

La formación del enlace de hidrogeno en el ADN como una propiedad emergente: aplicaciones en caminatas aleatorias

Rodrigo Rodríguez-Gutiérrez, Francisco Hernández-Cabrera,
Francisco Javier Almaguer-Martínez, María Esther Grimaldo-Reyna

Universidad Autónoma de Nuevo León,
Facultad de Ciencias Físico-Matemáticas,
México

francisco.hernandezcbr@uanl.edu.mx

Resumen. Los mecanismos celulares que sustentan la vida son comprendidos a partir de los estudios realizados en la molécula de ADN. Esta molécula está formada por secuencias antiparalelas de bases nitrogenadas estabilizadas por puentes de hidrógeno. Las propiedades fisicoquímicas de estos componentes del ADN son fundamentales para comprender su estructura en forma helicoidal. En este trabajo se representaron las propiedades fisicoquímicas de las nucleobases con las reglas de estructura (pirimidina o purina), grupo funcional (ceto o amino) y tipo de enlace de hidrógeno (débil o fuerte). Además, se utilizó una métrica binaria y operadores lógicos aplicados a los estados de las reglas, para representar el tipo de enlace de hidrógeno entre bases nitrogenadas. Se demostró que la regla de enlace de hidrógeno resulta de una transformación lineal entre las propiedades de estructura y grupo funcional de las nucleobases. Esta modelación permitió obtener caminatas aleatorias de ADN de una regla específica, cuando se tienen los estados de las otras dos reglas. Esta herramienta podría ser aplicada para realizar análisis filogenético en secuencias de ADN. Por otra parte, se propone una métrica matricial, la cual posee propiedades algebraicas semejantes a las utilizadas para describir los estados de observables en la mecánica cuántica. Estos hallazgos podrían conducir al desarrollo de nuevas herramientas para el análisis de secuencias y la comprensión de los procesos biológicos que involucran el ADN.

Palabras clave: Operadores lógicos, enlace de hidrógeno, propiedad emergente, transformación lineal, caminata aleatoria.

The Formation of the Hydrogen Bond in DNA as an Emergent Property; Applications in Random Walks

Abstract. The cellular mechanisms that sustain life are understood from studies carried out on the DNA molecule. This molecule is formed by antiparallel sequences of nitrogenous bases stabilized by hydrogen bonds. The physicochemical properties of these DNA components are fundamental to understanding its helical structure. In this work, the physicochemical properties of nucleobases were represented with the rules of structure (pyrimidine or purine), functional group (keto or amino) and type of hydrogen bond (weak or strong). In addition, a binary metric and logical operators applied to the states of the rules were used to represent the type of hydrogen bond between nitrogenous bases. It was shown that the hydrogen bonding rule results from a linear transformation between the structure and functional group properties of nucleobases. This

modeling allowed us to obtain random DNA walks of a specific rule, when the states of the other two rules are available. This tool could be applied to perform phylogenetic analysis on DNA sequences. On the other hand, a matrix metric is proposed, which has algebraic properties similar to those used to describe the states of observables in quantum mechanics. These findings could lead to the development of new tools for sequence analysis and understanding biological processes involving DNA.

Keywords: Logical operators, hydrogen bond, emergent property, linear transformation, random walks.

1. Introducción

La información que sustenta los mecanismos celulares para permitir el desarrollo y adaptación de los seres vivos se encuentra codificada en la molécula de ácido desoxirribonucleico (ADN). El ADN tiene una estructura de doble hélice formada por dos secuencias antiparalelas de nucleótidos. Estas secuencias son estabilizadas por enlaces de hidrógeno entre bases nitrogenadas complementarias, las cuales se localizan en hebras opuestas del ADN.

El tipo de apareamiento entre bases nitrogenadas fue descrito por primera vez por Watson y Crick, estableciendo que la Adenina (A) se une mediante dos enlaces de hidrógeno con la Timina (T), mientras que la Guanina (G) se une mediante tres enlaces de hidrógeno con Citosina (C) (figura 1).

Estas interacciones complementarias entre bases nitrogenadas se rigen por tres propiedades fisicoquímicas; estructura, grupo funcional y tipo de enlace, cada una de las cuales contiene dos estados. Así, se han establecido tres reglas de clasificación, la primera regla describe dos tipos de estructuras (E); la estructura química llamada Purina (R) tiene dos anillos que están presentes en la Adenina y Guanina. Por otra parte, la estructura llamada Pirimidina (Y) tiene un anillo simple y se encuentra en Timina y Citosina.

