## COMMUNICATION

Check for updates

# Structure-based DNA memory with Boolean random access

Yanyu Zhou,†[a] Lanyang Gao,†[ab] Chu Jiang,[a] Yuxi Li,[a] Qi Liu,[a] Chengguo Xu,[a] Yingying Liu,[c] Huajie Liu [ID] *[a] and Yinan Zhang [ID] *[a]

Structure-based DNA memory represents a paradigm shift from nucleotide-encoded storage, circumventing the costly repetitive synthesis and sequencing dependencies while harnessing the programmable architecture of DNA nanostructures. However, the absence of random-access capability has constrained practical implementation. Here, we implement a Boolean search-enabled random-access scheme for structural DNA memory, wherein data is encoded on DNA origami tiles with orthogonally dimensioned index strands (1D/2D/3D addressing). Boolean operations are executed by hybridizing biotinylated probes to target index combinations, enabling the magnetic extraction of specific files. Atomic force microscopy validation confirms the precise retrieval of data across single-, dual-, and triple-indexed libraries. This approach establishes a robust framework for enabling random access in complex structural DNA databases.

DNA has emerged as a transformative medium for next-generation archival storage, capitalizing on its inherent molecular stability, unparalleled information density (up to 455 EB g$^{-1}$), and biochemical scalability.[1–4] Conventional sequence-encoded strategies, which involve mapping binary data to nucleotide arrangements,[5–9] are hindered by reliance on sequencing techniques and significant costs associated with synthesis-driven updates.[10–13] To transcend these constraints, structure-based DNA memory that harnesses programmable DNA nanostructures has been developed, establishing a storage paradigm that is independent of sequence encoding.[14–17]

## New concepts

DNA offers a promising solution for information storage due to its high density, fidelity, and durability. While traditional sequence-encoded strategies have been thoroughly researched, they continue to face limitations with sequence synthesis and sequencing. Structure-based DNA memory has emerged to address these issues, but research on random access methods remains limited. In this study, we have developed a DNA origami data memory system that utilizes orthogonal dimension index strands, facilitating random access to specific indexed DNA origami files through Boolean logic operations. Our findings provide an innovative approach to random access in structural DNA memory and establish a foundational framework for logic-gated archive queries within this context.

Structural DNA memory exploits programmable Watson–Crick base pairing to construct spatially organized data matrices for data encoding.[18–21] Through scaffold-stapling self-assembly, DNA origami enables the deterministic arrangement of data-encoding moieties, with an approximate positional variance of 5 nm.[22–25] This engenders fully addressable molecular canvases, thereby allowing for the precise positioning of elements in user-defined patterns to encode data.[26–28]

Random access refers to the selective retrieval of specific files or subsets from a larger database. By enhancing read efficiency and accuracy, this approach significantly reduces the time and costs of data retrieval and analysis, making it essential for large-scale data storage and dynamic operations.[29,30] For sequence-encoded DNA storage, a variety of random-access strategies—covering PCR amplification,[31] molecular interactions,[32] physical separation,[33,34] and fluorescence-activated sorting (FAS)[35]—have been developed. However, as a nascent alternative to sequence-encoded approaches, dedicated random-access methodologies have yet to be systematically established for DNA-nanostructure archives, primarily attributed to its recent emergence.[36,37] With the rapid progress of structural DNA memory, developing tailored random-access strategies for structural DNA memory has become an urgent imperative.

