

Next Generation Molecular Data Storage (NEO)

EIC Pathfinder Challenge 2022: DNA-based digital data storage



Paderborn University is the coordinator of NEO and specialised in designing and folding of 2D DNA nanostructures and their stability



University of Surrey is an expert in DNA strand-displacement reactions and in charge of developing a mechanism for highly parallel update information on DNA nanostructures



Imperial College London is specialised in developing encoding, error correction mechanism and automation of data encoding



KU Leuven is specialised in high resolution AFM imaging of 2D structures. They are developing efficient immobilisation and fast imaging of DNA nanostructures.



TU Graz, as experts in computer vision, is developing computer vision pipeline for recognition of DNA nanostructures.

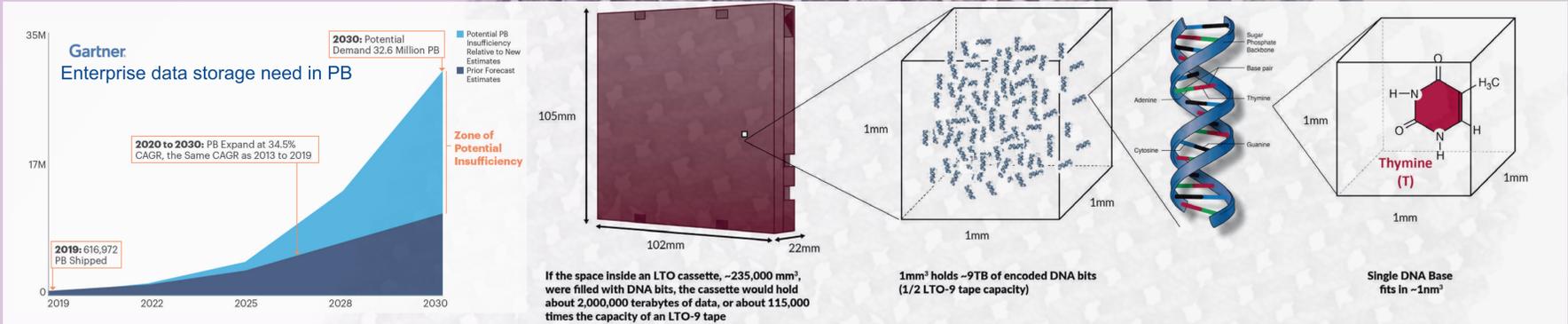


J. Heyrovsky Institute of Physical Chemistry is studying the long-term stability of DNA nanostructures due to radiations and temperatures. **HORIZON-WIDERA-2023-ACCESS**

Why DNA data storage?



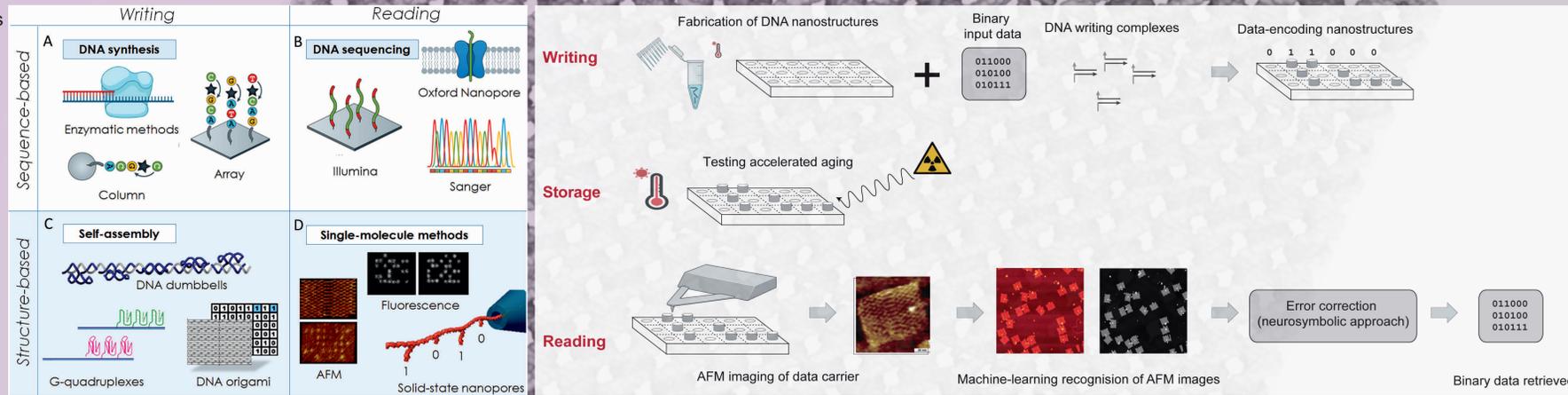
PRESERVING OUR DIGITAL LEGACY: AN INTRODUCTION TO DNA DATA STORAGE 2021 available at: <https://www.snia.org/groups/snia-dna-technology-affiliate>