La segunda regla relaciona el grupo funcional localizado en el anillo hexagonal de cada nucleobase; así, el grupo Amino (M) se encuentra en la Adenina y Citosina, mientras que el grupo Ceto (K) está presente en la Timina y Guanina.

La tercera regla establece que existe un triple enlace de hidrógeno (enlace S \rightarrow fuerte) entre la Guanina y Citosina, mientras que hay un doble enlace de hidrógeno (enlace W \rightarrow débil) entre la Adenina y Timina.

Estudios previos han relacionado la propiedad fisicoquímica de estructura (regla Y/R) con métricas binarias aplicadas a secuencias de ADN para proponer análisis de correlación y fractalidad en caminatas aleatorias [9, 13, 10].

Trabajos posteriores fueron agregando las reglas de grupo funcional (K/M) y tipo de enlace de hidrógeno (S/W) para realizar análisis filogenético en secuencias de ADN [5, 6, 14–16, 2]. También se ha propuesto el uso de operadores lógicos en secuencias de ADN para modelar el transporte de carga y en la búsqueda de patrones fractales [12, 8, 1, 4, 3]. Otros trabajos han realizado representaciones matriciales y binarias para estudiar la degeneración del código genético [11, 7].

Las investigaciones previas sugieren que las tres reglas relacionadas con las propiedades fisicoquímicas son independientes entre sí, además no consideran la

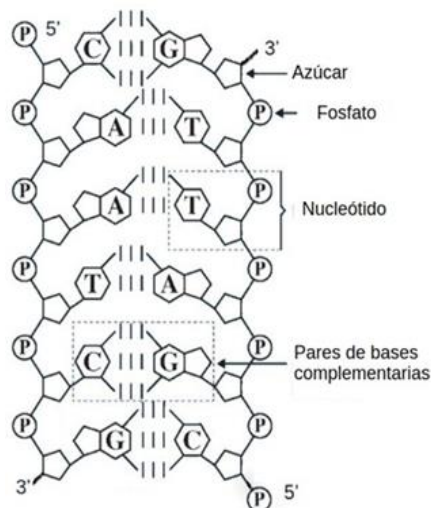


Fig. 1. La molécula de ADN se muestra representada por dos secuencias antiparalelas, cada una de ellas con los 4 tipos de nucleótidos (secuencia sentido 5' CAATCG 3' y secuencia anti sentido 3' GTTAGC 5'). Los enlaces de hidrógeno se muestran con líneas verticales entre las nucleobases complementarias. Imagen descargada bajo licencia libre de vecteezy.com.

posición de cada nucleobase en la secuencia sentido o antisentido en el ADN de doble cadena.

Por lo anterior es importante describir un modelo matemático que correlacione las reglas fisicoquímicas de estructura, grupo funcional y tipo de enlace.

Este modelo permitiría comprender coherentemente el significado biológico de la dependencia entre la estructura y grupo funcional, que permite la formación de enlaces de hidrógeno entre bases nitrogenadas complementarias en la doble cadena de ADN.

En este trabajo se propone un modelo de correlación entre las reglas fisicoquímicas. Nuestro modelo sugiere que la estructura y grupo funcional de cada nucleobase determinan las interacciones permitidas que conducen a la formación del tipo de enlace de hidrógeno correspondiente.

2. Metodología

Considerando las tres reglas binarias, representamos las características principales de las bases nitrogenadas en la tabla 1. Además, utilizamos las entradas lógicas falso (F) y verdadero (T) para los estados de cada una de las reglas que describen las bases nitrogenadas.

Así, la regla Y/R tiene valores lógicos $Y \rightarrow F$ y $R \rightarrow T$, para la regla K/M; $K \rightarrow F$ y $M \rightarrow T$, y la regla S/W; $S \rightarrow F$ y $A \rightarrow T$ (ver tabla 2). Con estas entradas lógicas se observa que los estados de las reglas de estructura (Y/R) y función (K/M) operados a través de un XOR, dan como resultado los estados de la regla de enlace (S/W) (ver figura 2).

Table 1. Tabla que resume las propiedades fisicoquímicas de las núcleo bases que forman al ADN, con sus respectivas abreviaciones.