Here, we present a random-access architecture for DNA structural data archives, facilitating the on-demand retrieval of arbitrary file subsets. Validated through a 16-file origami
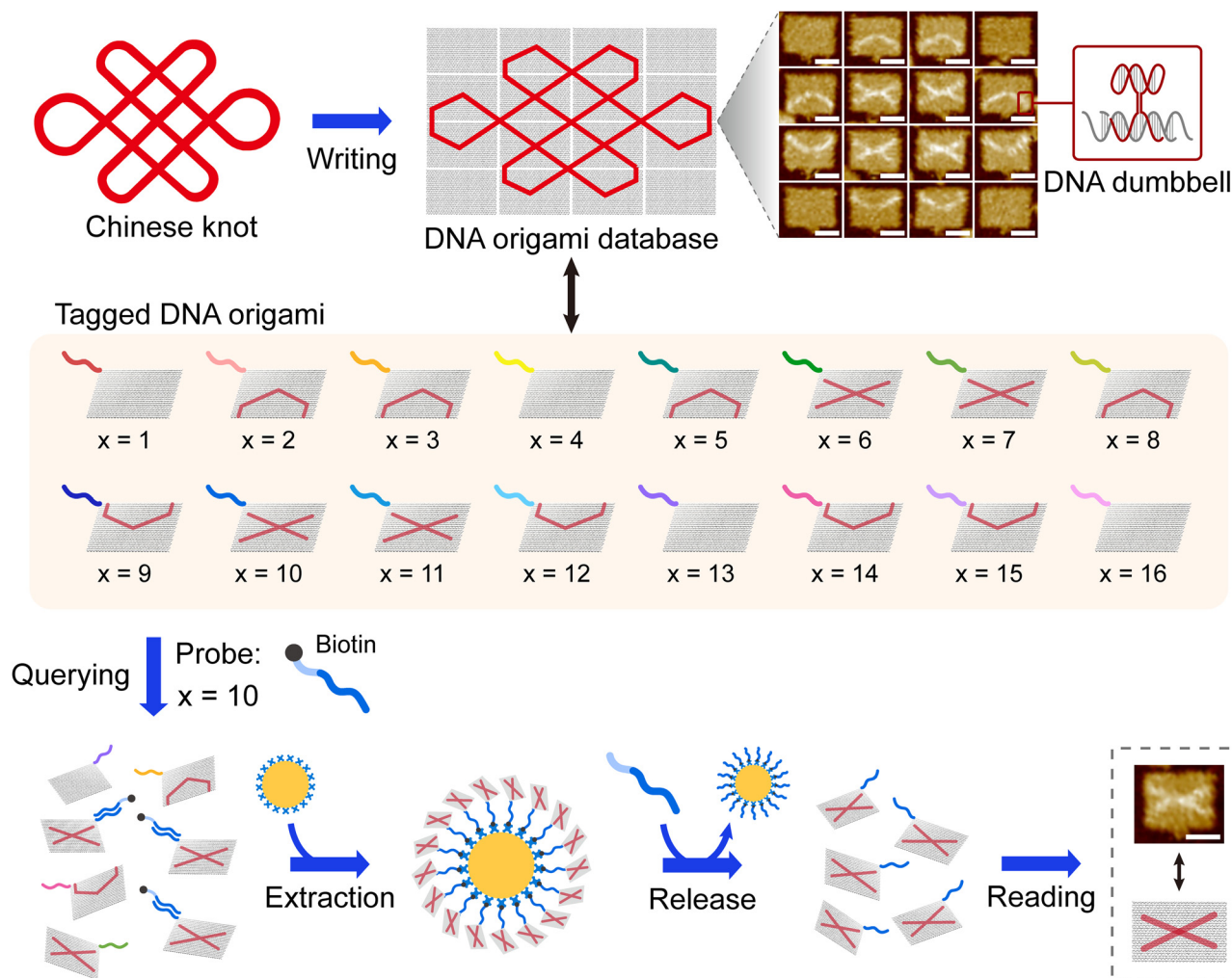
Fig. 1 Schematic diagram illustrating the random-access strategy for querying data, beginning with partitioning a traditional Chinese knot motif into 16 files converted to DNA patterns. Each pattern was inscribed on rectangular origami using dumbbell hairpin motifs, followed by the assignment of unique orthogonal indexes anchored *via* staple extensions. Biotinylated probes hybridized with corresponding indexes enabled streptavidin bead capture, after which toehold-mediated strand displacement triggered origami elution, and AFM imaging reconstructed patterns. Scale bars: 50 nm.

database encoding culturally significant Chinese knot motifs, our approach achieves a retrieval specificity exceeding 95%. We further enhanced the architecture by embedding orthogonal index sequences for each file, resulting in multi-tiered data addressing. By manipulating the index length, we demonstrate Boolean OR and AND operations, enabling flexible, logic-gated queries for arbitrary files. This work establishes a molecular foundation for logic-gated archival queries in structural DNA memory.

Fig. 1 illustrates the workflow for the coordinate-indexed random-access architecture. A traditional Chinese knot motif was partitioned into 16 discrete files, each converted into a DNA pattern inscribed onto a DNA origami structure using DNA dumbbells (Fig. S1). To enable precise retrieval, we assigned a unique orthogonal 20-nucleotide (nt) index to each DNA origami, corresponding to the coordinates of its respective file.[38] This yielded a DNA origami database containing 16 index-tagged origami structures. For selective retrieval, we designed a 28-nt biotinylated probe consisting of a 20-nt recognition

segment complementary to the target index and an 8-nt toehold domain. The probe hybridization facilitated the capture of the target DNA origami *via* streptavidin magnetic beads, after which toehold-mediated strand displacement triggered the release of the bound origami.[39–41] Finally, the retrieved DNA patterns were read out using atomic force microscopy (AFM) imaging.

