



Czy plastik może rozpocząć nową erę w archiwizacji danych?

Agnieszka Lizak

Sieć Badawcza Łukasiewicz – PORT Polskie Centrum Rozwoju Technologii / Lukaszewicz Research Network – PORT Polish Center for Technology Development (Poland)
agnieszka.lizak@port.lukasiewicz.gov.pl, ORCID 0000-0001-8599-8065

Róża Szweda

Sieć Badawcza Łukasiewicz – PORT Polskie Centrum Rozwoju Technologii / Lukaszewicz Research Network – PORT Polish Center for Technology Development (Poland)
Roza.Szweda@port.lukasiewicz.gov.pl, ORCID 0000-0003-2152-7656

STRESZCZENIE

Wraz z szybkim rozwojem technologii informatycznych wiele aspektów naszego życia przechodzi cyfrową transformację. Z roku na rok w sieci pojawia się coraz więcej użytkowników, a ciągle udoskonalana sztuczna inteligencja zyskuje na popularności, co prowadzi do wzrostu produkcji informacji. Informacje te przechowywane są głównie w centrach danych, które – przy nieustającym napływie nowych treści – są zmuszone powiększać swoją powierzchnię. Łączy się to z coraz większymi kosztami utrzymania i zapotrzebowaniem na energię. W 2021 r. zużycie energii elektrycznej przez centra danych wyniosło 220–320 TWh, co odpowiada około 0.9–1.3% globalnego zapotrzebowania. Bazy danych wymagające ciągłego dostarczania energii są czynnikiem generującym około 1% całkowitej emisji dwutlenku węgla. Ponadto raportowano już, że przy rosnącym wykładniczo zapotrzebowaniu, za 10 lat ilość dostępnego krzemu klasy mikroprocesorowej nie będzie wystarczająca do przechowywania wszystkich danych. Dlatego naukowcy poszukują alternatyw dla obecnych nośników danych i opracowują nowe technologie z wykorzystaniem cząsteczek chemicznych, a w ostatnim czasie plastiku. W artykule przedstawione zostały przykłady rozwijających się technologii przechowywania danych cyfrowych w polimerach. Omówiono polimery jako nośniki danych i porównano do obecnie stosowanych rozwiązań. Przedyskutowano także, czy plastik może stać się przyszłym materiałem do archiwizacji informacji.

SŁOWA KLUCZOWE

archiwizacja danych,
nośniki danych,
cyfrowe polimery,
makrocząsteczki
zawierające
informacje

Can plastic start a new era in data archiving?

ABSTRACT

With the rapid development of information technology, many aspects of our lives are undergoing a digital transformation. An increasing number of users are going online every year, and constantly improving artificial intelligence is gaining popularity, which leads to the growing production of

KEYWORDS

data archiving,
data carriers,
digital polymers,
macromolecules

information. Nowadays, information is usually stored in data centres, which will be forced to increase their space with the constant flow of new bits of information. Together with the increase in their space, energy consumption and associated maintenance costs are escalating. In 2021, global data centre power consumption was 220–320 TWh, which is about 0.9–1.3% of global power consumption. Continuous power supply for database operations is responsible for about 1% of total carbon dioxide emissions. Furthermore, it has already been reported that with the exponentially growing amount of data, in about 20 years, the amount of silicon for microprocessors will no longer be sufficient to store all the information. Therefore, scientists are looking for alternatives to the currently used data storage solutions and are developing new technologies using chemical molecules. Recently, even plastic has been explored as a data carrier. In this work, we present examples of new technologies for data storage in polymers. We have discussed polymers as data carriers in comparison with currently used solutions and deliberated whether plastic can become a future material for information archiving.

containing
information

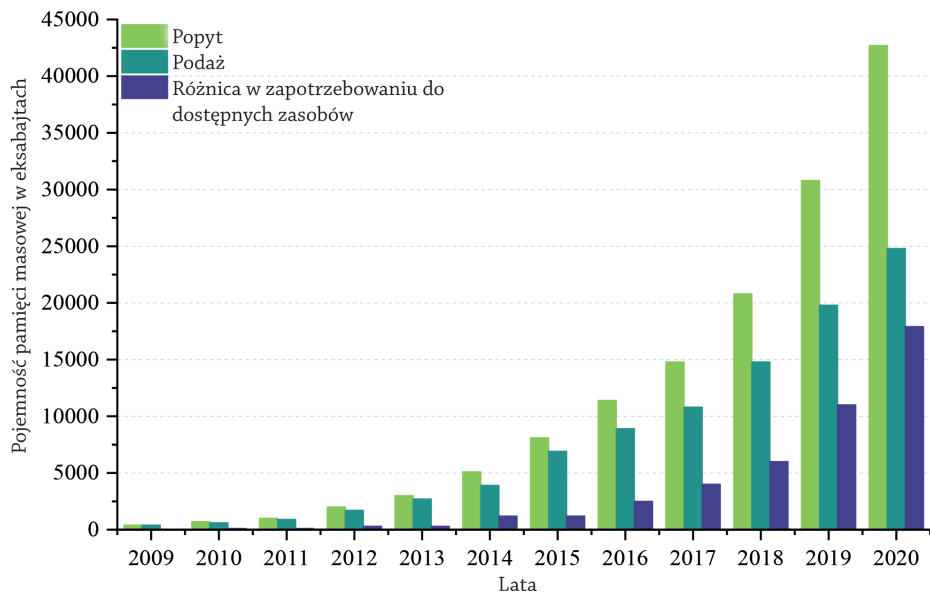
Wstęp¹

Już od czasów starożytnych ludzie musieli przechowywać informacje, aby przekazać zdobytą wiedzę następnym pokoleniom. Na przestrzeni wieków metodologia ich przechowywania znacznie się rozwinęła. Jednak, pomimo że dzisiejsze rozwiązania technologiczne pozwalające na przechowywane danych są bardzo zaawansowane, nie nadążają one za rosnącą ilością wytwarzanych informacji, w języku komputerów znaną jako liczba bitów². Obecnie ilość generowanych danych jest prawie dwukrotnie większa niż pojemność konwencjonalnych nośników, takich jak dyski twarde czy urządzenia pamięci flash, a tendencja ta bardzo szybko wzrasta (Rysunek 1)³.

¹ Praca powstała w wyniku realizacji projektu badawczego o nr. 2021/42/E/ST4/00010 finansowanego ze środków Narodowego Centrum Nauki.

² A. Extance, *How DNA could store all the world's data*, „Nature” 2016, t. 537, s. 22–24, DOI: 10.1038/537022a.

³ Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere, Nov 2018, <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf> [dostęp: 21.11.2022].



Rysunek 1. Popyt i podaż na przechowywanie danych na świecie⁴.

Jednym z najbardziej znaczących ograniczeń powszechnie używanych dysków twardych jest ich ograniczona stabilność⁵. W konsekwencji co mniej więcej 10 lat dane muszą być przegrywane na nowy nośnik⁶. Ponadto prognozy pokazują, że za niecałych 20 lat roczne zapotrzebowanie na przechowywanie danych cyfrowych przewyższy podaż krzemu klasy mikroprocesorowej, przy założeniu, że wszystkie dane będą przechowywane w pamięci flash⁷. Co więcej, utrzymanie centrów danych wymaga ogromnych ilości energii elektrycznej – jej zużycie przez centra danych w 2021 r. wyniosło 220–320 TWh, czyli około 0.9–1.3% globalnego zapotrzebowania⁸. Centra danych odpowiadają również za blisko 1% emisji gazów cieplarnianych. To wszystko powoduje, że obecne rozwiązania, w obliczu

⁴ Statista. Data storage supply and demand worldwide, from 2009 to 2020, <https://www.statista.com/statistics/751749/worldwide-data-storage-capacity-and-demand/> [dostęp: 09.11.2022]; L. Dhar, *Molecular Storage, Data Writ Small*, „ACS Central Science” 2019, t. 5, 753–754, DOI: 10.1021/acscentsci.9b00439.

⁵ L.C. Meiser, B.H. Nguyen, Y.J. Chen, J. Nivala, K. Strauss, L. Ceze, R.N. Grasset, *Synthetic DNA applications in information technology*, „Nature Communications” 2022, t. 13, 352, DOI: 10.1038/s41467-021-27846-9.

⁶ Ibidem.

⁷ V. Zhirnov, R.M. Zidegan, G.S. Sandhu, G.M. Church, W.L. Hughes, *Nucleic acid memory*, „Nature Materials” 2016, t. 15, s. 366–370, DOI: 10.1038/nmat4594.

⁸ IEA (2022). Data Centres and Data Transmission Networks, IEA, Paris, <https://www.iea.org/reports/data-centres-and-data-transmission-networks> [dostęp: 09.11.2022]; E. Masanet,

stale rosnącej potrzeby przechowywania coraz większej ilości informacji, nie są wystarczające. Zważywszy na te okoliczności, wkrótce będziemy mieli poważny problem z archiwizacją danych, dlatego konieczne jest rozwijanie nowych technologii, które pozwolą na przechowywanie danych w sposób energooszczędny, bezpieczny i długotrwały.

