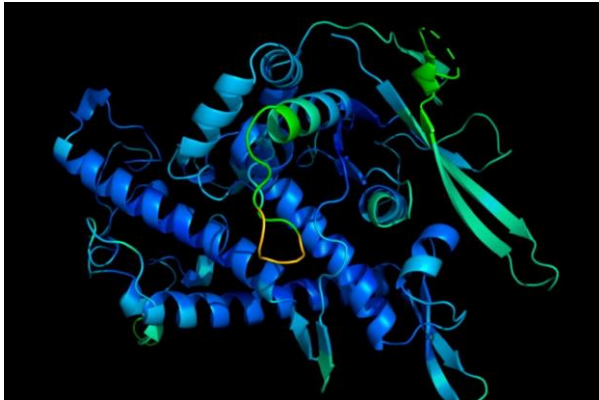


# Znamy kształt białek ludzkiego organizmu. To odkrycie na miarę poznania genomu człowieka

© Paulina Mozolewska

Gazeta Wyborcza, 23.07.2021

<https://wyborcza.pl/7,75400,27357801,firma-deepmind-opracowala-program-oparty-o-sztuczna-inteligencje.html>



Funkcja białka jest określona przez jego kształt 3D. (Fot. DeepMind)

Sztuczna inteligencja AlphaFold firmy DeepMind rozszyfrowała strukturę niemal wszystkich ludzkich białek, tzw. proteom człowieka - piszą naukowcy w specjalnym wydaniu "Nature".

Białko kojarzy się najczęściej ze składnikiem zbilansowanej diety, z białą częścią jajka albo odżywką dla sportowców. Kiedy jednak naukowo mówimy o białkach, mamy zazwyczaj na myśli podstawową jednostkę budulcową oraz funkcjonalną wszystkich organizmów.

## Kształt białek – aminokwasowe korale

Wyobraźmy sobie, że chcemy nawlec sznur koralu. Dysponujemy koralikami w 20 różnych barwach oraz wieloma długościami żyłki, na które będziemy je nawlekać.

W dużym uproszczeniu na tym właśnie polega proces tworzenia białek przez wszystkie ziemskie organizmy. Białka to struktury złożone z występujących po sobie aminokwasów, czyli najmniejszych elementów budulcowych, koralików w naszym naszyjniku. Aminokwasów budujących białka jest 20. Ostateczne "naszyjniki" mogą mieć różną konfigurację i długość, dlatego różnorodność białek w przyrodzie jest ogromna.

W przypadku koralu ostateczny efekt naszej pracy ograniczać będzie jedynie nasza wyobraźnia i cierpliwość do prac manualnych.

*Jeśli chodzi o białka, nie ma mowy o spontaniczności. Każde tworzone jest na podstawie precyzyjnej instrukcji, zapisanej w DNA.*

To właśnie materiał genetyczny niesie dokładny przepis na konstrukcję wszystkich białek budowanych przez skomplikowaną maszynę komórkową. Genom każdego organizmu to zatem opasła księga z przepisami. Aby zmieściła się w komórce, musi być solidnie

upakowana. W swojej najbardziej skondensowanej formie DNA tworzy chromosomy – struktury przypominające literę X.

Podobnie jest z białkami. Kiedy są jeszcze w formie sznura koralu, ich przydatność dla organizmu jest niewielka. Dopiero ich splątanie i ułożenie w precyzyjną trójwymiarową formę sprawia, że każde białko jest gotowe, aby pełnić przeznaczone mu funkcje. Np. enzymy są zaangażowane w metabolizm, a przeciwciała w zwalczanie infekcji.

### **Różnorodność form i kształtów**

Biorąc pod uwagę ogromną różnorodność białek na poziomie struktury prostego naszyjnika oraz mnogość form 3D, które mogą tworzyć, trudno sobie wyobrazić wielość architektur białkowych w przyrodzie.

