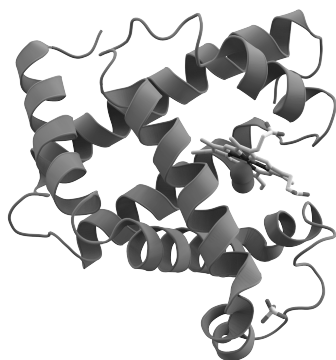


Nie tylko szachy i gra go...

W 1994 roku w środowisku naukowym pojawił się konkurs CASP (*Critical Assessment of Techniques for Protein Structure Prediction*) o oryginalnej formie. Laboratorium, w którym osiągnięto postęp w opracowaniu modelu struktury przestrzennej wybranego białka metodami fizycznymi (rentgenowska analiza kryształu, niskotemperaturowa mikroskopia elektronowa, NMR), udostępniało swoje wyniki organizatorom konkursu. Do wiadomości publicznej podawano tylko kolejność (sekwencję) aminokwasów ułożonych w łańcuch danego białka, strukturę przestrzenną białka pozostawiając jedynie do wiadomości organizatorów konkursu.

Dla aktywności biologicznej białka najważniejszy jest sposób zwinięcia jego łańcucha w przestrzeni. Według reguły sformułowanej jeszcze w latach 60. ubiegłego wieku przez Christiana Anfinsena (Nagroda Nobla z chemii w 1972 r.) sekwencja aminokwasów determinuje jednoznacznie strukturę przestrzenną łańcucha. Od tamtych czasów liczne grupy doświadczalne usiłowały znaleźć reguły opisujące tę strukturę. Zagadnienie błahe nie jest. Forma przestrzenna determinuje oddziaływanie białka z innymi cząsteczkami w komórce, decyduje o jego funkcji biologicznej. Na pewno wielu Czytelnikom pozwoliła ta informacja uznać, że wiedza o strukturze przestrzennej każdego białka jest najważniejsza dla zrozumienia normy i patologii w medycynie. Rosnąca lawinowo wiedza genomyczna pozwala obecnie na tworzenie molekularnych map ewolucyjnych, zagłębienie w historię rozwoju życia. Znajomość struktur przestrzennych białek, kodowanych przez geny, tę wiedzę pogłębia i rozszerza. Nawet w naszym genomie, poznanym już 20 lat temu, istnieje wiele tysięcy sekwencji kodujących białka, których funkcji wciąż nie znamy.

Wracamy do konkursu CASP. Polega on na tym, że uczestnicy jedynie na podstawie znajomości określonej sekwencji aminokwasów mają zaproponować jednoznacznie strukturę przestrzenną białka, którą jurorzy konkursu mogą porównać z ustaloną wcześniej doświadczalnie w laboratorium. Zgodność bywa różna, zależnie od zastosowanych metod teoretycznych, niekiedy jednak prawie doskonała. W konkursie CASP bierze udział około 100 zespołów z całego świata, rozstrzygnięcia zapadają co 2 lata. Organizatorem światowych konkursów związania białek jest dr Krzysztof Fidelis, absolwent Zakładu Biofizyki na Wydziale Fizyki Uniwersytetu Warszawskiego, aktualnie pracuje w Genome Center Uniwersytetu Kalifornijskiego w Davis.



W tej dziedzinie w Polsce działały i osiągały światowe wyniki zespoły stosujące metody heurystyczne wspomagane specjalizowanymi programami komputerowymi (Janusz Bujnicki – Międzynarodowy Instytut Biologii Molekularnej i Komórkowej, Warszawa, Krzysztof Ginalski – do 2019 roku pracujący w Centrum Nowych Technologii Uniwersytetu Warszawskiego, Leszek Rychlewski – BioinfoBank Institute, Poznań), a także zespoły realizujące modelowanie związania białek z wykorzystaniem własnych platform symulacyjnych, bazujących na fizyce i potencjałach oddziaływania między aminokwasami (Andrzej Koliński – Wydział Chemii, Uniwersytet Warszawski i Adam Liwo – Wydział Chemii, Uniwersytet Gdański).

DeepMind, spółka zależna od Google'a, zaprezentowała na konkursie CASP w 2018 roku program sztucznej inteligencji AlphaFold, opisujący trójwymiarowe struktury przestrzenne białek. Ulepszona wersja programu w roku 2020 zdecydowanie wygrała konkurs, osiągając wyniki bliższe doświadczeniu niż wszystkie zespoły ludzkie. Między innymi dzięki temu programowi udało się określić budowę jednego z białek bakteryjnych – rozpracowywaną doświadczalnie od dziesięciu lat. DeepMind działa w skali godzin. W początkach 2020 roku AlphaFold obliczył także strukturę przestrzenną paru białek SARS-CoV-2.

W ten sposób sztuczna inteligencja wkroczyła na pola biologii molekularnej i medycyny bardziej spektakularnie niż zwyciężając z ludźmi w szachy i go.

Można zapytać, czy mając takie osiągnięcia, AlphaFold spowoduje zamknięcie laboratoriów stosujących stare metody przewidywania i badania struktur? Zapewne jeszcze nie jutro. Zmieni się jednak zasadniczo skala czasowa znajdowania odpowiedzi na ważne pytania z zakresu nauk podstawowych i stosowanych.

Magdalena FIKUS (magda.fikus@gmail.com)