Przyglądając się żywej materii i organizmom biologicznym, które ewoluowały na przestrzeni miliardów lat, dążąc do optymalizacji ich adaptacji, można zauważyć, że do przechowywania informacji wykorzystują makrocząsteczki. Głównym nośnikiem danych w naturze są kwasy nukleinowe, w których informacja zakodowana jest na poziomie molekularnym w sekwencji nukleotydów. Sekwencja DNA zawiera instrukcje niezbędne do reprodukcji i rozmnażania wszystkich znanych organizmów żywych⁹. Jeśli spojrzymy na DNA z perspektywy chemika, to widzimy, że jest to makrocząsteczka zbudowana z powtarzających się podjednostek, w nomenklaturze naukowej zwanych merami. Każdy łańcuch DNA charakteryzuje się ściśle określoną sekwencją podjednostek¹⁰. Makrocząsteczki kwasów nukleinowych stały się inspiracją dla naukowców do opracowania nowej technologii przechowywania informacji na poziomie molekularnym. W 2012 r. opublikowano pionierskie prace, w których syntetyczne cząsteczki DNA wykorzystano jako nośniki danych¹¹. Wykazano, że DNA może być wykorzystane do przechowywania informacji cyfrowej, czytelnej dla komputera. Pokazano, że tekst, obrazy lub instrukcje procesora komputerowego można zapisać w sekwencji syntetycznego DNA przy zastosowaniu systemu binarnego¹². Kod binarny reprezentowany jest przez dwa symbole – „0” i „1”, które przypisano wybranym podjednostkom, np. adeninę (A) lub cytozynę (C) opisano jako cyfrowe „0”, natomiast guaninę (G) lub tyminę (T) jako „1”. Zapis danych odbywa się poprzez syntezę chemiczną, która w obecnych czasach jest

A. Shehabi, N. Lei, S. Smith, J. Koomey, *Recalibrating global data center energy-use estimates*, „Science” 2020, t. 367, s. 984–986, DOI: 10.1126/science.aba3758.

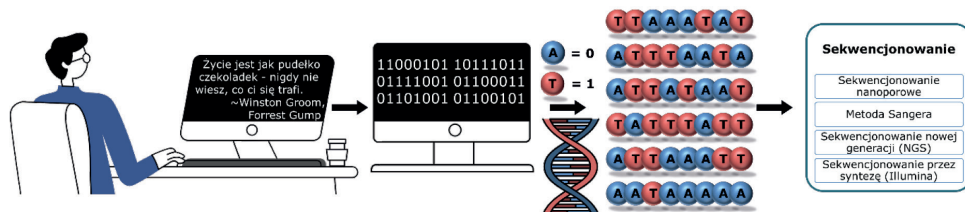
⁹ G.M. Cooper, *The Cell: A Molecular Approach*, 2nd edition, Sunderland (MA) 2000.

¹⁰ B. Alberts, A. Johnson, J. Lewis i in., *Molecular Biology of the Cell*, 4th edition, Nowy Jork 2002, The Structure and Function of DNA.

¹¹ N. Goldman i in., *Towards practical, high-capacity, low-maintenance information storage in synthesized DNA*, „Nature” 2013, t. 494, s. 77–80, DOI: 10.1038/nature11875; G.M. Church, Y. Gao, S. Kosuri, *Next-Generation Digital Information Storage in DNA*, „Science” 2012, t. 337, s. 1628, DOI: 10.1126/science.1226355.

¹² A. Żyła, *DNA: prawie niezniszczalny i najbardziej pojemny nośnik danych*, „Archeion” 2021, t. 122, s. 33–43, DOI: 10.4467/26581264ARC.21.014.14494.

w pełni zautomatyzowana. Natomiast odczyt danych może być wykonany przez dowolną technologię sekwencjonowania DNA (Rysunek 2)¹³.



Rysunek 2. Koncepcja zapisu informacji w sekwencji DNA. Cytat przekonwertowano na kod binarny przy wykorzystaniu języka ASCII. Następnie kod binarny jest zapisany w sekwencji podjednostek A i T odpowiadającym symbolom 0 i 1. Informacja może być odczytana dowolną techniką sekwencjonowania DNA.

Głównymi zaletami przechowywania informacji w syntetycznym DNA jest ulepszonej stabilność nośnika, brak konieczności dostarczania energii i duża gęstość przechowywania informacji. Dostępne dane pokazują, że zliofilizowane DNA w warunkach normalnych jest stabilne przez co najmniej 100 lat¹⁴. Teoretycznie, magazynowanie cyfrowego DNA nie wymaga dodatkowej energii. Makrocząsteczki DNA są tak małe, że wszystkie dostępne dane świata można zapisać i przechować w 1 gramie materiału¹⁵.

¹³ M.T. Pervez, M.J. Hasnain, S.H. Abbas, M.F. Moustafa, N. Aslam, S.S.M. Shah, *A Comprehensive Review of Performance of Next-Generation Sequencing Platforms*, „BioMed Research International” 2022, t. 2022, 3457806, DOI: 10.1155/2022/3457806.

¹⁴ E. Podivinsky, J.L. Love, L. van der Colff, L. Samuel, *Effect of storage regime on the stability of DNA used as a calibration standard for real-time polymerase chain reaction*, „Analytical Biochemistry” 2009, t. 394, s. 132–134, DOI: 10.1016/j.ab.2009.06.024; A. Baoutina, S. Bhat, L. Partis, K.R. Emslie, *Storage Stability of Solutions of DNA Standards*, „Analytical Chemistry” 2019, t. 91, s. 12268–12274, DOI: 10.1021/acs.analchem.9b02334; J. Bonnet i in., *Chain and conformation stability of solid-state DNA: implications for room temperature storage*, „Nucleic Acids Research” 2010, t. 38, s. 1531–1546, DOI: 10.1093/nar/gkp1060; N.V. Ivanova, M.L. Kuzmina, *Protocols for dry DNA storage and shipment at room temperature*, „Molecular Ecology Resources” 2013, t. 13, s. 890–898, DOI: 10.1111/1755-0998.12134.

¹⁵ R. Service, *DNA could store all of the world's data in one room*, „Science” 2017, s. 1–2, DOI: 10.1126/science.aal0852.

Jednakże, pomimo licznych hipotetycznych korzyści, technologia przechowywania danych z wykorzystaniem DNA pociąga za sobą wiele ograniczeń:

- brak możliwości sterowania właściwościami makrocząsteczek – do przechowywania danych w DNA można stosować jedynie nukleotydy, które determinują właściwości polimeru i ograniczają metody odczytu informacji do dostępnych technologii sekwencjonowania;
- ograniczona żywotność – kwasy nukleinowe tworzące podwójną helisę są relatywnie stabilne w postaci stałej, jednak do odczytu zakodowanej informacji wymagane jest ich rozpuszczenie. Długotrwałe przechowywanie DNA w roztworze często wymaga niskiej temperatury otoczenia oraz odpowiedniego medium, w zależności od metody enkapsulacji i metody odczytu informacji¹⁶. Częstotliwość odczytu informacji skraca czas stabilności DNA, dlatego – aby archiwizować dane – odczyt musi być ograniczony, np. raz na 10 lat¹⁷;
- ograniczona gęstość przechowywania danych do 2 bitów na zasadę¹⁸. Dotąd największa osiągnięta gęstość danych, którą można było uzyskać to 1.3610²⁰ bit/gram¹⁹. Wynika to z ograniczonej liczby podjednostek budujących łańcuch DNA do czterech nukleotydów;
- ograniczona możliwość edycji sekwencji – dla DNA przegrupowanie sekwencji nie jest łatwe do osiągnięcia i wymaga modyfikacji epigenetycznych lub użycia maszynarii bakteryjnej²⁰. Dlatego syntetyczne DNA nie pozwala na edycję zapisanych danych przy użyciu bodźców natury fizycznej, które mają potencjał do powszechnego zastosowania;
- sekwencjonowanie DNA jest zaawansowaną, kosztowną i obciążoną błędami metodą²¹.

¹⁶ H. Tateishi-Karimata, N. Sugimoto, *Structure, stability and behaviour of nucleic acids in ionic liquids*, „Nucleic Acids Research” 2014, t. 42, s. 8831–8844, DOI: 10.1093/nar/gku499.

¹⁷ K. Matange, J.M. Tuck, A.J. Keung, *DNA stability: a central design consideration for DNA data storage systems*, „Nature Communications” 2021, t. 12, s. 1358, DOI: 10.1038/s41467-021-21587-5.

¹⁸ Y. Hao, Q. Li, C. Fan, F. Wang, *Data Storage Based on DNA*, „Small Structures” 2021, t. 2, s. 2000046, DOI: 10.1002/ssstr.202000046.

¹⁹ L. Organick i in., *Probing the physical limits of reliable DNA data retrieval*, „Nature Communications” 2020, t. 11, s. 616, DOI: 10.1038/s41467-020-14319-8.

²⁰ J. Bonnet, P. Subsoontorn, D. Endy, *Rewritable digital data storage in live cells via engineered control of recombination directionality*, „Proceedings of the National Academy of Sciences” 2012, t. 109, s. 8884–8889, DOI: 10.1073/pnas.1202344109; C. Mayer, G.R. McInroy, P. Murat, P. van Delft, S. Balasubramanian, *An Epigenetics-Inspired DNA-Based Data Storage System*, „Angewandte Chemie” 2016, t. 128, s. 11310–11314, DOI: 10.1002/anie.201605531.

²¹ B. Langmead, S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2*, „Nature Methods” 2012, t. 9, s. 357–359, DOI: 10.1038/nmeth.1923; F. Pfeiffer i in., *Systematic evaluation of error rates*

Pomimo znacznego rozwoju metod sekwencjonowania na przestrzeni ostatnich lat, co pozwoliło na znaczne obniżenie kosztów odczytu, metody te są znacznie bardziej skomplikowane niż tradycyjny odczyt informacji z dysków twardych. W konsekwencji wydłużony czas odczytu, koszt i generowane błędy stanowią barierę dla ich powszechnego zastosowania.