*Poznanie dokładnych struktur białek w 3D jest jednak kluczowe – choćby po to, by móc projektować nowe terapie leczące wiele chorób.*

Do tej pory aby poznać dokładną budowę przestrzenną białka, naukowcy stosowali wymagające czasu, precyzji i cierpliwości metody oparte na krystalografii rentgenowskiej i mikroskopii krioelektronowej (cryo-EM).

Najpierw musieli wyprodukować wystarczająco dużą ilość białka, potem je oczyścić i przygotować do badania. Niepowodzenie na choćby jednym z tych etapów wiązało się z rozpoczęciem eksperymentów od nowa.

Kolejnym minusem była konieczność precyzji. Każda pomyłka w określaniu struktury ważnych białek mogła skierować dalsze badania w ślepą uliczkę.

### **DeepMind na tropie białek**

Od kilku lat ogromne nadzieje pokładano w zupełnie nowym podejściu do przewidywania struktur przestrzennych białek, opartym na sztucznej inteligencji (AI).

Spółka DeepMind należąca do Google'a opracowała sztuczną inteligencję AlphaFold. Obecnie firma współpracuje z Europejskim Laboratorium Biologii Molekularnej (EMBL), flagowym europejskim laboratorium nauk przyrodniczych.

*Opracowany przez firmę program wykorzystuje uczenie maszynowe. Oznacza to, że system sam uczy się na podstawie swoich wcześniejszych "doświadczeń".*

Podobne schematy znamy ze sposobu, w jakim przeglądarki internetowe dobierają dla nas reklamy. Jeśli akurat przeglądaliśmy oferty telefonów komórkowych, możemy się spodziewać, że to właśnie tego typu banery reklamowe będą nam towarzyszyć przez kolejne dni buszowania w internecie.

Na szczęście sztuczną inteligencję da się również wyszkolić do zadań mniej irytujących.

Na podstawie sekwencji aminokwasowej, czyli sznurka koralu, sztuczna inteligencja potrafi przewidzieć formę przestrzenną białka. System "uczy się" coraz lepiej opracowywać kształty białek.

## Poznanie ludzkiego proteomu zmieni medycynę

Program sztucznej inteligencji firmy DeepMind z roku na rok radził sobie z tym zadaniem coraz lepiej, a jego przewidywania były coraz pewniejsze. W listopadzie 2020 r. podczas konkursu CASP (Critical Assessment of Structure Prediction), czyli tzw. olimpiady białkowej,

*AlphaFold przyćmił konkurencję i w przewidywaniu kształtów białek osiągnął dokładność analogiczną do pracochłonnych metod laboratoryjnych.*

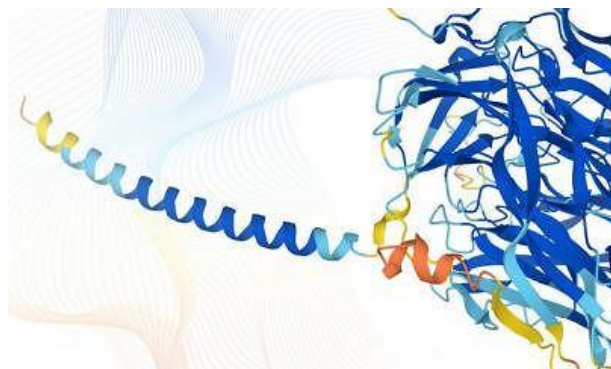
– To przełom w grze. To zmieni medycynę, zmieni badania, zmieni bioinżynierię. To wszystko zmieni – oceniał wtedy Andrei Lupas, biolog ewolucyjny z Instytutu Biologii Rozwoju im. Maxa Plancka w Tybindze w Niemczech.

W minionym tygodniu twórcy oprogramowania opublikowali pracę naukową i kod źródłowy wyjaśniające proces stworzenia swojego systemu.

