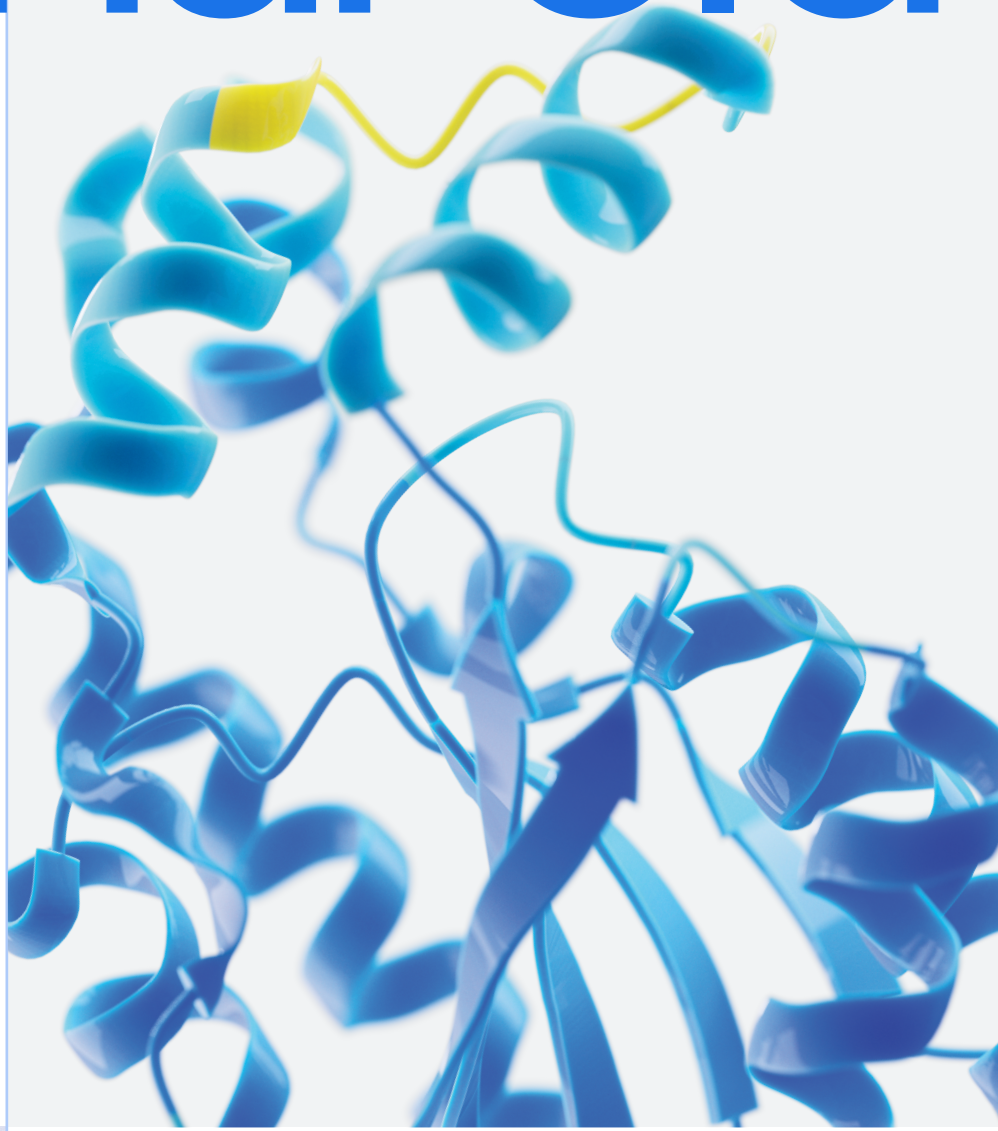


AlphaFold



AI 加速生物科學
突破

AlphaFold 2

如何揭開自然界中最堅韌的蛋白質之一

結構生物學家徐尚德博士 (Dr. Danny Hsu) 及其團隊主要探索分子的 3D 結構, 特別是其中的蛋白質組成。他們對基因突變或錯誤折疊所產生的結果特別感興趣, 因為其打結和摺疊的方式, 會直接影響如阿茲海默症、帕金森氏症或癌症等疾病的發生。



