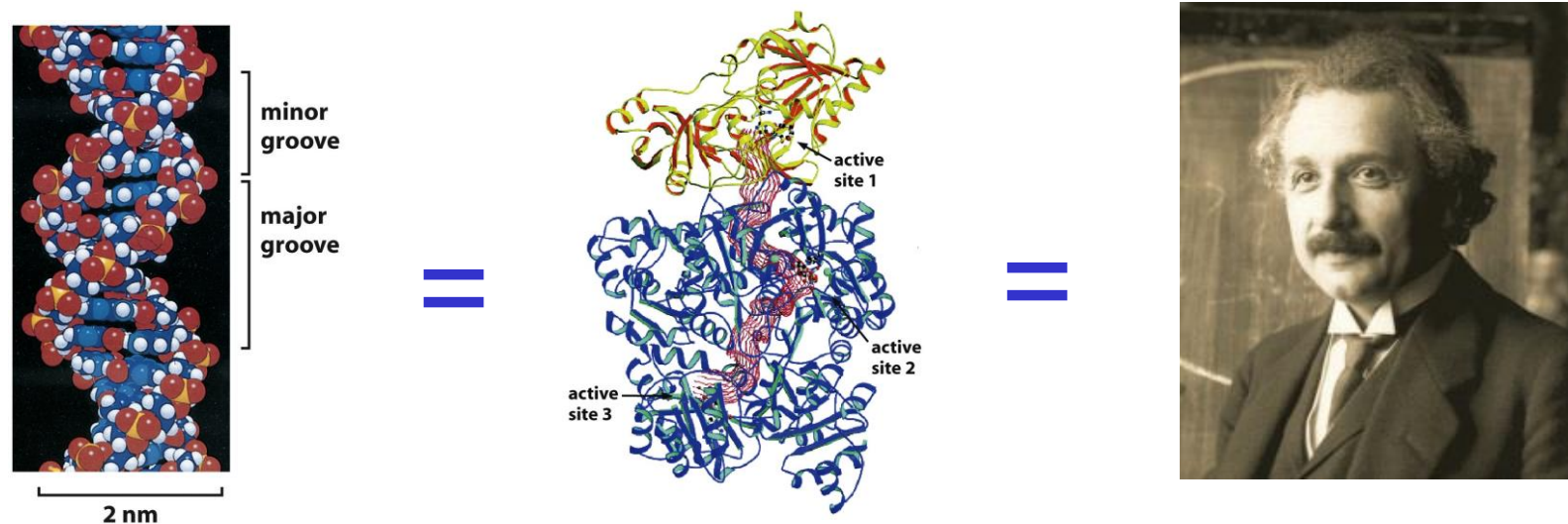


AlphaFold2: Von der DNA-Sequenz zur Proteinstruktur, ist der Mensch berechenbar?



Dozent: Prof. Dr. Werner Schiebler

WiTechWi 8.Juni 2022

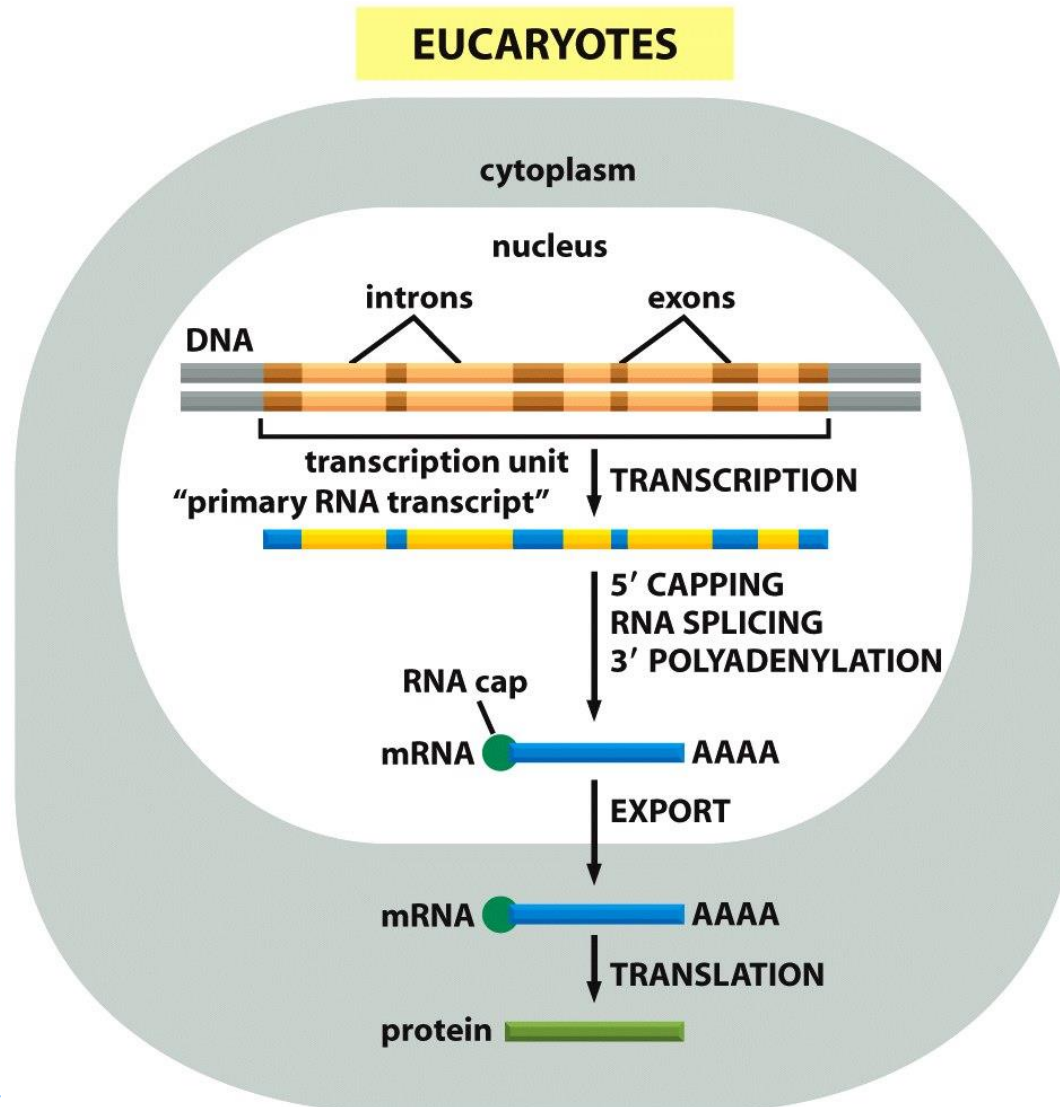
Was erwartet Sie?

- 1) Informationsfluss: DNA to RNA to Protein
- 2) Aminosäuresequenz bestimmt 3D Struktur eines Proteins
- 3) Proteinstruktur: Bedeutung, Bestimmung, Probleme
- 4) Das Proteinfaltungsproblem
- 5) Lösung: AlphaFold2
- 6) Anwendungsbeispiele
- 7) Mutationen und Berechnung des Menschen

Was erwartet Sie?

