Crypto-ncRNA: Non-coding RNA (ncRNA) Based Encryption Algorithm

Xu Wang[1, a*] **Yiquan Wang**[2, b*] **Tin-Yeh Huang**[3, c*]

[1]Tsinghua University-Peking University Joint Center for Life Sciences, Tsinghua University, Beijing, 100084, China
[2]College of Mathematics and System Science, Xinjiang University, Urumqi, Xinjiang, 830046, China
[3]Department of Industrial and Systems Engineering, Faculty of Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, 999077, China

[†a]wangxu2020@mails.jlu.edu.cn [b]ethan@stu.xju.edu.cn [c]tin-yeh.huang@connect.polyu.hk

*:Xu Wang, Yiquan Wang and Tin-Yeh Huang contributed equally to this work. They are co-first authors.

Abstract

In the looming post-quantum era, traditional cryptographic systems are increasingly vulnerable to quantum computing attacks that can compromise their mathematical foundations. To address this critical challenge, we propose crypto-ncRNA—a bio-convergent cryptographic framework that leverages the dynamic folding properties of non-coding RNA (ncRNA) to generate high-entropy, quantum-resistant keys and produce unpredictable ciphertexts. The framework employs a novel, multi-stage process: encoding plaintext into RNA sequences, predicting and manipulating RNA secondary structures using advanced algorithms, and deriving cryptographic keys through the intrinsic physical unclonability of RNA molecules. Experimental evaluations indicate that, although crypto-ncRNA's encryption speed is marginally lower than that of AES, it significantly outperforms RSA in terms of efficiency and scalability while achieving a 100% pass rate on the NIST SP 800-22 randomness tests. These results demonstrate that crypto-ncRNA offers a promising and robust approach for securing digital infrastructures against the evolving threats posed by quantum computing.

1 Introduction

Biomolecular cryptography has emerged as a potential breakthrough in post-quantum encryption (Balamurugan et al. (2021); Mondal & Ray (2023)). Moreover, with the rapid advancement of artificial intelligence, RNA-based research has gradually unfolded into a new realm of innovation (Townshend et al. (2021)). Recent studies showed that the dynamic folding processes of RNA molecules intrinsically exhibit physical unclonable functions (PUFs) characteristics (Herder et al. (2014); Li et al. (2022); Luescher et al. (2024); Zhou et al. (2021)), thereby establishing a pathway for designing post-quantum cryptography (PQC) systems (Arapinis et al. (2021); Cambou et al. (2021)).

In this paper, we introduce *Crypto-ncRNA*, an encryption framework harnessing the dynamic folding mechanisms of non-coding RNA (ncRNA) to address quantum-era security challenges. By exploiting ncRNA's intrinsic PUFs and high conformational entropy, enhanced through deep learning-based RNA secondary structure prediction, the scheme generates cryptographically robust keys and ciphertexts with enhanced stochas-

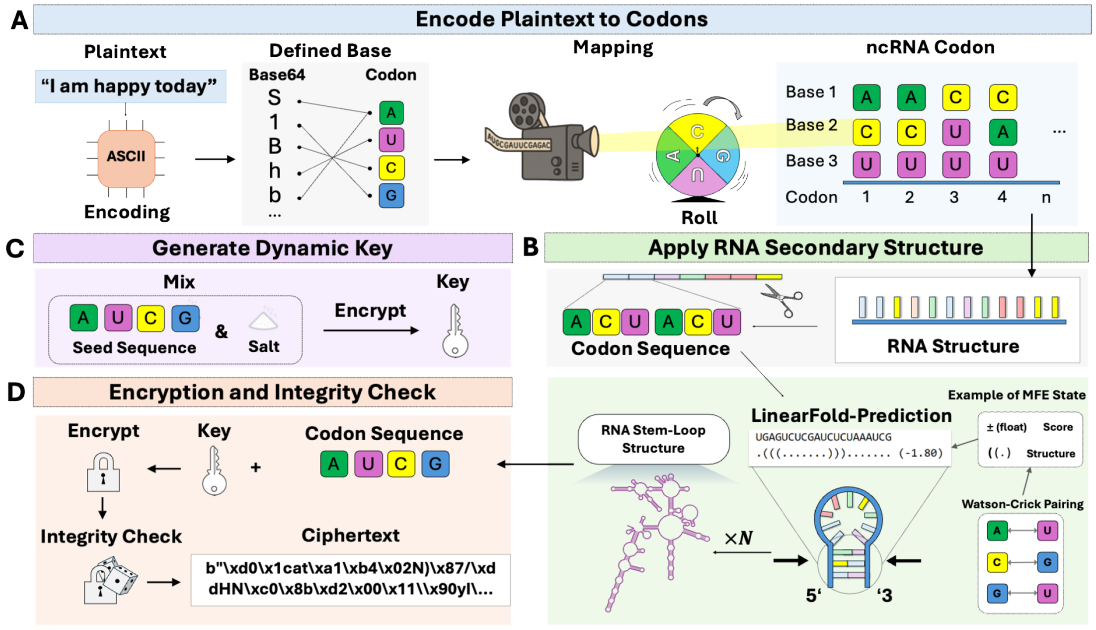

Figure 1: The Framework of Crypto-ncRNA Algorithm

ticity. This work provides a new direction for encryption technology in the quantum era by integrating RNA's PUFs with encryption algorithms.