We initially validated our approach using a two-file sub-database that contained DNA origami files $x = 1$ and $x = 2$ (Fig. S2a). AFM imaging confirmed the coexistence of both DNA origami at approximately equimolar ratios before selection (Fig. S2b and c). Upon introducing the $x = 2$ probe, the retrieved DNA origami predominantly exhibited the $x = 2$ pattern, achieving a specificity of 99.1% (Fig. S2b and d). To assess the scalability, we expanded the database to include four files. AFM verification revealed that the single-file query maintained a specificity of 98.7% (Fig. S3). These findings demonstrate our approach's ability for precise target file selection and its adaptability for complex DNA origami databases.
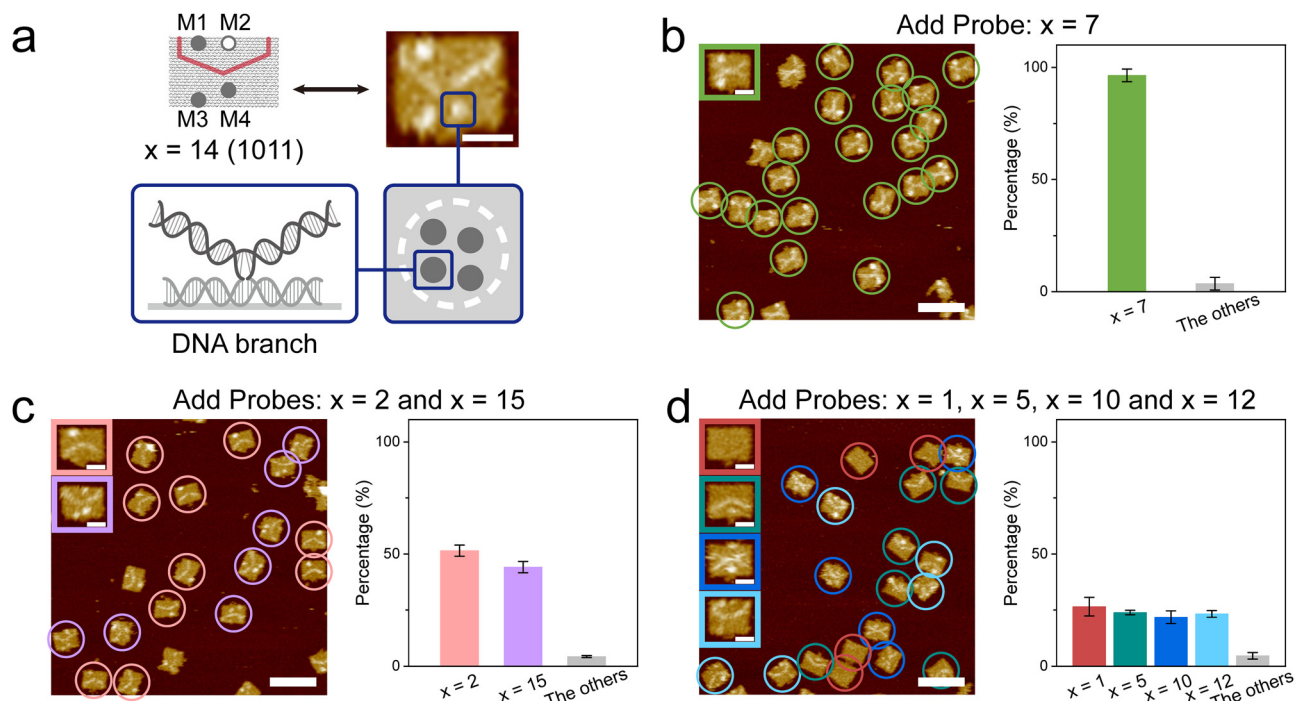
**Fig. 2** Querying files in the single-index Chinese knot database. (a) The arrangement of pixel markers composed of four DNA branches per pixel, exemplified by file $x = 14$ (solid/hollow = 1/0). (b)–(d) AFM images and statistical distribution of the retrieved structures, corresponding to querying one file (b), two files (c), and four files (d). Scale bar in (a): 50 nm. Scale bars in (b)–(d): 200 nm (insets: 50 nm).

