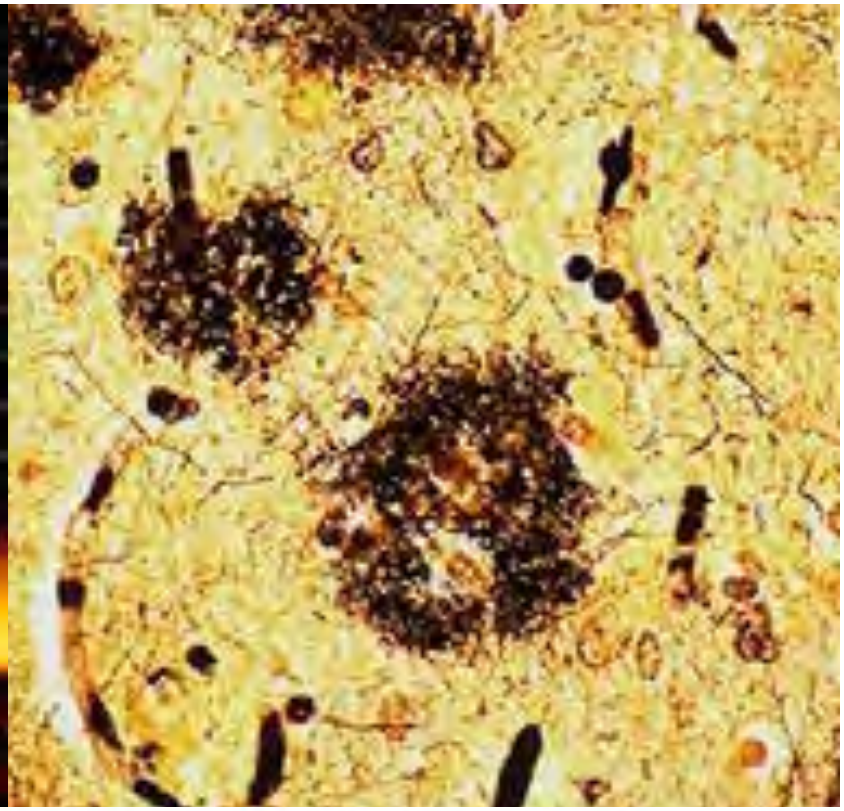
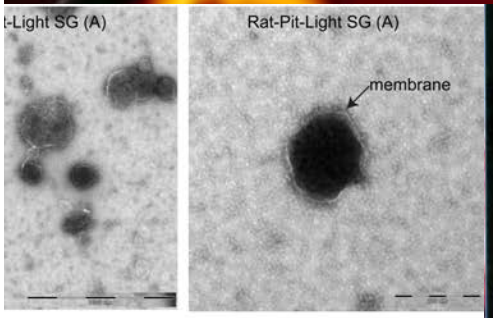
Grundlagenforschung: Vom Ursprung des Lebens aus der Proteinperspektive

Roland Riek
ETH Zurich

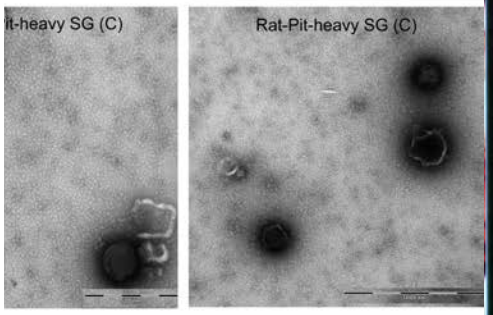
Biological NMR Group

- AI (Dr. Piotr Klukoswki, Dr. Peter Güntert, Henry Whetton, Luca Wenchel)
- NMR (Dr. Felix Torres, Matthias Bütikofer, Gabriela Stadler, Olivia Gampp)
- Parkinson's disease (Dr. Lukas Frey, Dr. Dhiman Ghosh, Dr Juan Gerez, David Rhyner)
- Origin of Life (Dr. Saroj Rout, Dr. Jason Greenwald, Dr. Witek Kwiatkowski)



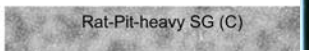

-Light SG (A)

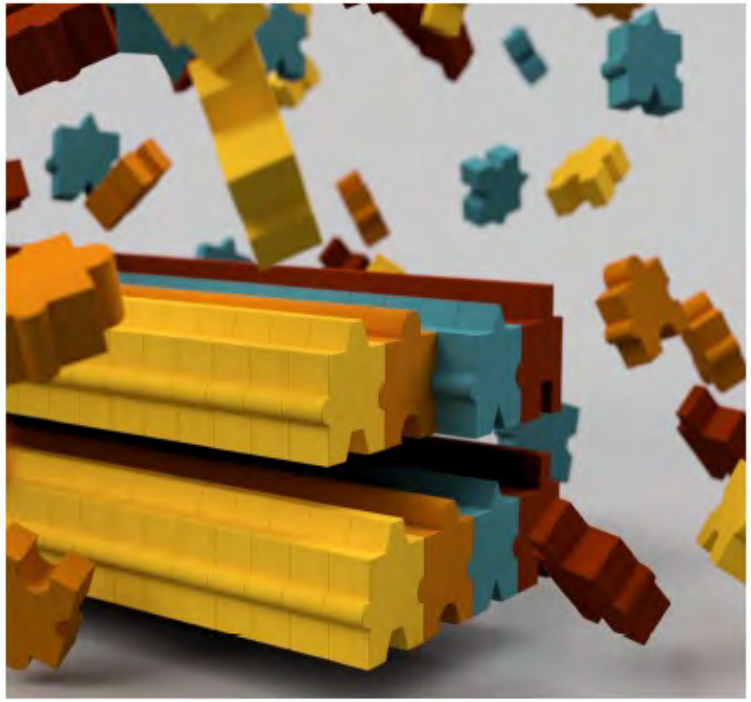
Rat-Pit-Light SG (A)
membrane


t-heavy SG (C)

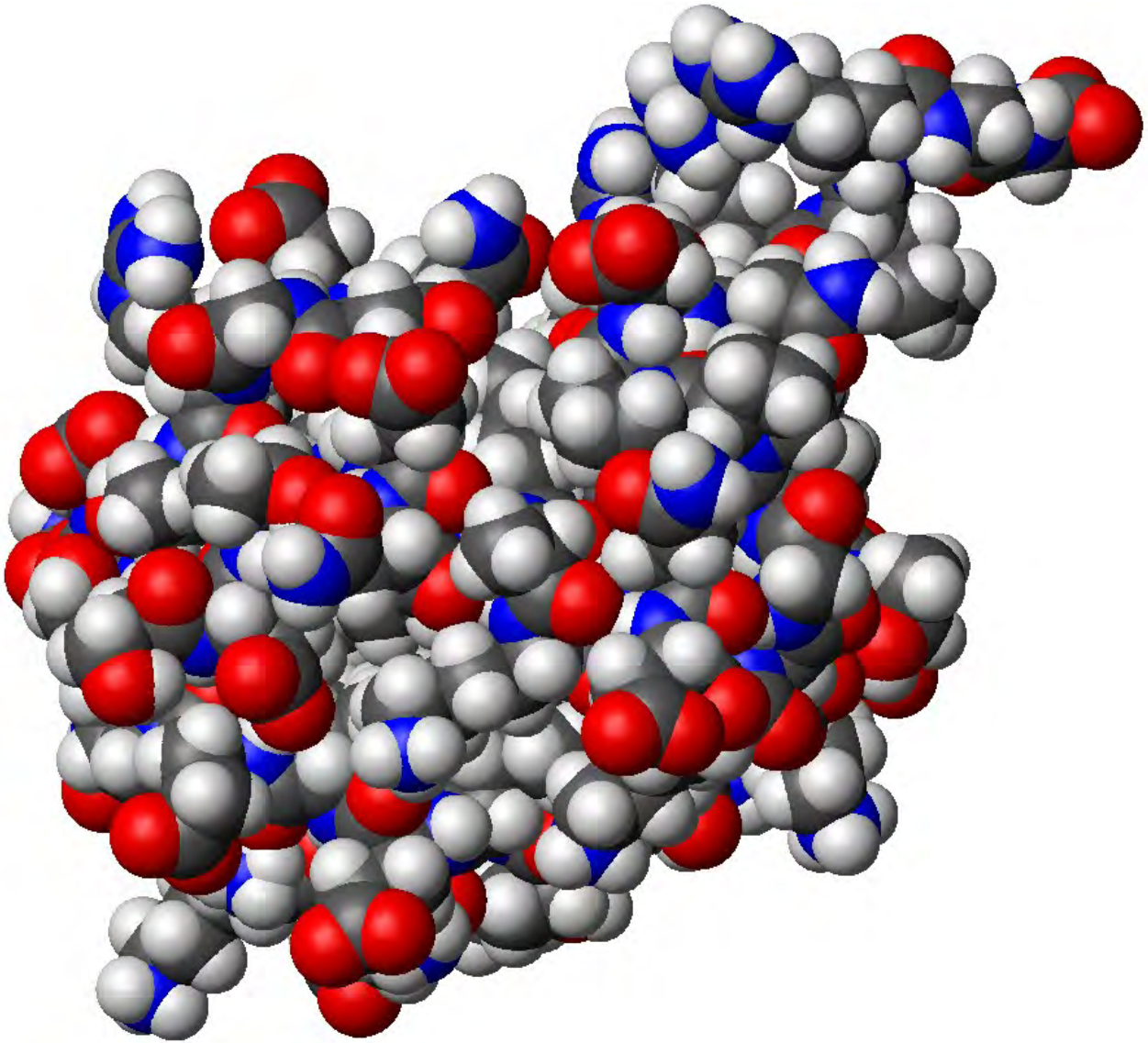
Rat-Pit-heavy SG (C)


-Pit-heavy SG (C)

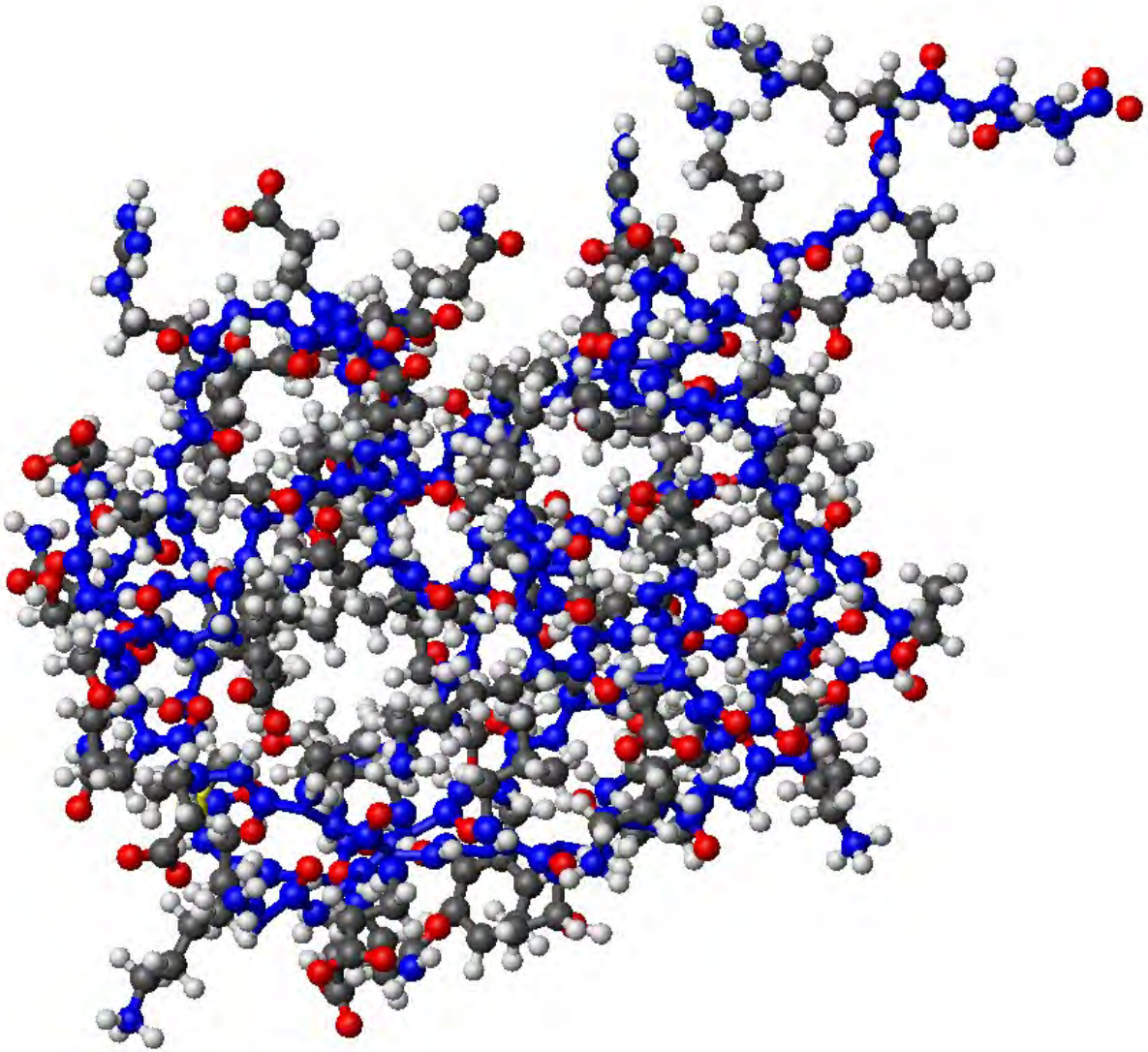

Rat-Pit-heavy SG (C)