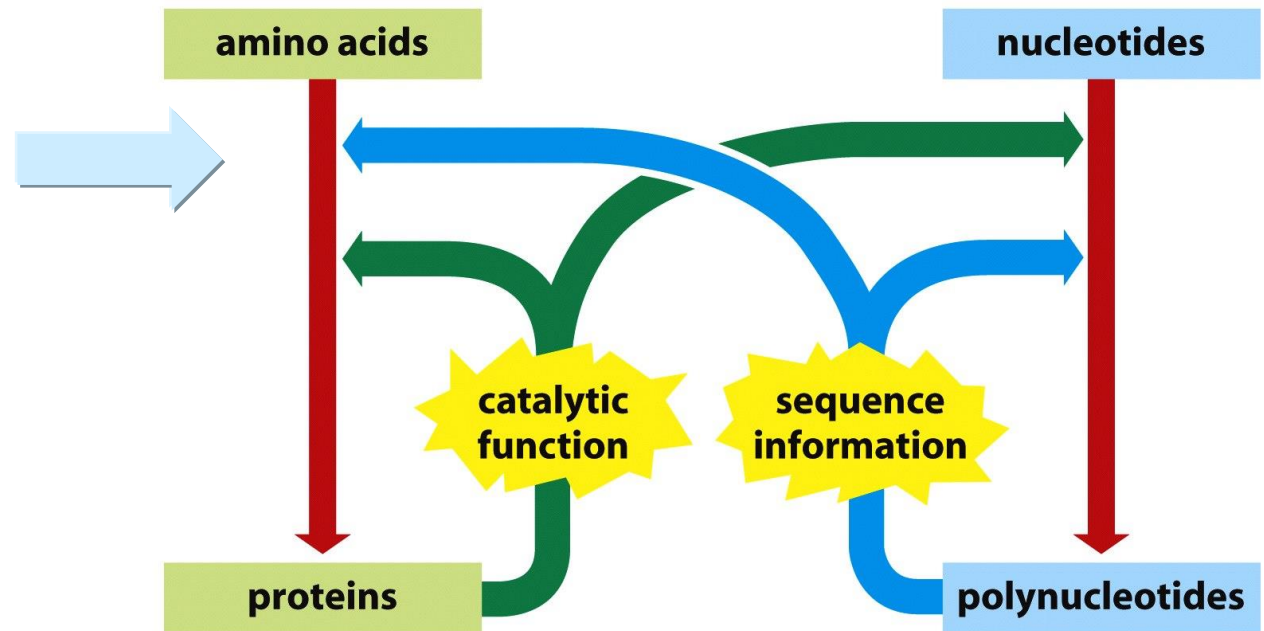
- 1) Tanz der Moleküle: Translation
- 2) Das Universum ist nicht alt genug
- 3) Wie strukturiert man sich?
- 4) „Deep Learning“ und neuronale Netze
- 5) Neue Katalysatoren für jedermann
- 6) Wie kann ich einen Krebs heilen?
- 7) Wieso gibt es die USA?

1) Informationsfluss: DNA zu RNA zu Protein



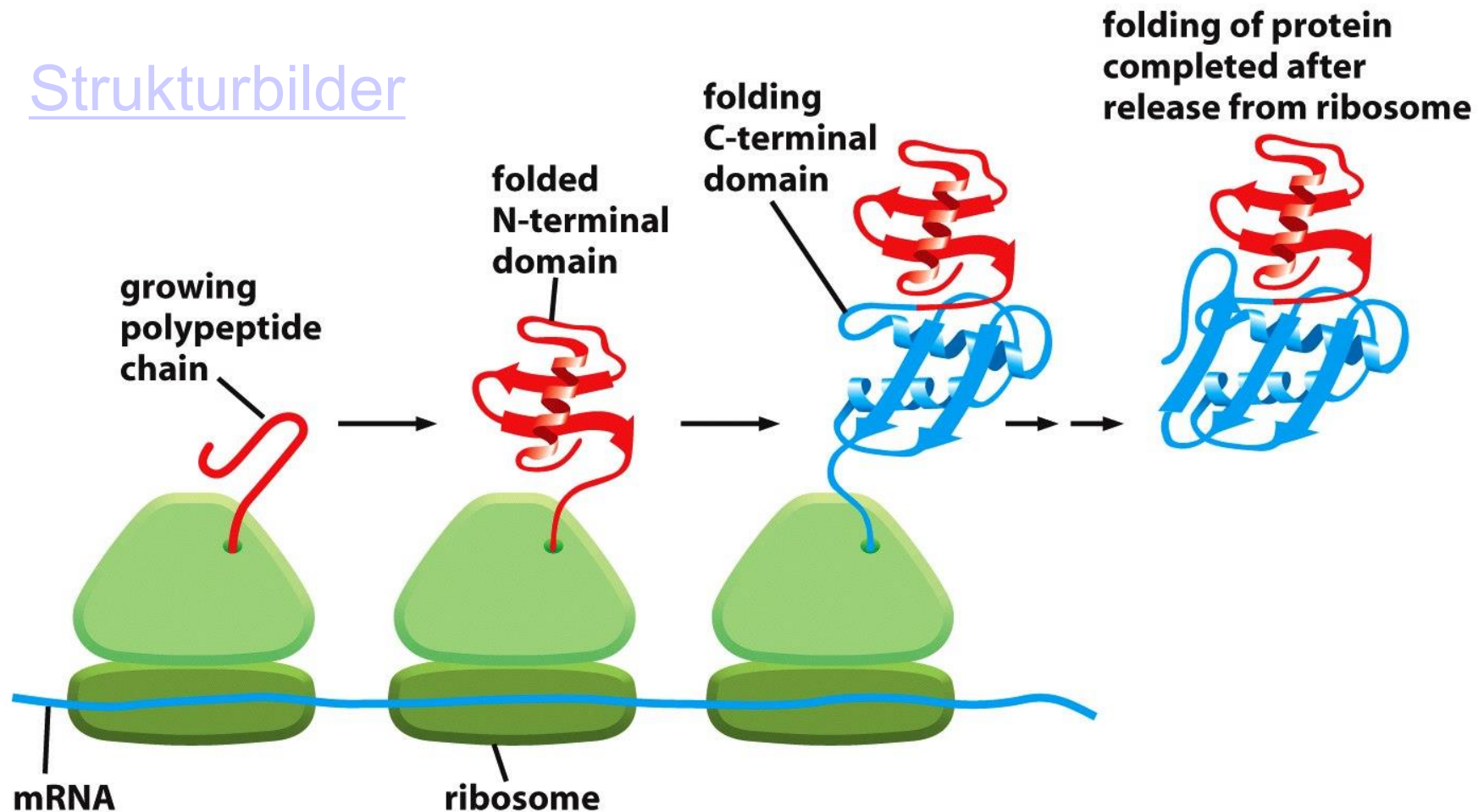
1) Informationsfluss: DNA zu RNA zu Protein

Translation



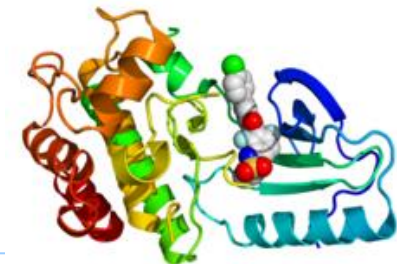
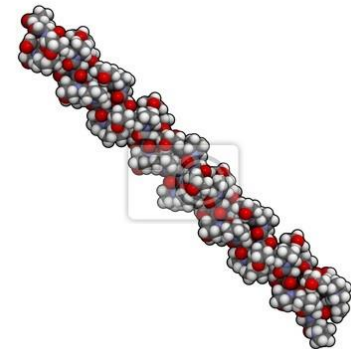
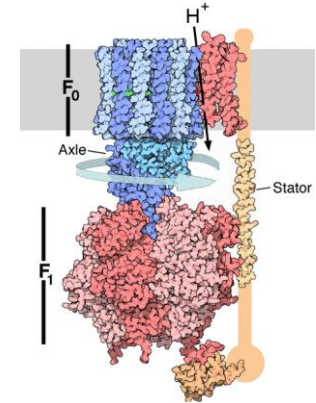
2) Aminosäuresequenz bestimmt Proteinstruktur

Strukturbilder



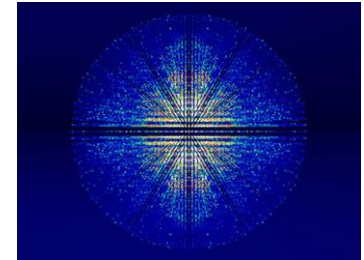
3) Proteinstruktur: Bedeutung

- 1) Die 3D-Struktur eines Proteins erklärt dessen biologische Funktion
- 2) Die 3D-Struktur ermöglicht die Vorhersage von Wechselwirkungen mit anderen Proteinen (Netzwerk, Makrostrukturen)
- 3) Die 3D-Struktur ermöglicht die gezielte chemische Synthese von therapeutisch wirksamen Molekülen

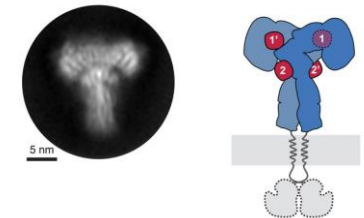


3) Proteinstruktur: Bestimmung

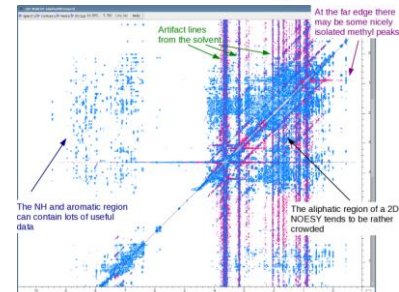
1) Röntgenstrahl-Beugung an Proteinkristallen