A w czwartek 22 lipca upublicznili na łamach specjalnego wydania tygodnika "Nature" prognozy dotyczące struktur 3D ok. 350 tys. białek 20 różnych organizmów – czyli ich tzw. proteomy (występujące w komórce pełne zestawy białek wraz z ich kształtem – odpowiedniki genomów, czyli przepisów na konstrukcję białek). Były to głównie tzw. organizmy modelowe, takie jak myszy, muszki owocowe czy bakteria E.coli, wykorzystywane w badaniach laboratoryjnych.

*Oprócz nich w bazie znalazło się aż 98 proc. (czyli ok. 20 tys.) wszystkich ludzkich białek, wspólnie określanych jako ludzki proteom.*

Nowa baza danych ponaddwukrotnie zwiększyła dotychczasową liczbę dostępnych dla naukowców struktur białkowych określonych z wysoką dokładnością.



Fot. Struktury białek reprezentujące dane uzyskane za pomocą AlphaFold Karen Arnott/EMBL-EBI

– W DeepMind naszym celem zawsze było budowanie sztucznej inteligencji, a następnie wykorzystywanie jej jako narzędzia do przyspieszenia tempa odkryć naukowych, a tym samym lepszego zrozumienia otaczającego nas świata – chwali się dr Demis Hassabis, założyciel i dyrektor generalny DeepMind.

*Użyliśmy AlphaFold, aby wygenerować najbardziej kompletny i dokładny obraz ludzkiego proteomu. Uważamy, że stanowi to największy wkład, jaki sztuczna inteligencja wniosła do postępu wiedzy naukowej i jest doskonałą ilustracją tego, jakie korzyści może przynieść społeczeństwu – dodaje szef DeepMind.*

Aby ocenić poziom pewności szacunków, AlphaFold przeanalizowała swoje obliczenia, porównując je z najlepiej do tej pory poznanymi strukturami białek. Wyniki wskazują, że program był w stanie przewidzieć kształt białka na dobrym poziomie w 95 proc. przypadków.

### **Naturalna inteligencja jeszcze nie umarła**

Mimo tej wysokiej pewności standardowe metody określania budowy przestrzennej białek jeszcze długo nie odejdą w zapomnienie. Przewidywania, nawet te o wysokim stopniu prawdopodobieństwa, to jednak wciąż tylko szacunki, które należy zweryfikować.

Jednocześnie prace eksperymentalne nie będą musiały dawać dokładnych wyników. AlphaFold pomoże przewidzieć kształt nowych białek już na podstawie skrawka dotychczas potrzebnych informacji.

*To, co nam zajęło miesiące i lata, AlphaFold był w stanie zrobić w jeden weekend – nie kryje prof. John McGeehan, dyrektor Centrum Innowacji Enzymów na Uniwersytecie w Portsmouth w Wielkiej Brytanii.*

Teraz rozpocznie się czas testowania rozwiązań oprogramowania oraz niezależna ocena dokładności ich przewidywań.

Cała baza danych oraz kod źródłowy programu zostały udostępnione w systemie otwartym. Oznacza to, że każdy może wykorzystać te informacje do swoich własnych celów. Trzeba oczywiście wiedzieć, jak je wykorzystać, no i odpowiednio oznaczyć ich pochodzenie.

Na pytanie, dlaczego DeepMind udostępnia tak dużą wiedzę za darmo, jej twórcy odpowiadają, że od samego początku taka była misja firmy. Poza tym w konstruowaniu i ulepszaniu systemu pracowano na dostępnych otwartych bazach danych, tworzonych przez inne zespoły, często za publiczne pieniądze.

*„Baza danych AlphaFold jest doskonałym przykładem pozytywnego sprzężenia zwrotnego otwartej nauki” – cieszy się dyrektor generalna EMBL Edith Heard.*

– Jeśli chcemy zrozumieć świat, dane muszą być ogólnodostępne – dodaje jej zastępca Ewan Birney.

### **Kropla w morzu możliwości**

DeepMind już rozpoczęła współpracę z naukowcami zajmującymi się projektami, w których system AlphaFold może przynieść przełom.