W ostatnich latach syntetyczne polimery wyłoniły się jako obiecująca alternatywa nośników do przechowywania danych na poziomie molekularnym²². Przełomowe prace grupy Jean-François Lutz²³ były jednymi z pierwszych przykładów pokazujących, że polimery o zdefiniowanej sekwencji (SDP, ang. *sequence-defined polymers*) rzeczywiście mogą przechowywać informacje cyfrowe. Koncepcja ta była rozwijana wraz z postępem w dziedzinie chemii polimerów²⁴.

Jak zapisać dane w polimerach?

Warunkiem koniecznym do zapisu informacji w polimerach jest pełna kontrola nad sekwencją ułożenia merów w łańcuchu. Istnieją różne strategie syntezy polimerów o zdefiniowanej sekwencji²⁵, spośród których najbardziej rozpowszechnionym podejściem jest synteza na nośniku stałym, gdzie – rozpoczynając od nośnika (np. żywicy polimerowej) – przyłączany jest kolejno mer po merze²⁶. Zastosowanie nośnika stałego w syntezie pozwala na prostą automatyzację pro-

and causes in short samples in next-generation sequencing, „Scientific Reports” 2018, t. 8, s. 10950, DOI: 10.1038/s41598-018-29325-6.

²² H.M. Colquhoun, J.F. Lutz, *Information-containing macromolecules*, „Nature Chemistry” 2014, t. 6, s. 455–456, DOI: 10.1038/nchem.1958; D. Andrieux, P. Gaspard, *Nonequilibrium generation of information in copolymerization processes*, „Proceedings of the National Academy of Sciences” 2008, t. 105, s. 9516–9521, DOI: 10.1073/pnas.0802049105.

²³ R.K. Roy i in., *Design and synthesis of digitally encoded polymers that can be decoded and erased*, „Nature Communications” 2015, t. 6, s. 7237, DOI: 10.1038/ncomms8237; J.F. Lutz, *Coding Macromolecules: Inputting Information in Polymers Using Monomer-Based Alphabets*, „Macromolecules” 2015, t. 48, s. 4759–4767, DOI: 10.1021/acs.macromol.5b00890.

²⁴ M.G.T.A. Rutten, F.W. Vaandrager, J.A.A.W. Elemans, R.J.M. Nolte, *Encoding information into polymers*, „Nature Reviews Chemistry” 2018, t. 2, s. 365–381, DOI: 10.1038/s41570-018-0051-5.

²⁵ P. Nanjan, M. Porel, *Sequence-defined non-natural polymers: synthesis and applications*, „Polymer Chemistry” 2019, t. 10, s. 5406–5424, DOI: 10.1039/C9PY00886A; J.F. Lutz, M. Ouchi, D.R. Liu, M. Sawamoto, *Sequence-Controlled Polymers*, „Science” 2013, t. 341, s. 1238149, DOI: 10.1126/science.1238149.

²⁶ R.B. Merrifield, *Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide*, „Journal of the American Chemical Society” 1963, t. 85, s. 2149–2154, DOI: 10.1021/ja00897a025.

cesu, podobnie jak w przypadku syntezy DNA, dzięki czemu proces może zostać łatwo skomercjalizowany²⁷. Dzięki kontroli sekwencji merów w łańcuchu możliwe jest zakodowanie informacji w polimerze, korzystając z analogicznej koncepcji jak zapis informacji w DNA. Wybrane mery oznaczają się jako symbole „0” oraz „1”. W ostatniej dekadzie opracowano szereg metod pozwalających na otrzymanie różnego rodzaju polimerów o zdefiniowanej sekwencji w szerokim przedziale właściwości chemicznych i fizycznych²⁸.

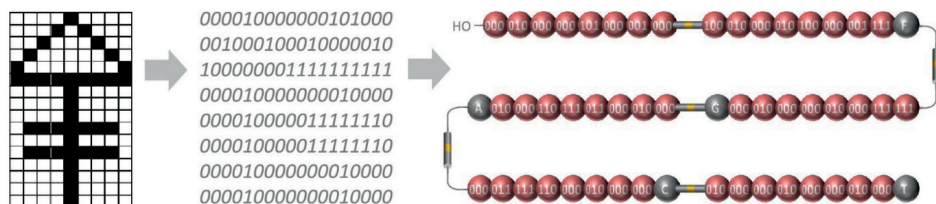
Wykorzystanie syntetycznych polimerów pozwala na powiększenie biblioteki bloków budulcowych, co daje możliwość zagęszczenia kodowanej informacji²⁹. Na przykład, wykorzystując bibliotekę zawierającą 8 monomerów, pokazano jak za pomocą automatycznej syntezy otrzymano „cyfrowe polimery”³⁰. Rozszerzenie alfabetu merów pozwoliło na zwiększenie gęstości danych do 3 bitów/monomer (Rysunek 3).

²⁷ J.M. Lee, J. Kwon, S.J. Lee, H. Jang, D.G. Kim, J. Song, K.T. Kim, *Semiautomated synthesis of sequence-defined polymers for information storage*, „Science Advances” 2022, t. 8, s. eabl8614, DOI: 10.1126/sciadv.abl8614; C. Mertens, M. Soete, M.L. Ślęczkowski, A.R.A. Palmans, E.W. Meijer, N. Badi, F.E. Du Prez, *Stereocontrolled, multi-functional sequence-defined oligomers through automated synthesis*, „Polymer Chemistry” 2020, t. 11, s. 4271–4280, DOI: 10.1039/d0py00645a; J.O. Holloway, C. Mertens, F.E. du Prez, N. Badi, *Automated Synthesis Protocol of Sequence-Defined Oligo-Urethane-Amides Using Thiolactone Chemistry*, „Macromolecular Rapid Communications” 2019, t. 40, s. 1800685, DOI: 10.1002/marc.201800685; A. Al Ouahabi, M. Kotera, L. Charles, J.F. Lutz, *Synthesis of Monodisperse Sequence-Coded Polymers with Chain Lengths above DP100*, „ACS Macro Letters” 2015, t. 4, s. 1077–1080, DOI: 10.1021/acsmacrolett.5b00606; J. Vandenberg, G. Reekmans, P. Adriaenssens, T. Junkers, *Synthesis of sequence-defined acrylate oligomers via photo-induced copper-mediated radical monomer insertions*, „Chemical Science” 2015, t. 6, s. 5753–5761, DOI: 10.1039/C5SC02035B.

²⁸ E. Laurent, R. Szweida, J.F. Lutz, *Synthetic Polymers with Finely Regulated Monomer Sequences: Properties and Emerging Applications*, [w:] *Macromolecular Engineering*, red. N. Hadjichristidis, Y. Gnanou, K. Matyjaszewski i M. Muthukumar, bm. 2022, s. 1–34, DOI: 10.1002/9783527815562.mme0041.

²⁹ S. Martens, A. Landuyt, P. Espeel, B. Devreese, P. Dawyndt, F. Du Prez, *Multifunctional sequence-defined macromolecules for chemical data storage*, „Nature Communications” 2018, t. 9, s. 4451, DOI: 10.1038/s41467-018-06926-3; E. Laurent, J.A. Amalian, M. Parmentier, L. Oswald, A. Al Ouahabi, F. Dufour, K. Launay, J.L. Clément, D. Gimes, M.A. Delsuc, L. Charles, J.F. Lutz, *High-Capacity Digital Polymers: Storing Images in Single Molecules*, „Macromolecules” 2020, t. 53, s. 4022–4029, DOI: 10.1021/acs.macromol.0c00666.

³⁰ E. Laurent, J.A. Amalian, M. Parmentier i in., op.cit.



Rysunek 3. Idea kodowania czarno-białego obrazu o rozmiarach od 80 do 144 pikseli poprzez przekonwertowanie na kod binarny i dalej na sekwencję polimerów. Każdy z ośmiu wykorzystanych monomerów koduje triadę (3 bity). Reprinted with permission from³¹. Copyright 2020 American Chemical Society.

Inne podejście pozwalające na zwiększenie gęstości przechowywania danych zaproponowała grupa Filipa Du Preza³², która zastosowała rozszerzony alfabet monomerów. Tekst, gdzie każda litera to 8 bitów, zapisano korzystając z biblioteki 19 monomerów. Każdy z monomerów odpowiadał jednej literze, co pozwoliło na zakodowanie 8 bitów/monomer. Zademonstrowano także, jak zapisać kod QR przy użyciu biblioteki 15 monomerów akrylanowych, które oznaczono jako A, B, C... O (Rysunek 4). Ciąg znaków binarnych reprezentujących kod QR rozdzielono na mniejsze części, a do każdej z nich dołączono fragment z informacją o pozycji (fioletowy) oraz długości (niebieski) ciągu bitów. Po określeniu kolejności wszystkich fragmentów są one sortowane, przycinane i sklejane przez program komputerowy. Na koniec sklejona informacja zamieniana jest na ciąg bitów, który rekonstruuje oryginalny kod QR.

Ponadto syntetyczne polimery można w łatwy sposób sekwencyjnie zorganizować i otrzymać „materiał cyfrowy”³³. Korzystając z techniki warstwa po warstwie (ang. Layer-by-Layer), w sposób kontrolowany ułożono sekwencję monowarstwowych filmów z polimerów zawierających informację cyfrową (Rysunek 5). W tym przypadku każdy polimer biblioteki zawierał 10 bajtów tekstu zapisanego w kodzie binarnym (tj. 80 zakodowanych monomerów, 1 bit/monomer). Metoda ta umożliwia przestrzenną organizację łańcuchów polimerowych zawierających informację w sposób kontrolowany, co pozwala na przezwycięzenie problemu

³¹ Ibidem.