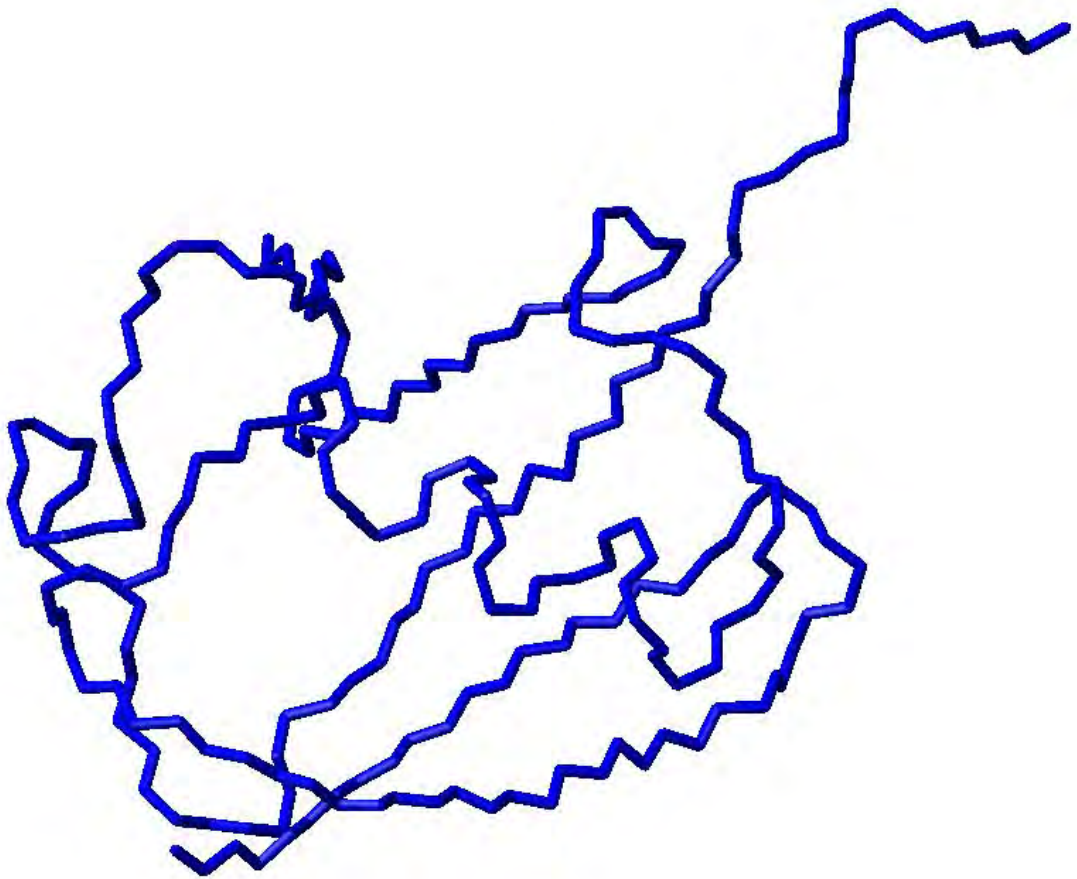
Protein structure



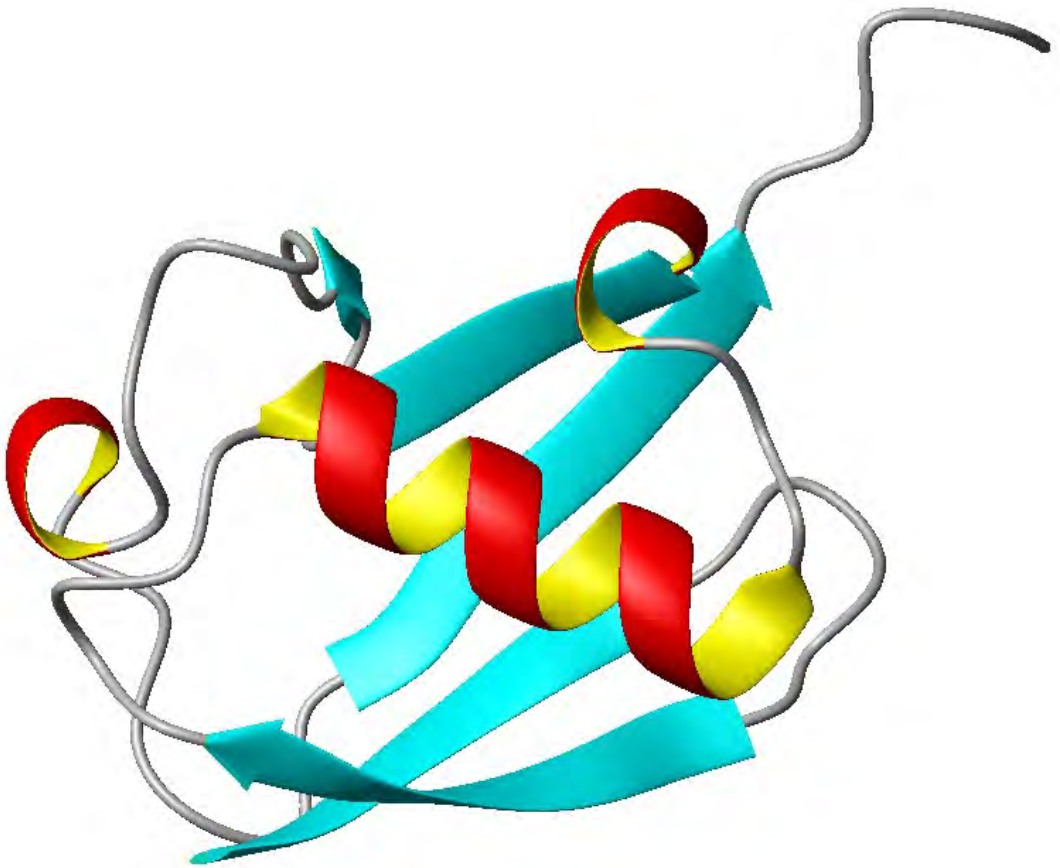
Protein structure



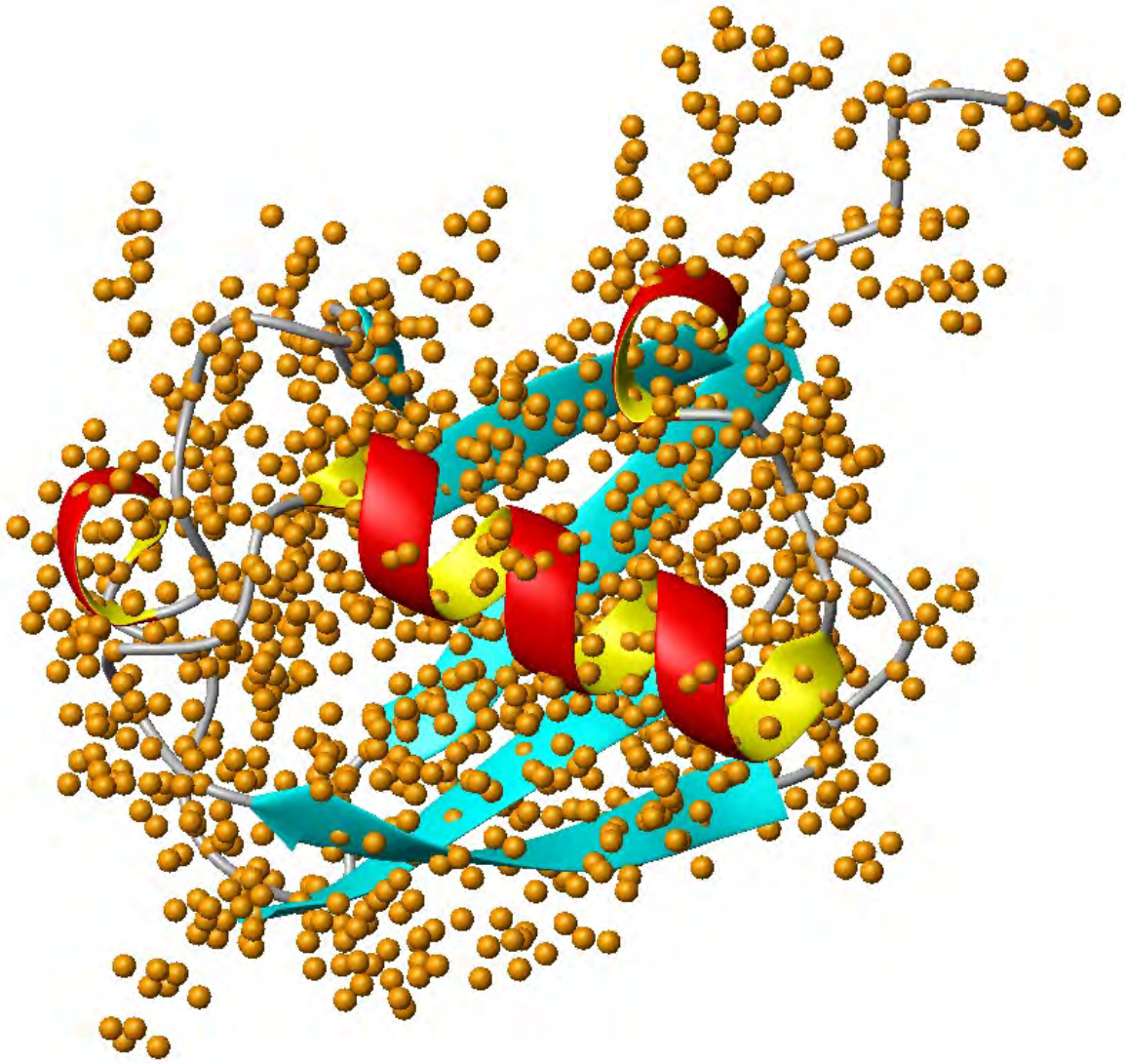
Protein structure

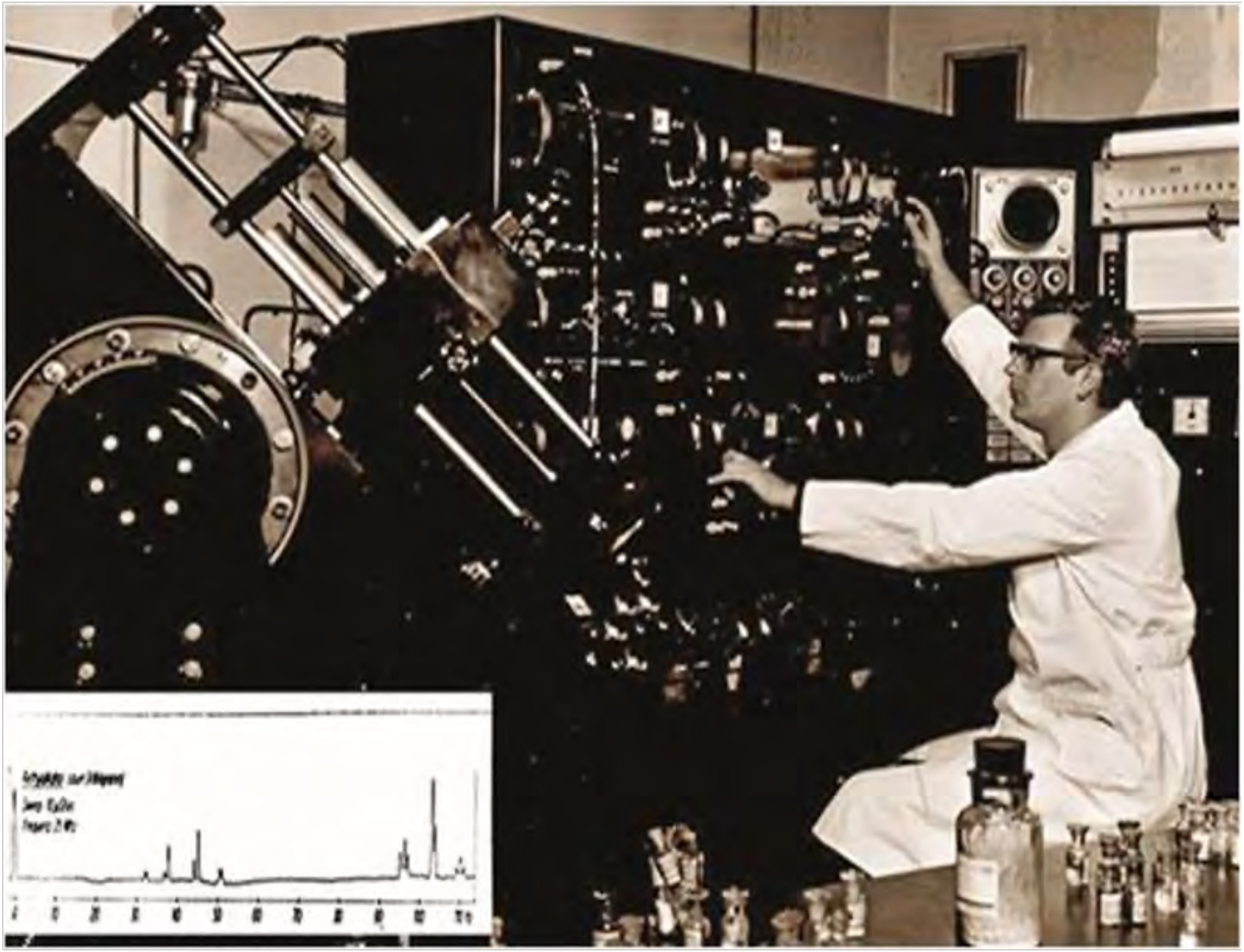


Protein structure



NMR: Kernspin-Resonanz-Spektroskopie

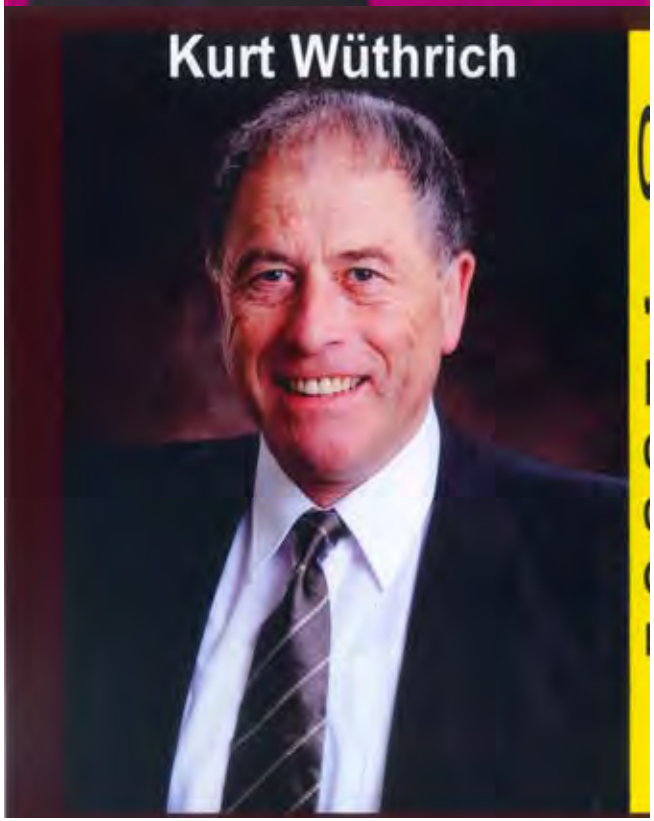




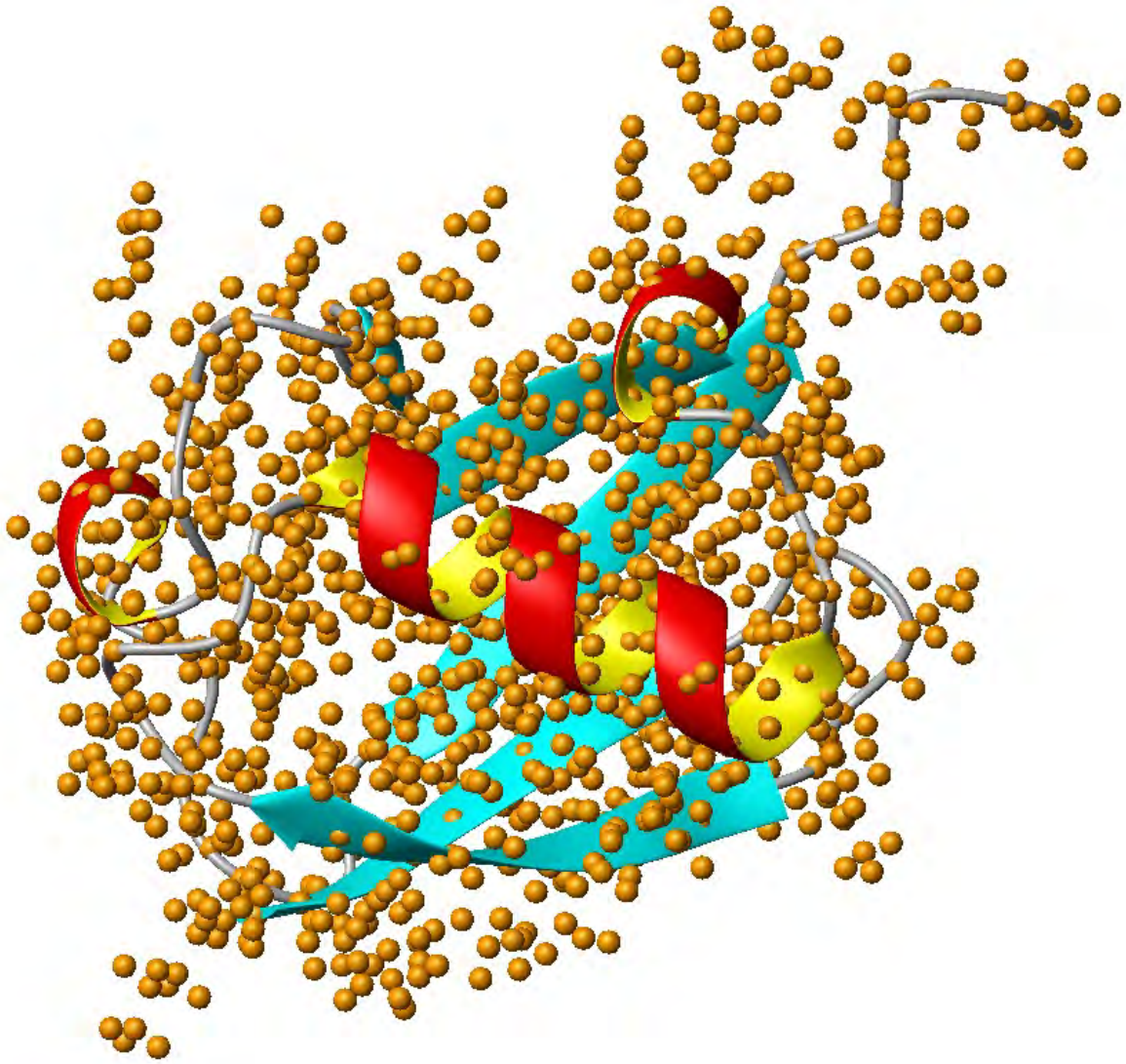

NITROGEN NITROGEN

NMR: Kernspin-Resonanz-Spektroskopie

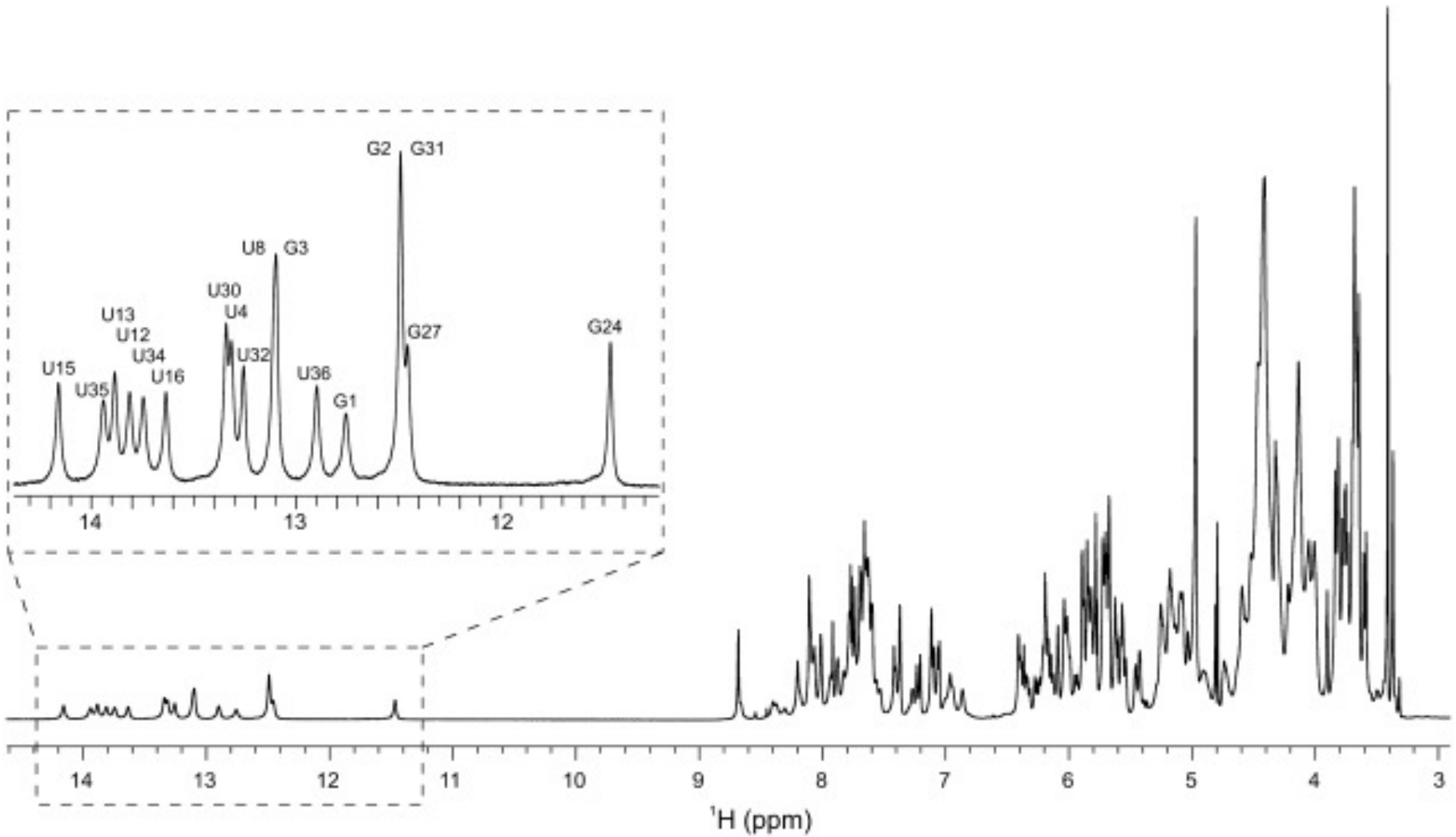
- 1938 NMR in molecular beams; Rabi (Nobelpreis in Physik 1943)
- 1946 NNR in bulk matter; Bloch & Purcell (Nobelpreis in Physik 1952)
- 1970 FT-NMR and multidimensional NMR; Ernst & Jener (Nobelpreis für R. Ernst in Chemie 1991)
- 1973 MRI; Lauterbur and Mansfield (Nobelpreis in Medizin 2003)
- 1984/85 NMR structure of a protein (Arseniev, Kaptein, Wagner, Braun, Wüthrich (Nobelpreis für K. Wüthrich in Chemie 2002)