Jest to między innymi poszukiwanie leków na choroby dotyczące głównie biedniejsze rejony świata, badania nad opornością bakterii na antybiotyki, problem zanieczyszczenia środowiska plastikiem czy biologia koronawirusa SARS-CoV-2 wywołującego COVID-19.

Osiągnięcia twórców programu z pewnością umożliwią naukowcom wielu dziedzin przyspieszenie ich pracy. Jednym z najciekawszych aspektów będzie wykorzystanie systemu do badania potencjalnych leków, szczepionek i enzymów. Ale to tylko kropla w morzu możliwości.

Ekspertcy uważają również, że otwierają się nowe możliwości użycia sztucznej inteligencji w badaniach biomedycznych. Być może wraz z rozwojem systemu powstaną nowe, niewyobrażalne jeszcze dziś dziedziny nauki.

*A ponieważ pole do zastosowania technologii jest nieograniczone, nie da się obecnie wycenić wartości tego odkrycia.*

Baza danych i system będą okresowo aktualizowane. W nadchodzących miesiącach ich twórcy planują znacznie rozszerzyć zasięg na prawie każde zsekwencjonowane białko znane nauce – czyli ponad 100 mln struktur!

– Będzie to jeden z najważniejszych zbiorów danych od czasu zmapowania ludzkiego genomu" – przyznaje zastępca dyrektora generalnej EMBL Ewan Birney.

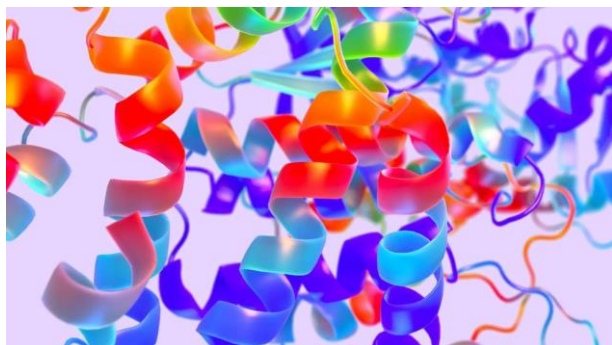
\*\*\*\*\* Wcześniejszy tekst na ten temat – poniżej

## **Google hakuje biologię. Sztuczna inteligencja DeepMind okazuje się niezrównana w odgadywaniu struktury białek**

© Wojciech Orliński

Gazeta Wyborcza, 08.12.2020

[https://wyborcza.pl/7,75400,26582381,google-hakuje-biologie-sztuczna-inteligencja-deepmind-jest.html#s=S.embed\\_article-K.C-B.1-L.5.zw](https://wyborcza.pl/7,75400,26582381,google-hakuje-biologie-sztuczna-inteligencja-deepmind-jest.html#s=S.embed_article-K.C-B.1-L.5.zw)



Symulacja trójwymiarowego kształtu białek za pomocą programu sztucznej inteligencji AlphaFold (DeepMind)

Wzrostywanym co dwa lata konkursie CASP wygrywa ten, kto najlepiej przewidzi strukturę przestrzenną białek. W tym roku wszystkich daleko w tyle pozostawiła sztuczna inteligencja AlphaFold zaprogramowana przez spółkę DeepMind należącą do Google'a. Biolodzy dostali do ręki potężne narzędzie. Co z nim zrobią?

„Życie jest formą istnienia białka, ale w kominie coś czasem załka” - śpiewali Skaldowie, parafrazując Fryderyka Engelsa (przetworzonego przez poetycki talent Agnieszki Osieckiej). Dzięki sztucznej inteligencji AlphaFold zaprogramowanej przez Google'a przybliżyliśmy się do rozumienia działania białek, choć nadal komin pozostaje pełen zagadek.

Za życia Engelsa chemicy jeszcze nawet nie rozumieli, co to jest białko. W dosłownym cytacie Engels nie użył tego słowa, tylko nieprecyzyjnego określenia „substancje albumiczne” (dziś

albuminami nazywamy białka rozpuszczalne w czystej wodzie w odróżnieniu od nierozpuszczalnych w niej globulin).