Subsequently, we conducted a comprehensive random-access evaluation of the complete single-index Chinese knot database comprising all 16 DNA origami files. The inherent symmetry of the Chinese knot can lead to identical DNA patterns emerging from different coordinate assignments, compromising query validation. To address this, we implemented unique coordinate-specific signatures for each DNA origami by employing four asymmetric pixel markers (four DNA branches per pixel; solid/hollow circles denoting active/inactive states) (Fig. 2a and Fig. S4 and S5), enabling AFM discrimination of all 16 origami (Fig. S6).[42] We targeted the file $x = 7$ using the corresponding probe, with AFM imaging confirming the correct DNA origami and pixel markers. Statistical analysis showed that 96.4% of the retrieved structures matched the intended pattern and marker array, indicating high-fidelity file recovery (Fig. 2b).

Furthermore, we executed multi-file queries, successfully retrieving subsets of two files ($x = 2$ and $x = 15$) and four files ($x = 1$, $x = 5$, $x = 10$, and $x = 12$). AFM imaging corroborated the presence of anticipated DNA patterns, with pixel marker arrays validating correct coordinate assignments. Target DNA origami constituted 95.7% of the structures recovered in the two-file query (Fig. 2c) and 95.4% in the four-file query (Fig. 2d), indicating the retention of specificity even amidst simultaneous multi-file retrieval. Excess probes were used for effective retrieval, demonstrating that the retrieved abundance reflected the initial stoichiometric equivalence across all 16 origami. Additional file-subset queries consistently achieved accurate recovery of the desired DNA pattern with comparable specificity (Fig. S7–S9).

The densely addressable surface of DNA origami enables incorporation of multiple orthogonal indices per file.[43–45] This combinatorial strategy enhances the quantity of uniquely addressable files while reducing the dependency on entirely orthogonal sequences, broadening the database capacity and complexity. We constructed a dual-index Chinese knot database in which each DNA origami carries two orthogonal index strands encoding its $x$ and $y$ coordinates for precise 2D addressing (Fig. 3a). Boolean logic queries were implemented by tailoring index length: longer indexes enabled OR operations (activation if either index matches), whereas shorter indexes enforced AND operations (activation only when both indexes match).

To validate this, we assembled a sub-database containing only the (1, 1) and (2, 1) DNA origami files. Each index strand was segmented into non-interacting OR/AND logic domains, with probes exclusively hybridizing to target coordinate-specific segments (Fig. 3b). We systematically evaluated 15-nt and 20-nt OR indexes, confirming both achieved high-specificity target capture with comparable extraction yields of 96.7% *vs.* 98.1% (Fig. S10). Given oligonucleotide length constraints, the 15-nt configuration was selected for subsequent validation. Employing this design, we observed that probe OR $x = 1$ selectively enriched the file (1, 1), while the probe OR $x = 2$ retrieved only the file (2, 1) (Fig. S11a and b). Simultaneous addition of both probes recovered both targets at roughly equimolar ratios, demonstrating effective parallel selection (Fig. 3c, e and Fig. S11c).