2) Kryoelektronenmikroskopie



3) 2D NMR an gelösten Proteinen

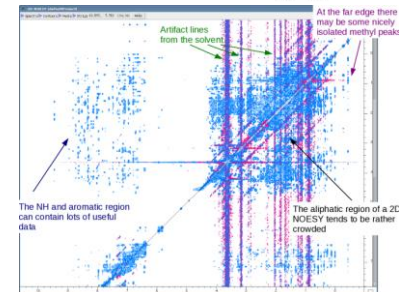
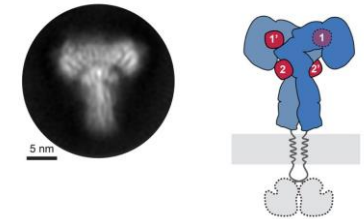
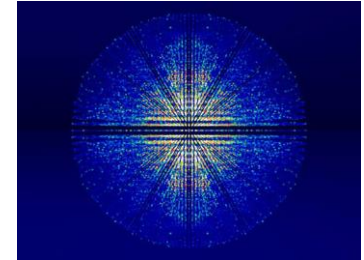


4) In silico Berechnung (Computer)



3) Proteinstruktur: Probleme

- 1) Röntgenstrahl-Beugung an Proteinkristallen
 - a) nicht alle Proteine kristallisieren
 - b) Kristallisationsvorgang komplex/langwierig
- 2) Kryoelektronenmikroskopie
 - a) instrumentell aufwändig
 - b) minimale Proteingröße 30 kDa
- 2) 2D NMR an gelösten Proteinen
 - a) funktioniert nur mit kleinen Proteinen (max 50 kDa)
 - b) erfordert häufig Isotopen-Synthese
- 2) In silico Berechnung (Computer)
 - a) bis vor kurzem nur zu ca. 50% erfolgreich



4) Das Proteinfaltungsproblem

Levinthals Paradox:

Zahl möglicher Faltungen eines
mittelgroßen Proteins (ca. 1000 Aminosäuren,
2 Faltungszustände pro Aminosäure): $2^n = >$
 10^{300}



Cyrus Levinthal
1922-1990

Proteinfaltung erfolgt nach ca. 1 msec

Findung richtiger Faltung nach „Trial und
Error“ ca. 5×10^{296} sec

Universum existiert „erst“ seit ca. 4×10^{16} sec!

4) Das Proteinfaltungsproblem

Thermodynamik der Proteinfaltung:

Faltung einer Kette von 100 Aminosäuren in H_2O :
Energiegewinn ca. 48 kJ/mol (ΔH) bzw.

0,48 kJ pro mol und Peptidbindung

Durchschnittliche thermische Energie bei 37 Grad Celsius (durchschn. kinetische Energie H_2O -Moleküle):

2,5 kJ pro mol und Peptidbindung!

Proteinfaltung: Ordnung am Rande des Chaos

4) Das Proteinfaltungsproblem

Warum faltet sich überhaupt ein Protein in kurzer Zeit?

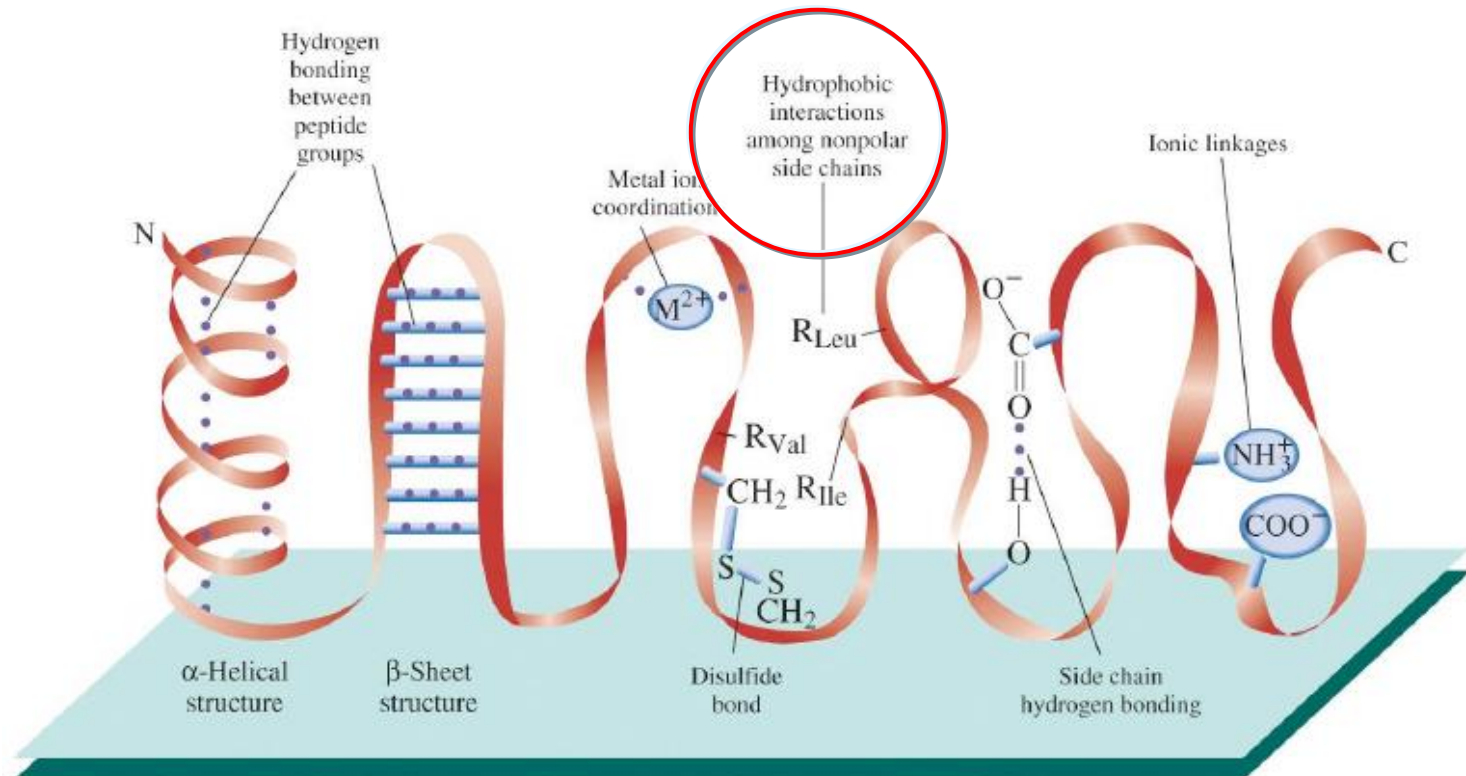
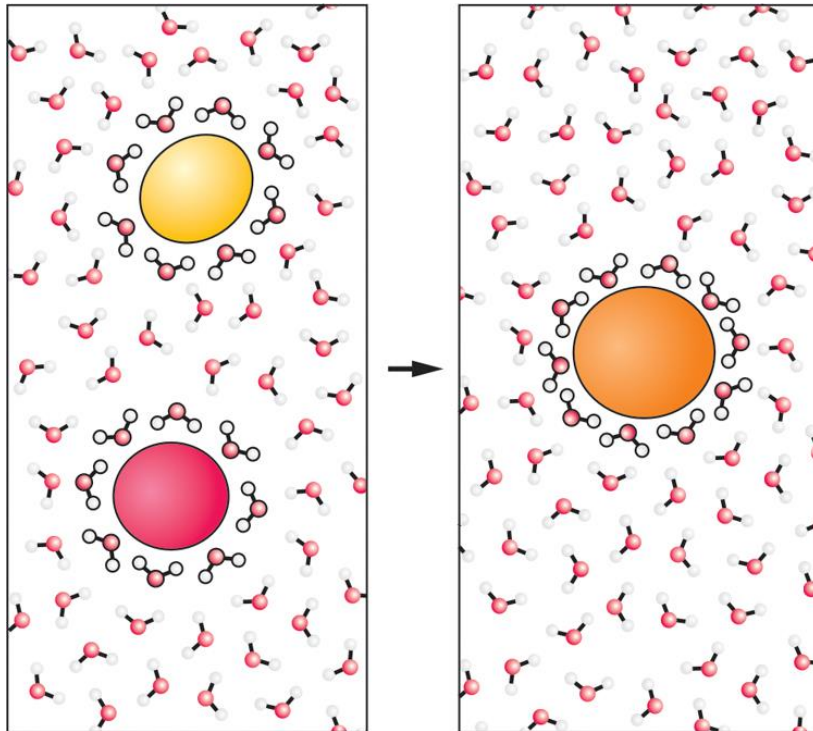