Nucleobase	Estructura (E)	Grupo funcional (F)	Tipo de enlace (H)
Adenina (A)	Purina (R)	Amino (M)	Weak (W)
Guanina (G)	Purina (R)	Keto (K)	Strong (S)
Citosina (C)	Pirimidina (Y)	Amino (M)	Strong (S)
Timina (T)	Pirimidina (Y)	Keto (K)	Weak (W)

Table 2. Representación de las propiedades fisicoquímicas con los valores lógicos aplicados a las reglas Y/R, K/M y S/W.

Regla	A	G	C	T
Y/R	T	T	F	F
K/M	T	F	T	F
S/W	T	F	F	T

2.1. Relación entre la caminata de ADN y circuitos lógicos

Otra manera de modelar las secuencias de ADN es utilizando series de tiempo. Así, se puede definir una secuencia de M nucleótidos como $Q = \{q_1q_2\dots q_M\}$, en donde $q_k \in \{A, C, G, T\}$ y $k, \in \{1, 2, \dots, M\}$. En este caso la secuencia se transforma en una caminata aleatoria aplicando la métrica (-1,1) a los estados de cada una de las reglas. Con las expresiones 2-4 se asignan valores a cada nucleótidos en la secuencia de ADN, y se obtiene la sumatoria en cada paso para representar cada una de las caminatas como se muestra en la tabla 4. Entonces posición del nésimo paso en la caminata de ADN está dada por:

$$X_n = x_0 + \sum_{i=1}^n x_i, \tag{1}$$

en donde $x_0 = 0$ es la posición inicial y X_n representa la resultante de la caminata en el paso n :

Regla Y/R:

$$x_k = \begin{cases} -1 & \text{si } q_k \in Y = \{C, T\} \\ 1 & \text{si } q_k \in R = \{A, G\} \end{cases}, \tag{2}$$

Regla K/M:

$$y_k = \begin{cases} -1 & \text{si } q_k \in K = \{G, T\} \\ 1 & \text{si } q_k \in M = \{A, C\} \end{cases}, \tag{3}$$

Regla S/W:

$$z_k = \begin{cases} -1 & \text{si } q_k \in S = \{C, G\} \\ 1 & \text{si } q_k \in W = \{A, T\} \end{cases}. \tag{4}$$

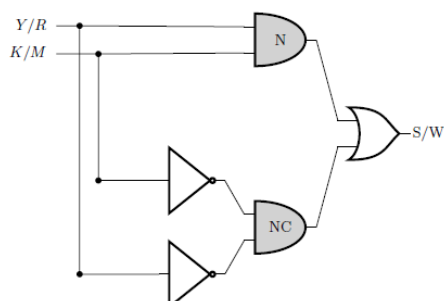


Fig. 2. Circuito lógico que modela la interacción entre las reglas Y/R y K/M para formar la regla S/W, como se muestra en la tabla 2. Las compuertas AND se muestran sombreadas en gris, en donde la salida de N representa a cualquier nucleobase (A, T, G, C) mientras que la salida NC a su nucleobase complementaria.

Table 3. Valores de las reglas Y/R, K/M y S/W para las nucleobases, sustituyendo los valores de las variables binarias V y F de la tabla 2, por las expresiones 2-4.

Regla	A	G	C	T
Y/R	1	1	-1	-1
K/M	1	-1	1	-1
S/W	1	-1	-1	1

La figura 3 es una representación de las 3 caminatas para las reglas Y/R, K/M y S/W utilizando una secuencia de 16 000 nucleobases.

2.2. Relación entre las caminatas de ADN y vectores ortogonales

Utilizando la tabla 4 junto con las expresiones 2-4 para las reglas Y/R, K/M y S/W, se puede encontrar una relación algebraica que relaciona las secuencias de estructura y función con la caminata correspondiente al tipo de enlace.

Para obtener esta relación, primero simplificamos las reglas como se muestra en las ecuaciones 5-7. En donde x_m representa el valor de la regla Y/R para una nucleobase en la m-ésima posición de la secuencia de ADN. Por lo tanto, los valores de x_m , y_m y z_m corresponden a las 3 reglas aplicadas a la misma nucleobase:

$$(Y/R)_m = x_m, \tag{5}$$

$$(K/M)_m = y_m, \tag{6}$$

$$(S/W)_m = z_m. \tag{7}$$

Extendiendo estos valores a los n nucleobases de la secuencia de ADN, obtenemos los vectores de la ecuación 8. Estos vectores representan los valores binarios de los estados en cada posición de la secuencia:

Table 4. Valores de las reglas Y/R, K/M y S/W con sus respectivas caminatas para una secuencia de nucleobases.