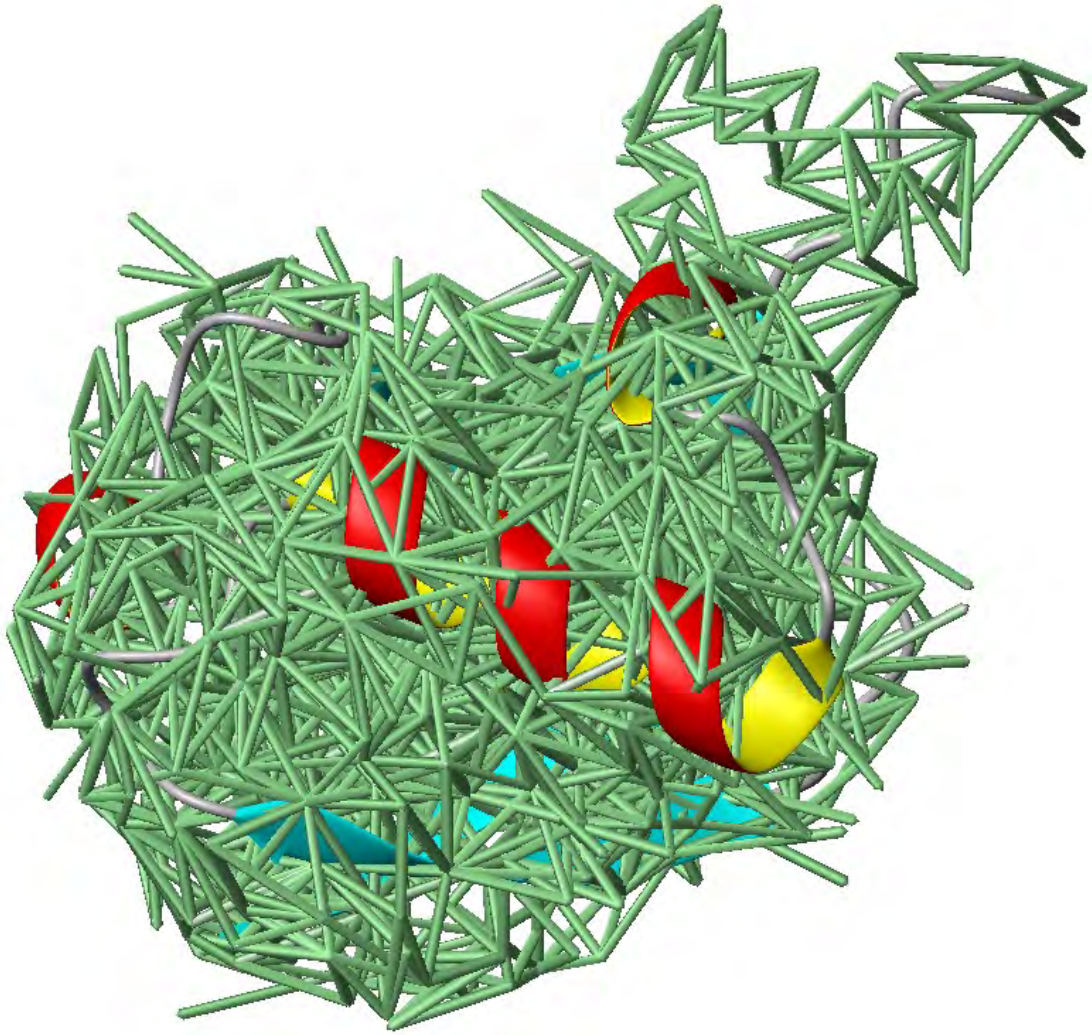
NMR: Kernspin-Resonanz-Spektroskopie



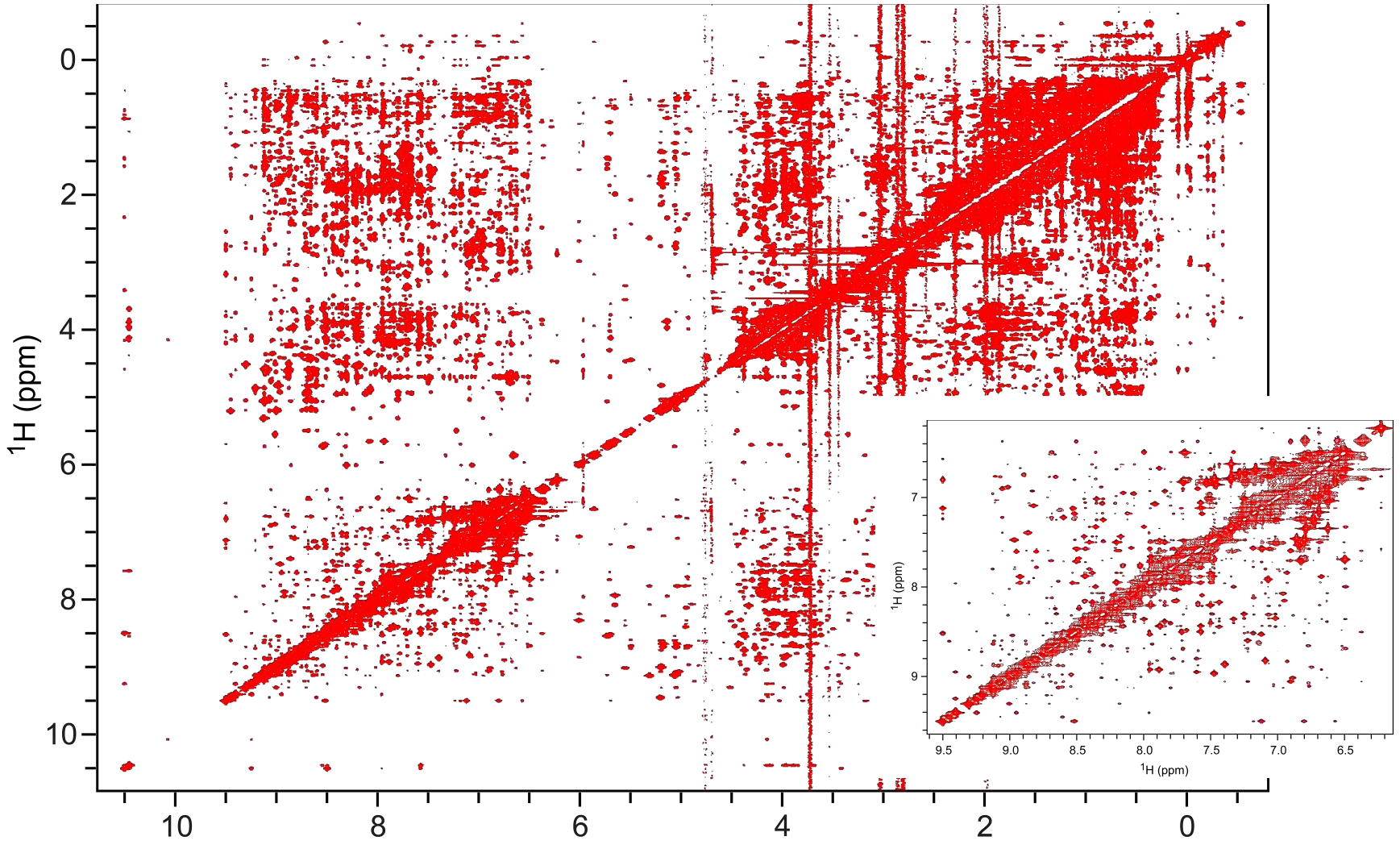
1D 1H NMR of RNA



Kommunikation zwischen Spins

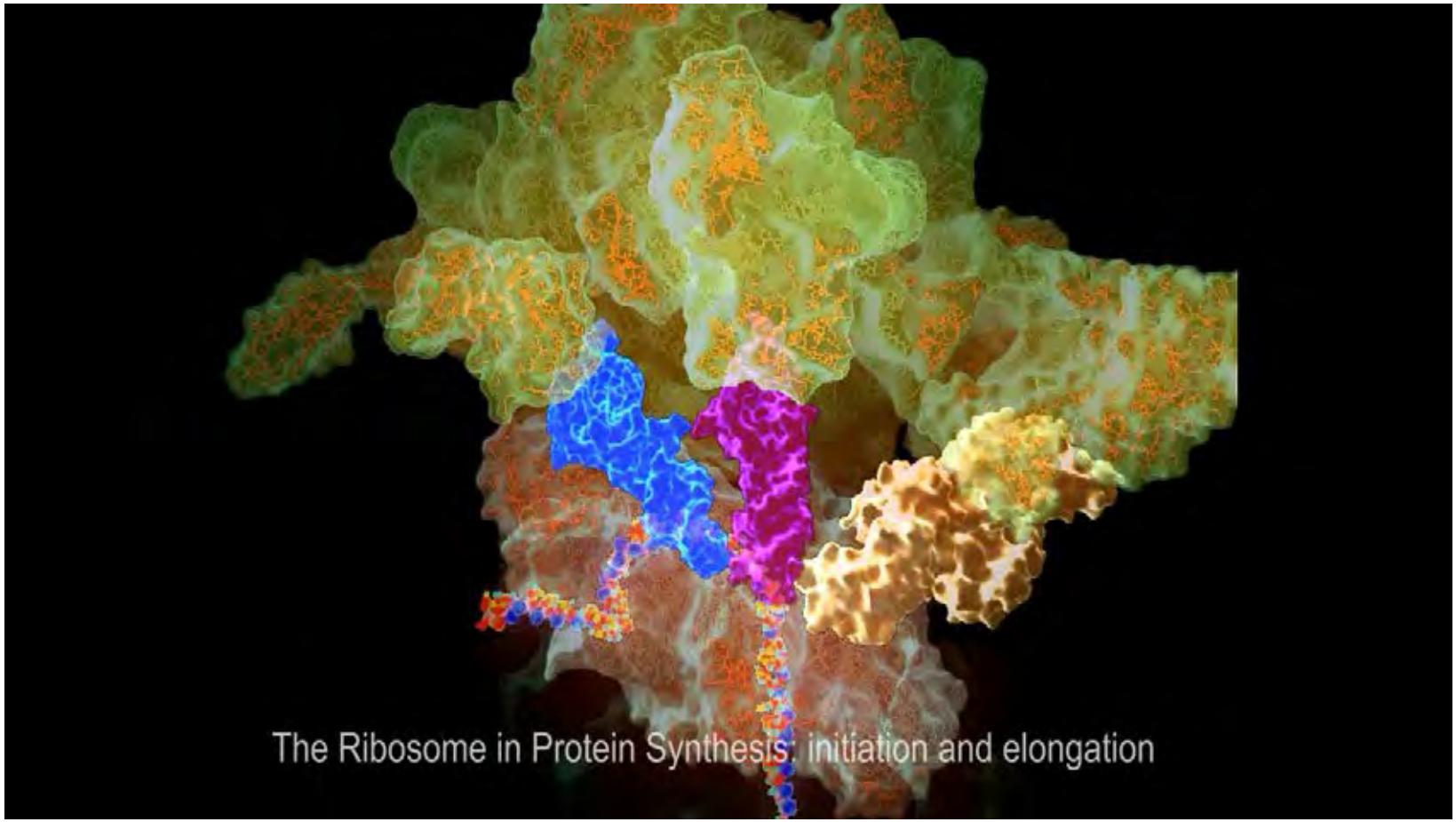


Kommunikation zwischen Spins

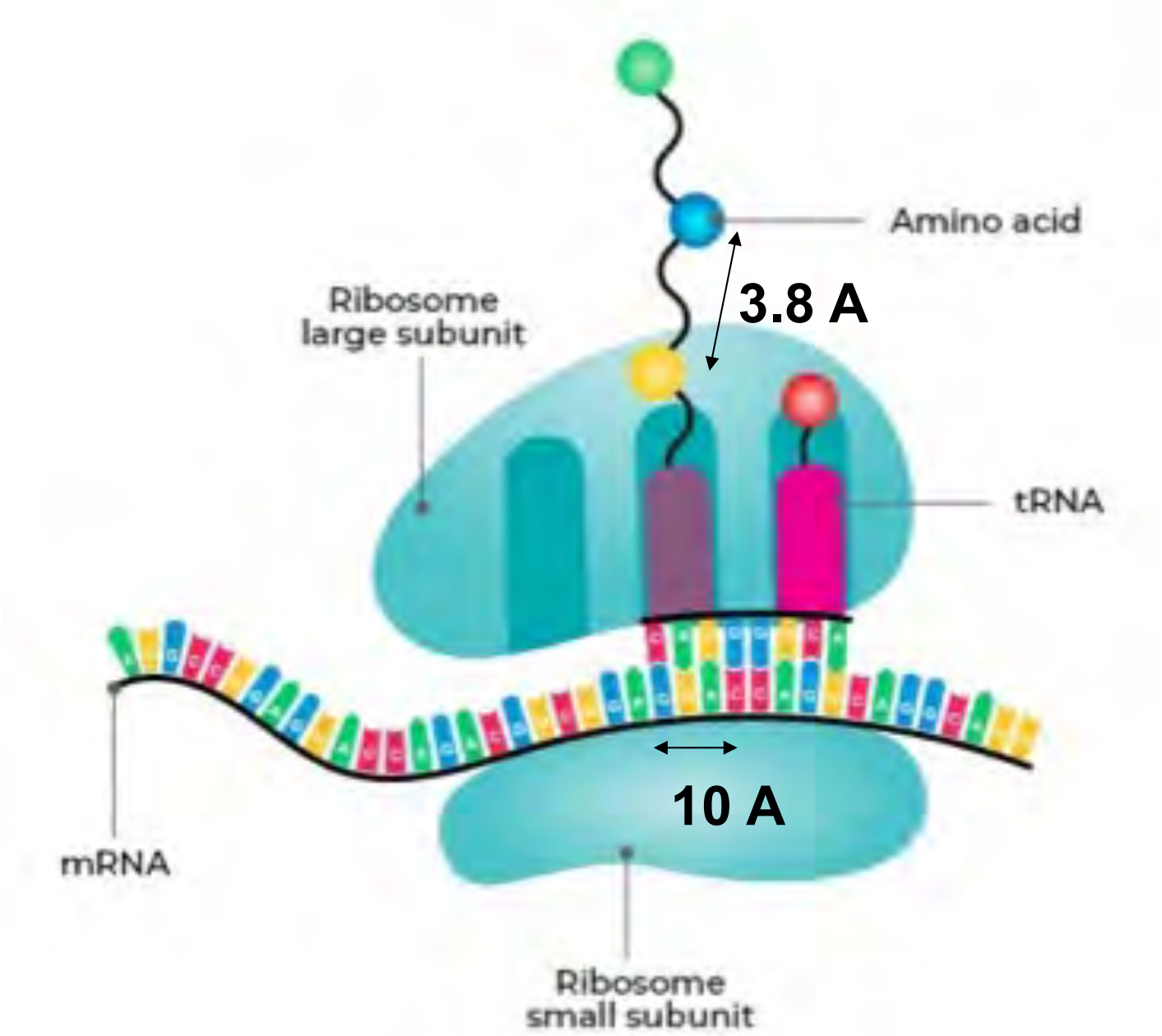


Protein Strukturen

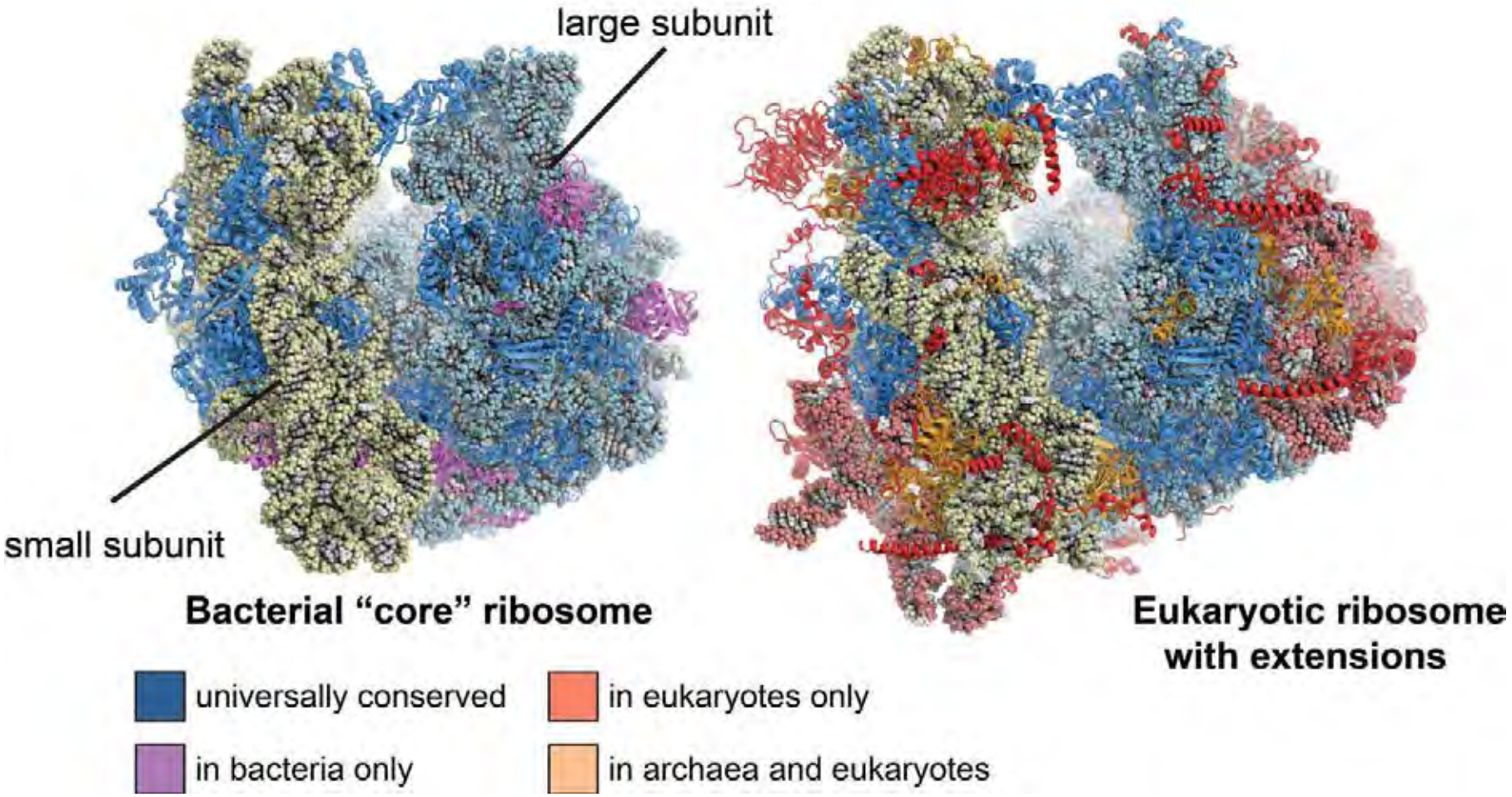
Ribosome


The Ribosome in Protein Synthesis: initiation and elongation

Translation machinery

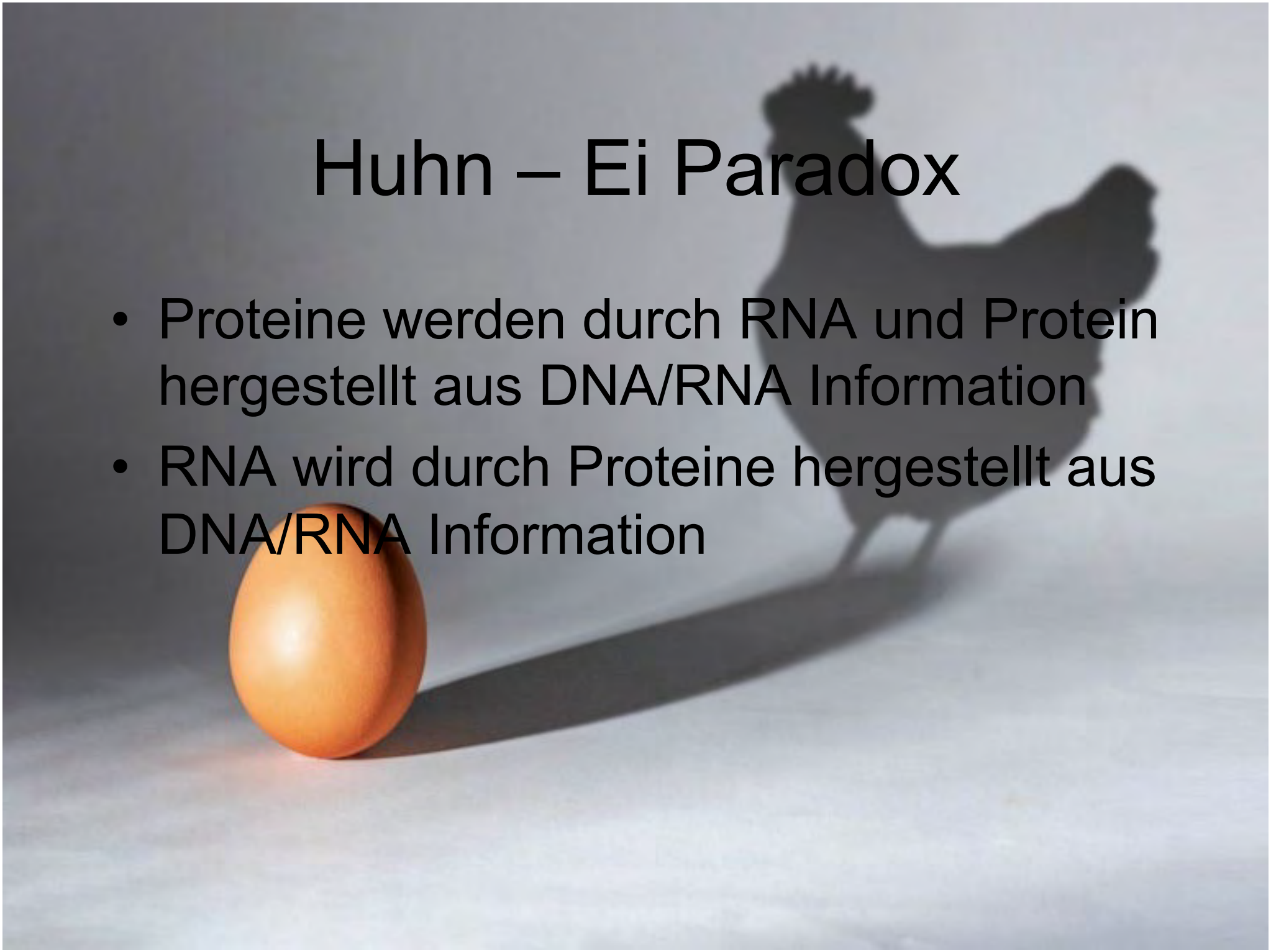


Translation machinery



Huhn – Ei Paradox

- Proteine werden durch RNA und Protein hergestellt aus DNA/RNA Information
- RNA wird durch Proteine hergestellt aus DNA/RNA Information



Leben aus der chemischen Perspektive

- Komplexität aus ”einfachen” Bausteinen
- Proteine mit Aminosäuren als Bausteine
- Membrane mit Lipide/Fettsäuren als Bausteine
- RNA/DNA mit Nucleotide als Bausteine

- Ein totes Bakterium ist tot

A first important question is what (if any) type of sequence selection is made in the synthesis process.

Number of unique sequences: $N=A^n$ (A = # of amino acid types, n = length of peptide)

4 amino acids, peptide length 6 **N ~ 4k**

4 amino acids, peptide length 10 **N ~ 1M**

20 amino acids, peptide length 100 **$N = 20^{100}$**

Mass calculation for a single copy of each possible 100-mer:

20^{100} x 10000 g mol^{-1} / 6.022 x 10^{23} mol^{-1} = 3 x 10^{81} g

Mass of observable universe = 3 x 10^{55} g

Only enough mass to make a single copy of 1 in 10^{26} sequences !!!!

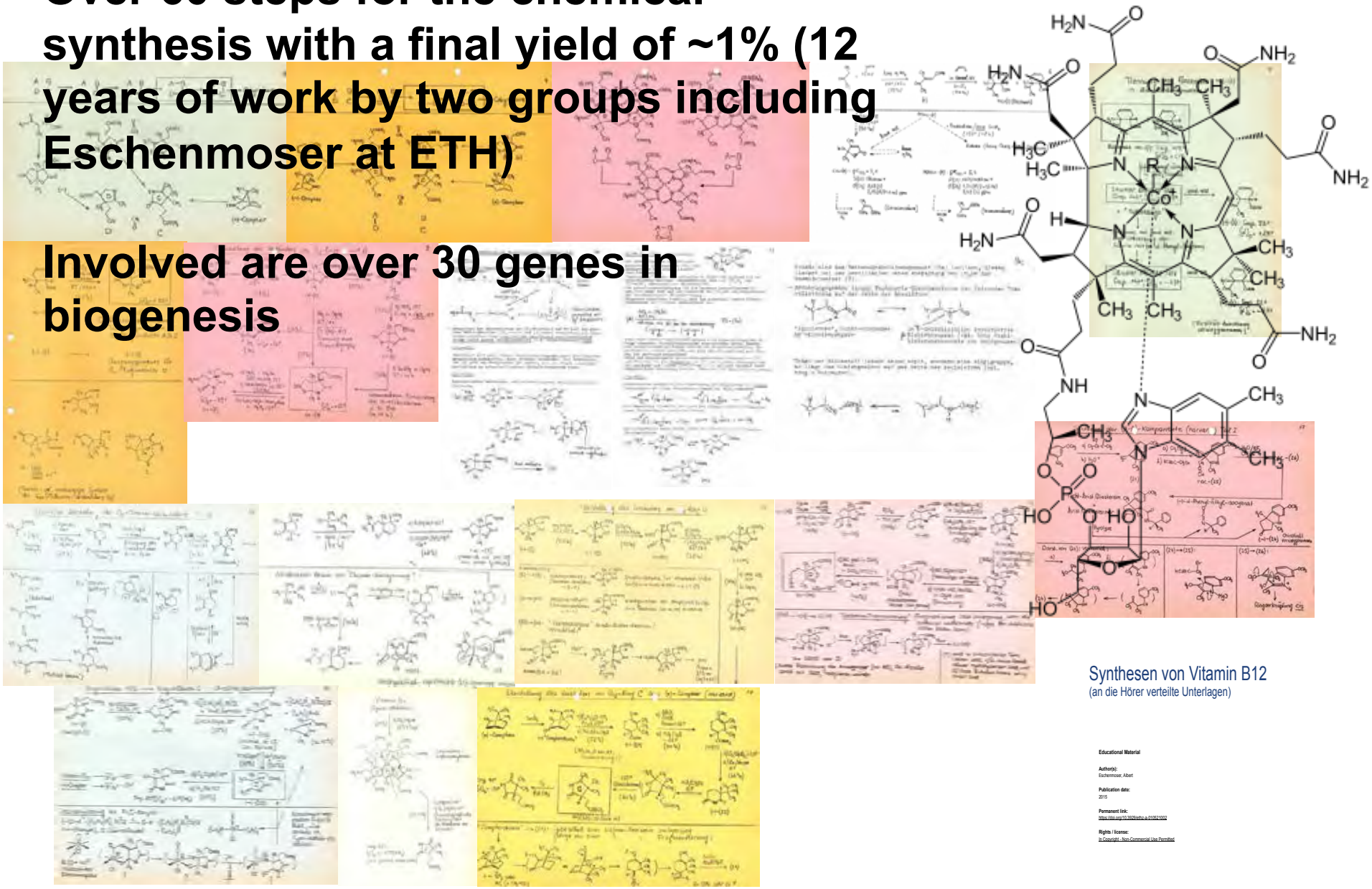
Putting this into perspective for life on Earth:

If the entire Mass of the earth contributed to making a single copy of every 100mer, only 1 in 5e53 of them could be made once.

Relative to an object the size of the Earth, the size of the space that has been searched is a 1/50th of a hydrogen atom.

Over 60 steps for the chemical synthesis with a final yield of ~1% (12 years of work by two groups including Eschenmoser at ETH)

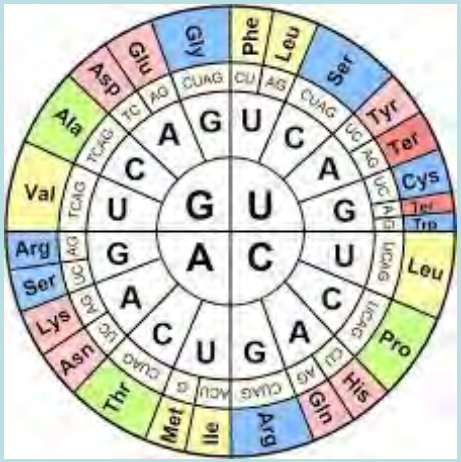
Involved are over 30 genes in biogenesis



Synthesen von Vitamin B12
(an die Hörer verteilte Unterlagen)

Chemische Information

- Genetischer Code
- Proteine/DNA/RNA/Lipide
- Genome Information

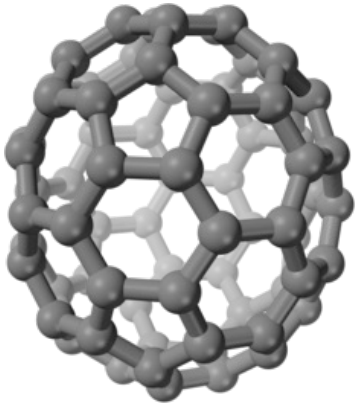


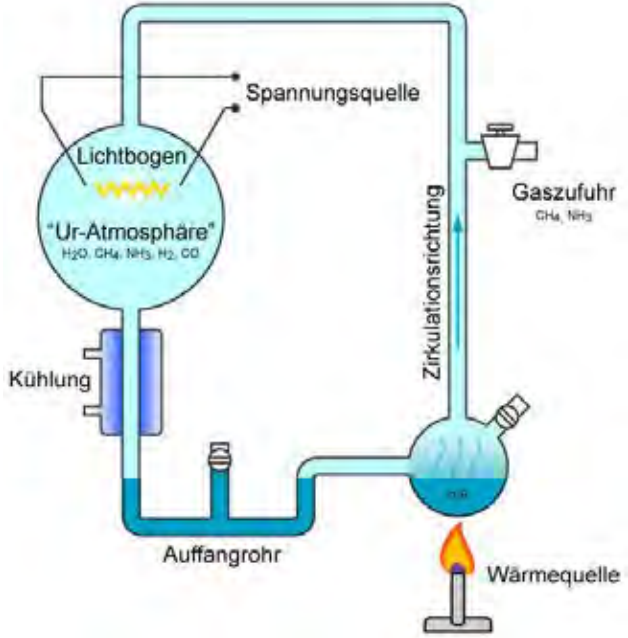
2nd base in codon

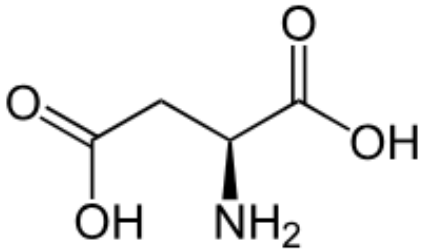
1st base in codon	U	C	A	G	3rd base in codon
U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U C A G
C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G

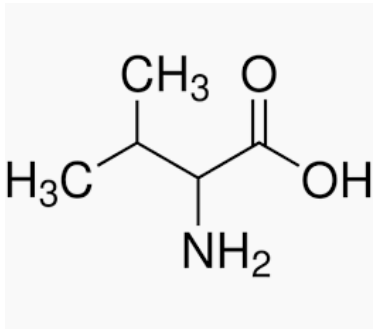
Einfache Moleküle im All (Ursuppe Experimente)

- Fullerene



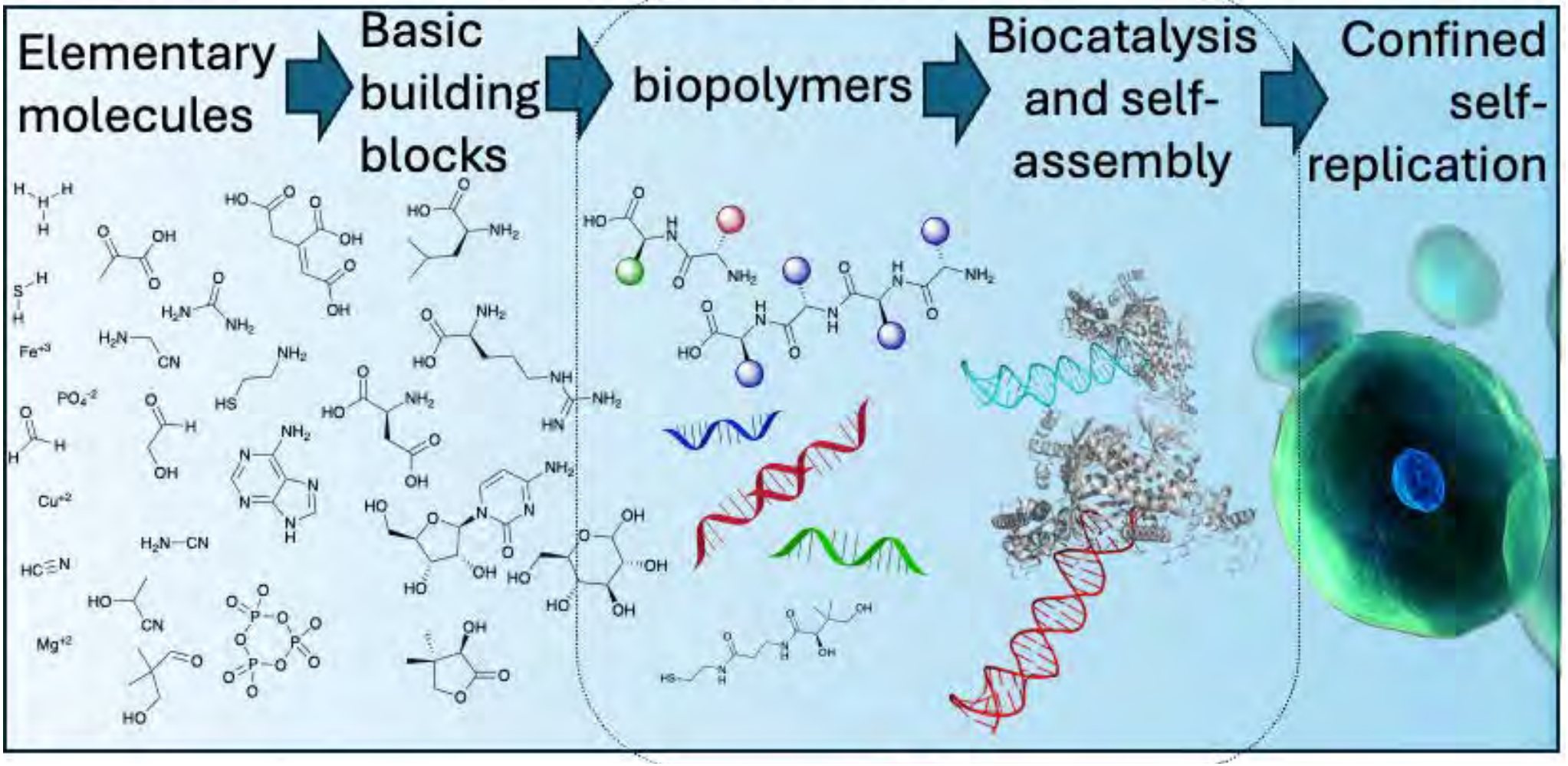






	M1	*M2*	*M3*	*I1*	*A1*
G - Gly	1.00	1.0	1.000	1.000	1.000
A - Ala	0.34	0.4	0.380	0.293	0.540
D - Asp	0.19	0.5	0.035	0.022	0.006
E - Glu	0.40	0.5	0.110	0.000	0.010
V - Val	0.19	0.3	0.100	0.012	0.000
S - Ser	0.00	0.0	0.003	0.072	0.000
I - Ile	0.13	0.0	0.060	0.000	0.000
L - Leu	0.04	0.0	0.035	0.000	0.000
P - Pro	0.29	0.1	0.000	0.001	0.000
T - Thr	0.00	0.0	0.003	0.000	0.000
K - Lys	0.00	0.0	0.000	0.000	0.000
F - Phe	0.00	0.0	0.000	0.000	0.000
R - Arg	0.00	0.0	0.000	0.000	0.000
H - His	0.00	0.0	0.000	0.000	0.000
NQCYMW	0.00	0.0	0.000	0.000	0.000