Next, we investigated AND logic by systematically testing 7-, 9-, and 11-nt index segments under three conditions: AND $x = 2$
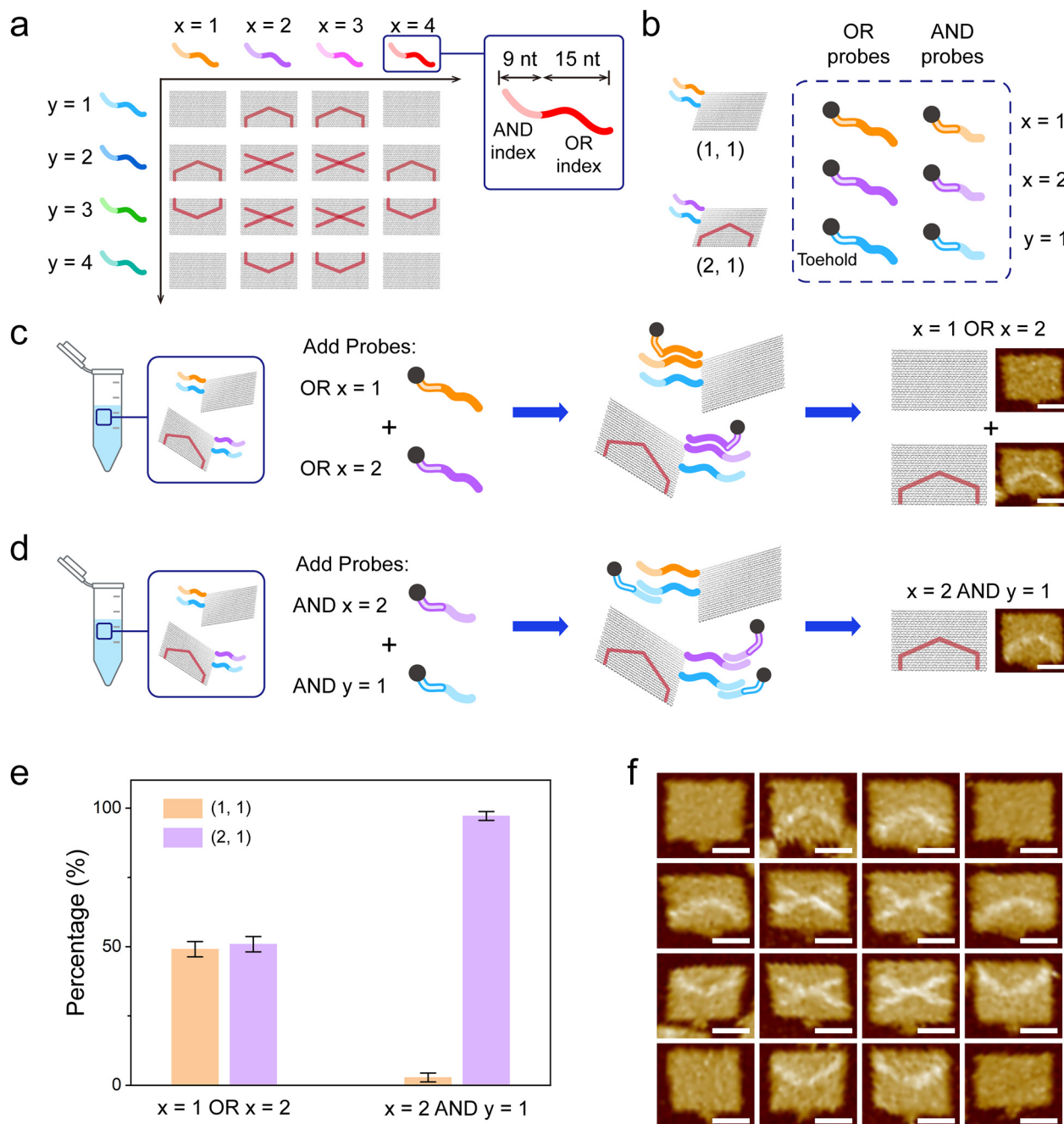
Fig. 3 Querying files in the dual-index Chinese knot database using Boolean logic operations. (a) Design of the dual-index Chinese knot database and the indexing scheme. (b) The two-file sub-database and its associated Boolean logic probes. (c)–(e) Schematic illustration of the OR operation $x = 1$ OR $x = 2$ (c) and the AND operation $x = 2$ AND $y = 1$ (d), along with the statistical distribution of the retrieved structures (e). (f) AFM images of file retrieval using the AND operation in the dual-index Chinese knot database. Scale bars: 50 nm.

only, AND $y = 1$ only, or both probes. AFM analysis indicated that 7-nt indexes exhibited insufficient binding, resulting in no file recovery even with both probes, while 11-nt indexes showed excessive binding, with a single probe sufficient for detection. Critically, the 9-nt index met the AND requirement, as file $(2, 1)$ was detected only with both probes, achieving a specificity of 97.2% (Fig. 3d, e and Fig. S12).

Applying this optimized 9-nt design for the dual-index Chinese knot database enabled unambiguous target retrieval,

though the increased non-target structures reduced the average specificity to 78.4% (Fig. 3f and Fig. S13). We omitted pixel markers to simplify AFM recognition. To enhance retrieval capacity, we developed a three-index database of 32 DNA origami featuring the Chinese knot ($z = 1$) and the Fangsheng pattern ($z = 2$) (Fig. 4a), the latter containing single-pixel identifiers (Fig. S14).