DNA data storage approaches

> Emerging Approaches to DNA Data Storage: Challenges and Prospects
A. Doricchi, et al., ACS Nano 2022 16 (11), 17552-17571
DOI: 10.1021/acsnano.2c06748

>> NEO consortium approach neodna.eu



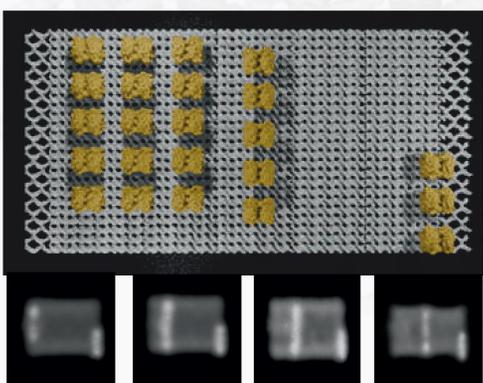
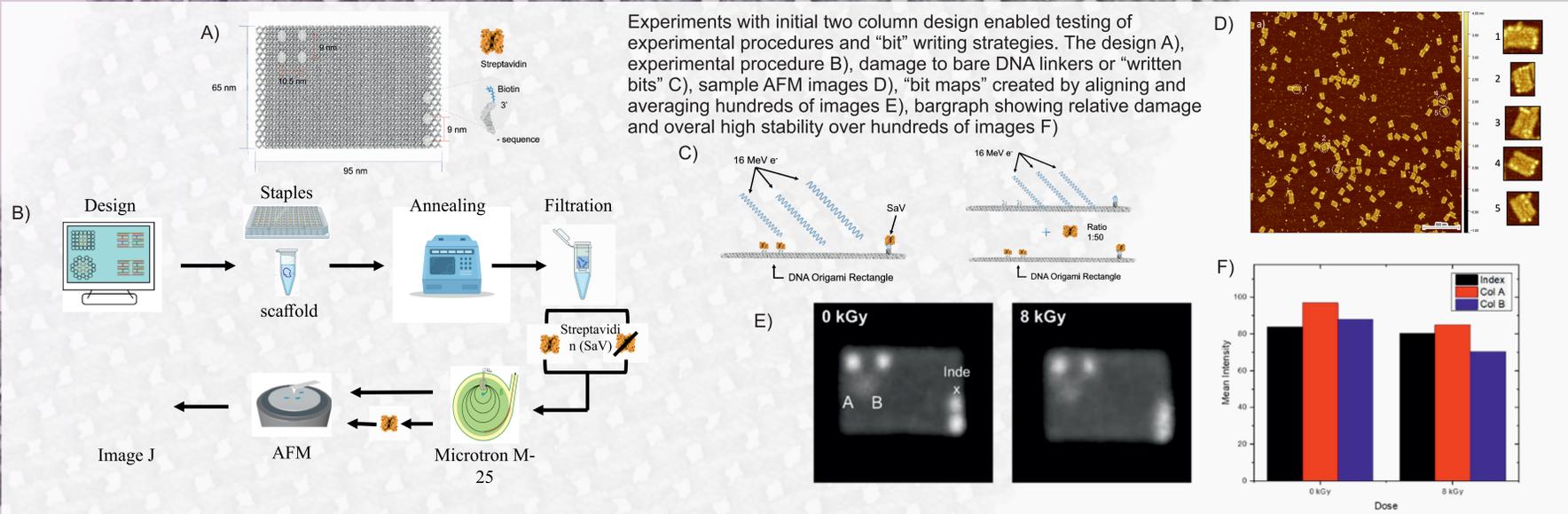
Experiments at J. Heyrovský Institute



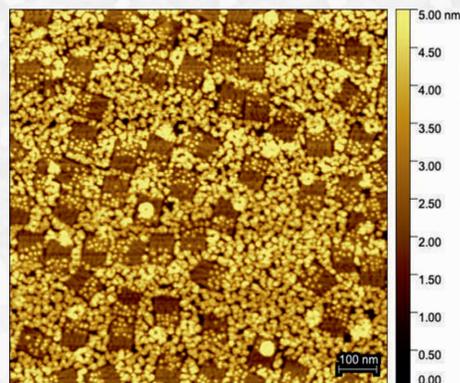
and Institute of Nuclear Physics



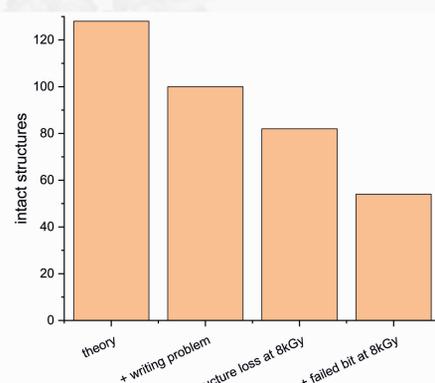
Czech Academy of Sciences



A) Current design and sample "bit maps" from tests of writing probabilities for individual columns.