	G	C	T	G	T	T	A	A	...
Y/R	1	-1	-1	1	-1	-1	1	1	...
$\sum Y/R$	1	0	-1	0	-1	-2	-1	0	...
K/M	-1	1	-1	-1	-1	-1	1	1	...
$\sum K/M$	-1	0	-1	-2	-3	-4	-3	-2	...
S/W	-1	-1	1	-1	1	1	1	1	...
$\sum S/W$	-1	-2	-1	-2	-1	0	1	2	...

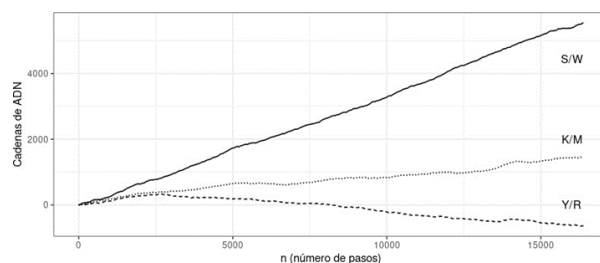


Fig. 3. Caminatas de ADN para las reglas Y/R (línea solida), K/M (línea punteada) y S/W (línea discontinua) en donde n representa el número de pasos de la caminata.

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \\ \vdots \\ x_n \end{pmatrix} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \\ \vdots \\ y_n \end{pmatrix} \quad Z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \\ \vdots \\ z_n \end{pmatrix}. \quad (8)$$

Para obtener la caminata correspondiente a la regla S/W, suponemos que las reglas de estructura y función son independientes, es decir, sus vectores correspondientes deben ser ortogonales.

Por lo tanto, utilizamos la operación XOR como la combinación de una compuerta NOT y 2 compuertas AND (multiplicación matricial), como se muestra en la figura 2. Entonces podemos obtener la relación para la caminata de la regla S/W con la ecuación 9:

$$XTY = X \cdot Y = Z, \quad (9)$$

donde el superíndice T implica la transpuesta del vector. Escribiendo de forma matricial la ecuación 9 tenemos:

$$(x_1, x_2, \dots, x_m) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \sum_{i=1}^m (x_i y_i) = \sum_{i=1}^m z_i. \quad (10)$$

También podemos expresar los vectores X y Y como se muestra en las ecuaciones 11-12, sin embargo, desde 1 punto de vista biológico solo es correcta la primera transformación. Lo anterior de justifica debido a que las propiedades fisicoquímicas de estructura y función son fundamentales para la formación del enlace de hidrógeno entre las bases nitrogenadas complementarias. Esta relación implica que el enlace de hidrógeno representa una propiedad emergente necesaria para la estabilidad de la estructura de la doble cadena de ADN.

$$Y^T Z = Y \cdot Z = X, \quad (11)$$

$$X^T Z = X \cdot Z = Y. \quad (12)$$

2.3. Representación como álgebra de Clifford

Generalizando las propiedades algebraicas de las reglas Y/R, K/M y S/W, para cada nucleobase en la posición m de la secuencia de ADN se tienen las transformaciones siguientes:

$$x_m y_m = z_m, \quad (13)$$

$$y_m z_m = x_m, \quad (14)$$

$$z_m x_m = y_m, \quad (15)$$

$$x_m^2 = y_m^2 = z_m^2 = 1, \quad (16)$$

$$x_m y_m y_m = 1. \quad (17)$$

Los elementos que conforman este conjunto de ecuaciones forman un álgebra de Clifford (3, 0), debido que, el cuadrado de cada uno de sus elementos es igual a 1 (ecuación 16). Además, utilizando el anticommutador $\{A, B\} = AB + BA$ para x_m, y_m y z_m , se obtienen las siguientes relaciones:

$$\{x_m, x_m\} = \{y_m, y_m\} = \{z_m, z_m\} = 2, \quad (18)$$

$$\{x_m, y_m\} = \{y_m, x_m\} = 2z_m, \quad (19)$$

$$\{y_m, z_m\} = \{z_m, y_m\} = 2x_m, \quad (20)$$

$$\{z_m, x_m\} = \{x_m, z_m\} = 2y_m. \quad (21)$$

Haciendo $\varepsilon_i = x_m, \varepsilon_j = y_m, \varepsilon_k = z_m$ y $1/\sqrt{2}$ como factor de normalización obtenemos:

$$\varepsilon_i \varepsilon_j = \begin{cases} 1 & \text{si } i = j \\ 1\varepsilon_k & \text{si } i \neq j \end{cases} \quad (22)$$