Obecnie wiemy już bardzo dużo, ale im więcej się dowiadujemy, tym więcej napotykamy zagadek. Podsumujmy więc na razie stan wiedzy w roku 2020, stan niewiedzy (owego „łkania w kominie”) oraz pospekulujmy, co się zmieni dzięki AlphaFold.

### **Prosty przepis na białko**

W szkole uczymy się, że białka składają się z aminokwasów. Jak sama nazwa wskazuje, aminokwas to substancja, która ma grupę kwasową ( $-\text{COOH}$ ) oraz grupę aminową ( $-\text{NH}_2$ ).

Te grupy chętnie ze sobą reagują, tworząc wiązanie C-N (i uwalniając cząsteczkę wody,  $\text{H}_2\text{O}$ ). Powstaje wtedy amid (najbardziej znanym amidem jest chyba dietyloamid kwasu lizergowego, czyli LSD).

Jeśli aminokwas tworzy takie wiązanie z drugim aminokwasem, dzieje się coś bardzo ciekawego. Grupa kwasowa (fachowo zwana karboksylową) jednego z nich łączy się z grupą aminową drugiego, ale ten pierwszy nadal ma gotową do kolejnej reakcji grupę aminową, a ten drugi grupę karboksylową - obie umieszczone teraz po przeciwnych stronach utworzonej molekuly.

Nic nie stoi na przeszkodzie, żeby do tej wolnej grupy  $\text{COOH}$  przyczepił się inny aminokwas swoją grupą  $\text{NH}_2$ , a z drugiej strony - by do tej wolnej grupy  $\text{NH}_2$  inny aminokwas przyczepił się swoją grupą  $\text{COOH}$ .

Tak można w nieskończoność, choć w praktyce kończy się to zazwyczaj na setkach czy tysiącach aminokwasów. Opisane powyżej wiązanie C-N-C nazywamy wiązaniem peptydowym, alternatywną nazwą dla białek są więc peptydy (zwyczajowo jednak przyjęło się nazywanie tak krótkich łańcuchów, złożonych zaledwie z kilku aminokwasów).

Najbardziej niesamowitym odkryciem dokonany przez chemików na początku XX wieku było to, że niemal wszystkie białka składają się z zaledwie 22 aminokwasów, zwanych proteinogenicznymi. Chemicy znają ich dużo więcej, bo jak pamiętamy ze szkoły, do węgla zawierającego grupy  $-\text{NH}_2$  i  $-\text{COOH}$  możemy dopisać sobie niemal dowolną resztę, skrótowo zapisywaną jako -R.

Przyrodzie wystarczają 22 reszty. W popularnych tekstach spotkamy też liczby 20 i 21 - wszystkie da się uzasadnić, ale nie wchodźmy w to teraz, bo jakkolwiek by patrzeć, to niedużo.

*Już zróżnicowanie tych białek, które każdy z nas ma w swoim własnym organizmie, jest ogromne. Keratyna w naszych włosach i paznokciach, hemoglobina w naszej krwi, insulina w naszej trzustce, kolagen w naszych ścięgnach - to wszystko białka.*

Na pierwszy rzut oka różnią się tak bardzo, że aż trudno uwierzyć, że należą do tej samej grupy związków chemicznych. Jedne rozpuszczają się w wodzie, inne nie. Jedne są przezroczyste, inne nie. Jedne są ciałem stałym, inne cieczą.

Jeszcze bardziej niesamowite jest to, że białka w innych organizmach żywych - roślinach, grzybach, pierwotniakach, bakteriach itd. - są zbudowane z tych samych aminokwasów. Czasem to nawet te same białka (przyroda jest leniwa).

Jak to możliwe? Chemicy długo tego nie mogli zrozumieć.

Od XIX wieku jednak wiedzieli jedno. Pod wpływem różnych czynników, takich jak temperatura (niezbyt zresztą wysoka), białka gwałtownie tracą swoje właściwości.