Figure 4-3 Concepts in Biochemistry, 3/e
© 2006 John Wiley & Sons

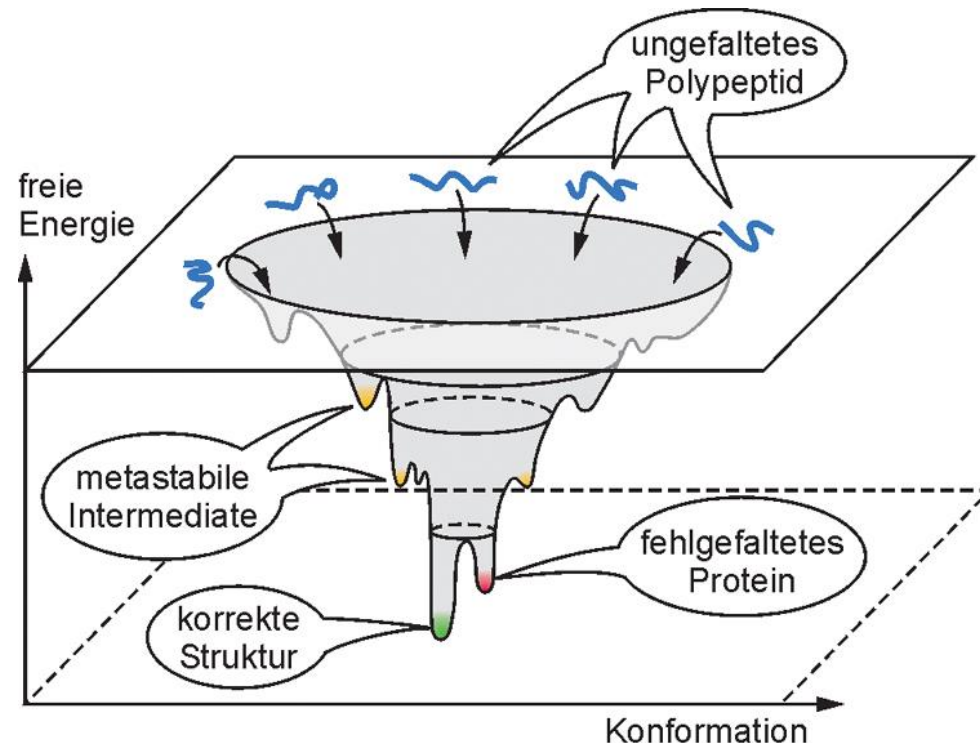
Es gibt zwischenmolekulare Kräfte in Subdomänen!

4) Das Proteinfaltungsproblem

Freigesetzte Enthalpie durch intramolekulare Kräfte und erhöhte Entropie von H_2O Molekülen treiben die Faltung voran (Hydrophober Effekt in Wasser!)



Aus Müller-Esterl, *Biochemie*, © 2004 Elsevier GmbH



Aus Müller-Esterl, *Biochemie*, © 2004 Elsevier GmbH

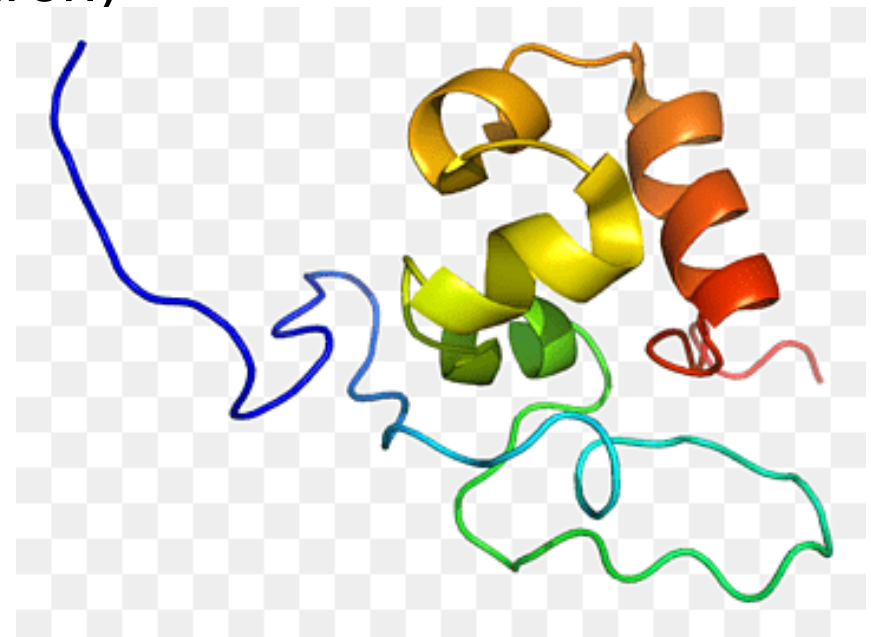
4) Das Proteinfaltungsproblem

Wie macht das die Natur?

Villin = Aktinbindungsprotein aus Darmepithelzellen

„Headpiece“ (70 Aminosäuren)

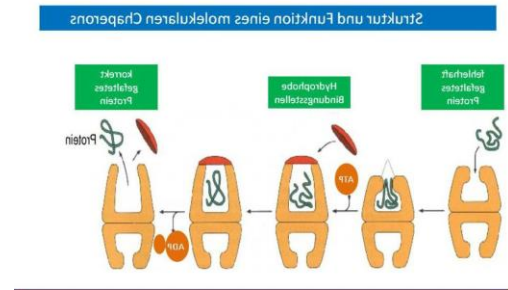
Faltung von Villin



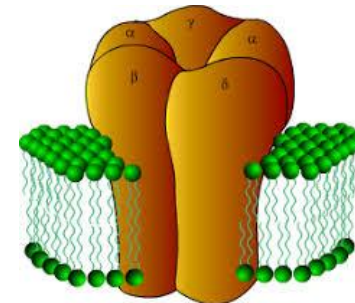
4) Das Proteinfaltungsproblem

Auch lebende Zellen haben Probleme mit der Proteinfaltung:

Die Mehrzahl der Proteine braucht Faltungshelfer (Chaperonine)



Der nikotinische Acetylcholin-Rezeptor (Pentamer $\alpha_2\beta\gamma\delta$):
Von 10 Syntheseversuchen wird durchschnittlich nur
1 Rezeptormolekül als korrekt gefaltet bewertet und in die
Muskelzellmembran eingebaut !