The rise of molecular complexity and the emergence on function (life)

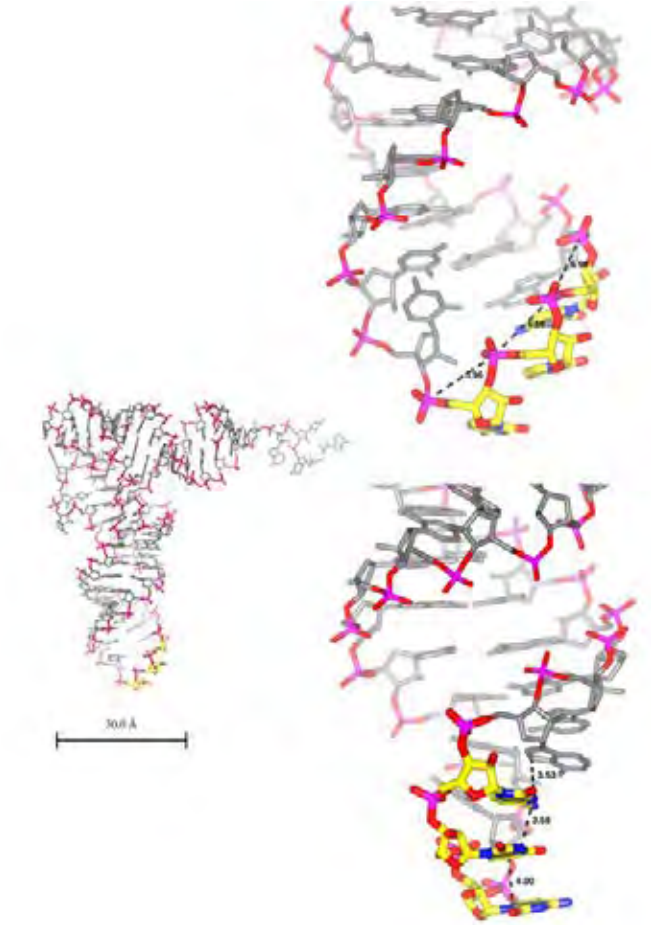


The “worlds”

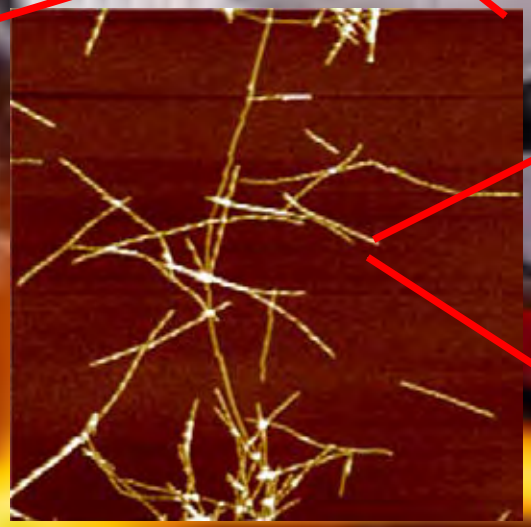
- RNA World hypothesis
- Membrane world hypothesis
- Protein/Peptid world hypothesis

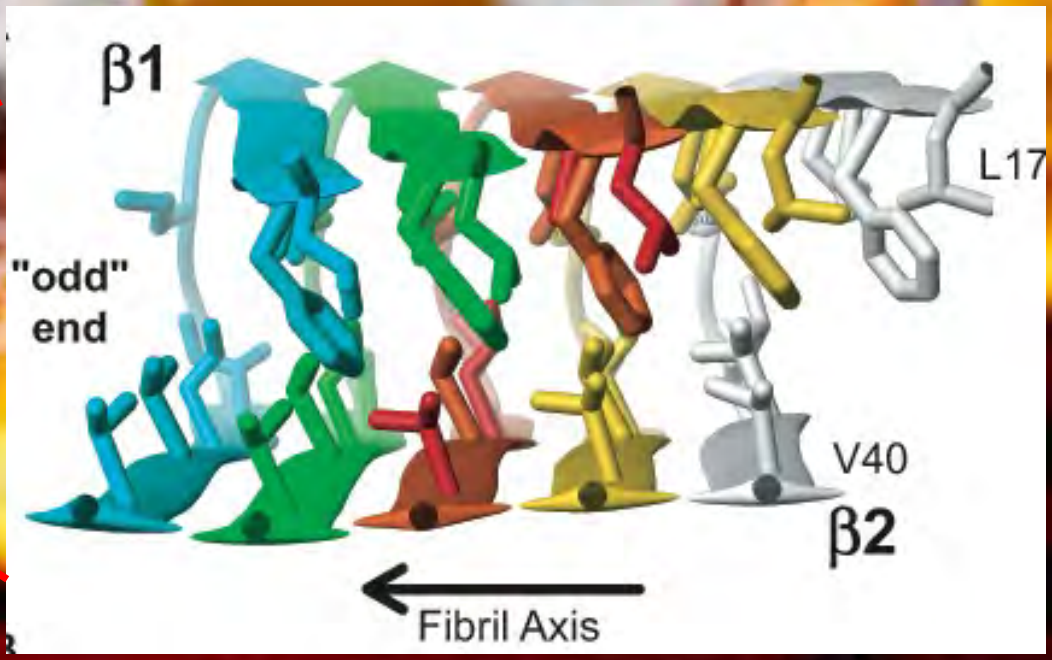
The RNA World Hypothesis

- Selbstvervielfachende RNA
- Ribzoyme
- Ribosome hat RNA als katalytisches Zentrum
- Genetische Information

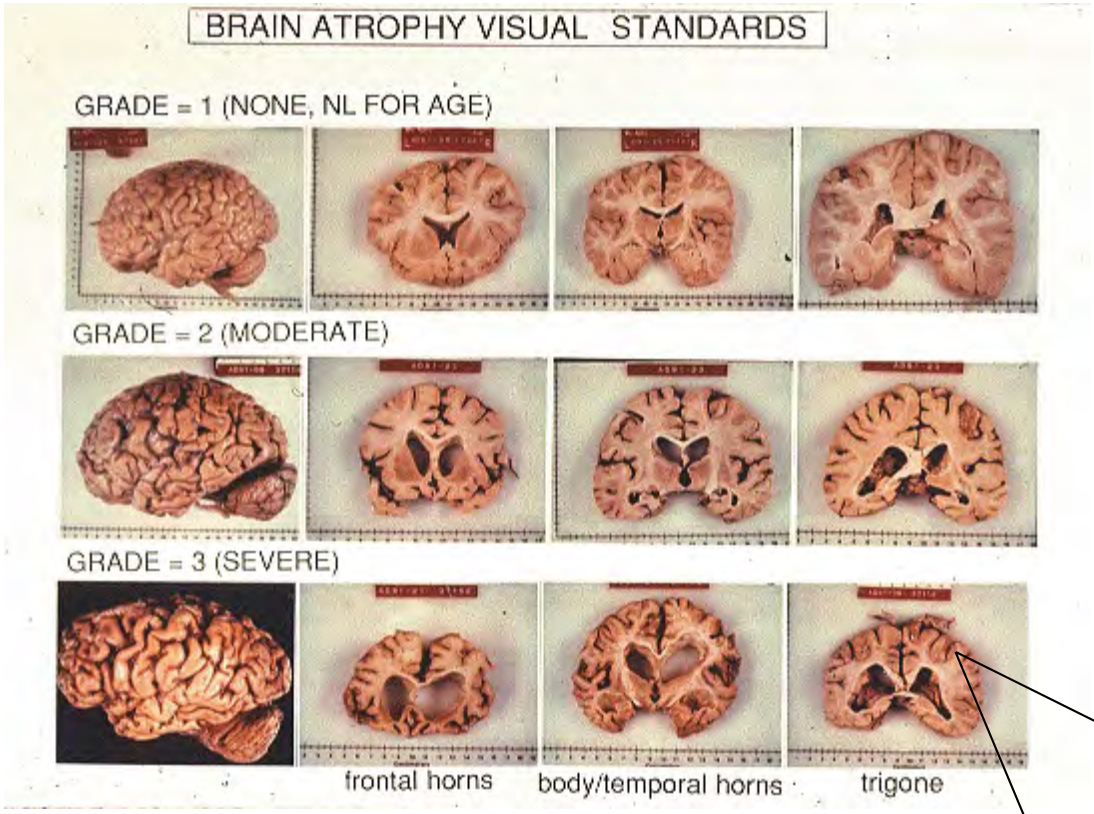


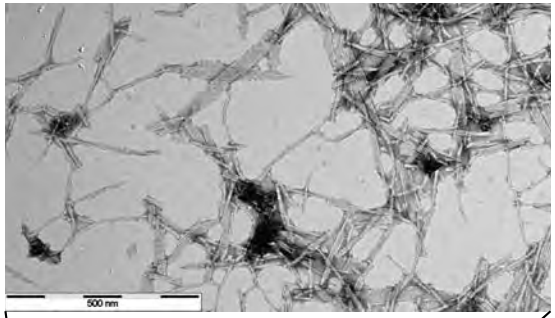
Amyloid Fibrillen

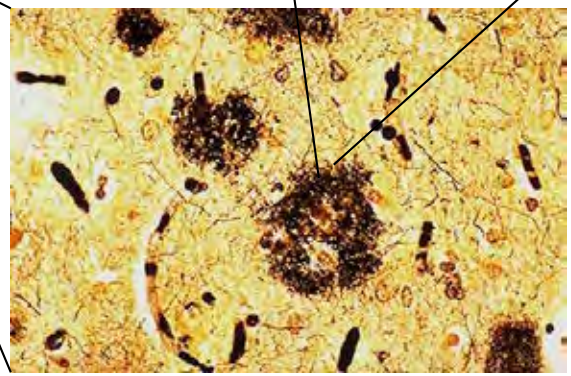




Alzheimer Erkankung

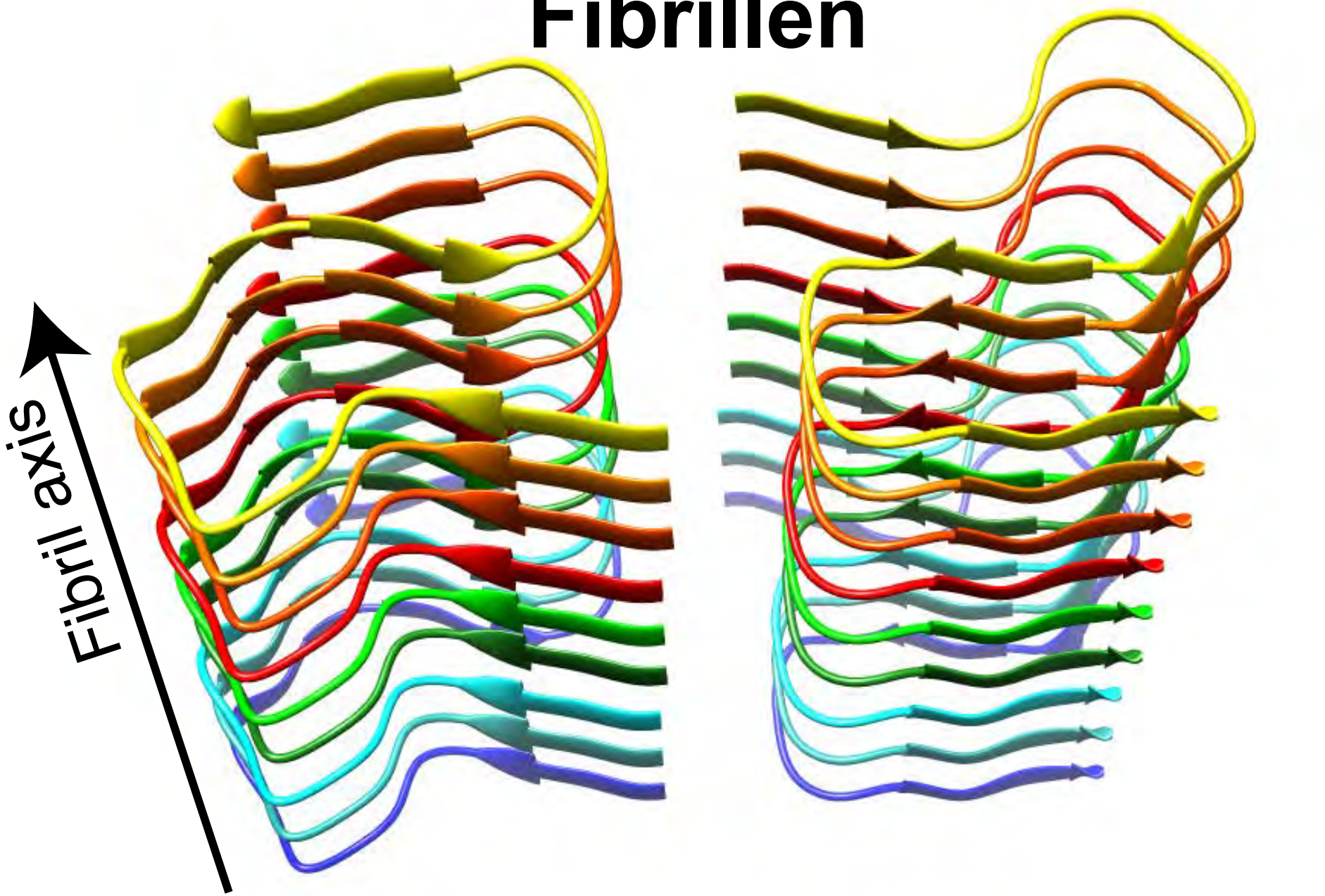




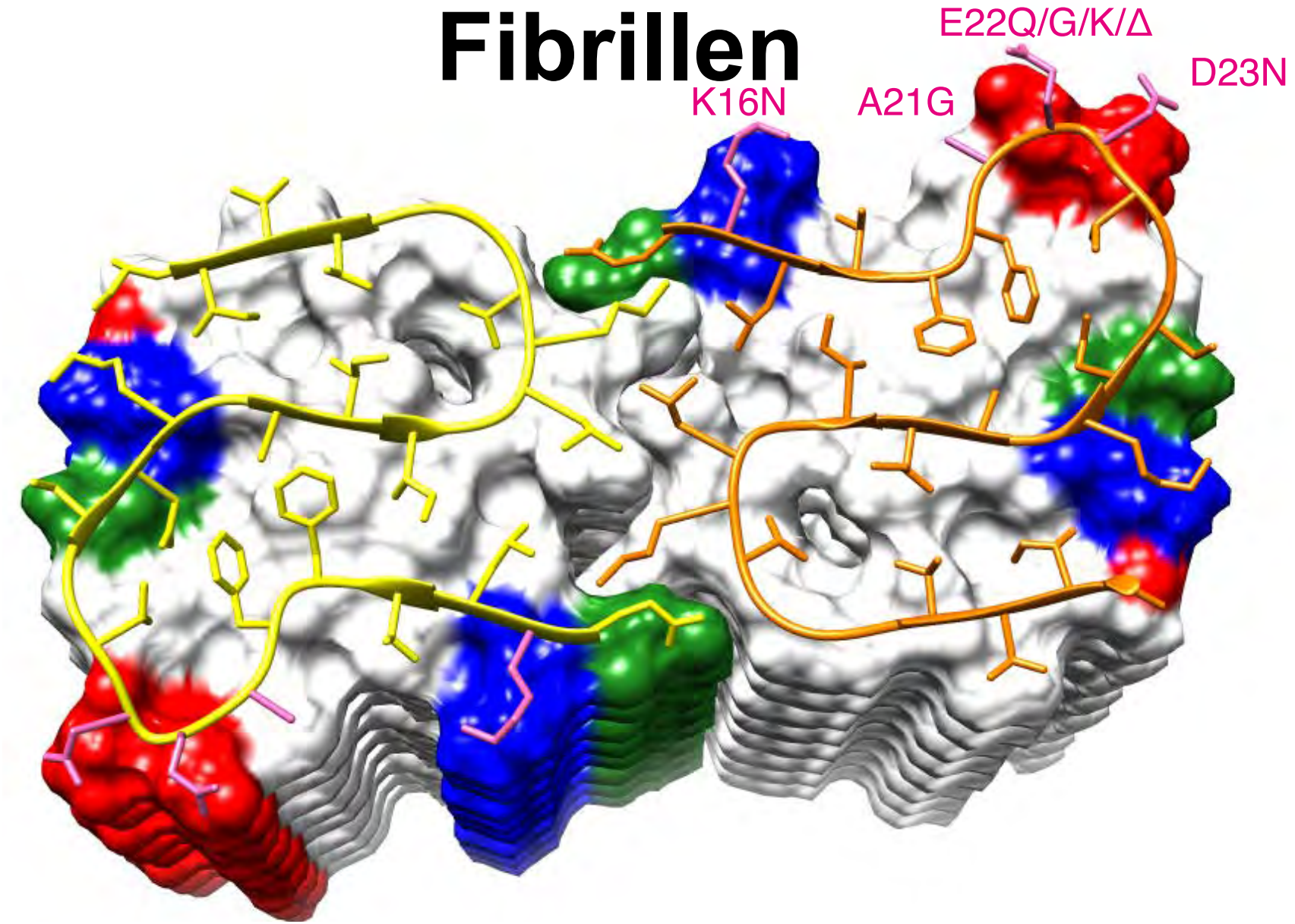


From Alzheimer's Disease Education and Referral (ADEAR) Center's Web site and library.med.utah.edu