Nazywano to denaturacją lub po prostu ścięciem. Typowy przykład to po prostu robienie jajecznicy.

### **Tajemnice jajecznicy: helisa alfa oraz harmonijka beta**

Co dokładnie się dzieje podczas smażenia jajek, zrozumiał ok. 70 lat temu Linus Pauling (dwukrotny laureat Nagrody Nobla). Okazało się, że powyższy opis białka jako łańcucha aminokwasów to dopiero wierzchołek góry lodowej.

Wyobraźmy sobie jakieś bardzo proste białko. Założmy, że składa się tylko z jednego aminokwasu, a do tego najprostszego: glicyny, dla której reszta R to po prostu atom wodoru, czyli mówiąc chemicznie,  $R=H$ .

Nasze białko ma fachową nazwę - poliglicyna. Skrótowo poli-Gly.

Zastanówmy się teraz, jak wygląda jego molekula. Czy to jest jedna prosta kreska Gly-Gly-Gly-Gly?

Już 100 lat temu chemicy wiedzieli, że to mało prawdopodobne. Wiedzieli, że atomy wodoru w grupach -OH mają skłonność do przyklejania się do wolnych par elektronowych w atomach tlenu i azotu. Nazwali to mostkami wodorowymi.

Nawet najprostszy aminokwas ma jedno i drugie, a więc oczekivalibyśmy po poliglicynie, że powstaną w niej jakieś mostki wodorowe. Ale co się przyklei do czego? Wodór w dwudziestej glicynie do azotu z pięćdziesiątej glicyny? A może odwrotnie?

Linus Pauling odkrył, że odpowiedź jest stosunkowo prosta. Większość białek w praktyce przyjmuje jedną z dwóch struktur, które nazwano helisą alfa i harmonijką beta. Owa poliglicyna w wersji podstawowej (takiej, jaka przyjedzie do nas ze sklepu) będzie na przykład mieć molekuly w formie harmonijki.

Mostki wodorowe, które stabilizują te struktury, są znacznie słabsze od wiązań chemicznych. Pękają od byle czego, np. od niezbyt wysokiej temperatury. I to jest właśnie denaturacja.

Od 70 lat wiemy więc, że strukturę białka można rozpatrywać na kilku poziomach. Po polsku mówi się o rzędach.

### **Bez przyzwyczajenia w życiu ani rusz**

Struktura pierwszorzędowa to kolejność aminokwasów w łańcuchu, np. Gly-Ile-Val-Glu-Gln... to pierwsze pięć aminokwasów insuliny (glicyna, izoleucyna, walina, kwas glutaminowy, glutamina...).

Struktura drugorzędowa to kształt, który ten łańcuch przyjmuje świeżo po wyprodukowaniu go np. w komórce trzustki. W przypadku insuliny to głównie helisa alfa.

Trzeciorzędowa to ostateczny trójwymiarowy wygląd całości białka. Wpływają na niego różne dodatkowe czynniki, np. środowisko otaczające białko. Molekuły białka potrafią się ruszać, działając jak małe biologiczne maszyny.

Jeśli cząsteczki białka łączą się w większe kompleksy, nazywamy je strukturą czwartorzędową. Nie każde białko ma zaszczyt ją posiadać - ale np. hemoglobina jest tak naprawdę tetramerem, czyli składa się z czterech jednakowych łańcuchów.

Struktura drugorzędowa wynika pierwszorzędowej. Trzeciorzędowa z drugorzędowej i warunków środowiskowych. W jej przybraniu pomagają inne białka, zwane po polsku opiekuńczymi, a po angielsku i w innych językach przyzwoitkami („chaperone”). Metafora przyzwoitki jest o tyle trafna, że ich zadaniem jest pilnowanie wyglądu i zachowania innych białek.