5) Lösung Alphafold2

CASP „Critical Assessment of Protein Structure Prediction“

Wettbewerb, alle 2 Jahre durchgeführt seit 1994

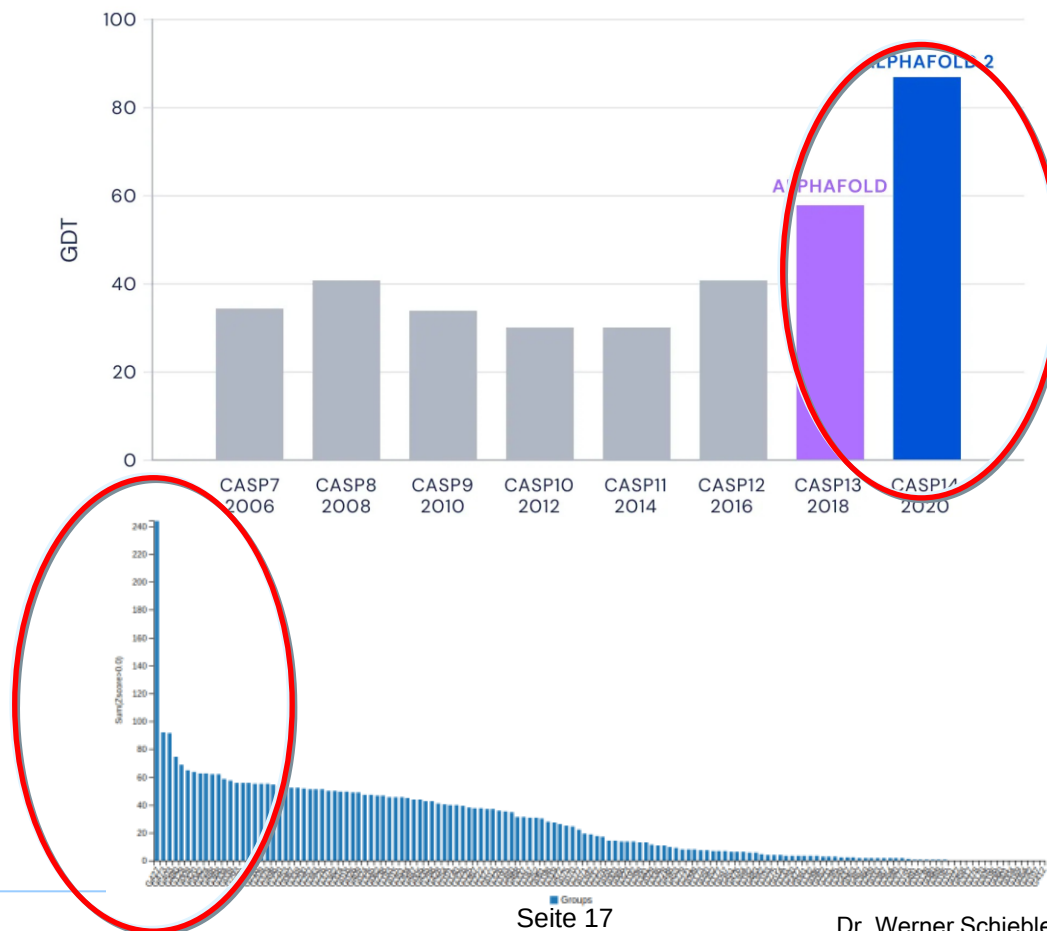
- a) Experimentell bestimmte Proteinstrukturen werden vor Publikation gesammelt.
- b) Die Aminosäure-Sequenzen werden interessierten Gruppen zur Berechnung einer 3D Struktur zur Verfügung gestellt.
- a) Errechnete und experimentell bestimmte Strukturen werden verglichen

Bewertung: GTD „Global Distance Test“ 0 bis 100 %
> 90 = Struktur entspricht Experiment

5) Lösung AlphaFold2

CASP „Critical Assessment of Protein Structure Prediction“

Median Free-Modelling Accuracy

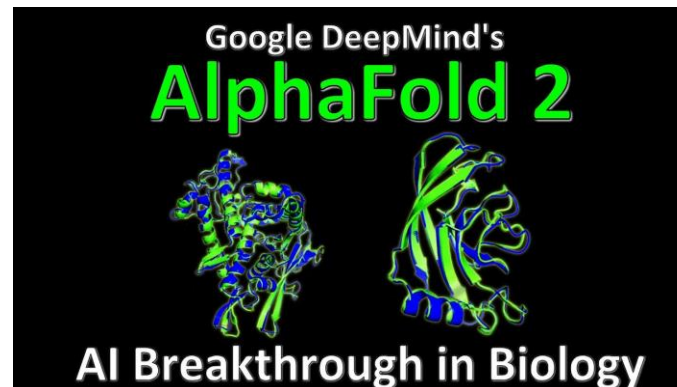


5) Lösung AlphaFold2

Google/DeepMind's AlphaFold2 Programm erreichte bei CASP14 einen Durchschnitts GDT von über 92 und war dreimal so erfolgreich wie das zweitbeste Programm.

Jeder kann jetzt Proteine falten
Turbo für die Biologie
Schneller zu neuen Medikamenten
Das Programm, das Omikron sichtbar macht
KI löst jahrzehntealtes Wissenschaftsproblem

Spektrum d. Wiss.
Süddeutsche Zeitung
Neue Züricher Zeitung
Die Welt
Finanznachrichten.de



5) Lösung AlphaFold2

Google/DeepMind`s AlphaFold2 Programm:
Was macht den Unterschied aus?

Deep Learning/Neuronale Netzwerkstruktur
Training an ca. 200000 3D Proteinstrukturen
Nutzung der größten Interspezies Datenbanken

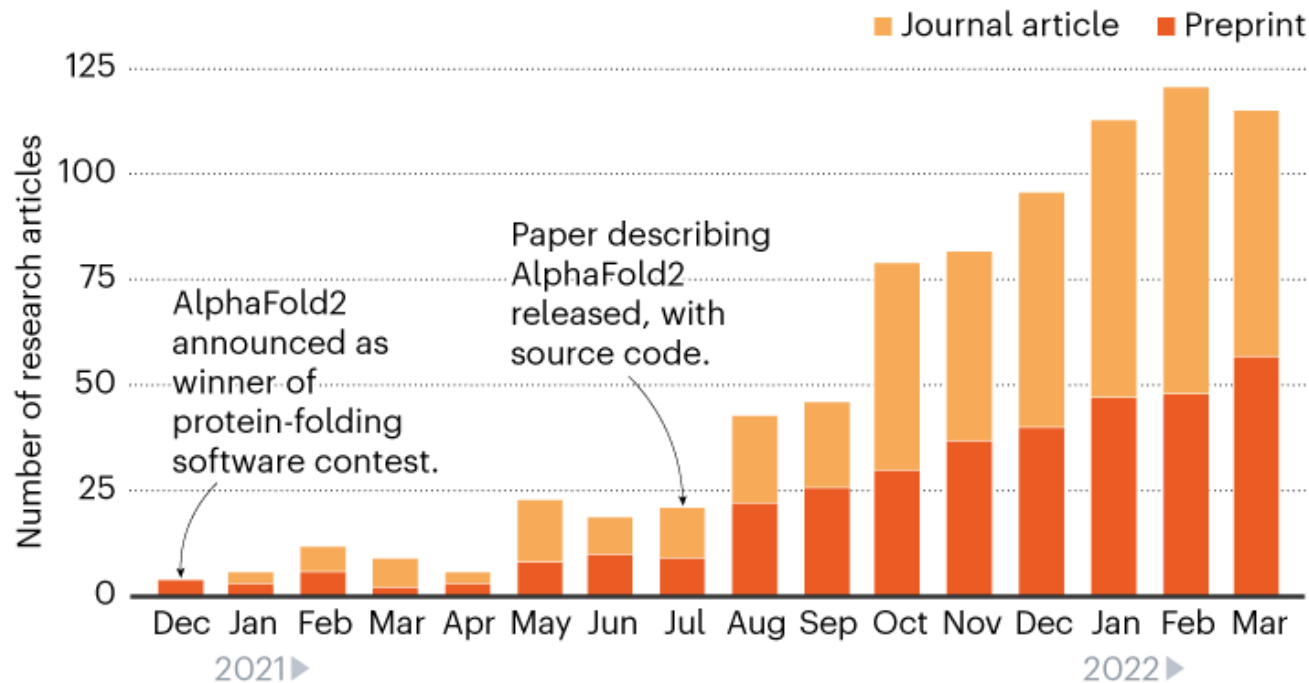
Zusammenführung von

- a) Sequenz Verwandtschaft (MSA)
- b) 3D Co-Evolution von Aminosäurepositionen
in Proteinen verschiedener Spezies
- c) Einführung einer „Attention“-Funktion
(Identifikation von „korrekten“ Teilstrukturen)

5) Lösung AlphaFold2

ALPHAFOLD MANIA

The number of research papers and preprints citing the AlphaFold2 AI software has shot up since its source code was released in July 2021*.



*Nature analysis using Dimensions database; removing duplicate preprints and papers/R. Van Noorden, E. Callaway.