Q9PR55 蛋白質是一個重大的謎團, 其分子結構異常複雜且神秘。徐博士任職於臺灣中央研究院生物化學所, 擔任研究員兼副所長, 多年來, 他和團隊系統性地分析帶有拓樸扭結骨架構型的蛋白質折疊機制和功能。這項工作需要仰賴其他結構生物學家共同分析不同蛋白質結構所共同建立的分子結構資料庫, 從中尋找帶有扭結構型的蛋白質。直到 2022 年 Google DeepMind AlphaFold 一口氣提供了兩億個蛋白質結構預測的龐大資料庫, 其中也預測了許多前所未見的扭結蛋白結構, 徐博士決定測試 Google DeepMind AlphaFold 2 的預測能力。

AlphaFold 2 提出了一個前所未見的 **7₁ 環面結** (7₁-torus knot) 拓樸結構, 一個從未有人聽過或見過的原型。徐博士說:「我們最初不確定 AlphaFold 2 的預測是否值得信賴, 因為我們認為這個工具必須使用現有數據來訓練。蛋白質的摺疊因其極度複雜而難以預測。所以, 如果這個工具提出一個如此錯綜複雜的新摺疊類型, 它會是真的嗎?」

然而, 事實證明, AlphaFold 2 準確地建構了迄今為止已知最複雜的蛋白質結構——而且它完全沒有參考任何關於該蛋白質的現有資訊或數據。



預測未知,填補空白

徐博士的團隊著手驗證 AlphaFold 2 這個大膽的預測。他們使用 X 射線晶體學和核磁共振光譜學等標準方法,來描繪出該蛋白質的真實形態,結果與 **7, 環面結** 的預測近乎完美吻合。這個結構被證實為迄今為止已知最小、最複雜的蛋白質結。徐博士表示:「這向我們證明,AlphaFold 不僅能預測從未見過的东西,其預測結果也是可靠的。這個工具有助於填補我們原本無法解決的空白,這非常重要。」



拼湊出拼圖的關鍵

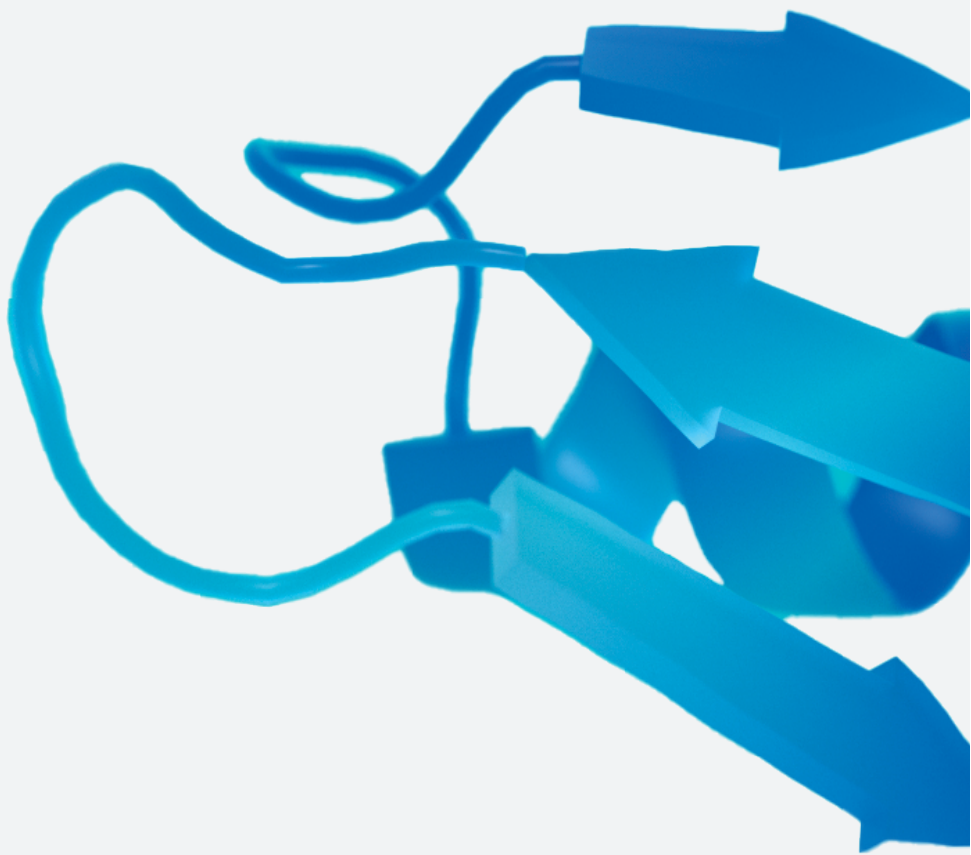
通常,這類規模的研究需要數月才能解決,有時甚至要花費數年,團隊仍可能一無所獲。各種技術瓶頸會阻礙進展,例如鑑定蛋白質所需的時間,這往往是研究中關鍵的第一步。在先進的 AI 出現之前,這項工作可能涉及蛋白質晶體學,一個更像是藝術而非科學的過程。舉例來說,徐博士的另一項研究曾試圖鑑定與許多癌症息息相關的蛋白質晶體結構,但十年後仍一無所獲。