Struktura czwartorzędowa wynika wreszcie z roli, jaką białko odgrywa w organizmie. Molekuła wytworzona w laboratorium może nie umieć jej przybrać z powodu braku jakiegoś mechanizmu działającego w żywej komórce (np. z powodu nieobecności przyzwoitki wejdzie w nie taki związek, w jaki powinno).

### **Co z tego wyniknie dla Jana Kowalskiego?**

Byłoby cudownie, gdyby biolodzy potrafili dzięki samej sekwencji aminokwasów (czyli struktury pierwszorzędowej) odgadnąć, że dane białko służy do trawienia cukru. Niestety, jesteśmy od tego kosmicznie daleko - i ten dystans nadal jest kosmiczny, choć dzięki AlphaFold zmienił się z takiego „jak stąd na Plutona” na taki „jak stąd na Marsa”.

Doświadczony biolog molekularny drugorzędową strukturę białka potrafi czasem odgadnąć, po prostu patrząc na pierwszorzędową. Jedne aminokwasy sprzyjają tworzeniu helisy alfa, inne harmonijki beta.

Od kilkudziesięciu lat wiadomo, które są które. Studenci pomagają sobie w zapamiętaniu tego mnemotechnicznymi regułami, takimi jak „zasada MALEK”, od literowych oznaczeń aminokwasów wyjątkowo chętnie tworzących helisę alfa. Zgodnie z tą zasadą możemy z dużym prawdopodobieństwem odgadnąć strukturę drugorzędową, jeśli w łańcuchu widzimy dużo cząstek metioniny, alaniny i leucyny.

Ze strukturą trzeciorzędową już nie jest tak prosto. Zbyt dużo czynników trzeba wziąć pod uwagę - może Chuck Norris dałby radę, gdyby studiował biologię. Reszcie pozostają programy komputerowe, które to próbują przewidzieć.

*Do skuteczności Chucka Norrisa im jednak daleko.*

Film na YT (engl.) Czym jest zwijanie białka? <https://youtu.be/KpedmJdrTpY>

Od 1994 roku naukowcy organizują konkurs krytycznej oceny przewidywania struktury białek CASP (Critical Assessment of Protein Structure Prediction), w którym modeluje się białka o strukturze ponad wszelką wątpliwość ustalonej krystalograficznie.

Ekspertcy przyznają tym modelom ocenę punktową, która w ogromnym uproszczeniu mówi, na ile procent model zgadzał się z rzeczywistością. W tym pierwszym konkursie z 1994 roku symulacje dawały zgodność bliską 100 proc. w przypadku bardzo prostych białek, ale szybko spadała poniżej 20 proc. w miarę narastającej złożoności.

Ówczesne programy komputerowe działały zgodnie z dominującym wtedy podejściem do sztucznej inteligencji, zwanym systemami eksperckimi. Maszynę programuje się wtedy tak, żeby jej zachowanie przypominało ludzkiego eksperta, czyli na przykład znane ludziom reguły, takie jak "zasada MALEK", przekłada się na język programowania (upraszczając, w pseudokodzie: IF metionina or alanina THEN helisa).

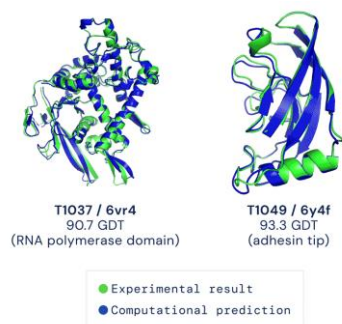
Google ze swoim systemem DeepMind od lat rozwija inne podejście - uczenie maszynowe. Komputer „karmiony” jest w nim jakimś zespołem danych, np. bazą danych o wszystkich znanych białkach, a potem „proszony” o to, żeby sam sobie w tym szukał prawidłowości.

W ten sposób od trzech lat działa Google Translate. Czasami może nas zaskakiwać rezultat jakiegoś przekładu - i wtedy nikt, nawet programiści z Google’a nie odgadną, dlaczego maszyna przełożyła coś tak, a nie inaczej. Jeśli jednak oceniać punktowo jakość przekładów, w 2017 roku skokowo się poprawiła - a z roku na rok jest coraz lepsza.

