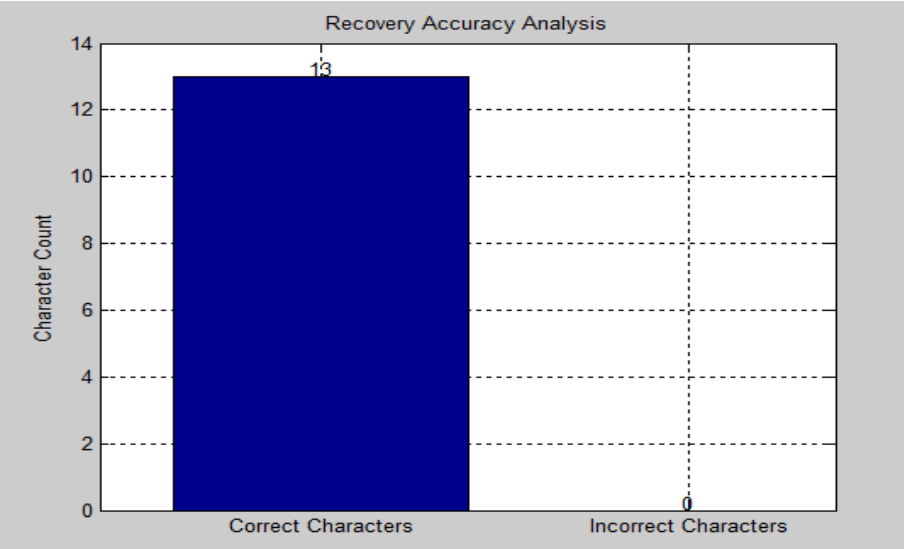


fig 3: The results of the Recovery Accuracy

*7.3 THE EMBEDDING RATIO*

The Embedding Ratio is used to measure the ratio of hidden data (such as encrypted text) to the total amplitude of a hidden medium (such as a DNA string or image). In this study, the embedding ratio was calculated by dividing the length of the DNA-encoded ciphertext by the length of the total DNA cover sequence. Experimental results revealed an embedding ratio of 50.00%, meaning that half of the available cover sequence was used to embed the secret information.This ratio reflects a balanced trade-off between payload capacity and imperceptibility, ensuring that the hidden data remains concealed within the cover without overloading it.Such a ratio is considered optimal in steganographic systems, as it maintains a sufficient level of cover redundancy to resist detection while maximizing the embedding efficiency. Figure (4) shows the distribution of the embedding ratio within the concealment medium