2 Method (details in Appendix B)

The *Crypto-ncRNA* implements a four-tiered encryption architecture. The research framework and workflow of the *Crypto-ncRNA* are illustrated in Figure 1.

1. **Codon Mapping and RNA Sequence Generation (Fig. 1a):** Textual data is encoded into RNA codons (e.g., AUG) via Base64-derived 6-bit indices, enabling 50% higher information density than binary systems.
2. **RNA Secondary Structure Folding (Fig. 1b):** RNA codons are partitioned into Watson-Crick-paired stem regions and unpaired loops using LinearFold-predicted minimum free energy (MFE) structures. Dynamic codon permutation within structural constraints generates combinatorial complexity (4^N configurations, N: dynamic positions).
3. **Dynamic Key Generation (Fig. 1c):** Quantum-resistant keys are synthesized using PBKDF2-HMAC-SHA256, leveraging RNA quaternary fingerprints (A/U/G/C positional entropy) and 256-bit cryptographic salts.
4. **Ciphertext Packaging and Integrity Verification (Fig. 1d):** ChaCha20 encrypts payloads using dynamic RNA keys. SHA-256 hashing and enzyme markers ensure ciphertext integrity and physical binding.

The process emphasizes sequence-dependent structural transformations and environmental noise integration to ensure cryptographic security and hardware-bound key uniqueness.

3 Results (details in Appendix D)

Crypto-ncRNA demonstrates robust performance across heterogeneous computing environments, validated through comprehensive benchmarking against classical algorithms (RSA, AES). The following visualizations (Figures 2) highlight its efficiency, reliability, and adaptability under varying workloads.

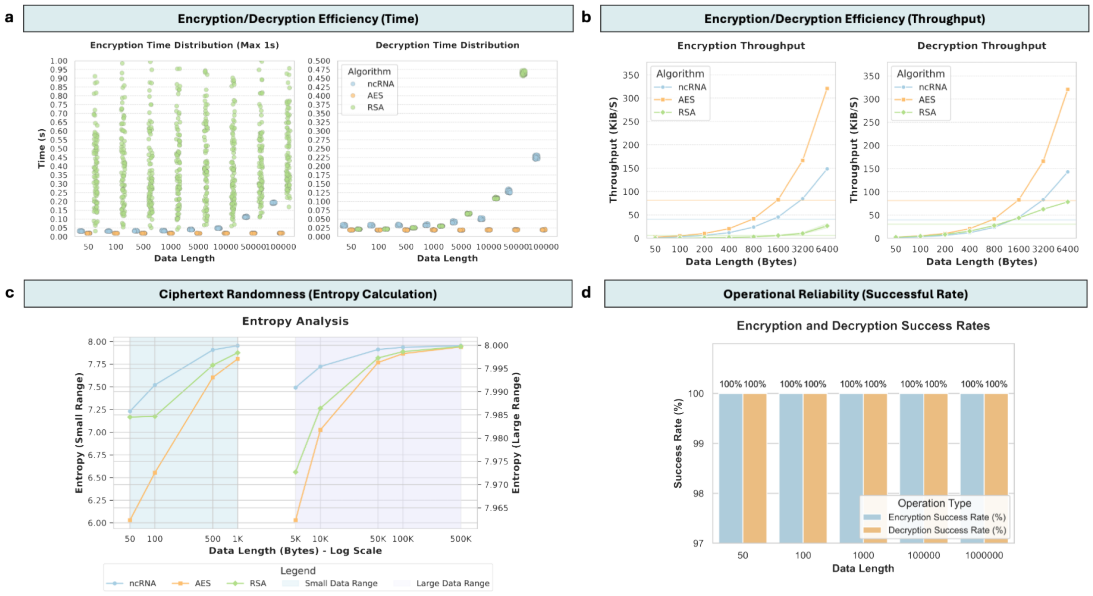

Figure 2: Summary of Algorithm(s) Comparation and Tesing Results

1. **Encryption/Decryption Efficiency** (Fig. 2A): The proposed method slightly underperforms AES in speed but surpasses RSA, achieving near-AES efficiency at smaller parameters.
2. **Encryption/Decryption Throughput Performance** (Fig. 2B): Throughput trends align closely with time efficiency results.
3. **Ciphertext Randomness** (Fig. 2C): Cryptographic average entropy consistently outperforms AES and RSAs.
4. **Operational Reliability** (Fig. 2D): Demonstrates 100% success rate regardless of data volume.
5. **Statistical Randomness** (Appendix Table 1): Passes all NIST SP 800-22 criteria (Rukhin et al. (2010)), confirming cryptographic robustness.