AlphaFold 2 對 **7, 環面結** 拓撲結構的預測,顯著地將研究的起點向前推進。徐博士分享道:「在不到六個月的時間裡,我們從零開始啟動研究,透過 AlphaFold 2 預測所利用的蛋白質序列進行基因合成和蛋白質生產,分子 3D 結構也很快就確定了。一旦結構被確定,後續的工作就相當直接了。」

7, 環面結的意義

7, 環面結之所以特別有趣,在於其打結的形成方式。蛋白質鏈的每一部分都參與了打結,形成了所謂的「緊結」(tight knot)。7, 環面結究竟是如何形成如此複雜的結構仍是個謎,但它顯然經歷了一個高度精密的摺疊過程才達到目前狀態。徐博士的意外發現為生物學界提供了探索許多其他研究領域的新起點,包括 7, 環面結的可承受外力拉扯不易解開折疊結構,這個特性可能與它獨特且複雜的打結模式有關。

儘管 7, 環面結的生物功能尚待發現,但它突如其來的發現是極具顛覆性的。徐博士說:「展望未來,我們可以利用 AlphaFold 2 給我們一個非常接近的蛋白質結構猜測,這為我們提供了一個很好的起點。特別是它的**置信度分數**(confidence score),在我們的實驗分析上證實非常有用,因為它能讓我們知道在某個結構域的預測比另一個結構域更有信心。這有助於我們在解讀預測時引導分析,也應該讓更多 AlphaFold 的使用者注意這個指標的重要性。」



從 7, 環面結到祖先序列

AlphaFold 2 能夠準確預測像 7, 環面結這樣前所未見的蛋白質結構,且無需事先數據,這為結構生物學領域曾被認為是科幻般的可能性開啟了一扇大門。這證明了 AI 不只是一個插值工具或花俏的資料庫搜尋,更是一個已掌握蛋白質摺疊基本原理的深度學習引擎。這個突破開啟了一個全新的世界。這個工具對一個從未見過的結構所做的大膽預測,激勵徐博士和他的實驗室去追求看似不可能的任務,最終導致了迄今為止最複雜的蛋白質結被發現。

對徐博士而言,這股新興的力量讓研究人員能夠透過回溯時間來解決過去長期懸而未決的遺傳難題。他表示:「憑藉 AlphaFold 的力量,我們可以探索**祖先序列重建**(ancestral sequence reconstruction)。我們可以回到過去,看看一種蛋白質可能長什麼樣子,預測缺失的結構,甚至洞察一種蛋白質是如何演化的。」

透過這種方式,AlphaFold 已成為計算生物學的理论世界與實驗室科學的現實世界之間的終極橋樑。它將曾是費力且碰運氣的過程,轉變為一個通往突破性發現的高效率、高針對性的旅程。