©nature

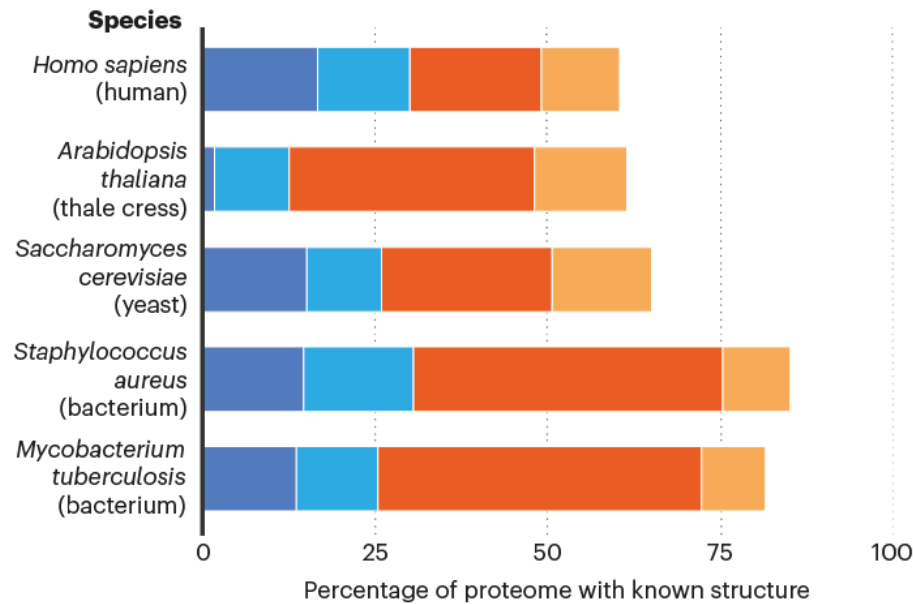
5) Lösung Alphafold2

WHAT'S KNOWN ABOUT PROTEOMES

AlphaFold's predictions have greatly increased the proportion of confidently known structures in the human proteome — the collection of all human proteins. The software is even more useful for other species.

Source of knowledge about proteome

- High-quality experimental structures in the PDB*
- Structural knowledge derived from related proteins in the PDB*
- Knowledge from AlphaFold models only (high confidence)
- Knowledge from AlphaFold models only (intermediate confidence)



*PDB: Protein Data Bank. AlphaFold can also be used to calculate these structures — but doesn't add significantly to what's already known.

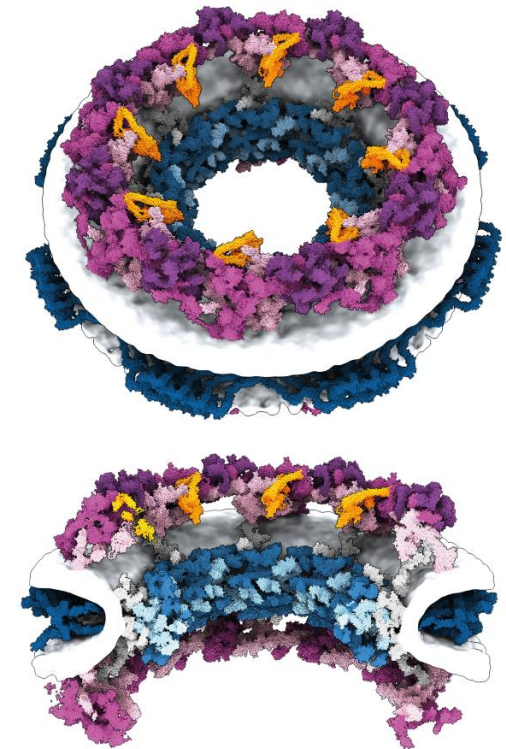
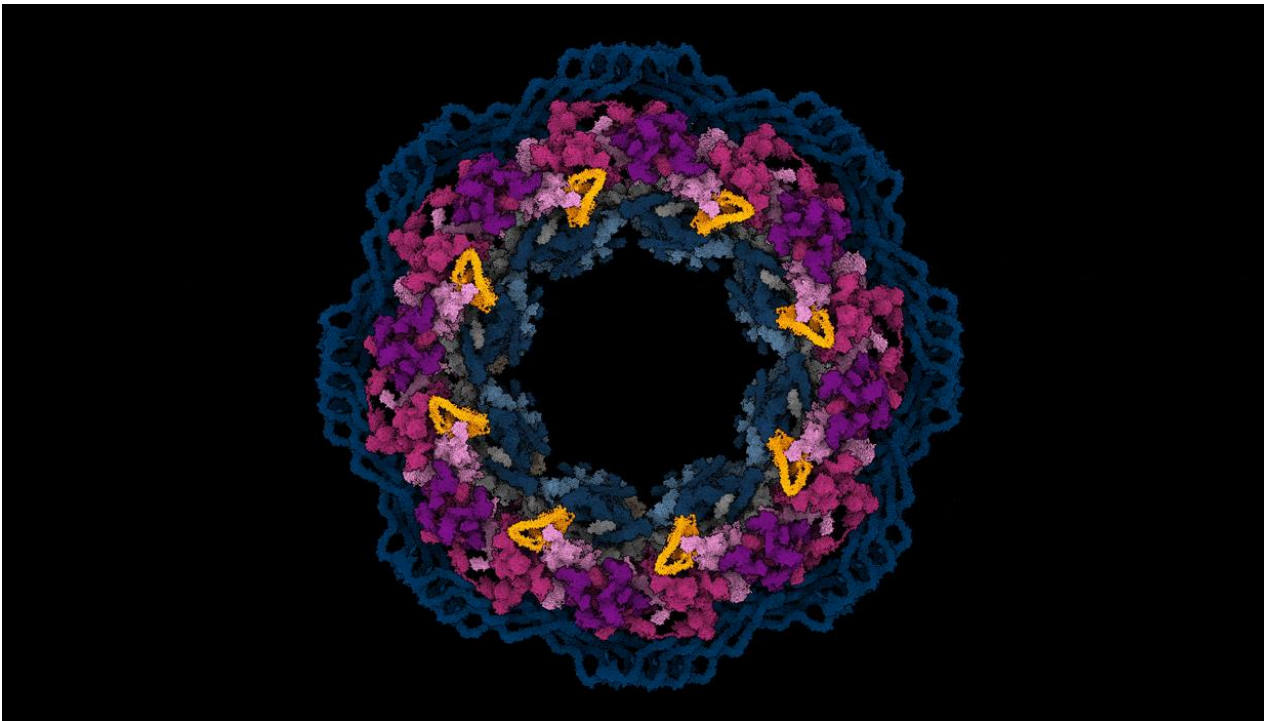
©nature

6) Anwendungsbeispiele

AlphaFold2 Modell des Kernmembran-Porenkomplex (60% aller Proteine berücksichtigt)

GENOME GATEWAY

Two views of the human nuclear pore complex show how this massive assembly embeds into the nuclear membrane (white).



©nature

6) Anwendungsbeispiele

DREAMING UP PROTEINS

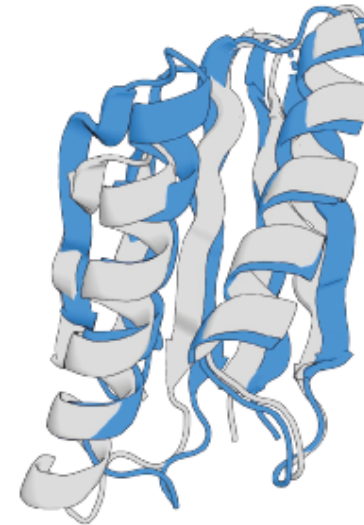
Researchers used deep neural networks to invent, or 'hallucinate', sequences of amino acids that could fold into proteins; in some cases they have synthesized these proteins to compare their actual structures with predictions.



'Hallucinated'
protein (software
prediction)



Actual structure
(experimentally
determined)