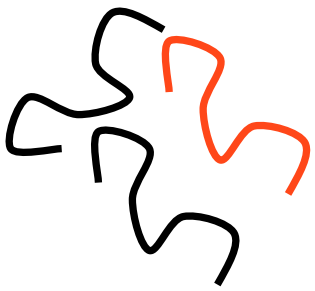
3D Struktur von Aβ(1-42) Amyloid Fibrillen

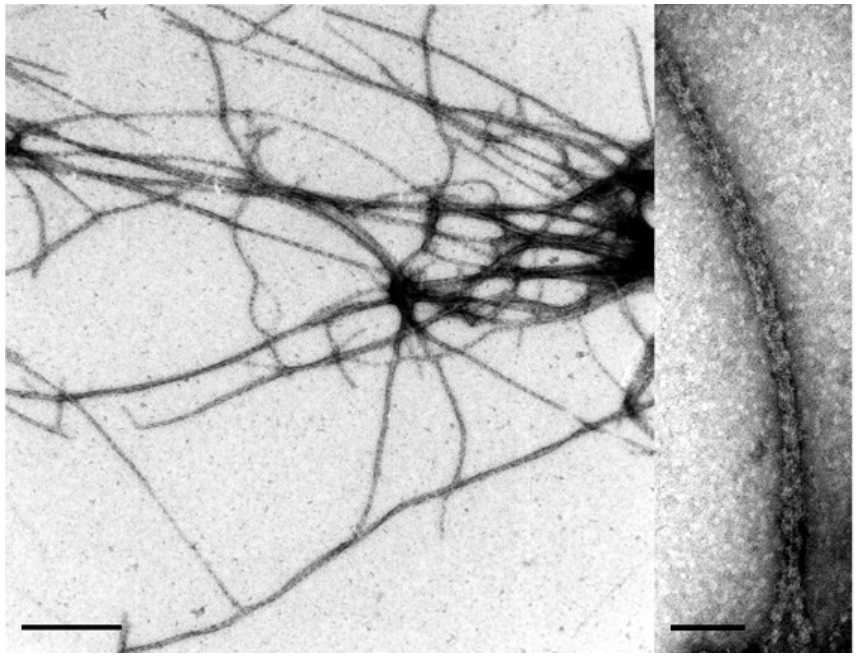


3D Struktur von Aβ(1-42) Amyloid Fibrillen

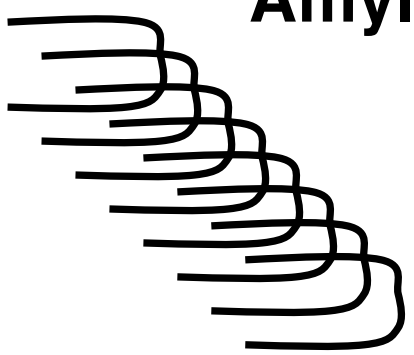


Protein-only Hypothese und Amyloidogenesis

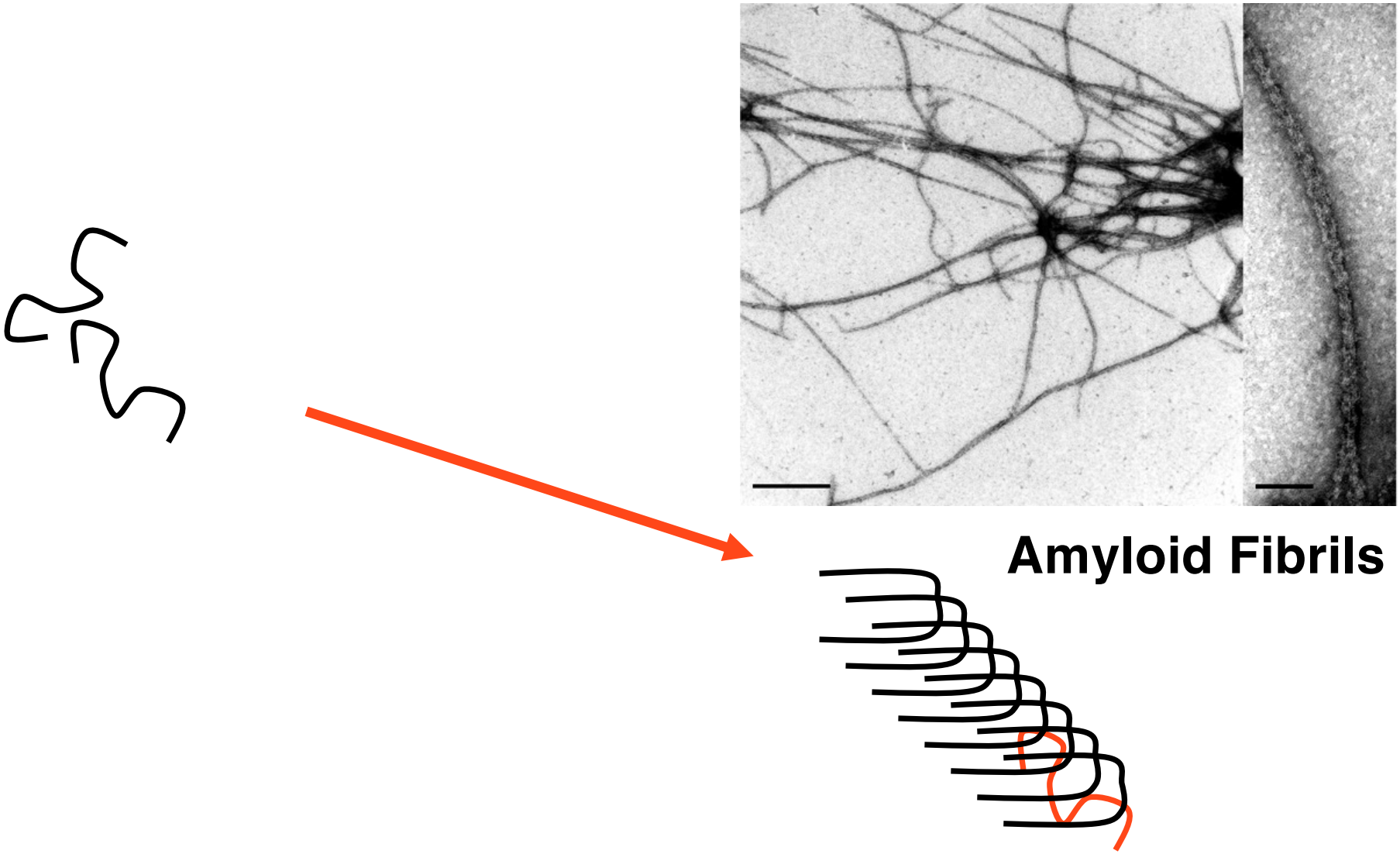




Amyloid Fibrils

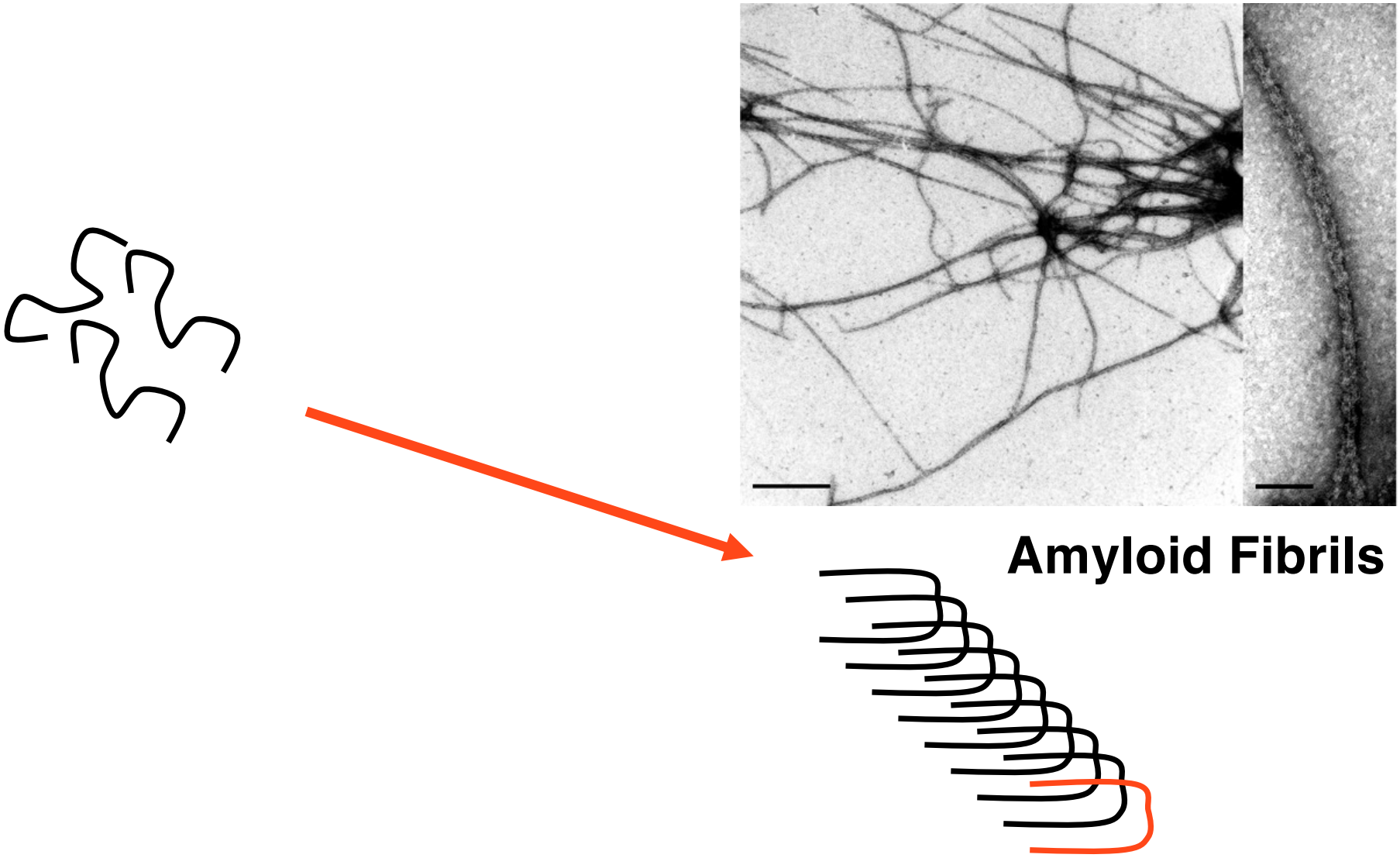


Protein-only Hypothese und Amyloidogenesis

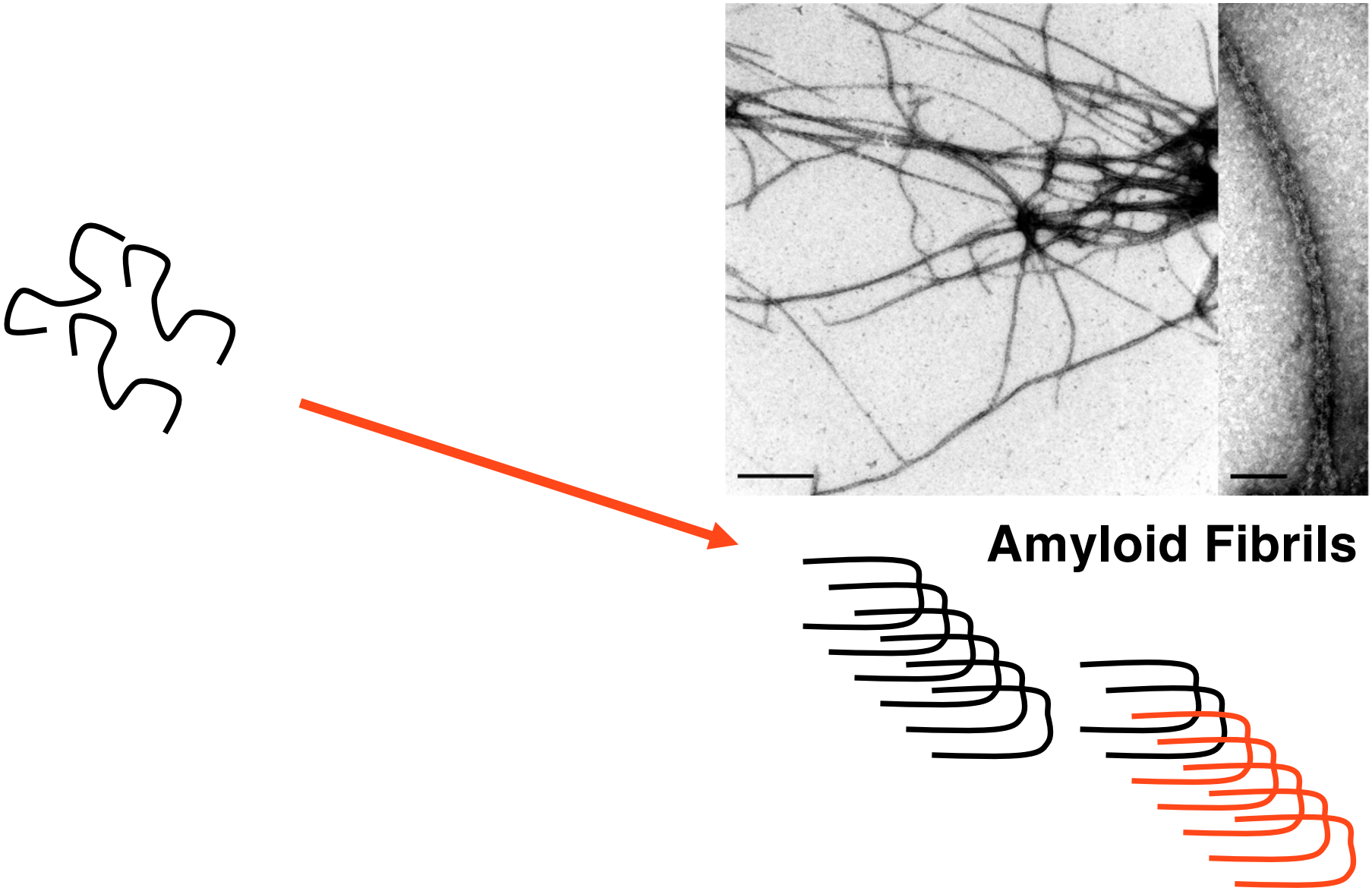


Amyloid Fibrils

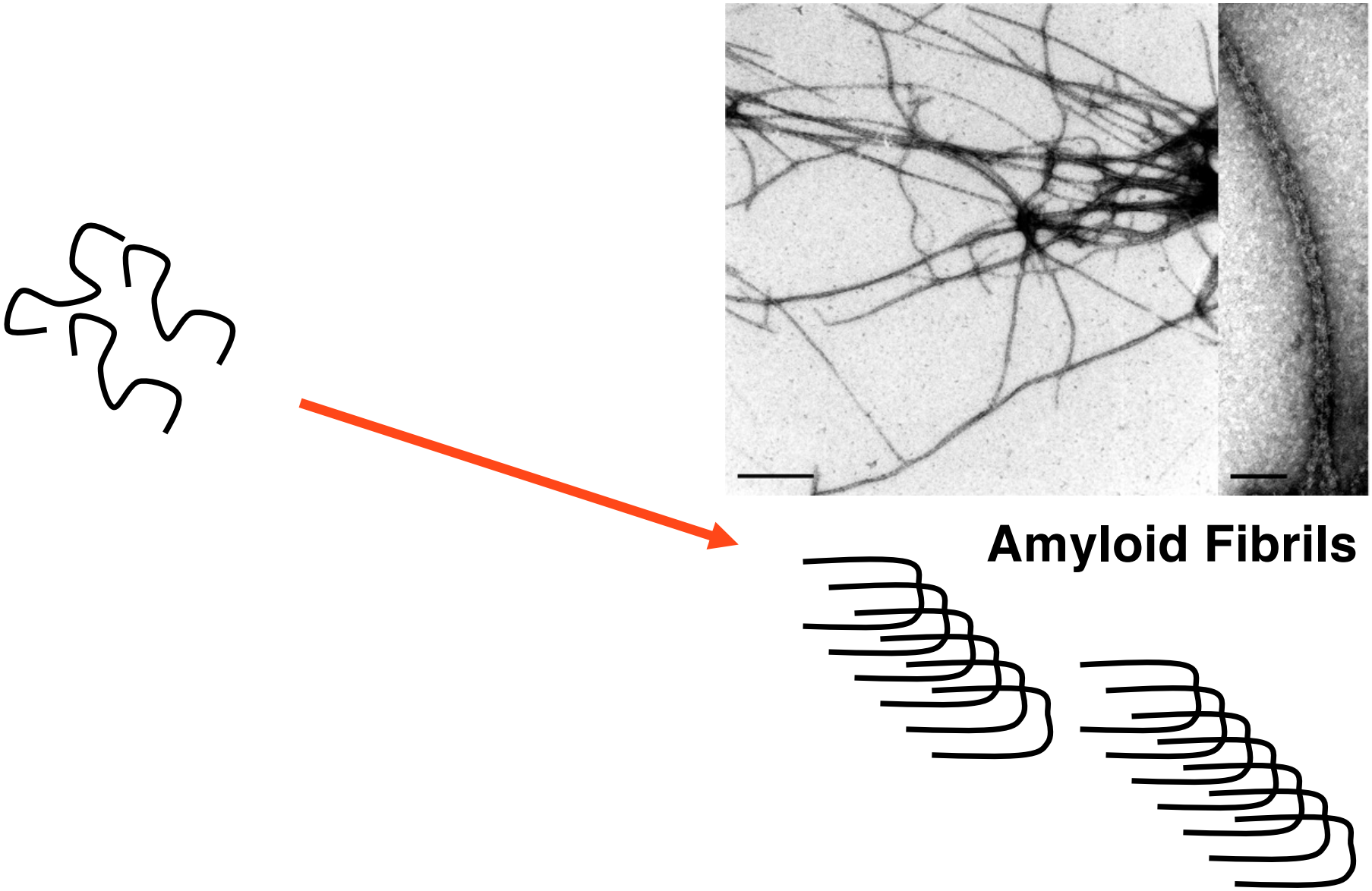
Protein-only Hypothese und Amyloidogenesis