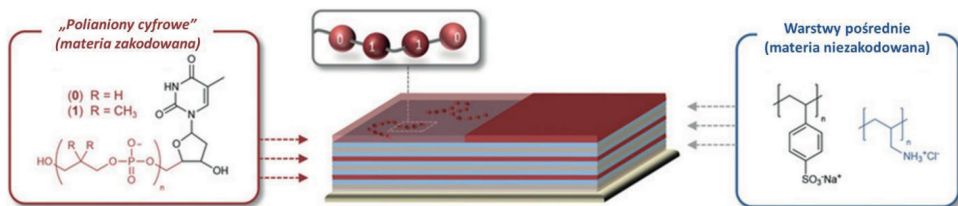
³² S. Martens, A. Landuyt, P. Espeel, B. Devreese, P. Dawyndt, F. Du Prez, op.cit.

³³ R. Szweda, M. Tschopp, O. Felix, G. Decher, J.F. Lutz, *Sequences of Sequences: Spatial Organization of Coded Matter through Layer-by-Layer Assembly of Digital Polymers*, „Angewandte Chemie International Edition” 2018, t. 57, s. 15817–15821, DOI: 10.1002/anie.201810559.

ograniczonej długości łańcucha. Podejście to pokazuje także, że polimery nie muszą być ekstremalnie długimi łańcuchami, aby przechowywać duże ilości danych, ponieważ można je sekwencyjnie upakować.



Rysunek 4. Schematy kodowania (po lewej) i dekodowania (po prawej) kodu QR zaproponowane przez grupę Du Preza³⁴. Reprinted with permission from³⁵. Copyright 2018 Creative Commons.



Rysunek 5. Koncepcja produkcji „materiałów cyfrowych” przy wykorzystaniu metody warstwa po warstwie. Polimery syntetyczne (polielektrolity) z zakodowaną informacją są w sposób kontrolowany zorganizowane na powierzchni, dzięki czemu możemy zwiększyć pojemność przechowywanych informacji. Reprinted with permission from³⁶. Copyright 2018 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.

³⁴ Ibidem.
³⁵ Ibidem.
³⁶ Ibidem.

Jakimi metodami można odczytać informacje z polimerów?

Techniki sekwencjonowania zostały pierwotnie opracowane do analizy biopolimerów, takich jak białka³⁷ i DNA³⁸. Aktualnie proces ten można przeprowadzić w sposób wydajny przy użyciu automatycznych sekwenserów, jednak często jest to bardzo skomplikowany sposób, który wymaga zaangażowania doświadczonego operatora. Metody sekwencjonowania polimerów abiotycznych są znacznie mniej ugruntowane. Dotychczas zademonstrowano jedynie kilka metod odczytywania informacji zakodowanej w polimerach abiotycznych, w tym spektrometrię mas, jądrowy rezonans magnetyczny czy sekwencjonowanie nanoporowe³⁹. Spośród tych metod najczęściej stosowana jest spektrometria mas.

W odróżnieniu od naturalnych makrocząsteczek, struktury polimerów mogą być odpowiednio dostosowane, co umożliwia łatwiejszy odczyt sekwencji⁴⁰. Na przykład do struktury polimerów wprowadzane są wiązania chemiczne, które w łatwy sposób fragmentują podczas sekwencjonowania z wykorzystaniem spektrometrii mas. Pozwala to na minimalizację błędów odczytu oraz upraszcza analizę. Ponadto, dzięki dekonwolucji wyników analizy do kodu binarnego przy użyciu oprogramowania, odczyt informacji poprzez fragmentację z wykorzystaniem spektrometrii mas może być w pełni zautomatyzowany i przeprowadzony przez osobę bez specjalistycznego doświadczenia (Rysunek 6)⁴¹. Jednakże spektrometria mas ma wiele wad z punktu widzenia praktycznego

³⁷ L. Restrepo-Pérez, C. Joo, C. Dekker, *Paving the way to single-molecule protein sequencing*, „Nature Nanotechnology” 2018, t. 13, s. 786–796, DOI: 10.1038/s41565-018-0236-6.

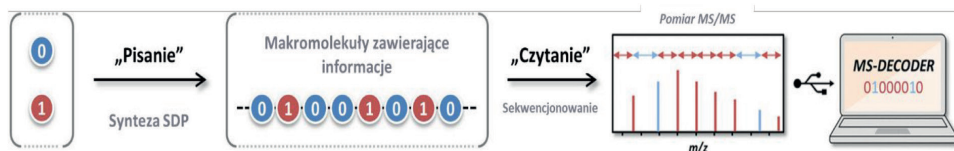
³⁸ J. Shendure, S. Balasubramanian, G.M. Church, W. Gilbert, J. Rogers, J.A. Schloss, R.H. Waterston, *DNA sequencing at 40: past, present and future*, „Nature” 2017, t. 550, s. 345–353, DOI: 10.1038/nature24286.

³⁹ H. Mutlu, J.F. Lutz, *Reading Polymers: Sequencing of Natural and Synthetic Macromolecules*, „Angewandte Chemie International Edition” 2014, t. 53, s. 13010–13019, DOI: 10.1002/anie.201406766.

⁴⁰ L. Charles, G. Cavallo, V. Monnier, L. Oswald, R. Szweda, J.F. Lutz, *MS/MS-Assisted Design of Sequence-Controlled Synthetic Polymers for Improved Reading of Encoded Information*, „Journal of the American Society for Mass Spectrometry” 2017, t. 28, s. 1149–1159, DOI: 10.1007/s13361-016-1543-5; A. Al Ouahabi, J.A. Amalian, L. Charles, J.F. Lutz, *Mass spectrometry sequencing of long digital polymers facilitated by programmed inter-byte fragmentation*, „Nature Communications” 2017, t. 8, s. 967, DOI: 10.1038/s41467-017-01104-3.

⁴¹ A. Burel, C. Carapito, J.F. Lutz, L. Charles, *MS-DECODER: Milliseconds Sequencing of Coded Polymers*, „Macromolecules” 2017, t. 50, s. 8290–8296, DOI: 10.1021/acs.macromol.7b01737.

wykorzystania. Przede wszystkim metoda ta jest destrukcyjna – po odczytaniu informacji próbka i dane są na zawsze tracone. Ponadto wymaga ona zaawansowanej i drogiej aparatury, co eliminuje ją z powszechnego użycia.



Rysunek 6. Schemat kodowania i odczytywania zakodowanych danych z użyciem spektrometrii mas przy zastosowaniu algorytmu MC-DECODER, stworzonego przez grupę Lutz do automatycznego sekwencjonowania syntetycznych makrocząsteczek. Adapted with permission from⁴². Copyright 2015, 2017 American Chemical Society.

Inną metodą umożliwiającą sekwencjonowanie polimerów jest technologia nanoporowa⁴³. Sekwencjonowanie nanoporowe jest unikalną, skalowalną technologią, która pozwala na bezpośrednią analizę w czasie rzeczywistym długich fragmentów makrocząsteczek. Działa ona w oparciu o monitorowanie zmian prądu elektrycznego, gdy różne sekwencje są przepuszczane przez białkowy nanopor, co daje możliwość dekodowania zapisanej informacji. Sekwencjonowanie nanoporowe, początkowo opracowane do analizy DNA, zostało dostosowane do sekwencjonowania „polimerów cyfrowych”⁴⁴. Wraz ze wzrostem gęstości informacji binarnej zapisanej w polimerze, rośnie błąd metody. Na przykład przy gęstości 4 bit/monomer poziom błędów wzrasta już do 13,5%. Jednak metoda ta jest bardzo obiecująca i ma wiele zalet w porównaniu ze spektrometrią mas, np. jest szybka, urządzenie jest proste w obsłudze i może być używane przez niedoświadczonego użytkownika. Jednak ma ona tendencję do podatności na błędy (poziom błędów w sekwencjonowaniu na-

⁴² J.F. Lutz., op.cit.; A. Burel, C. Carapito, J.F. Lutz, L. Charles, op.cit.

⁴³ R. Bowden i in., *Sequencing of human genomes with nanopore technology*, „Nature Communications” 2019, t. 10, s. 1–9, DOI: 10.1038/s41467-019-09637-5.

⁴⁴ C. Cao, L.F. Krapp, A. Al Ouahabi, N.F. König, N. Cirauqui, A. Radenovic, J.F. Lutz, M.D. Peraro, *Aerolysin nanopores decode digital information stored in tailored macromolecular analytes*, „Science Advances” 2020, t. 6, DOI: 10.1126/sciadv.abc2661.

noporowym może wynosić nawet 15%⁴⁵) i z praktycznego punktu widzenia odzyskanie próbki może być pracochłonne. Ponadto została ona zoptymalizowana tylko dla wybranych monomerów, w przypadku innych dane sekwencjonowania są zaszumione i trudne do interpretacji, co komplikuje odczyt zakodowanej informacji⁴⁶. Sekwencjonowanie nanoporowe jest niewątpliwie atrakcyjną techniką odczytu informacji zakodowanej w polimerach, ale w przypadku polimerów o dużej gęstości danych nadal wymaga rozwoju i optymalizacji.

Niekonwencjonalne metody magnetycznego rezonansu jądrowego mogą być również wykorzystane do sekwencjonowania polimerów. Duży potencjał ma np. opracowana technika „molekularnej pęsety”⁴⁷. Cząsteczki „pęsety” mogą wiązać specyficzne motywy sekwencji i w ten sposób umożliwić ich odczyt na postawie sygnałów widma. Technika ta była dotychczas stosowana głównie do analizy losowych kopolimerów i wymaga dostosowania i optymalizacji, aby umożliwić sekwencjonowanie syntetycznych makrocząsteczek zawierających informację.