Aplicando la propiedad de isomorfismo al algebra de Clifford podemos representar las ecuaciones de 5-7 por matrices $M_{2 \times 2}$ como sigue:

$$x_m = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (23)$$

$$y_m = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (24)$$

$$z_m = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}. \quad (25)$$

Para la cadena complementaria (antisentido), es necesario rotar 180° para obtener la regla S/W a partir de las reglas Y/R y K/M. Entonces tenemos las ecuaciones 26-28:

$$x'_m y'_m = z'_m, \quad (26)$$

$$y'_m z'_m = x'_m, \quad (27)$$

$$z'_m x'_m = y'_m, \quad (28)$$

$$(x'_m)^2 = (y'_m)^2 = (z'_m)^2 = 1, \quad (29)$$

$$z'_m y'_m x'_m = -1. \quad (30)$$

Estos elementos también forman un Algebra de Clifford $(3, 0)$. Utilizando el anticonmutador y normalizando con la constante $1/\sqrt{2}$, obtenemos las siguientes relaciones:

$$\varepsilon'_i \varepsilon'_j = \begin{cases} 1 & \text{si } i = j, \\ 1\varepsilon'_k & \text{si } i \neq j. \end{cases} \quad (31)$$

Ahora las ecuaciones 26-28 se pueden representar en forma matricial, como se muestra en las ecuaciones 32-34:

$$x'_m = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (32)$$

$$y'_m = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (33)$$

$$z'_m = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}. \quad (34)$$

3. Resultados

3.1. Operadores lógicos y caminatas de ADN

Trabajos previos han estudiado las propiedades fisicoquímicas de las nucleobases a través de sus reglas de estructura, grupo funcional y tipo de enlace como variables independientes. En este estudio se propone un modelo algebraico, que determina la relación entre los estados de las reglas asociadas a las propiedades fisicoquímicas fundamentales (estructura y función). Con estos resultados es posible modelar una caminata de ADN correspondiente al tipo de enlace de hidrógeno, a partir de los estados

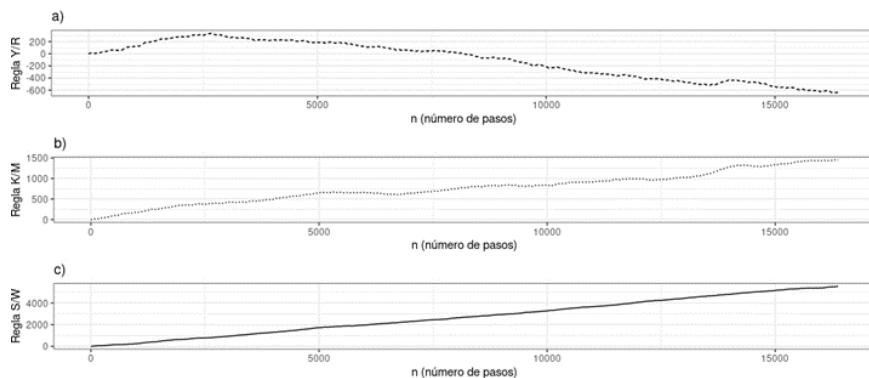


Fig. 4. Se muestran las caminatas de ADN para las 3 reglas. a) regla Y/R, b) regla K/M, c) regla S/W, para una secuencia de 16 000 bases nitrogenadas.

de las caminatas de estructura y grupo funcional de acuerdo con la ecuación 10 (ver figura 4).

3.2. Caminatas de ADN y álgebra de Clifford

Del modelo propuesto surgen propiedades algebraicas entre las reglas asociadas a las propiedades fisicoquímicas. Como se mostró anteriormente, este modelo cumple con un álgebra de Clifford (3, 0) para las representaciones en las cadenas sentido y anti sentido. De lo anterior se llega a las siguientes expresiones:

$$\xi_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (35)$$

$$-\xi_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (36)$$

$$\xi_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (37)$$

$$-\xi_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad (38)$$

donde las matrices $\pm\xi_1$ pertenece al grupo ortogonal especial SO (2), mientras que $\pm\xi_2$ al grupo ortogonal O (2).