4 Conclusion

Crypto-ncRNA establishes a bio-convergent security framework that synergizes the biophysical complexity of non-coding RNA with cryptographic principles. Looking ahead, deep learning offers a pathway to more robust RNA encryption methods through enhanced RNA structure prediction. By offering intrinsic unclonability via dual resistance, this architecture provides a future-proof solution for securing digital infrastructures in the post-quantum landscape.

URM Statement

The authors acknowledge that at least one key author of this work meets the URM criteria of ICLR 2025 AI4NA Tiny Papers Track.

References

M. Arapinis, M. Delavar, M. Doosti, and E. Kashefi. Quantum physical unclonable functions: Possibilities and impossibilities. *Quantum*, 5:475, 2021.

C. Balamurugan, K. Singh, G. Ganesan, and M. Rajarajan. Post-quantum and code-based cryptography—some prospective research directions. *Cryptography*, 5(4):38, 2021.

C. Cachin. *Entropy measures and unconditional security in cryptography*. PhD thesis, ETH Zurich, 1997.

B. Cambou, M. Gowanlock, B. Yildiz, D. Ghanaimiandoab, K. Lee, S. Nelson, C. Philabaum, A. Stenberg, and J. Wright. Post quantum cryptographic keys generated with physical unclonable functions. *Applied Sciences*, 11(6):2801, 2021.

C. Herder, M.-D. Yu, F. Koushanfar, and S. Devadas. Physical unclonable functions and applications: A tutorial. *Proceedings of the IEEE*, 102(8):1126–1141, 2014.

L. Huang, H. Zhang, D. Deng, K. Zhao, K. Liu, D. A. Hendrix, and D. H. Mathews. Linearfold: Linear-time approximate rna folding by 5'-to-3' dynamic programming and beam search. *Bioinformatics*, 35(14): i295–i304, 2019.

I. Kimsey, K. Petzold, B. Sathyamoorthy, and H. M. Al-Hashimi. Visualizing transient watson–crick-like mispairs in dna and rna duplexes. *Nature*, 519(7543):315–320, 2015.

Y. Li, M. M. Bidmeshki, T. Kang, C. M. Nowak, Y. Makris, and L. Bleris. Genetic physical unclonable functions in human cells. *Science Advances*, 8:eabm4106, 2022.

A. M. Luescher, A. L. Gimpel, W. J. Stark, R. Heckel, and R. N. Grass. Chemical unclonable functions based on operable random dna pools. *Nature Communications*, 15:2955, 2024.

M. Mondal and K. S. Ray. Review on dna cryptography. *International Journal of Bioinformatics and Intelligent Computing*, 2(1):44–72, 2023.

D. Pribnow. Nucleotide sequence of an rna polymerase binding site at an early t7 promoter. *Proceedings of the National Academy of Sciences of the United States of America*, 72(3):784–788, 1975.

V. Rijmen and J. Daemen. Advanced encryption standard, 2001. Proceedings of Federal Information Processing Standards Publications, National Institute of Standards and Technology, 19, 22.

R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert, J. Dray, and S. Vo. A statistical test suite for random and pseudorandom number generators for cryptographic applications (nist special publication 800-22 rev. 1a). Technical report, National Institute of Standards and Technology, 2010.

A. Rényi. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pp. 547–562. University of California Press, January 1961.

K. Sato, M. Akiyama, and Y. Sakakibara. Rna secondary structure prediction using deep learning with thermodynamic integration. *Nature Communications*, 12:941, 2021.

I. Tinoco Jr and C. Bustamante. How rna folds. *Journal of Molecular Biology*, 293(2):271–281, 1999.

R. J. Townshend, S. Eismann, A. M. Watkins, R. Rangan, M. Karelina, R. Das, and R. O. Dror. Geometric deep learning of rna structure. *Science*, 373(6558):1047–1051, 2021.

P. Yakovchuk, E. Protozanova, and M. D. Frank-Kamenetskii. Base-stacking and base-pairing contributions into thermal stability of the dna double helix. *Nucleic Acids Research*, 34(2):564–574, 2006.

W. Zhou, D. Melamed, G. Banyai, et al. Expanding the binding specificity for rna recognition by a puf domain. *Nature Communications*, 12:5107, 2021.

M. Zuker and D. Sankoff. Rna secondary structures and their prediction. *Bulletin of Mathematical Biology*, 46(4):591–621, 1984.