

# Proteinstruktur und KI

Wie künstliche Intelligenz Proteinstrukturen vorhersagt

Fares Albakri, Adnan Memic, Anna Petrosyan

Fachbereich Informatik

11. Februar 2025

# Gliederung

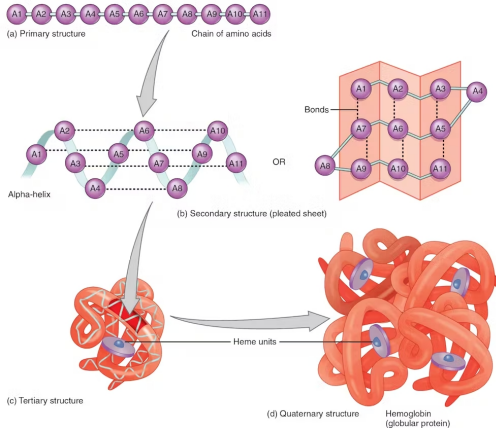
- 1 Einführung und Grundlagen
- 2 KI-Ansätze
- 3 Anwendungen und Herausforderungen

# Was sind Proteine?

## Proteine: Die Bausteine des Lebens

- Essenzielle Makromoleküle für biologische Prozesse
- Funktionen:
  - **Enzyme:** Katalysieren biochemische Reaktionen (z.B. Fettabbau)
  - **Hormone:** Regulieren Körperfunktionen (z.B. Insulin)
  - **Strukturelle Komponenten:** Bilden Haare, Haut und Muskeln
- Aufbau: Lange Ketten von Aminosäuren, verknüpft durch Peptidbindungen

# Proteinstruktur: Ein Überblick



Quelle: Zimmer, 2020. AI Makes Huge Progress Predicting How

# Ebenen der Proteinstruktur

## Die vier Ebenen der Proteinstruktur:

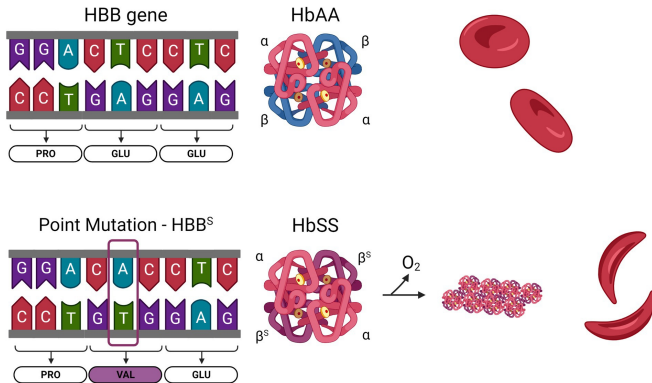
- **Primärstruktur:** Die Aminosäuresequenz
- **Sekundärstruktur:** Faltung in Alpha-Helices (Spiralen) und Beta-Faltblätter (Zickzackmuster)
- **Tertiärstruktur:** 3D-Faltung einer einzelnen Polypeptidkette zu einem funktionellen Protein
- **Quartärstruktur:** Zusammenschluss mehrerer Polypeptidketten (z.B. Hämoglobin)

# Beispiel für Fehlfaltung: Sichelzellanämie

## Sichelzellanämie: Eine einzige Aminosäureänderung

- **Veränderung:** Glutaminsäure (hydrophil) wird durch Valin (hydrophob) ersetzt
- **Folgen:**
  - Hämoglobin-Moleküle verklumpen
  - Sichelförmige rote Blutkörperchen blockieren den Blutfluss
  - Schmerzen, Müdigkeit und Organschäden
- **Fazit:** Eine kleine Änderung kann die Proteinfunktion stark beeinträchtigen

# Beispiel von Sichelzellanämie



Quelle: Ramadas Sparkenbaugh, 2023. The APC-EPCR-PAR1 Axis in Sickle Cell Disease.

# Die Komplexität der Proteinfaltung

## Warum ist Proteinfaltung schwer vorherzusagen?

### ■ Levinthals Paradoxon:

- Ein kleines Protein mit 100 Aminosäuren hat mehr mögliche Faltungen als Atome im Universum
- Trotzdem falten sich Proteine in Millisekunden

### ■ Herausforderungen:

- Wechselwirkungen zwischen Aminosäuren (z.B. Wasserstoffbrücken, hydrophobe Effekte)
- Einfluss der Umgebung (z.B. Temperatur, pH-Wert)



# Technische Herausforderungen und Lösungen

## Wie unterstützt die Computerwissenschaft die Proteinmodellierung?

- **Simulationen:** Rechnergestützte Modelle helfen, Proteinfaltung vorherzusagen
- **KI und maschinelles Lernen:** Systeme wie AlphaFold revolutionieren die Proteinforschung
- **Zukunftsausblick:** Fortschritte in der Informatik ermöglichen neue medizinische Entdeckungen