Implementing a three-condition AND gate required re-optimized index lengths. We compared binding segments of
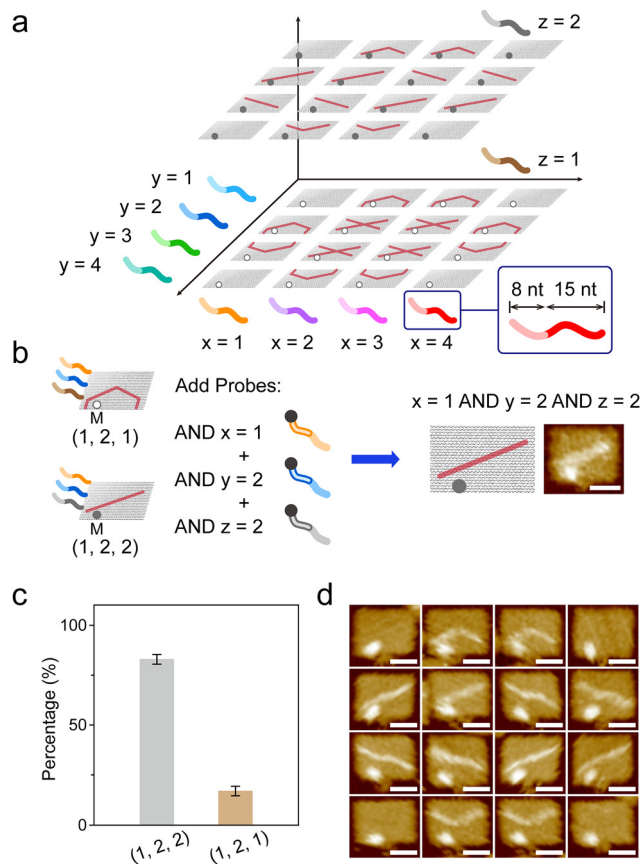
**Fig. 4** Querying files using three-condition AND operations in the three-index DNA origami database. (a) Design of the three-index DNA origami database and its indexing scheme. (b) and (c) Schematic illustration (b) and statistical distribution (c) of querying file (1, 2, 1) in the two-file sub-database. (d) AFM images of querying the Fangsheng pattern in the three-index DNA origami database using three-condition AND operations. Scale bars: 50 nm.

6-, 7-, and 8-nt designs in a two-file sub-database, in which only the 8-nt indexes yielded detectable signals, retrieving file (1, 2, 2) with 82.9% specificity (Fig. 4b, c and Fig. S15). Using these 8-nt indexes, we successfully isolated all Fangsheng files from the full database, demonstrating precise multi-index selection at scale (Fig. 4d).

## Conclusions

In summary, we developed a random-access structural DNA storage system that employs orthogonal index sequences for file tagging and selective retrieval. Index-specific probes hybridize with minimal crosstalk, while the integrated magnetic extraction and AFM readout enables file recovery with over 95% specificity. By scaling to multi-index databases and optimizing index lengths for Boolean logic operations, we demonstrated precise logic-gated queries while maintaining high selectivity. This strategy provides a robust random-access solution for structure-based DNA memory and establishes a framework for complex molecular database retrieval.

Additionally, increasing the number of mutually orthogonal indexes within a DNA origami structure can facilitate the execution of more intricate Boolean logic operations and enable multi-condition retrieval. However, extraction yields of extremely short indexes may be compromised even at higher multiplicities, necessitating systematic evaluation of the length-quantity balance. Simultaneously, more complex random-access operations can be supported without altering index sequences through adjusted probe modifications and optimized magnetic bead selection, where parameters such as biotin affinity tags should be systematically adjusted to maximize hybridization efficiency.

## Author contributions

Y. Z., and Y. Z. conceived the project. Y. Z. and L. G. conducted the experiments, analyzed the data, and interpreted the results. C. J., Y. L., Q. L. and C. X. assisted in AFM analysis. Y. Z., Y. L., H. L. and Y. Z. drafted the manuscript. Y. Z. supervised the research. All authors have approved the final version of the manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Data availability

Data supporting this work are available within this article and its supplementary information (SI). Supplementary information: materials, methods, Fig. S1–S15 and DNA sequences. See DOI: **https://doi.org/10.1039/d5nh00502g**.

## Acknowledgements

## References

1  L. Ceze, J. Nivala and K. Strauss, *Nat. Rev. Genet.*, 2019, **20**, 456–466.
2  V. Zhirnov, R. M. Zadegan, G. S. Sandhu, G. M. Church and W. L. Hughes, *Nat. Mater.*, 2016, **15**, 366–370.
3  M. Yu, X. Tang, Z. Li, W. Wang, S. Wang, M. Li, Q. Yu, S. Xie, X. Zuo and C. Chen, *Chem. Soc. Rev.*, 2024, **53**, 4463–4489.
4  W. Chen, M. Han, J. Zhou, Q. Ge, P. Wang, X. Zhang, S. Zhu, L. Song and Y. Yuan, *Natl. Sci. Rev.*, 2021, **8**, nwab028.
5  L. Anavy, I. Vaknin, O. Atar, R. Amit and Z. Yakhini, *Nat. Biotechnol.*, 2019, **37**, 1229–1236.