3.3. Representaciones de las propiedades fisicoquímicas del ADN y su relación con las matrices de Pauli

El conjunto ξ es una representación matricial de las reglas para las nucleobases. Este conjunto de matrices genera un álgebra de Clifford Cl (3, 0), de manera similar a las matrices de Pauli. Esto sugiere que puede haber alguna relación matemática entre las observables de la mecánica cuántica y propiedades fisicoquímicas de las bases nitrogenadas en el ADN de doble cadena.

4. Discusión

La molécula de ADN es el resultado de un largo proceso evolutivo de interacciones entre moléculas con diversas estructuras y grupos funcionales. La afinidad molecular entre nucleobases complementarias permitió el auto ensamblaje a través de enlaces de hidrógeno e incrementó la estabilidad en secuencias antiparalelas. Esto ha favorecido los procesos de almacenamiento de información, transcripción y replicación.

Los estudios realizados sobre el ADN han permitido comprender mejor las interacciones moleculares que permiten la conformación de la doble compuerta lógicas de ADN para modelar procesos biológicos [12, 8, 1, 4, 3]. Así, se han desarrollado nuevas estrategias de modelación y herramientas de análisis de secuencias de ADN.

En este trabajo, se han utilizado operadores lógicos para representar las propiedades fisicoquímicas de las bases nitrogenadas. Esto permite modelar el apareamiento entre las bases de Watson y Crick a través de los estados correspondientes a la regla de enlace de hidrógeno. Así, demostramos por primera vez que el enlace de hidrógeno entre nucleobases complementarias representa una propiedad emergente, que surge como resultado de dos propiedades independientes; estructura y grupo funcional.

También se analizaron de manera conjunta las métricas asociadas a las caminatas de ADN. Lo anterior permitió desarrollar un método para obtener la caminata de la regla S/W a partir de los estados de reglas Y/R y K/M aplicadas a una secuencia de ADN. Esta dependencia se basa en la representación vectorial de las reglas Y/R y K/M, para sus caminatas correspondientes.

Estos vectores deben ser ortogonales y su multiplicación vectorial da como resultado los pasos de la caminata para la regla S/W. La métrica propuesta puede ser representada por un álgebra de Clifford $Cl(3, 0)$, de manera similar a las matrices de Pauli. Nuestros resultados sugieren que podría haber una conexión entre el desarrollo de las métricas asociadas a las caminatas de ADN y los fundamentos matemáticos de la mecánica cuántica.

Lo anterior concuerda con las ideas desarrolladas por Petukov [11], debido a que las matrices obtenidas en este trabajo son simétricas (hermitianas reales) y unitarias. Esta conexión con la mecánica cuántica podría ser útil para el desarrollo de nuevas herramientas en bioinformática y el estudio de las secuencias de ADN.

5. Conclusión

El ADN es la molécula compleja más estudiada en los últimos 70 años, la comprensión de su estructura y función resulta de gran importancia para el estudio de los sistemas biológicos. En este contexto la modelación matemática desarrollada en este trabajo puede contribuir a la representación matricial de los estados de las nucleobases dentro de las secuencias de ADN de doble cadena.

Además, los resultados sugieren que el modelo propuesto podría ser utilizado para mejorar los estudios en caminatas de ADN. También, se proporcionan nuevas perspectivas para el desarrollo de herramientas de análisis, basadas en métricas matriciales que cumplen un álgebra de Clifford $(3, 0)$. Finalmente, se sugiere que podría existir una conexión entre la métrica propuesta y la representación de estados, de manera similar a la mecánica cuántica.