# Wie funktioniert AlphaFold?

## 1 Datenanalyse mit Multiple Sequence Alignment (MSA)

AlphaFold sucht ähnliche Proteine in Datenbanken und erkennt Muster

Vergleich wie Sprachverarbeitung bei GPT zur Mustererkennung

# Wie funktioniert AlphaFold?

## 1 Datenanalyse mit Multiple Sequence Alignment (MSA)

AlphaFold sucht ähnliche Proteine in Datenbanken und erkennt Muster

Vergleich wie Sprachverarbeitung bei GPT zur Mustererkennung

## 2 Transformer-Netzwerk

Berechnet, welche Aminosäuren interagieren

Nutzt Self-Attention-Mechanismen ähnlich wie BERT

# Wie funktioniert AlphaFold?

## 1 Datenanalyse mit Multiple Sequence Alignment (MSA)

AlphaFold sucht ähnliche Proteine in Datenbanken und erkennt Muster

Vergleich wie Sprachverarbeitung bei GPT zur Mustererkennung

## 2 Transformer-Netzwerk

Berechnet, welche Aminosäuren interagieren

Nutzt Self-Attention-Mechanismen ähnlich wie BERT

## 3 3D-Faltungsmodellierung mit Gradient Descent

Optimiert die Struktur, um die wahrscheinlichste Faltung zu finden

# Wie sicher ist die Vorhersage?

- AlphaFold gibt eine Vertrauensbewertung für jedes Atom
- **Predicted Local Distance Difference Test-Score (pLDDT):**
  - Blau: Sehr genau über 90
  - Gelb: Mittlere Genauigkeit zwischen 70 und 90
  - Rot: Unsichere Vorhersage unter 70
- **Leukemia Inhibitory Factor (LIF)**  
Kann Tumorwachstum fördern oder hemmen

## AlphaFold-Datenbank

# Beispiel von AlphaFold

Quelle: Google DeepMind, AlphaFold

# Warum ist AlphaFold revolutionär?

- **Schneller** – Proteinfaltung in **Stunden statt Jahren**

# Warum ist AlphaFold revolutionär?

- **Schneller** – Proteinfaltung in **Stunden statt Jahren**
- **Günstiger** – Kosteneffizient durch KI-gestützte Simulationen



# Warum ist AlphaFold revolutionär?

- **Schneller** – Proteinfaltung in **Stunden statt Jahren**
- **Günstiger** – Kosteneffizient durch KI-gestützte Simulationen
- **Präzise** – Hohe Genauigkeit, vergleichbar mit Labormethoden

# Zukunft von KI in der Proteinforschung

- **Präzisere Vorhersagen** – Optimierte Deep-Learning-Modelle

# Zukunft von KI in der Proteinforschung

- **Präzisere Vorhersagen** – Optimierte Deep-Learning-Modelle
- **KI in der Medikamentenentwicklung** – Schnellere und personalisierte Wirkstoffanalyse

# Zukunft von KI in der Proteinforschung

- **Präzisere Vorhersagen** – Optimierte Deep-Learning-Modelle
- **KI in der Medikamentenentwicklung** – Schnellere und personalisierte Wirkstoffanalyse
- **Hochleistungsrechnen und Cloud** – Skalierbare KI-Modelle für komplexe Berechnungen von Proteinen

# Anwendungen: Medikamentenentwicklung

- **Ziel:** Wirkstoff finden, der an ein Zielprotein bindet
- **Problem:** Traditionelle Methoden sind zeitaufwendig und teuer
- **Virtuelles Screening:**
  - Computergestützte Methode zur Analyse chemischer Verbindungen
  - Filtert Millionen von Molekülen, um die besten Kandidaten zu finden
  - **Vorteil:** Spart Zeit und Geld durch Reduzierung aufwendiger Laborexperimente

# Virtuelles Screening

## 1. Zielproteinanalyse:

- Analyse der Struktur eines Zielproteins, z.B.: mit **AlphaFold**
- **Vorteil:** Präzise Vorhersage der 3D-Struktur eines Zielproteins
- Entscheidende Grundlage für das virtuelle Screening und die Wirkstoffentwicklung

## 2. Auswahl der Moleküle:

- Moleküle werden aus großen Datenbanken wie ZINC und PubChem ausgewählt
- Filterung geeigneter Moleküle anhand chemischer und physikalischer Eigenschaften

# Virtuelles Screening

## 3. Docking-Simulation:

- Simulation der Bindung zwischen einem Molekül und dem Zielprotein
- Es werden verschiedene Konformationen und Positionen getestet

## 4. Scoring:

- Bewertung der Bindungsstärke
- Je höher der Bindungswert desto höher die Wahrscheinlichkeit, dass das Molekül bindet

# Anwendungen: Krankheitsforschung

- Fehlfaltungen von Proteinen führen zu neurodegenerativen Erkrankungen wie:
  - **Alzheimer:** Bildung toxischer Aggregate in Nervenzellen
  - **Parkinson:** Ansammlung fehlgefalteter Proteine
- Verständnis der Faltung und Fehlfaltung essenziell für die Ursachenforschung

## AlphaFold:

- Verwendung zur Vorhersage der Proteinstruktur (3D-Modell)
- Die "vorhergesagte" Proteinstruktur wird mit der Fehlgefalteten verglichen



# Anwendungen: Design von Proteinen

- Neue Proteine mit spezifischen Eigenschaften entwerfen
- Können eingesetzt werden in folgende Bereiche:
  - **Medizin**
  - **Industrie**
  - **Umweltschutz**

## Rolle von KI:

- Nutzung von KI-Algorithmen wie neuronale Netzwerke
- **Vorteil:** Design von Proteinen wird schneller und präziser

**Vielen Dank für Ihre  
Aufmerksamkeit!**