Podsumowując, wszystkie wymienione wyżej techniki sekwencjonowania polimerów mają wiele ograniczeń i na obecnym etapie rozwoju tych technologii nie nadają się do praktycznego wykorzystania. Dlatego wciąż istnieje potrzeba rozwijania nowych metod, które umożliwią łatwy odczyt informacji zakodowanych w polimerach.

Dlaczego to polimery mogą być przyszłością w przechowywaniu danych?

Polimery zawierające informacje zakodowane w sekwencji monomerów, zwane „polimerami cyfrowymi”, oferują wiele zalet w porównaniu z DNA i tradycyjnymi nośnikami pamięci. Właściwości syntetycznych polimerów mogą być precyzyjnie modulowane i dostosowane do określonych wymagań poprzez wybór odpowiednich komponentów budulcowych z szerokiej biblio-

⁴⁵ F.J. Rang, W.P. Kloosterman, J. de Ridder, *From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy*, „Genome Biology” 2018, t. 19, s. 90, DOI: 10.1186/s13059-018-1462-9.

⁴⁶ J. Bonnet, P. Subsoontorn, D. Endy, op.cit.

⁴⁷ Z. Zhu, C.J. Cardin, Y. Gan, H.M. Colquhoun, *Sequence-selective assembly of tweezer molecules on linear templates enables frameshift-reading of sequence information*, „Nature Chemistry” 2010, t. 2, s. 653–660, DOI: 10.1038/nchem.699.

teki komercyjnie dostępnych monomerów⁴⁸. W ten sposób można zwiększyć stabilność, a tym samym wydłużyć czas życia nośnika danych. Ponadto struktura polimerów może być łatwo zoptymalizowana dla pożądaných celów, np. w celu uproszczenia odczytu informacji lub edycji danych⁴⁹. Ostatnio pokazano, że zakodowane w polimerach dane mogą być łatwo edytowane pod wpływem światła, dzięki zastosowaniu monomerów zawierających fotoczułe podstawniki⁵⁰.

Największymi zaletami polimerów w porównaniu do dysków twardych i pamięci flash są niewielkie zużycie energii, duża gęstość zapisu danych oraz długi czas eksploatacji⁵¹. Zapisane dane można by bezpiecznie archiwizować nawet przez setki lat (Rysunek 7)⁵². Przykładem tak wysokiej stabilności są polimery termoplastyczne, zwane tworzywami sztucznymi, i powszechnie plastikiem, np. polichlorek winylu (PVC), polipropylen (PP), polistyren (PS) i polietylen (HDPE, LDPE i LLDPE)⁵³. Pomimo że są to tak „pospolite” związki, każdy z nich jest wysoce zaawansowanym materiałem o precyzyjnych właściwościach fizykochemicznych – można je dowolnie formować, są obojętne chemicznie i biologicznie oraz są odporne na czynniki środowiskowe. Ponadto, aby jeszcze bardziej przedłużyć ich żywotność, w trakcie ich produkcji używane są różne przeciwutleniacze i stabilizatory.

⁴⁸ E. Laurent, R. Szweda, J.F. Lutz, op.cit.

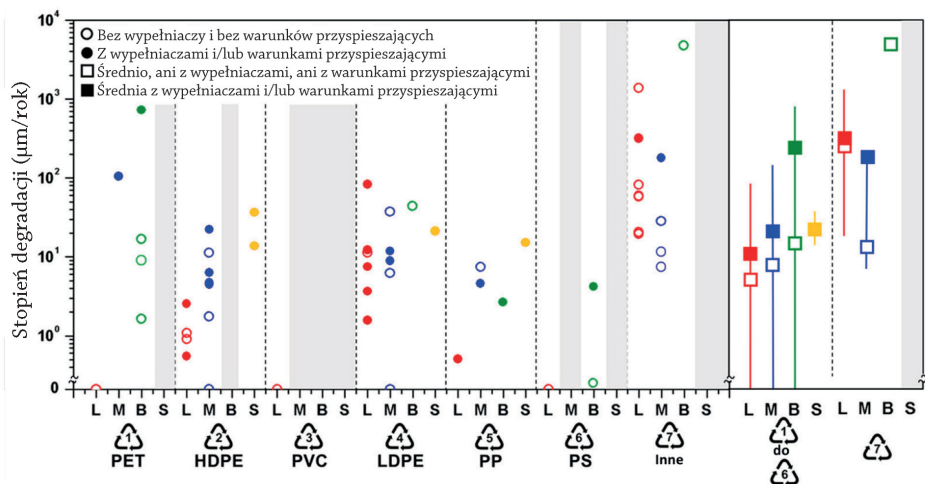
⁴⁹ L. Charles, G. Cavallo, V. Monnier, L. Oswald, R. Szweda, J.F. Lutz, *MS/MS-Assisted Design of Sequence-Controlled Synthetic Polymers for Improved Reading of Encoded Information*, „Journal of the American Society for Mass Spectrometry” 2017, t. 28, s. 1149–1159, DOI: 10.1007/s13361-016-1543-5; M. Boukhet, N.F. König, A. Al Ouahabi, G. Baaken, J.F. Lutz, J.C. Behrends, *Translocation of Precision Polymers through Biological Nanopores*, „Macromolecular Rapid Communications” 2017, t. 38, s. 1700680, DOI: 10.1002/marc.201700680.

⁵⁰ N.F. König, A. Al Ouahabi, L. Oswald, R. Szweda, L. Charles, J.F. Lutz, *Photo-editable macromolecular information*, „Nature Communications” 2019, t. 10, s. 3774, DOI: 10.1038/s41467-019-11566-2.

⁵¹ J.F. Lutz., op.cit.; S. Martens, A. Landuyt, P. Espeel, B. Devreese, P. Dawyndt, F. Du Prez, op.cit.

⁵² B.J. Cafferty, A.S. Ten, M.J. Fink, S. Morey, D.J. Preston, M. Mrksich, G.M. Whitesides, *Storage of Information Using Small Organic Molecules*, „ACS Central Science” 2019, t. 5, s. 911–916, DOI: 10.1021/acscentsci.9b00210.

⁵³ A. Chamas, H. Moon, J. Zheng, Y. Qiu, T. Tabassum, J.H. Jang, M. Abu-Omar, S.L. Scott, S. Suh, *Degradation Rates of Plastics in the Environment*, „ACS Sustainable Chemistry & Engineering” 2020, t. 8, s. 3494–3511, DOI: 10.1021/acssuschemeng.9b06635.



Rysunek 7. Wskaźniki degradacji powierzchni dla różnych tworzyw sztucznych, w μm na rok. Kolumny pionowe reprezentują różne warunki środowiskowe (L – gleba, M – morskie, B – biologiczne, S – światło słoneczne) i rodzaje tworzyw sztucznych (reprezentowane przez ich kody identyfikacyjne). Plastik typu 7, „inne”, odpowiada różnym nominalnie biodegradowalnym tworzywom sztucznym. Szare kolumny reprezentują kombinacje, dla których nie znaleziono danych⁵⁴. Copyright 2020 American Chemical Society.

W związku z bardzo dużą stabilnością, tworzywa sztuczne stały się rosnącym zagrożeniem dla środowiska, gdyż w zależności od tego, gdzie zostały porzucone, oraz z jakiego materiału zostały wykonane, mogą rozkładać się nawet ponad 1000 lat. Jednak w kontekście przechowywania danych wadę tę można przekuć w zaletę, gdyż w archiwizacji najważniejszym parametrem jest właśnie długa żywotność nośnika. Polimery do przechowywania informacji powinny być zaprojektowane m.in. z myślą o długotrwałości.

Rozszerzony alfabet syntetycznych monomerów pozwala sięgnąć po dużo większą gęstość i pojemność przechowywania informacji binarnej⁵⁵. Gęstość informacji binarnej (χ), którą można przechowywać w łańcuchach makrocząstek, zależy zarówno od długości łańcucha, jak i od liczby różnych bloków budulcowych (n), zgodnie z równaniem⁵⁶:

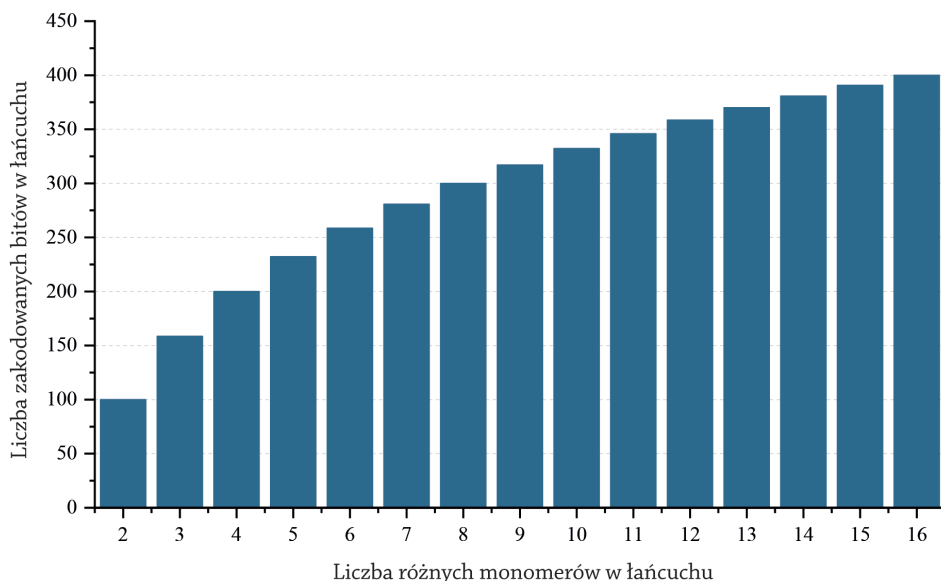
⁵⁴ Ibidem.