Referencias

- 1 Chao, Q., Zhang, Y., Li, Q., Jiao, L., Sun, X., Chen, X., Zhu, L., Yang, Q., Shang, C., Kong, R. M., Fan, G. C., Song, Z. L., Luo, X.: Compute and releaselogic gated DNA cascade circuit for accurate cancer cell imaging. *Analytical Chemistry* vol. 95, no. 19, pp. 7723–7734 (2023) doi: 10.1021/acs.analchem.3c00898
- 2 Corona-Ruiz, M., Hernandez-Cabrera, F., Cantú-González, J. R., González-Amezcuca, O., Javier Almaguer, F.: A stochastic phylogenetic algorithm for mitochondrial DNA analysis. *Frontiers in Genetics*, vol. 10 (2019) doi: 10.3389/fgene.2019.00066
- 3 Feng, C., Chen, T., Mao, D., Zhang, F., Tian, B., Zhu, X.: Construction of a ternary complex-based DNA logic nanomachine for a highly accurate imaging analysis of cancer cells. *ACS Sensors*, vol. 5, no. 10, pp. 3116–3123 (2020) doi: 10.1021/acssensors.0c01166
- 4 Gong, H., Dai, Q., Peng, P.: Cell-membrane-anchored DNA logic gated nanoassemblies for in situ extracellular bioimaging. *ACS Applied Materials & Interfaces*, vol. 14, no. 38, pp. 43026–43034 (2022) doi: 10.1021/acsami.2c13735, pMID: 36053489
- 5 Liao, B., Wang, T. M.: New 2D graphical representation of DNA sequences. *Journal of Computational Chemistry*, vol. 25, no. 11, pp. 1364–1368 (2004)
- 6 Liu, X. Q., Dai, Q., Xiu, Z., Wang, T.: PNN-curve: A new 2D graphical representation of DNA sequences and its application. *Journal of Theoretical Biology*, vol. 243, no. 4, pp. 555–561 (2006)
- 7 Nikolajewa, S., Friedel, M., Beyer, A., Wilhelm, T.: The new classification scheme of the genetic code, its early evolution, and tRNA usage. *Journal of Bioinformatics and Computational Biology*, vol. 4, no. 2, pp. 609–620 (2006)
- 8 Okamoto, A., Tanaka, K., Saito, I.: DNA logic gates. *Journal of the American Chemical Society*, vol. 126, no. 30, pp. 9458–9463 (2004) doi: 10.1021/ja047628k, pMID: 15281839
- 9 Peng, C. K., Buldyrev, S. V., Goldberger, A. L., Havlin, S., Sciortino, F., Simons, M., Stanley, H. E.: Long-range correlations in nucleotide sequences. *Nature*, vol. 356, no. 6365, pp. 168–170 (1992)
- 10 Peng, C.K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., Goldberger, A. L.: Mosaic organization of DNA nucleotides. *Physical Review. E, statistical physics, plasmas, fluids, and related interdisciplinary topics*, vol. 49, no. 2, pp. 1685–1689 (1994)
- 11 Petoukhov, S., Petukhova, E.: On Genetic Unitary Matrices and Quantum- Algorithmic Genetics. In: Hu, Z., Petoukhov, S., He, M. (eds) *Advances in Artificial Systems for Medicine and Education II, AIMEE2018 2018, Advances in Intelligent Systems and Computing*, Springer, vol. 902, pp. 103–115 (2019) doi: 10.1007/978-3-030-12082-5_10
- 12 Rothmund, P., Papadakis, N., Winfree, E.: Algorithmic self-assembly of DNA sierpinski triangles. *PLoS Biology*, vol. 2, no. 12, pp. e424 (2005). doi: 10.1371/journal.pbio.0020424
- 13 Stanley, H. E., Buldyrev, S. V., Goldberger, A. L., Havlin, S., Peng, C. K., Sciortino, F., Simons, M.: Scaling concepts and complex fluids: long-range power-law correlations in DNA. *Journal de Physique Archives*, vol. 3, no. C1, pp. C1–15, C1–25 (1993) doi: 10.1051/jp4:1993102
- 14 Xie, G., Mo, Z.: Three 3D graphical representations of DNA primary sequences based on the classifications of DNA bases and their applications. *Journal of theoretical biology*, vol. 269, no. 1, pp. 123–130 (2011) doi: 10.1016/j.jtbi.2010.10.018h
- 15 Zhang, C. T., Zhang, R., Ou, H. Y.: The Z curve database: a graphic representation of genome sequences. *Bioinformatics*, vol. 19, no. 5, pp. 593–599 (2003)
- 16 Zhang, Y., Liao, B., Ding, K.: On 3dd-curves of DNA sequences. *Molecular Simulation*, vol. 32, no. 1, pp. 29–34 (2006) doi: 10.1080/08927020500517223