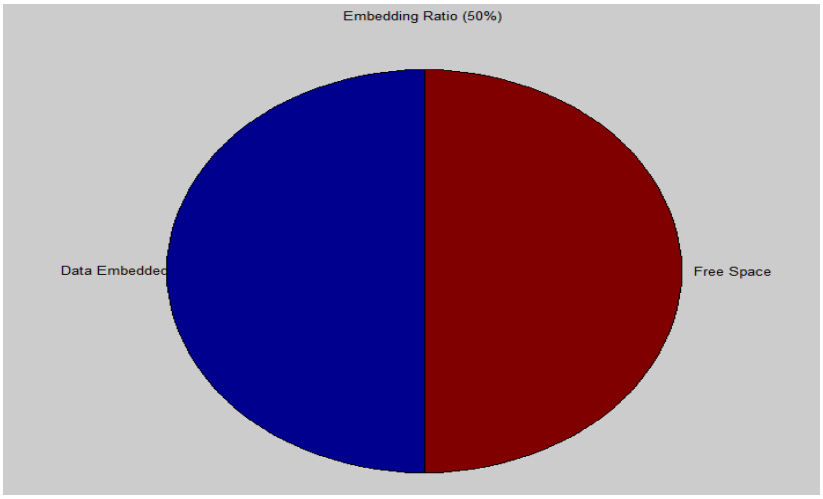


**fig 4**: The results of the Embedding Ratio

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

*7.4 PADDING COUNT*

Padding means the number of additional characters inserted to the end of the original text in order to adapt its length to suit the requirements of the encryption algorithm. In the proposed system, since the Hill Cipher algorithm relies on word processing according to blocks in size of 3 characters (as a result of using a key in the form of a 3×3 matrix), any text that is not in length of 3 multiples needs to be added to be able to go through the encryption process. According to the experimental results, two (2) letters were added to the original text to ensure that it was consistent with the required block size. This type of filler does not affect the accuracy of the recall, as it is ignored or deleted after decryption, but it is necessary to ensure the safety of mathematical processes within the algorithm. This metric indicates the efficiency of the system's adaptation to the limitations of the algorithm, and also helps to evaluate the amount of additional data resulting from the encryption process. The figure 5 shows a comparison of the size of the original message and the number of padding units added during the concealment process within the masking medium (DNA series). The first column represents the size of the message to be hidden, which amounted to 18 units, while the second column shows the number of padding units, which was very low in value of only two units. This result indicates that the size of the message was very close to the size of the capacity available in the hiding medium, which reduces the need to add a large amount of filling data that may affect the efficiency of concealment or increase the likelihood of detection. The lack of padding also indicates the effectiveness of the design in distributing and dividing data in proportion to the masking media, which enhances the level of security and the effectiveness of the process in general. Therefore, the masking system used in this research can be considered effective in terms of exploiting capacity and reducing waste resulting from additional data, which supports the reliability and effectiveness of the approved masking mechanism.

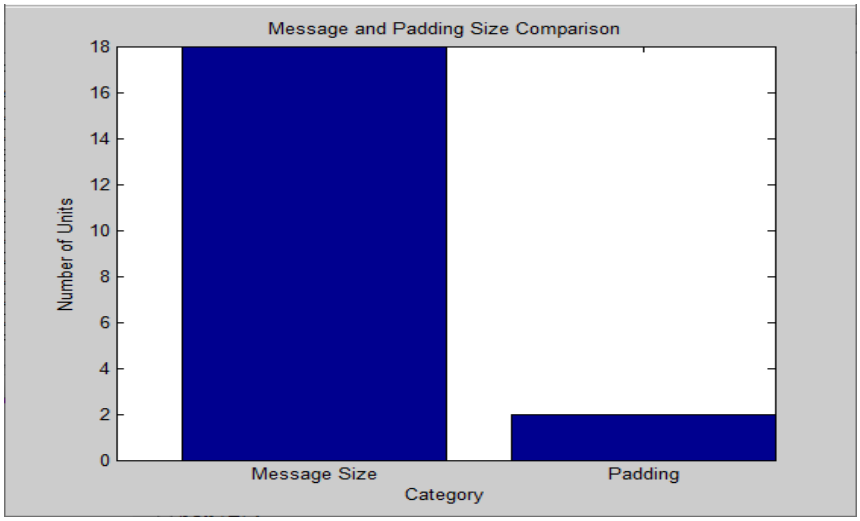


**fig 5**: The results of the *Padding Count*

## 8- DISCUSSION

The Proposed System Achieves High Security And Efficient Performance Through The Integration Of Hill Cipher Encryption and DNA Steganography. With respect to the computational complexity, the Hill Cipher process runs in $O(N^3)$ due to the matrix multiplication while DNA conversion and embedding are O(N) operations so the overall computational complexity is low.
Concerning trade-offs, the system has a well-balanced compromise between embedding capacity and quality of concealment. We understand that the embedding rate can be increased but it may compromise stealth. The chosen masking ratio provides sufficient information hiding while keeping a slight distortion to the information.
In the current system design, the use of pseudo-random DNA sequence generation and the use of a classical cipher (Hill Cipher) are limiting in that the use more secure forms of encryption or biologically realistic sequences can be used in in future work for improved robustness and scalability.

## 9. CONCLUSIONS

A hybrid text encryption system was designed and implemented for this research by integrating the Hill Cipher algorithm and DNA-based steganography as an extra protection mechanism for sensitive data. As for the system performance evaluation, multiple tests were conducted, and the findings include:

• The proposed system succeeds in attaining confidentiality, thus protecting data from unauthorized access visibility.
• It maintains data integrity and quality as evidenced by the low error rates in detection and recovery.
• The combined use of the Hill Cipher and DNA-based steganography proves to be attack resilient, thanks to the dual-layered encryption and concealment mechanisms. There are however certain limitations to the study:
• Depending on the complexity or size of the given data sequence, the speed, and computational efficiency may vary.
• Other combinations of the Hill Cipher and DNA-based steganography have not been tested. Other cryptographic forms have certainly not been researched.