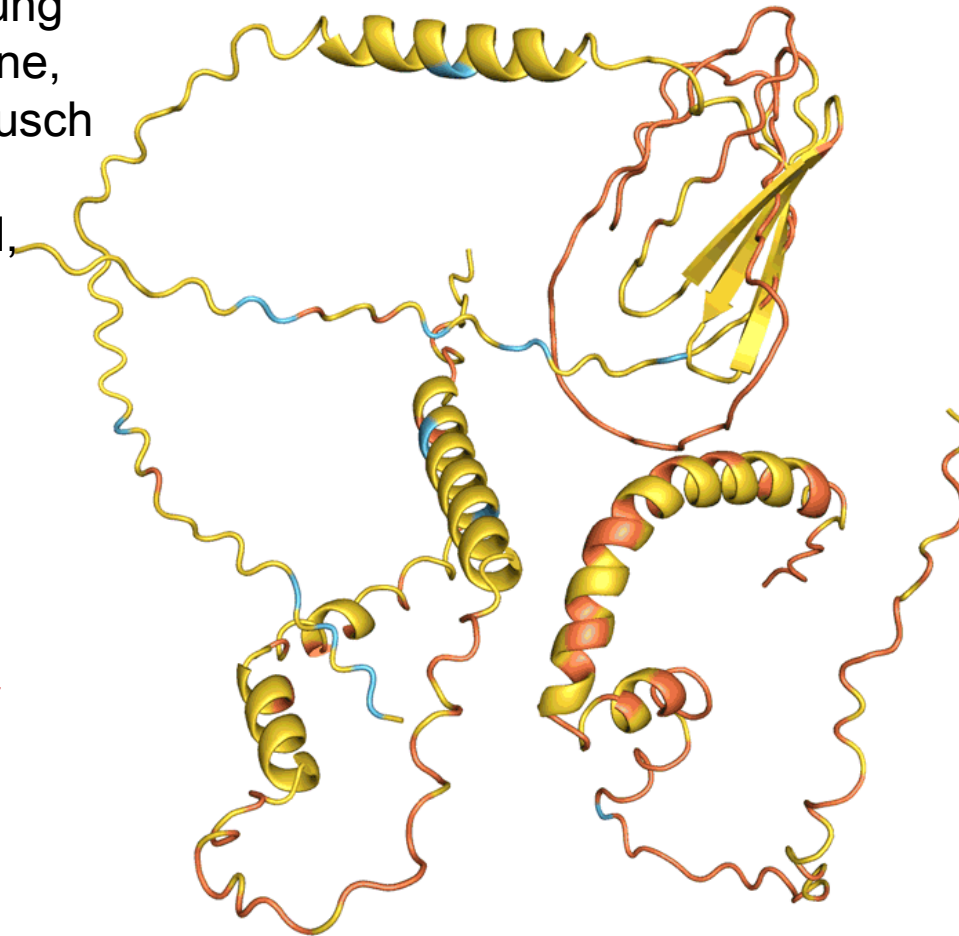
Superposition of
hallucinated (blue) and
actual (grey) structures

©nature

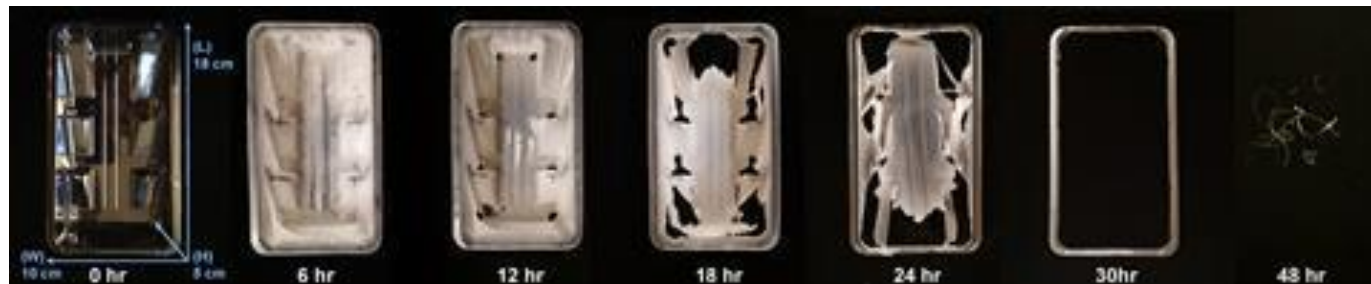
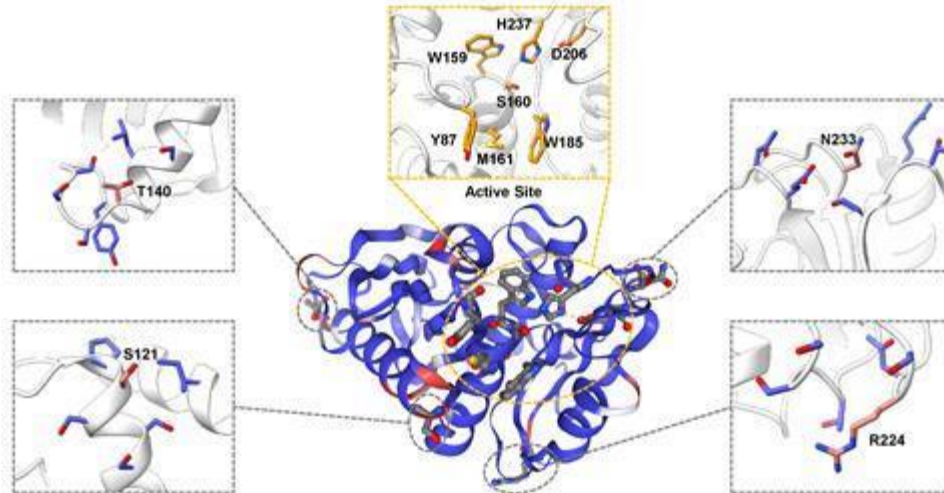
6) Anwendungsbeispiele

Faltungsberechnung künstlicher Proteine, bei der der Austausch von Aminosäuren mitberechnet wird, um stabile End-Strukturen zu erreichen

Ein Traum:
In Silico Design
von beliebigen
Katalysatoren zur
stereochemisch
kontrollierten
Synthese