⁵⁵ J.F. Lutz., op.cit.; S. Martens, A. Landuyt, P. Espeel, B. Devreese, P. Dawyndt, F. Du Prez, op.cit.

⁵⁶ Y. Choi, T. Ryu, A.C. Lee i in., *High information capacity DNA-based data storage with augmented encoding characters using degenerate bases*, „Scientific Reports” 2019, t. 9, s. 6582, DOI: 10.1038/s41598-019-43105-w.

$$\log_2 n = x \therefore 2^x = n$$

gdzie n = liczba znaków kodujących – monomerów, x = liczba zakodowanych bitów na monomer. Przykładowo, jeśli zastosujemy 16 różnych monomerów w polimerze o długości łańcucha 100, maksymalna liczba zakodowanych bitów wynosi 400 (4 bity/monomer), podczas gdy z 2 różnych monomerów w polimerze o tej samej długości łańcucha można zakodować wyłącznie 100 bitów (1 bit/monomer) (Rysunek 8).



Rysunek 8. Predykcja liczby zakodowanych bitów w monomerze w zależności od liczby wykorzystanych merów dla polimeru o długości łańcucha równej 100 jednostek.

Więcej niż archiwizacja...

Polimery o zdefiniowanej sekwencji mogą w przyszłości być wykorzystane do przechowywania danych, jednak nie jest to jedyne możliwe zastosowanie w nowoczesnych technologiach. Innym obszarem, w którym można się nimi posłużyć, jest kryptografia i steganografia na poziomie molekularnym⁵⁷.

⁵⁷ A.C. Boukis, K. Reiter, M. Frölich, D. Hofheinz, M.A.R. Meier, *Multicomponent reactions provide key molecules for secret communication*, „Nature Communications” 2018, t. 9, s. 1439, DOI: 10.1038/s41467-018-03784-x.

Na przykład Anslyn i wsp. w swojej pracy zsyntezowali 256-bitowy klucz szyfrujący, który został wykorzystany do zaszyfrowania dokumentu zawierającego powieść *Czarnoksiężnik z Krainy Oz* L. Franka Bauma (1900)⁵⁸. Aby zsyntetyzować i zakodować klucz szyfrujący w oligouretanach, opracowano skrypt szyfrujący i deszyfrujący zdolny do generowania i wykorzystywania klucza. Po wygenerowaniu klucz został przekształcony na szesnastkowy system liczbowy, a następnie znaki szesnastkowe przypisano do każdego z 16 monomerów. Aby zaprezentować rzeczywisty potencjał polimerów w steganografii, autorzy rozpuścili zsyntezowany klucz w atramencie, który następnie wykorzystali do napisania listu. Adresaci, po otrzymaniu „dyskretnych” instrukcji dotyczących dekonwolucji sekwencji, już za pierwszą próbą zdołali poprawnie odczytać sekwencję i po wprowadzeniu do algorytmu deszyfrującego ujawnili zakodowany dokument.

Wśród innych pojawiających się zastosowań, polimery były również wykorzystane jako znaczniki chemiczne w celu zwalczania fałszerstwa⁵⁹. Wykazano, że poliuretany mogą być używane jako znaczniki m.in. w materiałach polimerowych⁶⁰, implantach wewnątrzgałkowych⁶¹ oraz w modelach implantów *in vivo*⁶². Aby oznaczyć określony materiał, polimerowe znaczniki są albo fizycznie adsorbowane na jego powierzchni, albo wtapiane w jego wnętrze⁶³. Grupa F. Du Preza pokazała możliwość syntezy „molekularnych kodów pin”. Zademonstrowali oni sposób, w jaki można je zastosować w praktyce, umieszczając zaszyfrowane oli-

⁵⁸ S.D. Dahlhauser, C.D. Wight, S.R. Moor, R.A. Scanga, P. Ngo, J.T. York, M.S. Vera, K.J. Blake, I.M. Riddington, J.F. Reuther, E.V. Anslyn, *Molecular Encryption and Steganography Using Mixtures of Simultaneously Sequenced, Sequence-Defined Oligourethanes*, „ACS Central Science” 2022, t. 8, s. 1125–1133, DOI: 10.1021/acscentsci.2c00460.

⁵⁹ M. Soete, J. van Hoorde, F.E. du Prez, *Discrete, self-immolative N-substituted oligourethanes and their use as molecular tags*, „Polymer Chemistry” 2022, t. 13, s. 4178–4185, DOI: 10.1039/D2PY00630H.

⁶⁰ D. Karamessini, S. Poyer, L. Charles, J.F. Lutz, *2D Sequence-Coded Oligourethane Barcodes for Plastic Materials Labeling*, „Macromolecular Rapid Communications” 2017, t. 38, s. 1700426, DOI: 10.1002/marc.201700426.

⁶¹ D. Karamessini, B.E. Petit, M. Bouquoy, L. Charles, J.F. Lutz, *Identification-Tagging of Methacrylate-Based Intraocular Implants Using Sequence Defined Polyurethane Barcodes*, „Advanced Functional Materials” 2017, t. 27, s. 1604595, DOI: 10.1002/adfm.201604595.

⁶² D. Karamessini i in., *Abiotic Sequence-Coded Oligomers as Efficient In Vivo Taggants for the Identification of Implanted Materials*, „Angewandte Chemie International Edition” 2018, t. 57, s. 10574–10578, DOI: 10.1002/anie.201804895.

⁶³ I. Youssef, I. Carvin-Sergent, E. Konishcheva, S. Kebe, V. Greff, D. Karamessini, M. Matloubi, A. Al Ouahabi, J. Moesslein, J.A. Amalian, S. Poyer, L. Charles, J.F. Lutz, *Covalent Attachment and Detachment by Reactive DESI of Sequence-Coded Polymer Taggants*, „Macromolecular Rapid Communications” 2022, t. 43, s. 2200412, DOI: 10.1002/marc.202200412.

gomery na plastikowych banknotach, po czym dokonali ekstrakcji i odczytali je za pomocą spektrometrii masowej⁶⁴. Podejście to potencjalnie może być wykorzystane w przypadku oficjalnych dokumentów, w celu oznakowania oryginalnej wersji lub ich kopii.

Jakie problemy trzeba przezwyciężyć aby wykorzystać polimery do przechowywania danych w praktyce?

Podsumowując, w tabeli 1 porównano główne parametry nośników danych dla rozwijających się technologii przechowywania informacji z wykorzystaniem DNA, polimerów oraz dysków twardych. Zestawienie obrazuje, jakie parametry nośników mogą zostać usprawnione, gdy do zapisu danych wykorzystamy makrocząsteczki, tj. czas przechowywania danych oraz gęstość informacji. Jednakże głównym wyzwaniem tych rozwijających się technologii jest trudne i kosztowne sekwencjonowanie, innymi słowy: odczyt informacji, co na dzień dzisiejszy nie pozwala na ich powszechne zastosowanie. Ze względu na ograniczoną szybkość zapisu, którą determinuje synteza chemiczna, a także odczytu, polimery raczej nie wyprą stosowanych dysków twardych. Jednakże świetnie nadają się do archiwizacji informacji, dzięki czemu mogą w dużej mierze zminimalizować problem przechowywania danych.


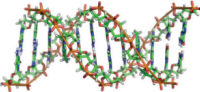
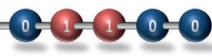
W kontekście przyszłych zastosowań, fundamentalnym parametrem hamującym wdrożenie polimerowych nośników danych jest brak metod łatwego odczytu informacji⁶⁵. Opracowano zaledwie kilka, które pozwalają na określenie sekwencji monomerów, a żadna z nich nie ma znaczenia dla zastosowań komercyjnych. Prosty i szybki odczyt „polimerów cyfrowych” przy użyciu metod nieniszczących pozostaje głównym wyzwaniem utrudniającym rozwój syntetycznych materiałów opartych na polimerach jako nośników danych. Dlatego kluczowym wyzwaniem dla naukowców jest opracowanie struktury polimeru dostosowanego do łatwego i taniego odczytu zakodowanych danych cyfrowych metodami nieniszczącymi. Istnieje wiele czułych i – przede wszystkim – nienisz-

⁶⁴ J.O. Holloway, F. van Lijsebetten, N. Badi, H.A. Houck, F.E. du Prez, *From Sequence-Defined Macromolecules to Macromolecular Pin Codes*, „Advanced Science” 2020, t. 7, s. 1903698, DOI: 10.1002/adv.201903698.

⁶⁵ R. Szweda, M. Tschopp, O. Felix, G. Decher, J.F. Lutz, op.cit; N.F. König, A. Al Ouahabi, L. Oswald, R. Szweda, L. Charles, J.F. Lutz, op.cit.

czących metod spektroskopowych, które potencjalnie można by było wykorzystać, jak np. fluorescencja czy spektroskopia wibracyjna, jednak nie opracowano jeszcze struktur chemicznych polimerów, które mogłyby być sekwencjonowane wymienionymi technikami.

Tabela 1. Porównanie parametrów przechowywania danych dla różnych nośników.