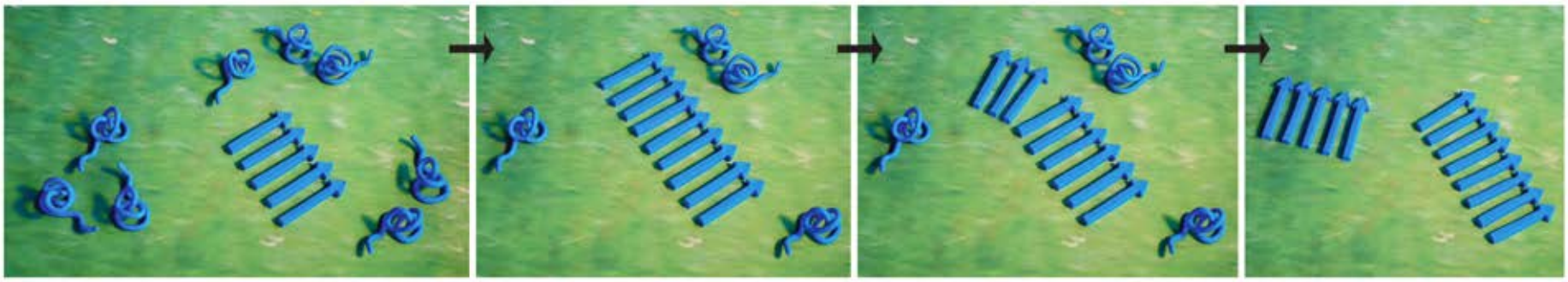
Protein-only Hypothese und Amyloidogenesis



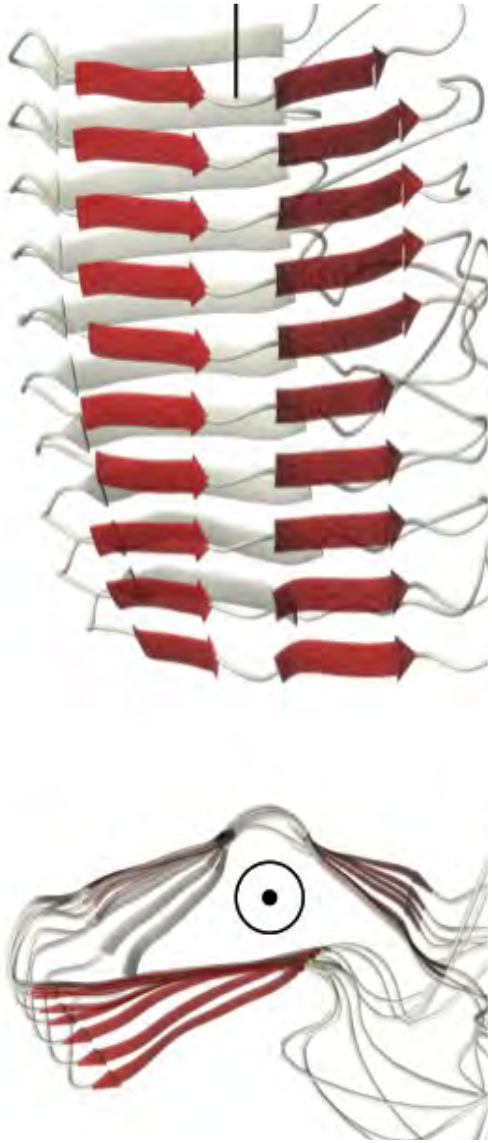
Protein-only Hypothese und Amyloidogenesis

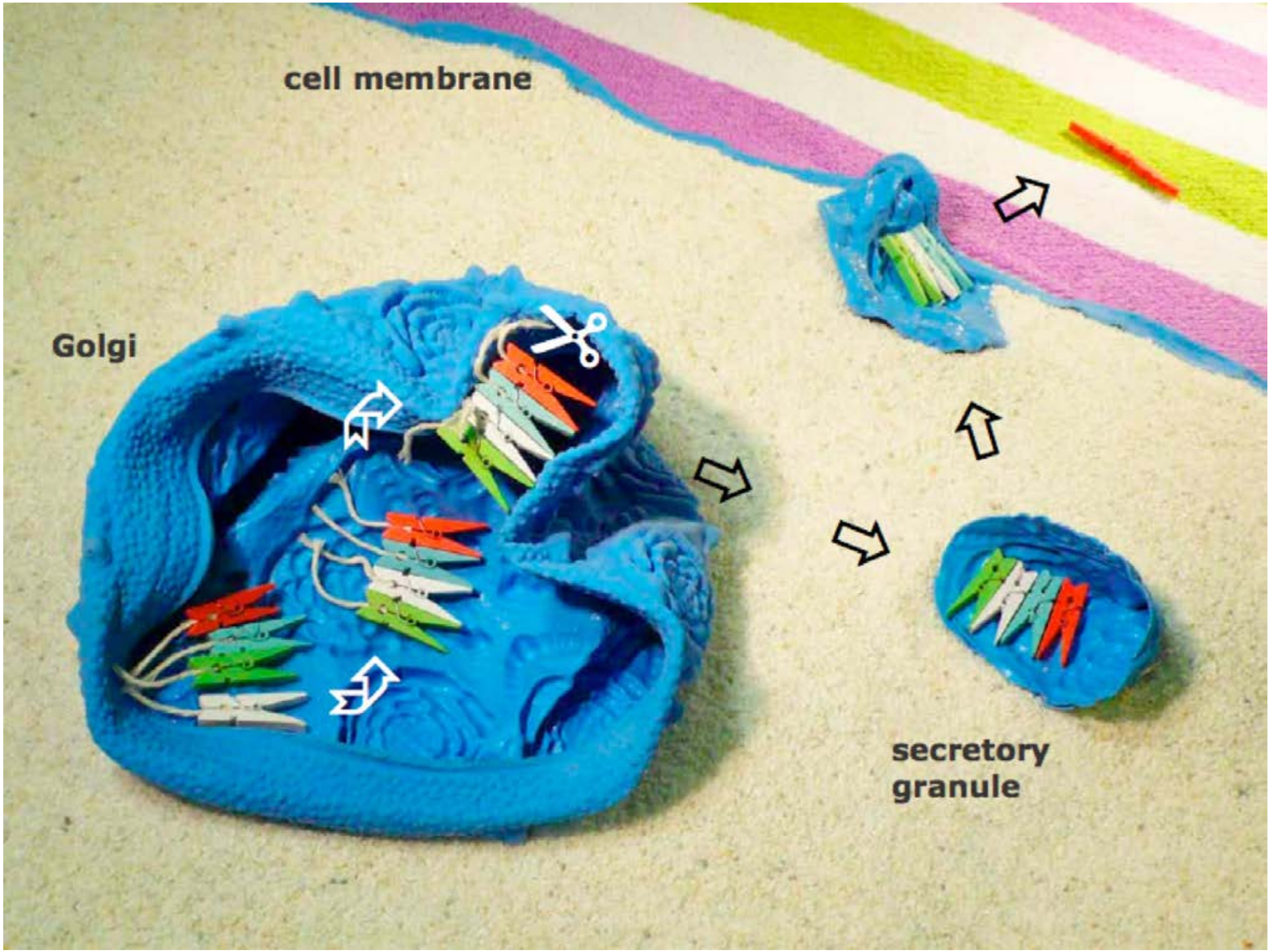


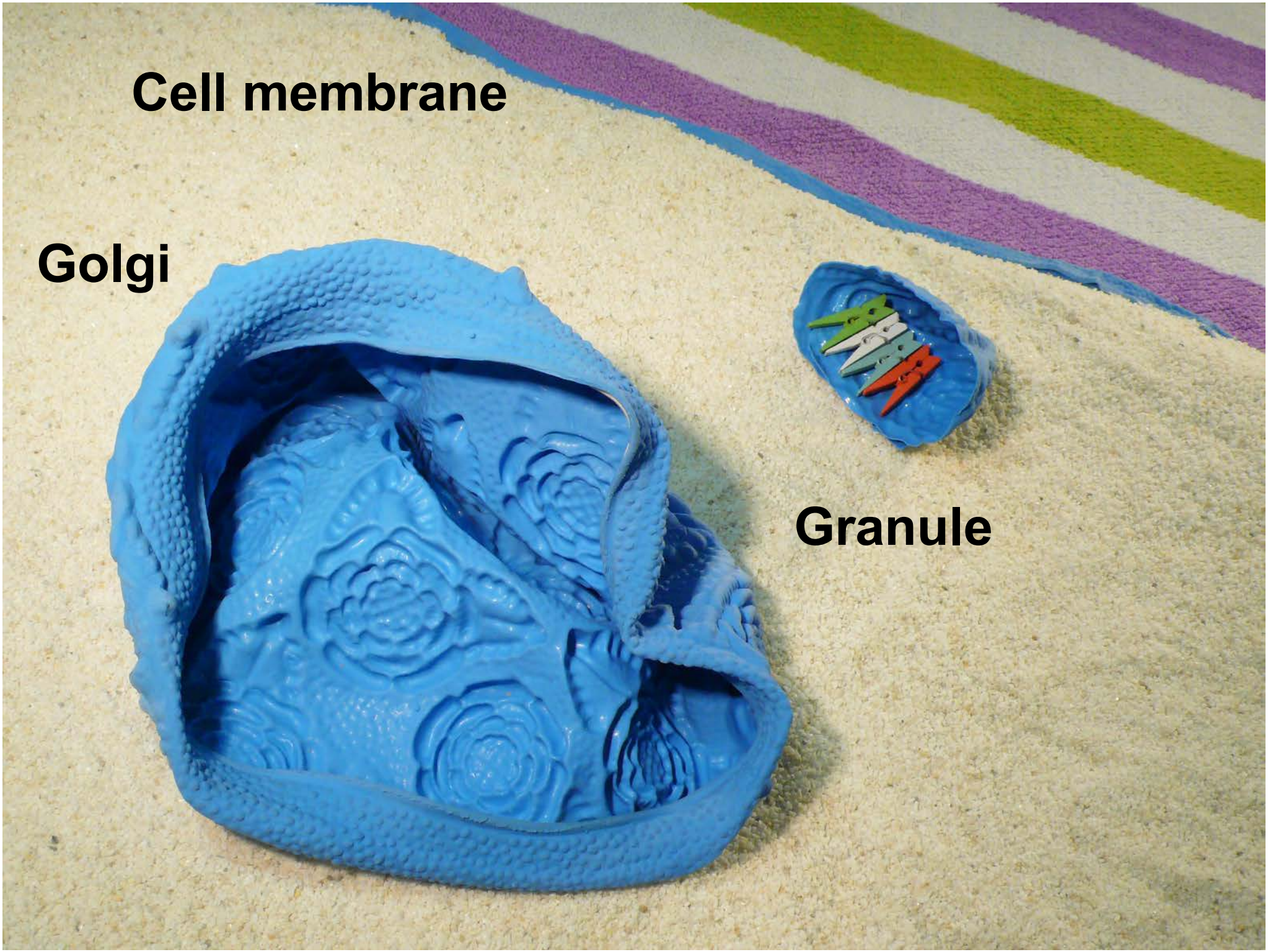
Protein-only Hypothese und Amyloidogenesis