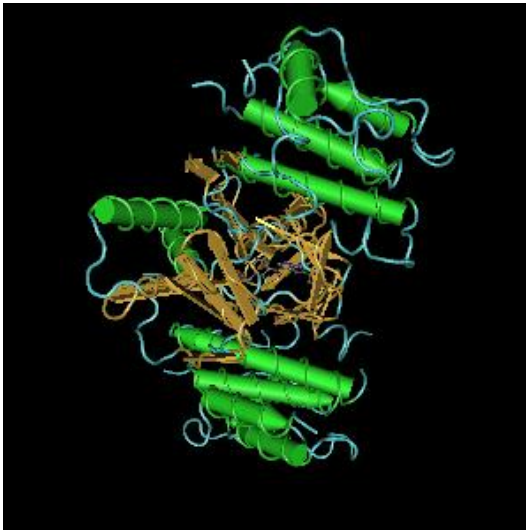


6) Anwendungsbeispiel Plastikabfall PETase



6) Anwendungsbeispiel Malignes Melanom

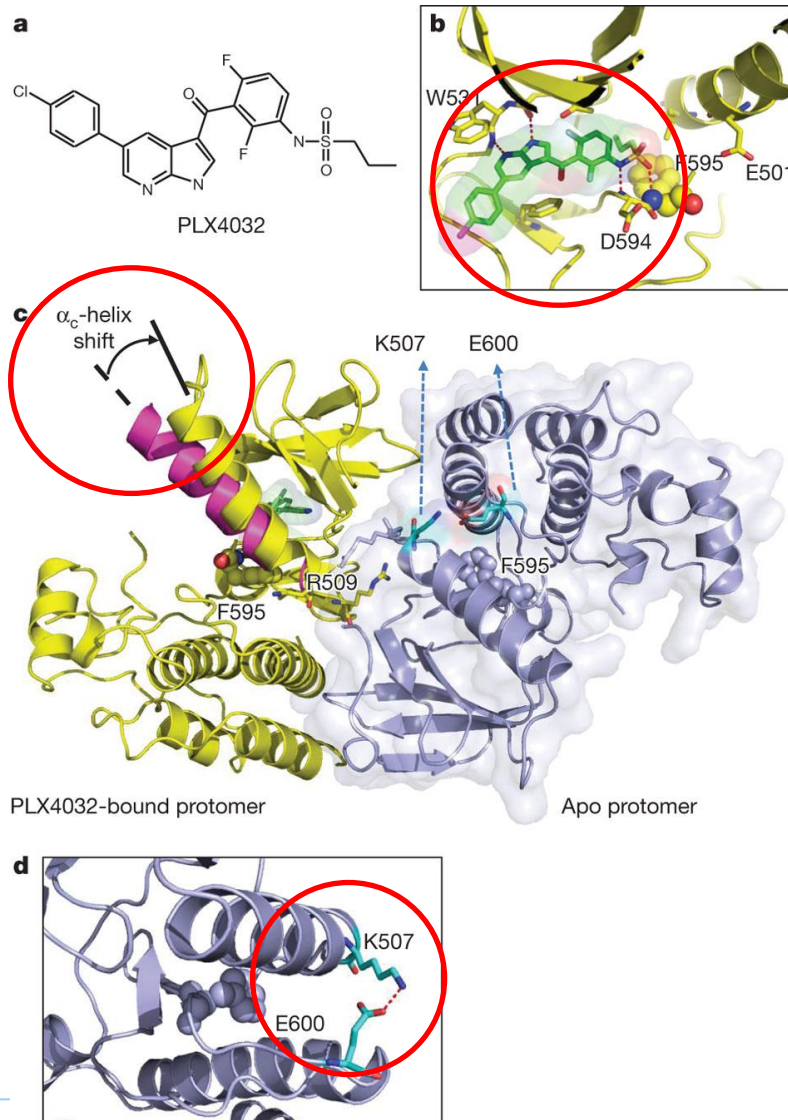
B-RAF wild type dimer



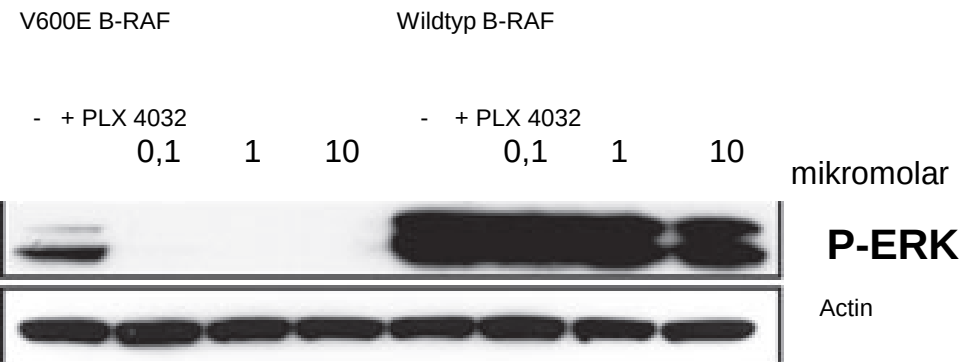
In ca. 40% aller malignen Melanome ist die Aminosäure **600 Valin (V)** der B-RAF Proteinkinase zu **Glutaminsäure (E)** mutiert (**V600E B-RAF**). Dies ist eine erworbene (somatische) Mutation, zB. durch UV-Licht. Die Mutation verhindert eine Valin-Phenylalanin Interaktion und führt dadurch zu einer Konformationsänderung, die die B-RAF Proteinkinase permanent aktiviert.



6) Anwendungsbeispiel Malignes Melanom



PLX 4032 Aktivität in Melanoma cell lines mit aktiver B-RAF Signaltransduktion:



Immunoblot mit P-ERK spezifischen Antikörpern.

6) Anwendungsbeispiel Malignes Melanom



Die BRIM 3 Studie wurde vorzeitig abgebrochen um die Chemotherapie-Vergleichspatienten rechtzeitig mit Vemurafenib behandeln zu können.

7) Mutationen und Berechnung des Menschen

<https://alphafold.ebi.ac.uk/>

7) Mutationen und Berechnung des Menschen

<https://alphafold.ebi.ac.uk/>

7) Mutationen und Berechnung des Menschen



7) Mutationen und Berechnung des Menschen

George III, König von
Grossbritannien und Irland 1760-1801,
Regent während des amerikanischen
Unabhängigkeitskriegs



Litt seit den 1760iger Jahren sporadisch und später permanent an einer Stoffwechselkrankheit, mit hoher Wahrscheinlichkeit an PORPHYRIE.

Bekannt für kompromisslose Haltung zur Behandlung der englischen Kolonie in Amerika:

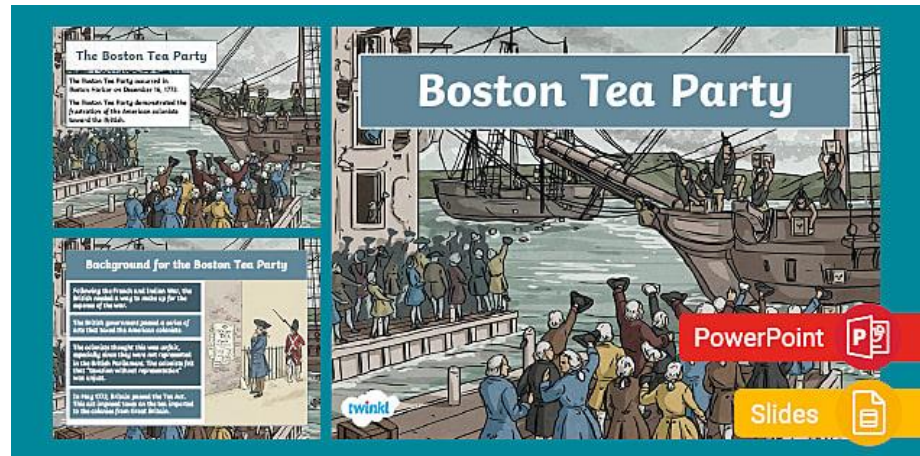
Weigert sich, die Teesteuer abzuschaffen
Wollte laut Historiker George Trevelyan
„die Rebellen drangsalieren, verängstigen
und verarmen lassen und nie die
Unabhängigkeit anerkennen“

7) Mutationen und Berechnung des Menschen

Folgen der Kompromisslosigkeit:



Boston Tea Party 16. Dez. 1773



Amerikanische Unabhängigkeits-
erklärung 4. Juli 1776

7) Mutationen und Berechnung des Menschen

Symptome und Ursachen der Porphyrrie:



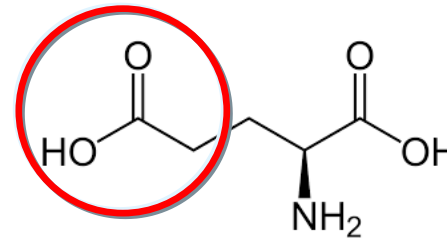
Kolikartige Darmkrämpfe
Psychiatrische Probleme,
Stimmungsschwankungen

Porphyrien sind eine Klasse von
Stoffwechselerkrankungen der
Hämoglobin-Biosynthese

Häufige Mutation:
Enzym UroD G281E
Position 281 Glycin zu Glutamat

7) Mutationen und Berechnung des Menschen

Die USA sind entstanden, weil
zum falschen Zeitpunkt an der
falschen Stelle eine
Carboxylgruppe bzw. mit Glutamat
ein Geschmacksverstärker stand



?

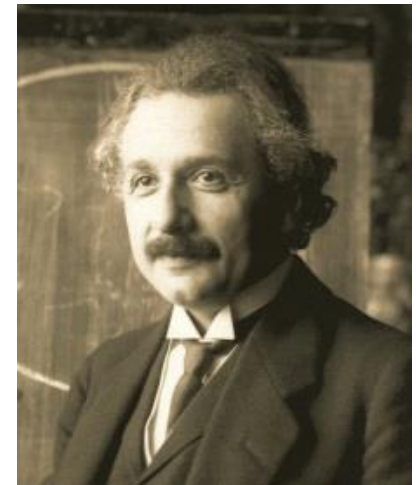
7) Mutationen und Berechnung des Menschen

Sind der Mensch und sein Verhalten
nun in Zukunft berechenbar?

Sind zukünftig alle Medikamente im
Computer planbar?

Ist komplett synthetisch-biologisches
Leben irgendwann im Computer
planbar?

Ich freue mich auf die Diskussion



Gott würfelt nicht!?

Quellen

[Molecular Biology of the Cell Alberts etal 5th Edition](#)

[Röntgenstruktur-Analyse: Proteine und Viren im Blitzgewitter- Physik & Mehr –FAZ](#)

[Forscher entdecken eine neue Weise, wie Insulin mit seinem Rezeptor interagiert—
Paul Langerhans Institut Dresden — TU Dresden \(tu-dresden.de\)](#)

[H-H NOESY | Protein NMR \(proteinmr.org.uk\)](#)

[20.000+ Computer Bilder und Fotos · Kostenlos Downloaden · Pexels Stock-Fotos](#)

[What's next for AlphaFold and the AI protein-folding revolution \(nature.com\)](#)

[VillinProtein Data Bank Actin-binding protein ABLIM3, others, text, biology, human
png | PNGWing](#)

Chemgapedia

source=Nature+ Briefing&utm_campaign=25384cfc13-briefing-dy-
20220510&utm_medium= email&utm_term=0_c9dfd39373-25384cfc13-43622797

Deepmind.com

[Home - CASP14 \(predictioncenter.orgBoston Tea Party Presentation- PowerPoint &
Google Slides \(twinkl.de\)\)](#)