Nośniki pamięci	Dysk twardy	DNA	SDP
Parametry			
Prędkość zapisu (bity/s)	10^9 ⁽¹⁾	$7 \cdot 10^{-2}$ ⁽²⁾	$7 \cdot 10^{-2}$ ⁽²⁾
Prędkość odczytu (bity/s)	10^9 ⁽¹⁾	$5 \cdot 10^3$ ⁽³⁾	$9 \cdot 10^{-2}$ ⁽⁴⁾
Przechowywanie danych (lata)	~10	~100	>100
Gęstość danych (bity/cm ³)	$\sim 4 \cdot 10^{11}$	$\sim 4 \cdot 10^{13}$ ⁽⁶⁾	$> 10^{20}$
Energia (TWh)	220–320 ⁽⁵⁾	~0	~0
* ⁽¹⁾ HDD; ⁽²⁾ przy użyciu syntezy ABI; ⁽³⁾ sekwencjonowanie Illumina ⁶⁶ ; ⁽⁴⁾ sekwencjonowanie MS/MS; ⁽⁵⁾ ref. ⁶⁷ ; ⁽⁶⁾ ref. ⁶⁸ ; pozostałe dane z ref. ⁶⁹			

Bibliografia

- Alberts B., Johnson A., Lewis J. i in., *Molecular Biology of the Cell*, 4th edition, Nowy Jork 2002.
- Al Ouahabi A., Amalian J.A., Charles L., Lutz J.F., *Mass spectrometry sequencing of long digital polymers facilitated by programmed inter-byte fragmentation*, „Nature Communications” 2017, t. 8, s. 967, DOI: 10.1038/s41467-017-01104-3.

⁶⁶ M.T. Pervez, M.J. Hasnain, S.H. Abbas, M.F. Moustafa, N. Aslam, S.S.M. Shah, op.cit.

⁶⁷ IEA (2022), Data Centres and Data Transmission... op.cit.; N. Jones, *How to stop data centres from gobbling up the world's electricity*, „Nature” 2018, t. 561, s. 163–166, DOI: 10.1038/d41586-018-06610-y.

⁶⁸ L. Organick i in., op.cit.

⁶⁹ G.M. Church, Y. Gao, S. Kosuri, op.cit.

- Al Ouahabi A., Kotera M., Charles L., Lutz J.F., *Synthesis of Monodisperse Sequence-Coded Polymers with Chain Lengths above DP100*, „ACS Macro Letters” 2015, t. 4, s. 1077–1080, DOI: 10.1021/acsmacrolett.5b00606.
- Andrieux D., Gaspard P., *Nonequilibrium generation of information in copolymerization processes*, „Proceedings of the National Academy of Sciences” 2008, t. 105, s. 9516–9521, DOI: 10.1073/pnas.0802049105.
- Baoutina A., Bhat S., Partis L., Emslie K.R., *Storage Stability of Solutions of DNA Standards*, „Analytical Chemistry” 2019, t. 91, s. 12268–12274, DOI: 10.1021/acs.analchem.9b02334.
- Bonnet J. i in., *Chain and conformation stability of solid-state DNA: implications for room temperature storage*, „Nucleic Acids Research” 2010, t. 38, s. 1531–1546, DOI: 10.1093/nar/gkp1060.
- Bonnet J., Subsoontorn P., Endy D., *Rewritable digital data storage in live cells via engineered control of recombination directionality*, „Proceedings of the National Academy of Sciences” 2012, t. 109, s. 8884–8889, DOI: 10.1073/pnas.1202344109.
- Boukhet M., König N.F., Al Ouahabi A., Baaken G., Lutz J.F., Behrends J.C., *Translocation of Precision Polymers through Biological Nanopores*, „Macromolecular Rapid Communications” 2017, t. 38, s. 1700680, DOI: 10.1002/marc.201700680.
- Boukis A.C., Reiter K., Frölich M., Hofheinz D., Meier M.A.R., *Multicomponent reactions provide key molecules for secret communication*, „Nature Communications” 2018, t. 9, s. 1439, DOI: 10.1038/s41467-018-03784-x.
- Bowden R. i in., *Sequencing of human genomes with nanopore technology*, „Nature Communications” 2019, t. 10, s. 1–9, DOI: 10.1038/s41467-019-09637-5.
- Burel A., Carapito C., Lutz J.F., Charles L., *MS-DECODER: Milliseconds Sequencing of Coded Polymers*, „Macromolecules” 2017, t. 50, s. 8290–8296, DOI: 10.1021/acs.macromol.7b01737.
- Cafferty B.J., Ten A.S., Fink M.J., Morey S., Preston D.J., Mrksich M., Whitesides G.M., *Storage of Information Using Small Organic Molecules*, „ACS Central Science” 2019, t. 5, s. 911–916, DOI: 10.1021/acscentsci.9b00210.
- Cao C., Krapp L.F., Al Ouahabi A., König N.F., Cirauqui N., Radenovic A., Lutz J.F., Peraro M.D., *Aerolysin nanopores decode digital information stored in tailored macromolecular analytes*, „Science Advances” 2020, t. 6, DOI: 10.1126/sciadv.abc2661.
- Chamas A., Moon H., Zheng J., Qiu Y., Tabassum T., Jang J.H., Abu-Omar M., Scott S.L., Suh S., *Degradation Rates of Plastics in the Environment*, „ACS Sustainable Chemistry & Engineering” 2020, t. 8, s. 3494–3511, DOI: 10.1021/acssuschemeng.9b06635.
- Charles L., Cavallo G., Monnier V., Oswald L., Szweda R., Lutz J.F., *MS/MS-Assisted Design of Sequence-Controlled Synthetic Polymers for Improved Reading of Encoded Information*, „Journal of the American Society for Mass Spectrometry” 2017, t. 28, s. 1149–1159, DOI: 10.1007/s13361-016-1543-5.

- Charles L., Laure C., Lutz J.F., Roy R.K., *MS/MS Sequencing of Digitally Encoded Poly(alkoxyamine amide)s*, „*Macromolecules*” 2015, t. 48, s. 4319–4328, DOI: 10.1021/acs.macromol.5b01051.
- Choi Y., Ryu T., Lee A.C. i in., *High information capacity DNA-based data storage with augmented encoding characters using degenerate bases*, „*Scientific Reports*” 2019, t. 9, s. 6582, DOI: 10.1038/s41598-019-43105-w.
- Church G.M., Gao Y., Kosuri S., *Next-Generation Digital Information Storage in DNA*, „*Science*” 2012, t. 337, s. 1628, DOI: 10.1126/science.1226355.
- Colquhoun H.M., Lutz J.F., *Information-containing macromolecules*, „*Nature Chemistry*” 2014, t. 6, s. 455–456, DOI: 10.1038/nchem.1958.
- Cooper G.M., *The Cell: A Molecular Approach*, 2nd edition, Sunderland (MA) 2000.
- Dahlhauser S.D., Wight C.D., Moor S.R., Scanga R.A., Ngo P., York J.T., Vera M.S., Blake K.J., Riddington I.M., Reuther J.F., Anslyn E.V., *Molecular Encryption and Steganography Using Mixtures of Simultaneously Sequenced, Sequence-Defined Oligourethanes*, „*ACS Central Science*” 2022, t. 8, s. 1125–1133, DOI: 10.1021/acscentsci.2c00460.
- Dhar L., *Molecular Storage, Data Writ Small*, „*ACS Central Science*” 2019, t. 5, 753–754, DOI: 10.1021/acscentsci.9b00439.
- Extance A., *How DNA could store all the world's data*, „*Nature*” 2016, t. 537, s. 22–24, DOI: 10.1038/537022a.
- Goldman N. i in., *Towards practical, high-capacity, low-maintenance information storage in synthesized DNA*, „*Nature*” 2013, t. 494, s. 77–80, DOI: 10.1038/nature11875.
- Hao Y., Li Q., Fan C., Wang F., *Data Storage Based on DNA*, „*Small Structures*” 2021, t. 2, s. 2000046, DOI: 10.1002/sstr.202000046.
- Holloway J.O., van Lijsebetten F., Badi N., Houck H.A., du Prez F.E., *From Sequence-Defined Macromolecules to Macromolecular Pin Codes*, „*Advanced Science*” 2020, t. 7, s. 1903698, DOI: 10.1002/advs.201903698.
- Holloway J.O., Mertens C., du Prez F.E., Badi N., *Automated Synthesis Protocol of Sequence-Defined Oligo-Urethane-Amides Using Thiolactone Chemistry*, „*Macromolecular Rapid Communications*” 2019, t. 40, s. 1800685, DOI: 10.1002/marc.201800685.
- Ivanova N.V., Kuzmina M.L., *Protocols for dry DNA storage and shipment at room temperature*, „*Molecular Ecology Resources*” 2013, t. 13, s. 890–898, DOI: 10.1111/1755-0998.12134.
- Jones N., *How to stop data centres from gobbling up the world's electricity*, „*Nature*” 2018, t. 561, s. 163–166, DOI: 10.1038/d41586-018-06610-y.
- Karamessini D. i in., *Abiotic Sequence-Coded Oligomers as Efficient In Vivo Taggants for the Identification of Implanted Materials*, „*Angewandte Chemie International Edition*” 2018, t. 57, s. 10574–10578, DOI: 10.1002/anie.201804895.