It can be stated that the Hill Cipher encryption and DNA-based steganography integration methods proposed in the study are simple yet powerful new methods to strengthen security in digital forms of communication while offering complete data concealment and protection.

## 11. REFERENCES

1- R. E. Vinodhini and P. Malathi, “DNA Based Image Steganography,” in Computational Vision and Bio Inspired Computing, D. Hemanth and S. Smys, Eds. Cham: Springer, 2018, Lecture Notes in Computational Vision and Biomechanics, vol. 28.
2- M. S. Taha, A. M. Ali, and H. A. Ahmed, “Combination of steganography and cryptography: A short survey,” IOP Conf. Ser.: Mater. Sci. Eng., vol. 518, no. 5, 2019.
3- M. Tayana, Image Steganography Applications for Secure Communication, Ph.D. dissertation, 2013.
4- P. P. Aung and T. M. Naing, “A Novel Secure Combination Technique of Steganography and Cryptography,” Int. J. Inf. Technol., Model. Comput. (IJITMC), vol. 2, no. 1, Feb. 2014.

5- K. A. Ameen, W. K. Abdulwahab, and Y. N. A. Taher, "Encryption Technique Using a Mixture of Hill Cipher and Modified DNA for Secure Data Transmission," Int. J. Comput. Digit. Syst., vol. 17, no. 1, pp. 1-9, 2025.
6- X. Huang et al., "Towards next-generation DNA encryption via an expanded universal base pair," Nat. Sci. Rev., vol. 12, no. 4, 2025.
7- S. Li et al., "A Hyperchaotic Lorenz and DNA-Encoded Image Encryption Algorithm," J. Electr. Eng. Technol., vol. 19, no. 6, pp. 2287-2297, 2024.
8- M. A. Nasr et al., "A robust audio steganography technique based on image encryption," Sci. Rep., vol. 14, no. 1, 2024.
9- D. E. Mfungo et al., "Enhancing Image Encryption with the Kronecker xor Product, the Hill Cipher, and the Sigmoid Logistic Map," Appl. Sci., vol. 13, no. 6, 2023.
10- I. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Digital watermarking," J. Electron. Imaging, vol. 11, no. 3, pp. 414–414, 2002.
11- A. K. Hmood, M. A. H. Khalaf, and H. A. Jalab, "An overview on hiding information technique in images," J. Appl. Sci., vol. 10, no. 18, pp. 2094–2100, 2010.
12- G. Hamed, M. M. Hassan, S. M. Mohamed, and H. S. Hassanein, "Comparative study for various DNA based steganography techniques with the essential conclusions about the future research," in Proc. 11th Int. Conf. Comput. Eng. Syst. (ICCES), IEEE, 2016.
13- G. J. Ibrahim, T. A. Rashid, and A. T. Sadiq, "Improving DNA computing using evolutionary techniques," Int. J. Adv. Comput. Sci. Appl., vol. 7, no. 3, 2016.
14- C. Guo, C.-C. Chang, and Z.-H. Wang, "A new data hiding scheme based on DNA sequence," Int. J. Innov. Comput. Inf. Control, vol. 8, no. 1, pp. 139–149, 2012.
15- S. Marwan, A. Shawish, and K. Nagaty, "Utilizing DNA Strands for Secured Data-Hiding with High Capacity," Int. J. Interact. Mob. Technol., vol. 11, no. 2, 2017.
16- L. S. Hill, "Cryptography in an Algebraic Alphabet," Amer. Math. Monthly, vol. 36, no. 6, pp. 306–312, Jun.–Jul. 1929.
17- L. S. Hill, "Concerning Certain Linear Transformation Apparatus of Cryptography," Amer. Math. Monthly, vol. 38, pp. 135–154, 1931.