



Este esquema representa el proceso que usa la inteligencia artificial para predecir la estructura tridimensional de una proteína a partir de su secuencia de aminoácidos, un avance que fue reconocido en el Premio Nobel de Química.

Aquí está el desglose del flujo de trabajo representado:

1. **Secuencia de proteínas:** El proceso comienza con la secuencia de aminoácidos de la proteína (los componentes básicos de la proteína están representados por diferentes símbolos de colores en el diagrama).
2. **Búsqueda genética y generación de embeddings (MSA embedding):** La secuencia de la proteína se compara con una base de datos de secuencias genéticas para encontrar proteínas similares. Esto genera una "alineación de secuencias múltiples" (MSA, por sus siglas en inglés), que permite al modelo ver patrones de conservación evolutiva entre proteínas similares. Estos patrones ayudan a predecir qué aminoácidos están más probablemente conectados o cercanos en la estructura tridimensional.
3. **Edges de residuos y secuencias (Sequence-residue edges):** Los residuos (aminoácidos individuales) de la proteína y sus posiciones se representan en un mapa de interacciones, que indica qué residuos están cerca entre sí en la estructura final. Esto se representa en una matriz que muestra las posibles distancias entre pares de residuos.
4. **Residue-residue edges y distancias por pares (Pairwise distances):** La información de interacción entre residuos se usa para calcular distancias entre cada par de aminoácidos, generando un mapa de distancias. Esto se convierte en una representación de la proximidad de aminoácidos en el espacio tridimensional.
5. **Módulo de estructura y puntaje de confianza (Structure module and Confidence Score):** Usando el modelo de IA, se calcula una estructura 3D de la proteína y se asigna un puntaje de confianza que indica cuán probable es que la estructura predicha sea precisa.
6. **Estructura tridimensional (3D structure):** Finalmente, el modelo genera la estructura tridimensional de la proteína. Esta representación es fundamental para comprender cómo funciona la proteína en el cuerpo y cómo puede interactuar con otras moléculas, lo cual es crucial para el diseño de fármacos y estudios biológicos.

Este proceso, facilitado por inteligencia artificial, permite obtener predicciones de estructuras de proteínas de manera rápida y precisa, lo cual solía ser un proceso extremadamente lento y costoso antes del uso de estos algoritmos avanzados.