6  S. K. Tabatabaei, B. Wang, N. B. M. Athreya, B. Enghiad, A. G. Hernandez, C. J. Fields, J.-P. Leburton, D. Soloveichik, H. Zhao and O. Milenkovic, *Nat. Commun.*, 2020, **11**, 1742.

7  S. Hoshika, N. A. Leal, M.-J. Kim, M.-S. Kim, N. B. Karalkar, H.-J. Kim, A. M. Bates, N. E. Watkins Jr., H. A. SantaLucia, A. J. Meyer, S. DasGupta, J. A. Piccirilli, A. D. Ellington, J. SantaLucia Jr., M. M. Georgiadis and S. A. Benner, *Science*, 2019, **363**, 884–887.

8  Y. Choi, H. J. Bae, A. C. Lee, H. Choi, D. Lee, T. Ryu, J. Hyun, S. Kim, H. Kim, S.-H. Song, K. Kim, W. Park and S. Kwon, *Adv. Mater.*, 2020, **32**, 2001249.

9  H. H. Lee, R. Kalhor, N. Goela, J. Bolot and G. M. Church, *Nat. Commun.*, 2019, **10**, 2383.

10 L. Song, F. Geng, Z.-Y. Gong, X. Chen, J. Tang, C. Gong, L. Zhou, R. Xia, M.-Z. Han, J.-Y. Xu, B.-Z. Li and Y.-J. Yuan, *Nat. Commun.*, 2022, **13**, 5361.

11 N. Goldman, P. Bertone, S. Chen, C. Dessimoz, E. M. LeProust, B. Sipos and E. Birney, *Nature*, 2013, **494**, 77–80.

12 K. Chen, J. Zhu, F. Bošković and U. F. Keyser, *Nano Lett.*, 2020, **20**, 3754–3760.

13 G. D. Dickinson, G. M. Mortuza, W. Clay, L. Piantanida, C. M. Green, C. Watson, E. J. Hayden, T. Andersen, W. Kuang, E. Graugnard, R. Zadegan and W. L. Hughes, *Nat. Commun.*, 2021, **12**, 2371.

14 N. C. Seeman and H. F. Sleiman, *Nat. Rev. Mater.*, 2017, **3**, 17068.

15 C. Jiang, R. Tan, W. Li, Y. Zhang and H. Liu, *Small*, 2024, **20**, 2406470.

16 Y. Song, S. Kim, M. J. Heller and X. Huang, *Nat. Commun.*, 2018, **9**, 281.

17 S. Fan, D. Wang, J. Cheng, Y. Liu, T. Luo, D. Cui, Y. Ke and J. Song, *Angew. Chem. Int. Ed.*, 2020, **59**, 12991–12997.

18 Y. Zhang, F. Wang, J. Chao, M. Xie, H. Liu, M. Pan, E. Kopperger, X. Liu, Q. Li, J. Shi, L. Wang, J. Hu, L. Wang, F. C. Simmel and C. Fan, *Nat. Commun.*, 2019, **10**, 5469.

19 A. R. Chandrasekaran, O. Levchenko, D. S. Patel, M. MacIsaac and K. Halvorsen, *Nucleic Acids Res.*, 2017, **45**, 11459–11465.

20 J. Zhu, N. Ermann, K. Chen and U. F. Keyser, *Small*, 2021, **17**, 2100711.

21 H. Talbot, K. Halvorsen and A. R. Chandrasekaran, *ACS Synth. Biol.*, 2023, **12**, 978–983.

22 S. Dey, C. Fan, K. V. Gothelf, J. Li, C. Lin, L. Liu, N. Liu, M. A. D. Nijenhuis, B. Saccà, F. C. Simmel, H. Yan and P. Zhan, *Nat. Rev. Methods Primers*, 2021, **1**, 13.