B) Sample AFM image of the current design with all bits written.



C) "damage budget" for the data carrying nanostructures. At the present state, we are able to correctly write and read around 80% of the predicted data writing bits. This level represents a base for our ionizing radiation studies and is represented by 100. Radiation dose was 8kGy, which is equivalent to approximately 16 million years of cumulative background radiation, considering the common Earth's surface radiation background of around 5mGy/year. At such dose, 20% of the structures are lost completely due to the structural damage to the DNA origami rectangle. Problems with failed bits are then observed at approximately 30%.





Featuring: NEO PROJECT

Interview with Jaroslav Kočíšek (J. Heyrovsky Institute of Physical Chemistry)

The Hop On tool of the WIDERA program enables applicants from widening countries to join a project funded under the Horizon Europe programme in whose consortium there is currently no partner from such a country. The team led by Jaroslav Kočíšek at the J. Heyrovský Institute (JHI) of Physical Chemistry of the Academy of Sciences of the Czech Republic has joined such a project, NEO (Next Generation Molecular Data Storage, <https://neodna.eu/>), among the projects funded in the EIC Pathfinder Challenges 2022 call: DNA-based digital data storage. The NEO project focuses on the development of new memory characterized by significantly higher recording density and durability as well as a smaller carbon footprint of production and operation. In connection with the invitation to present experiences with the Hop On tool at the Czech Days for European Research (CZEDER) conference, we asked the team leader of the Czech partner for some more information.



The team led by Jaroslav Kočíšek; Source: NEO Project

Can you please describe in more detail this new technology, what its advantages are and, on the contrary, the weaknesses that the research could solve?

DNA can be considered a centuries-old proven information carrier. The advantage of DNA is high recording density at the level of tens of TB per mm³, i.e. 1 000x higher than the record-setting 3D NAND memories and about 100 000x higher than magnetic tapes, which are commonly used for archiving. Another advantage is the high stability of DNA and the low energy consumption of this data storage technology. The disadvantages currently being addressed are the high price and time-consuming writing and reading, since DNA needs to be artificially synthesized and subsequently sequenced. Our consortium is trying to circumvent these disadvantages by storing data not in DNA sequences, but by using patterns written on DNA origami nanostructures, which we can picture as QR codes. These can be retrieved using available microscopic methods.

Can you please tell us how the original intention of the consortium is extended by the participation of the JHI team and what the team gains from the collaboration with the consortium?

The consortium aims to demonstrate the complete process of writing and reading data using DNA nanostructures. The consortium, therefore, includes experts in recording information, reading images, or microscopy, as well as experts in the field of DNA nanotechnology. Our task is to study the damage to these nanostructures by ionizing radiation. This allows us to predict their long-term stability and, at the same time, prepare sample data sets with different degrees of damage for training the reading algorithms.

During the project execution, we are establishing ourselves well also in other consortium activities, such as in the design of the resulting nanostructures or the writing process itself.

How did you establish contact with the consortium and convince the coordinator of the 'added value' of your participation?

In our group, we started using DNA origami nanostructures as a tool for studying the interaction of ionizing radiation with precisely defined DNA sequences. However, the implementation of this methodology required a large number of supporting experiments studying the stability of DNA nanostructures themselves. The stability of DNA nanostructures is also the focus of Adrian Keller, who is the NEO project coordinator. During the discussion of joint research, Adrian mentioned that our experiments with ionizing radiation could be interesting for the aforementioned EIC project. A great advantage was that the start of the NEO project was postponed due to the participation of partners from the United Kingdom. This fact allowed us to prepare the WIDERA hop on proposal with only half a year's delay compared to the original project. Otherwise, Hop On would probably not have made sense.

Does this project generate new ideas and contacts for possible further and broader international cooperation?

Obviously, as I have already mentioned, within the consortium we started researching DNA hybridization kinetics, image processing, and high-resolution scanning probe microscopy, i.e. areas in which we were previously more active as users of existing technology. We are also part of the DNA Data Storage Alliance, which brings together major players in the field, such as Microsoft and Western Digital. In summer of 2025, we organized a portfolio meeting of 9 EIC projects on this topic, combined with an international conference, where we gained further contacts. We hope to be involved in other projects in the area so that this interesting research direction can be further developed in the Czech Republic.