- Karamessini D., Petit B.E., Bouquey M., Charles L., Lutz J.F., *Identification-Tagging of Methacrylate-Based Intraocular Implants Using Sequence Defined Polyurethane Barcodes*, „Advanced Functional Materials” 2017, t. 27, s. 1604595, DOI: 10.1002/adfm.201604595.
- Karamessini D., Poyer S., Charles L., Lutz J.F., *2D Sequence-Coded Oligourethane Barcodes for Plastic Materials Labeling*, „Macromolecular Rapid Communications” 2017, t. 38, s. 1700426, DOI: 10.1002/marc.201700426.
- König N.F., Al Ouahabi A., Oswald L., Szweda R., Charles L., Lutz J.F., *Photo-editable macromolecular information*, „Nature Communications” 2019, t. 10, s. 3774, DOI: 10.1038/s41467-019-11566-2.
- Langmead B., Salzberg S.L., *Fast gapped-read alignment with Bowtie 2*, „Nature Methods” 2012, t. 9, s. 357–359, DOI: 10.1038/nmeth.1923.
- Laurent E., Amalian J.A., Parmentier M., Oswald L., Al Ouahabi A., Dufour F., Launay K., Clément J.L., Gigmes D., Delsuc M.A., Charles L., Lutz J.F., *High-Capacity Digital Polymers: Storing Images in Single Molecules*, „Macromolecules” 2020, t. 53, s. 4022–4029, DOI: 10.1021/acs.macromol.0c00666.
- Laurent E., Szweda R., Lutz J.F., *Synthetic Polymers with Finely Regulated Monomer Sequences: Properties and Emerging Applications*, [w:] *Macromolecular Engineering*, red. N. Hadjichristidis, Y. Gnanou, K. Matyjaszewski i M. Muthukumar, bm. 2022, s. 1–34, DOI: 10.1002/9783527815562.mme0041.
- Lee J.M., Kwon J., Lee S.J., Jang H., Kim D.G., Song J., Kim K.T., *Semiautomated synthesis of sequence-defined polymers for information storage*, „Science Advances” 2022, t. 8, s. eabl8614, DOI: 10.1126/sciadv.abl8614.
- Lutz J.F., *Coding Macromolecules: Inputting Information in Polymers Using Monomer-Based Alphabets*, „Macromolecules” 2015, t. 48, s. 4759–4767, DOI: 10.1021/acs.macromol.5b00890.
- Lutz J.F., Ouchi M., Liu D.R., Sawamoto M., *Sequence-Controlled Polymers*, „Science” 2013, t. 341, s. 1238149, DOI: 10.1126/science.1238149.
- Martens S., Landuyt A., Espeel P., Devreese B., Dawyndt P., Du Prez F., *Multifunctional sequence-defined macromolecules for chemical data storage*, „Nature Communications” 2018, t. 9, s. 4451, DOI: 10.1038/s41467-018-06926-3.
- Masanet E., Shehabi A., Lei N., Smith S., Koomey J., *Recalibrating global data center energy-use estimates*, „Science” 2020, t. 367, s. 984–986, DOI: 10.1126/science.aba3758.
- Matange K., Tuck J.M., Keung A.J., *DNA stability: a central design consideration for DNA data storage systems*, „Nature Communications” 2021, t. 12, s. 1358, DOI: 10.1038/s41467-021-21587-5.
- Mayer C., McInroy G.R., Murat P., van Delft P., Balasubramanian S., *An Epigenetics-Inspired DNA-Based Data Storage System*, „Angewandte Chemie” 2016, t. 128, s. 11310–11314, DOI: 10.1002/anie.201605531.

- Meiser L.C., Nguyen B.H., Chen Y.J., Nivala J., Strauss K., Ceze L., Grasset R.N., *Synthetic DNA applications in information technology*, „Nature Communications” 2022, t. 13, 352, DOI: 10.1038/s41467-021-27846-9.
- Merrifield R.B., *Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide*, „Journal of the American Chemical Society” 1963, t. 85, s. 2149–2154, DOI: 10.1021/ja00897a025.
- Mertens C., Soete M., Ślęczkowski M.L., Palmans A.R.A., Meijer E.W., Badi N., Du Prez F.E., *Stereocontrolled, multi-functional sequence-defined oligomers through automated synthesis*, „Polymer Chemistry” 2020, t. 11, s. 4271–4280, DOI: 10.1039/d0py00645a.
- Mutlu H., Lutz J.F., *Reading Polymers: Sequencing of Natural and Synthetic Macromolecules*, „Angewandte Chemie International Edition” 2014, t. 53, s. 13010–13019, DOI: 10.1002/anie.201406766.
- Nanjan P., Porel M., *Sequence-defined non-natural polymers: synthesis and applications*, „Polymer Chemistry” 2019, t. 10, s. 5406–5424, DOI: 10.1039/C9PY00886A.
- Organick L. i in., *Probing the physical limits of reliable DNA data retrieval*, „Nature Communications” 2020, t. 11, s. 616, DOI: 10.1038/s41467-020-14319-8.
- Pervez M.T., Hasnain M.J., Abbas S.H., Moustafa M.F., Aslam N., Shah S.S.M., *A Comprehensive Review of Performance of Next-Generation Sequencing Platforms*, „BioMed Research International” 2022, t. 2022, 3457806, DOI: 10.1155/2022/3457806.
- Pfeiffer F. i in., *Systematic evaluation of error rates and causes in short samples in next-generation sequencing*, „Scientific Reports” 2018, t. 8, s. 10950, DOI: 10.1038/s41598-018-29325-6.
- Podivinsky E., Love J.L., van der Colff L., Samuel L., *Effect of storage regime on the stability of DNA used as a calibration standard for real-time polymerase chain reaction*, „Analytical Biochemistry” 2009, t. 394, s. 132–134, DOI: 10.1016/j.ab.2009.06.024.
- Rang F.J., Kloosterman W.P., de Ridder J., *From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy*, „Genome Biology” 2018, t. 19, s. 90, DOI: 10.1186/s13059-018-1462-9.
- Restrepo-Pérez L., Joo C., Dekker C., *Paving the way to single-molecule protein sequencing*, „Nature Nanotechnology” 2018, t. 13, s. 786–796, DOI: 10.1038/s41565-018-0236-6.
- Roy R.K. i in., *Design and synthesis of digitally encoded polymers that can be decoded and erased*, „Nature Communications” 2015, t. 6, s. 7237, DOI: 10.1038/ncomms8237.
- Rutten M.G.T.A., Vaandrager F.W., Elemans J.A.A.W. & Nolte R.J.M., *Encoding information into polymers*, „Nature Reviews Chemistry” 2018, t. 2, s. 365–381, DOI: 10.1038/s41570-018-0051-5.
- Service R., *DNA could store all of the world's data in one room*, „Science” 2017, s. 1–2, DOI: 10.1126/science.aal0852.

- Shendure J., Balasubramanian S., Church G.M., Gilbert W., Rogers J., Schloss J.A., Waterston R.H., *DNA sequencing at 40: past, present and future*, „Nature” 2017, t. 550, s. 345–353, DOI: 10.1038/nature24286.
- Soete M., van Hoorde J., du Prez F.E., *Discrete, self-immolative N-substituted oligourethanes and their use as molecular tags*, „Polymer Chemistry” 2022, t. 13, s. 4178–4185, DOI: 10.1039/D2PY00630H.
- Szweda R., Tschopp M., Felix O., Decher G., Lutz J.F., *Sequences of Sequences: Spatial Organization of Coded Matter through Layer-by-Layer Assembly of Digital Polymers*, „Angewandte Chemie International Edition” 2018, t. 57, s. 15817–15821, DOI: 10.1002/anie.201810559.
- Tateishi-Karimata H., Sugimoto N., *Structure, stability and behaviour of nucleic acids in ionic liquids*, „Nucleic Acids Research” 2014, t. 42, s. 8831–8844, DOI: 10.1093/nar/gku499.
- Vandenbergh J., Reekmans G., Adriaensens P., Junkers T., *Synthesis of sequence-defined acrylate oligomers via photo-induced copper-mediated radical monomer insertions*, „Chemical Science” 2015, t. 6, s. 5753–5761, DOI: 10.1039/C5SC02035B.
- Wang S., Toreini E., Hao F., *Anti-Counterfeiting for Polymer Banknotes Based on Polymer Substrate Fingerprinting*, „IEEE Transactions on Information Forensics and Security” 2021, t. 16, s. 2823–2835, DOI: 10.48550/arXiv.2103.06184.
- Youssef I., Carvin-Sergent I., Konishcheva E., Kebe S., Greff V., Karamessini D., Matloubi M., Al Ouahabi A., Moesslein J., Amalian J.A., Poyer S., Charles L., Lutz J.F., *Covalent Attachment and Detachment by Reactive DESI of Sequence-Coded Polymer Taggants*, „Macromolecular Rapid Communications” 2022, t. 43, s. 2200412, DOI: 10.1002/marc.202200412.
- Zhirnov V., Zadegan R.M., Sandhu G.S., Church G.M., Hughes W.L., *Nucleic acid memory*, „Nature Materials” 2016, t. 15, s. 366–370, DOI: 10.1038/nmat4594.
- Zhu Z., Cardin C. J., Gan Y., Colquhoun H.M., *Sequence-selective assembly of tweezer molecules on linear templates enables frameshift-reading of sequence information*, „Nature Chemistry” 2010, t. 2, s. 653–660, DOI: 10.1038/nchem.699.
- Żyła A., *DNA: prawie niezniszczalny i najbardziej pojemny nośnik danych*, „Archeion” 2021, t. 122, s. 33–43, DOI: 10.4467/26581264ARC.21.014.14494.

Netografia

Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere, Nov 2018, <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf> [dostęp: 21.11.2022].

IEA (2022). Data Centres and Data Transmission Networks, IEA, Paris, <https://www.iea.org/reports/data-centres-and-data-transmission-networks> [dostęp: 09.11.2022].

Statista. Data storage supply and demand worldwide, from 2009 to 2020, <https://www.statista.com/statistics/751749/worldwide-data-storage-capacity-and-demand/> [dostęp: 09.11.2022].