Maszynowe uczenie się cały czas uczy się na swoich błędach, a systemy eksperckie nie. Cóż, w końcu ich zadaniem jest udawanie zachowania ludzkich ekspertów.

AlphaFold wystartował w poprzedniej edycji konkursu i już wtedy wygrał ze średni wynikiem 87 proc. Jego obecny wynik (92 proc.) to chyba wyrok śmierci na tradycyjne metody symulowania - google'owski system jest bezspornie lepszy.

Poniżej: przykłady dwóch białek, których trójwymiarową strukturę odgadł AlphaFold (na zielono - wynik laboratoryjnych pomiarów, na niebiesko - wynik komputerowej symulacji struktury białka)



*Czy to koniec biologii? Czy można teraz przewidzieć strukturę dowolnego białka? Po trzykroć nie.*

Nie - bo nadal mamy te ok. 8 proc. wymykające się symulacji. Nie - bo białka często wymagają do działania jakiejś składowej niebiałkowej, a nawet nieorganicznej (np. kation żelaza w hemoglobinie).

I przede wszystkim nie - bo to tylko struktury drugo- i trzeciorzędowe, a w biologii najważniejsze są te czwartorzędowe. Do ich przewidywania potrzebujemy już tylu czynników

(np. uwzględnienia roli „opiekuńczych przyzwoitek”), że nie poradzi sobie Chuck Norris z superkomputerem. Na pewno nie poradzi sobie w każdym razie AlphaFold.

Biolodzy i chemicy zyskali natomiast cenne narzędzie pracy. Gdy kiedyś dostali mikroskop, aparat rentgenowski czy spektrometr, to nie oznaczało dla nich utraty pracy, tylko przeciwnie, stworzenie nowych instytutów (spektroskopii, krystalografii, mikrobiologii itd.).

Tak samo będzie teraz. 20 lat temu przewidywanie struktury białek było żmudnym i niedokładnym procesem. Teraz w typowych przypadkach odpowiedź dokładną na 99,9 proc. będziemy dostawać w ciągu sekundy.

Nowe narzędzie otworzy nowe możliwości. Trudno jeszcze powiedzieć jakie - wynalazca mikroskopu też przecież nie wiedział, że 200 lat później wyjdzie z tego mikrobiologia.

Jedna możliwość nasuwa się jako najbardziej oczywista - teraz będzie można odwrócić ten proces. Narysujemy, „jak białko ma wyglądać”, i pójdziemy z tym do firmy biotechnicznej, żeby nam zrealizowała zamówienie „pod klucz”. Otrzymamy fiolkę z bakterią, która dostanie sztuczny gen produkujący to białko.

Skoro teraz AlphaFold potrafi w ciągu sekund na podstawie pierwszorzędowej struktury odgadnąć strukturę trzeciorzędową - to ile mu zajmie wyszukanie takiej struktury pierwszorzędowej, żeby pasowała do zadanej trzeciorzędowej? A niechby nawet i tysiąc sekund, nagle z niemożliwego to będzie możliwe.

*Będziemy to mogli robić nawet dla kaprysu. Elon Musk będzie mógł wyhodować drożdże produkujące białko o kształcie napisu „TESLA”. A do tego wytwarzające prąd.*

Czy coś z tego wyniknie dla Jana Kowalskiego? Na pewno, choć oczywiście nie od razu. W swojej kryształowej kuli już jednak widzę odbywającą się za parę lat dyskusję o tym, czy NFZ powinien refundować terapię wynalezioną przy pomocy AlphaFold.

Obraz jest rozmazany, więc nie widzę, co w ten sposób chcą leczyć. Cóż, kryształowe kule też nie chcą działać na 100 proc.

Film na YT (engl.) AlphaFold: The making of a scientific breakthrough

<https://youtu.be/gg7WjuFs8F4>