23 Y. Zhang, Q. Li, X. Liu, C. Fan, H. Liu and L. Wang, *Small*, 2020, **16**, 2000793.

24 L. Xiao, X. Hu, Z. Zhou, X. Xie, S. Huang, M. Ji, A. Xu and Y. Tian, *Nanoscale*, 2025, **17**, 10411–10432.

25 S. Jia, J. Wang, M. Xie, J. Sun, H. Liu, Y. Zhang, J. Chao, J. Li, L. Wang, J. Lin, K. V. Gothelf and C. Fan, *Nat. Commun.*, 2019, **10**, 5597.

26 Y. Zhang, F. Li, M. Li, X. Mao, X. Jing, X. Liu, Q. Li, J. Li, L. Wang, C. Fan and X. Zuo, *J. Am. Chem. Soc.*, 2019, **141**, 17861–17866.

27 K. Chen, J. Kong, J. Zhu, N. Ermann, P. Predki and U. F. Keyser, *Nano Lett.*, 2019, **19**, 1210–1215.

28 C. Lin, R. Jungmann, A. M. Leifer, C. Li, D. Levner, G. M. Church, W. M. Shih and P. Yin, *Nat. Chem.*, 2012, **4**, 832–839.

29 L. Organick, S. D. Ang, Y.-J. Chen, R. Lopez, S. Yekhanin, K. Makarychev, M. Z. Racz, G. Kamath, P. Gopalan, B. Nguyen, C. N. Takahashi, S. Newman, H.-Y. Parker, C. Rashtchian, K. Stewart, G. Gupta, R. Carlson, J. Mulligan, D. Carmean, G. Seelig, L. Ceze and K. Strauss, *Nat. Biotechnol.*, 2018, **36**, 242–248.

30 C. Bee, Y.-J. Chen, M. Queen, D. Ward, X. Liu, L. Organick, G. Seelig, K. Strauss and L. Ceze, *Nat. Commun.*, 2021, **12**, 4764.

31 C. Plesa, A. M. Sidore, N. B. Lubock, D. Zhang and S. Kosuri, *Science*, 2018, **359**, 343–347.

32 K. J. Tomek, K. Volkel, A. Simpson, A. G. Hass, E. W. Indermaur, J. M. Tuck and A. J. Keung, *ACS Synth. Biol.*, 2019, **8**, 1241–1248.

33 S. Newman, A. P. Stephenson, M. Willsey, B. H. Nguyen, C. N. Takahashi, K. Strauss and L. Ceze, *Nat. Commun.*, 2019, **10**, 1706.

34 B. W. A. Bögels, B. H. Nguyen, D. Ward, L. Gascoigne, D. P. Schrijver, A.-M. Makri Pistikou, A. Joesaar, S. Yang, I. K. Voets, W. J. M. Mulder, A. Phillips, S. Mann, G. Seelig, K. Strauss, Y.-J. Chen and T. F. A. de Greef, *Nat. Nanotechnol.*, 2023, **18**, 912–921.

35 J. L. Banal, T. R. Shepherd, J. Berleant, H. Huang, M. Reyes, C. M. Ackerman, P. C. Blainey and M. Bathe, *Nat. Mater.*, 2021, **20**, 1272–1280.

36 S. Wang, X. Mao, F. Wang, X. Zuo and C. Fan, *Adv. Mater.*, 2024, **36**, 2307499.

37 Y. Hao, Q. Li, C. Fan and F. Wang, *Small Struct.*, 2021, **2**, 2000046.

38 J. Ye, J. Teske, U. Kemper and R. Seidel, *Small*, 2021, **17**, 2007218.

39 S. Ranallo, D. Sorrentino and F. Ricci, *Nat. Commun.*, 2019, **10**, 5509.

40 T. Luo, S. Fan, Y. Liu and J. Song, *Nanoscale*, 2021, **13**, 2100–2112.

41 R. Li, A. S. Madhvacharyula, Y. Du, H. K. Adepu and J. H. Choi, *Chem. Sci.*, 2023, **14**, 8018–8046.

42 Y. Ke, S. Lindsay, Y. Chang, Y. Liu and H. Yan, *Science*, 2008, **319**, 180–183.

43 J. M. Weck and A. Heuer-Jungemann, *Nat. Commun.*, 2025, **16**, 1556.

44 M. T. Strauss, F. Schueder, D. Haas, P. C. Nickels and R. Jungmann, *Nat. Commun.*, 2018, **9**, 1600.

45 L. M. Wassermann, M. Scheckenbach, A. V. Baptist, V. Glembockyte and A. Heuer-Jungemann, *Adv. Mater.*, 2023, **35**, 2212024.