3D Struktur von β-endorphin Amyloid Fibrillen

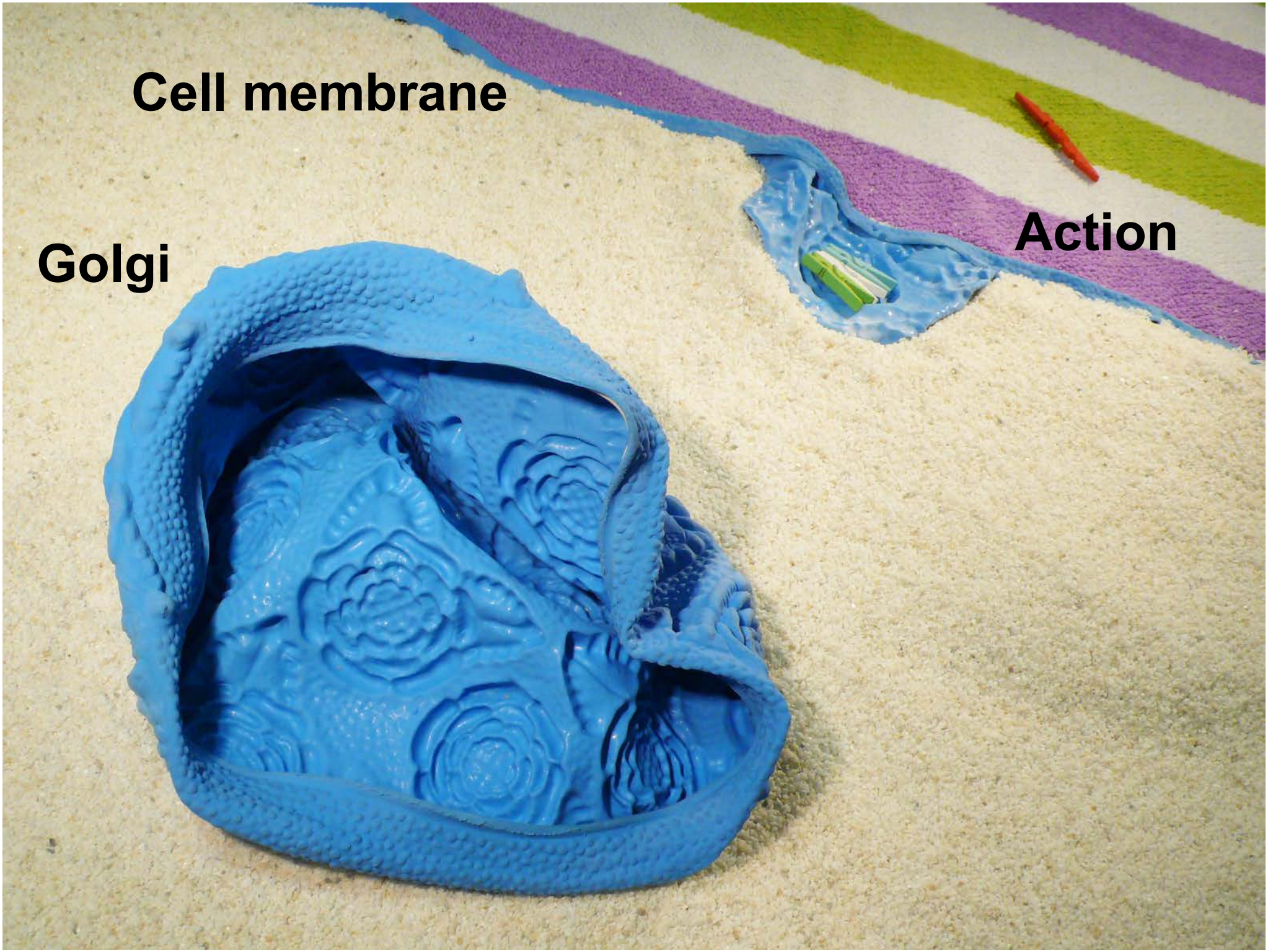


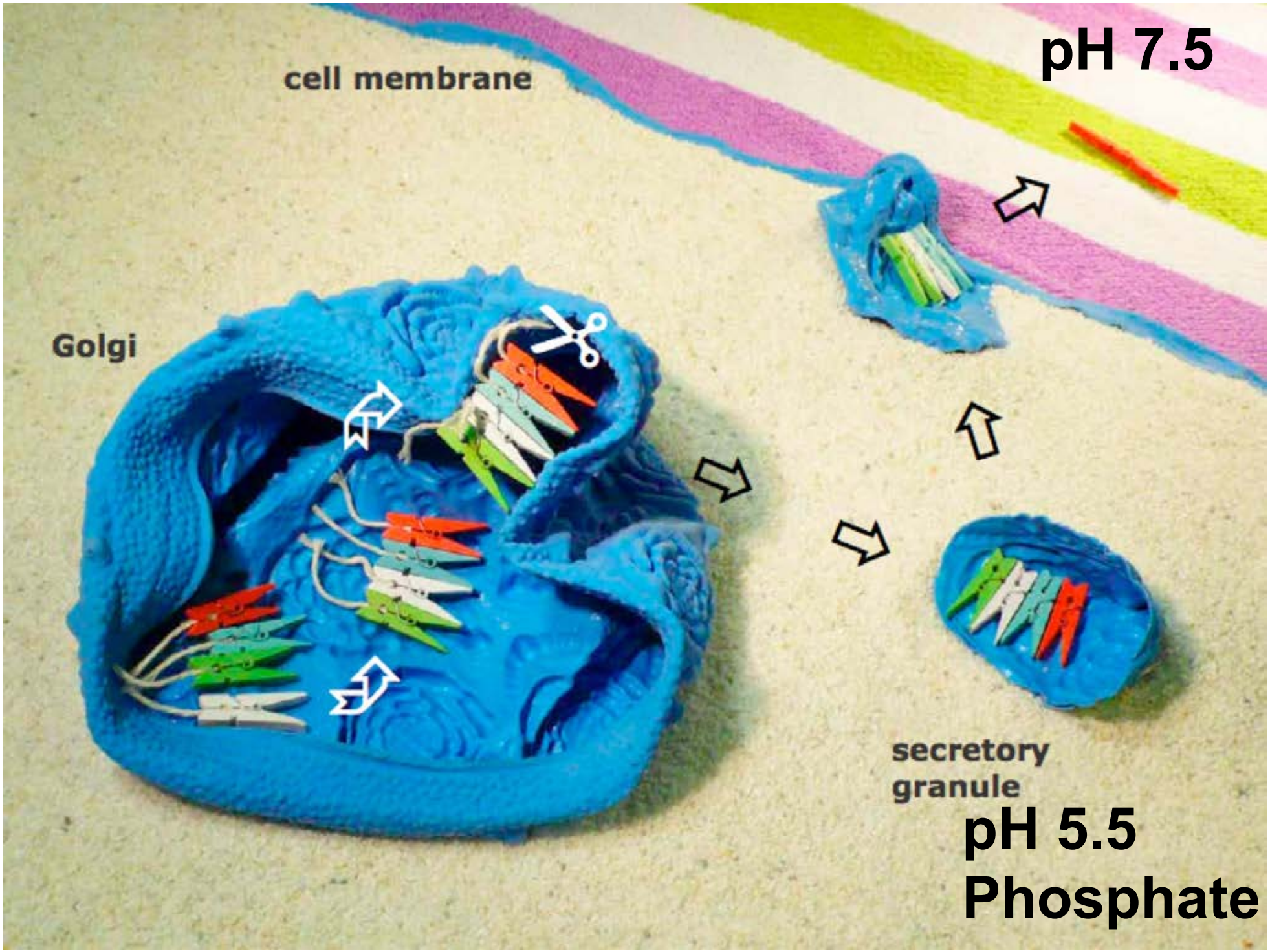

cell membrane
Golgi
secretory
granule

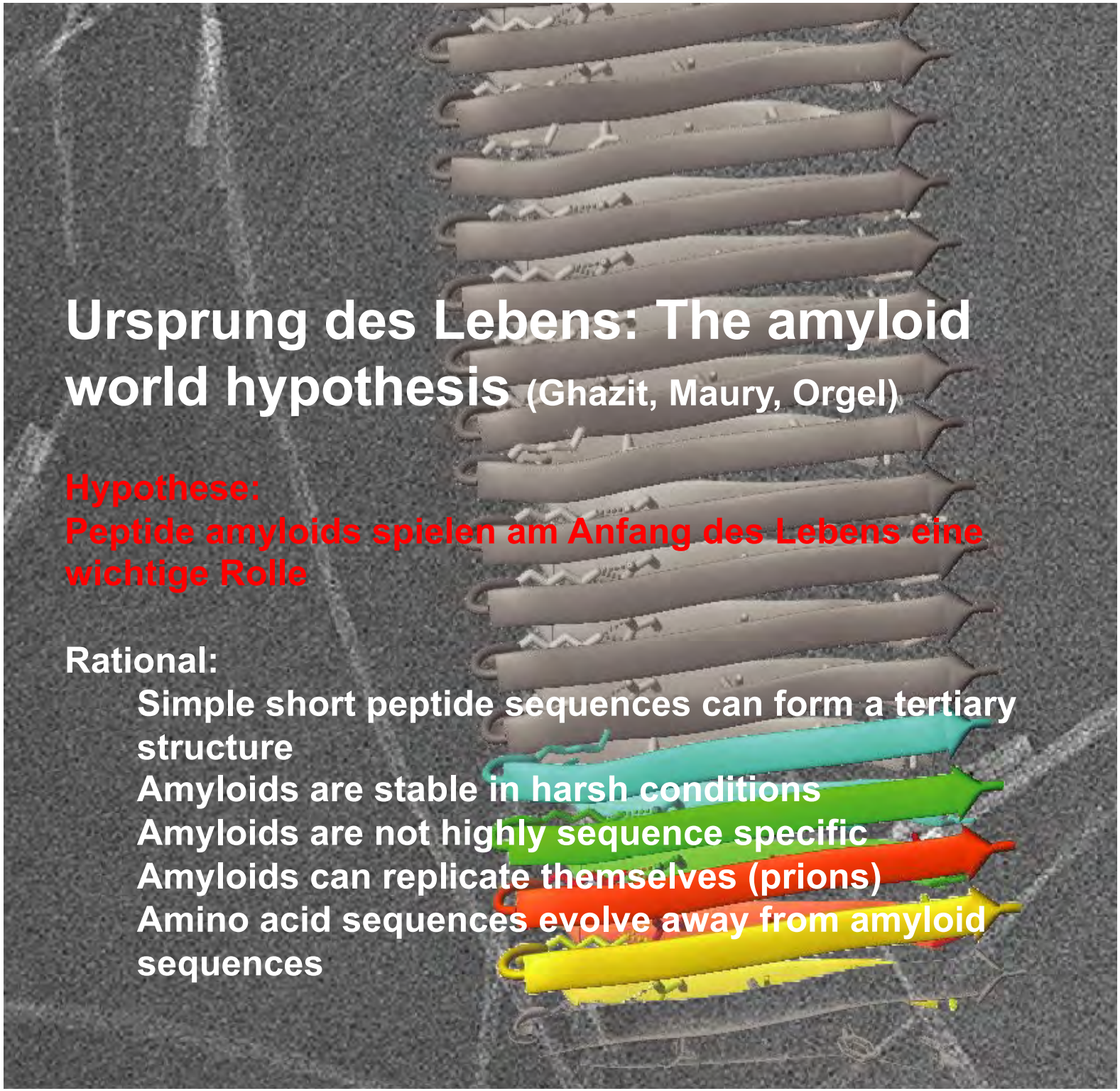

Cell membrane
Golgi
Granule


Cell membrane
Golgi


Cell membrane
Golgi


Cell membrane
Golgi
Action


pH 7.5
cell membrane
Golgi
secretory
granule
pH 5.5
Phosphate



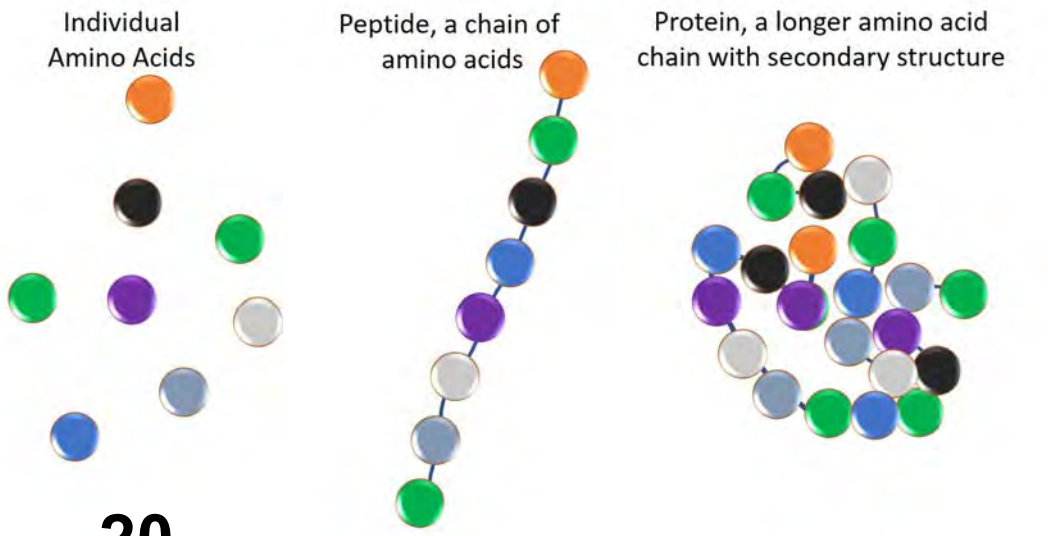
Ursprung des Lebens: The amyloid world hypothesis (Ghazit, Maury, Orgel)

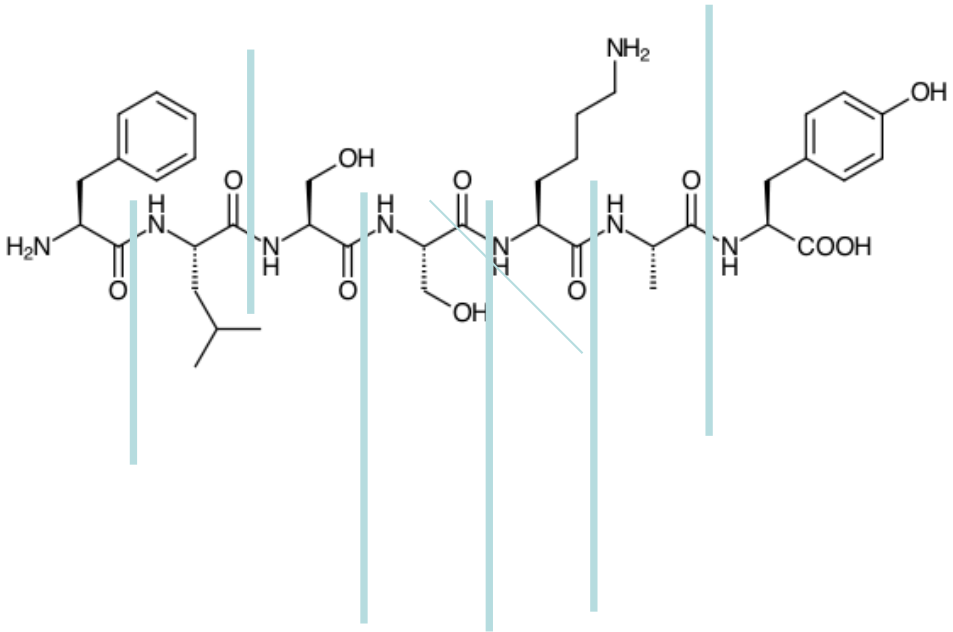
Hypothese:
Peptide amyloids spielen am Anfang des Lebens eine wichtige Rolle

Rational:

- Simple short peptide sequences can form a tertiary structure
- Amyloids are stable in harsh conditions
- Amyloids are not highly sequence specific
- Amyloids can replicate themselves (prions)
- Amino acid sequences evolve away from amyloid sequences

Peptide and proteins

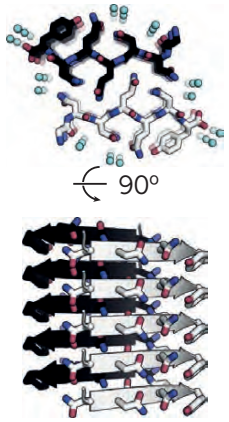




Peptide Amyloids

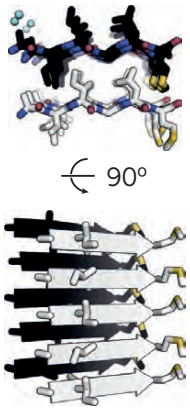
Class 1

GNNQQNY, yeast prion Sup35 (1YJP)



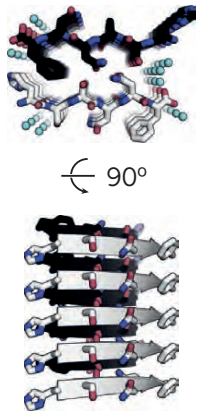
Class 2

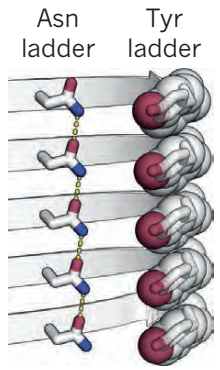
AIIGLM, amyloid-β peptide (2Y3J)



Class 4

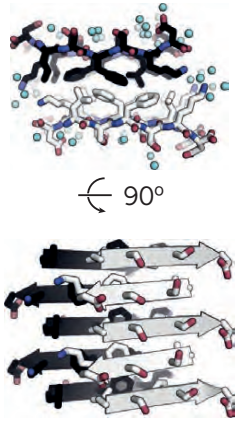
HSSNNF, IAPP (3FPO)





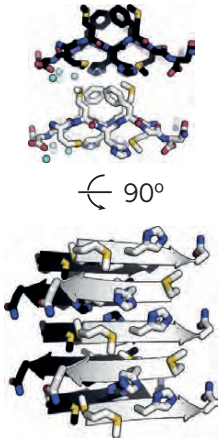
Class 5

LSFSKD, β_2 microglobulin (3LOZ)



Class 6

MMHFGN, Syrian hamster prion (3NVE)

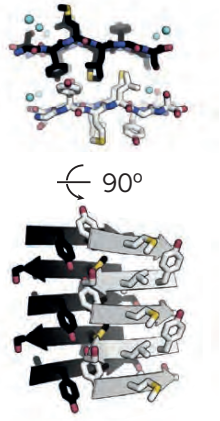


Class 7

MVGGVVIA, amyloid-β peptide (2Y3K)

Class 8

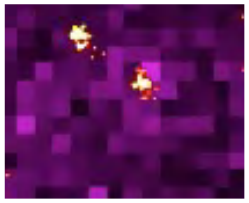
GYMLGS, human prion (3NHC)

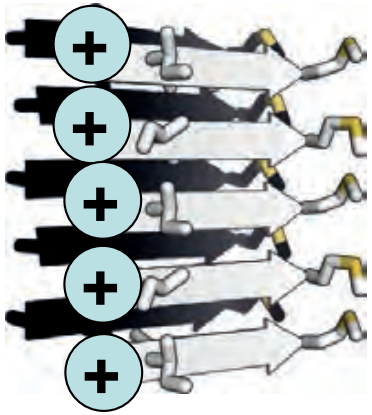


Eisenberg & Riek, Nature 2016

Eigenschaften von Amyloiden

- **Kurze Peptide machen eine 3D Struktur**
- **Amyloide sind stabil**
- **Interaktionen werden gestärkt durch Kooperativität**





A first important question is what (if any) type of sequence selection is made in the synthesis process.

Number of unique sequences: $N=A^n$ (A = # of amino acid types, n = length of peptide)

4 amino acids, peptide length 6 **N ~ 4k**

4 amino acids, peptide length 10 **N ~ 1M**

20 amino acids, peptide length 100 **$N = 20^{100}$**

Mass calculation for a single copy of each possible 100-mer:

20^{100} x 10000 g mol^{-1} / 6.022 x 10^{23} mol^{-1} = 3 x 10^{81} g

Mass of observable universe = 3 x 10^{55} g

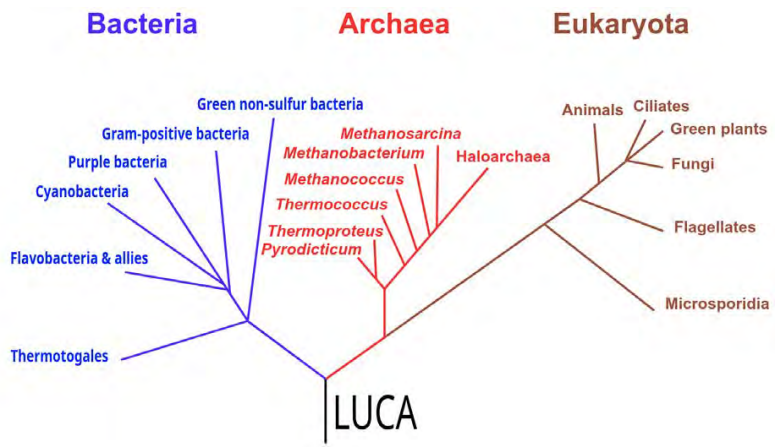
Only enough mass to make a single copy of 1 in 10^{26} sequences !!!!

Putting this into perspective for life on Earth:

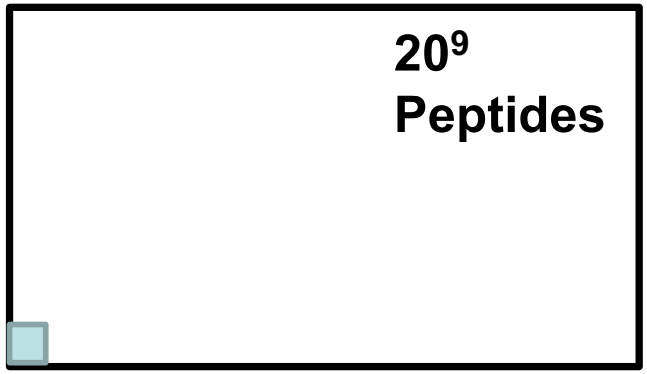
If the entire Mass of the earth contributed to making a single copy of every 100mer, only 1 in 5e53 of them could be made once.

Relative to an object the size of the Earth, the size of the space that has been searched is a 1/50th of a hydrogen atom.

Overall sequence Coverage



- Sequence coverage of 6, 7, 8, 9-mers: 99.8, 79.8, 20.9, 1.7%

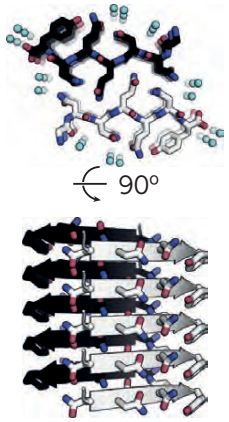


(9aa 512 billions arrangements)

Peptide Amyloids

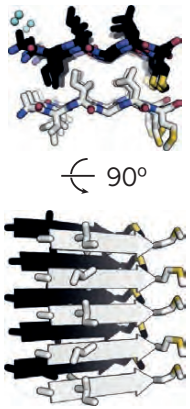
Class 1

GNNQQNY, yeast prion Sup35 (1YJP)



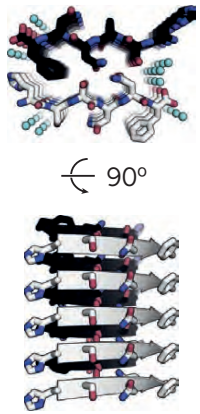
Class 2

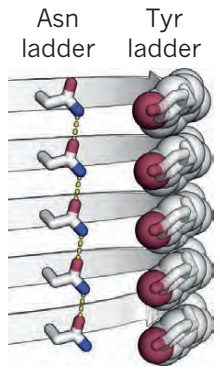
AIIGLM, amyloid-β peptide (2Y3J)



Class 4

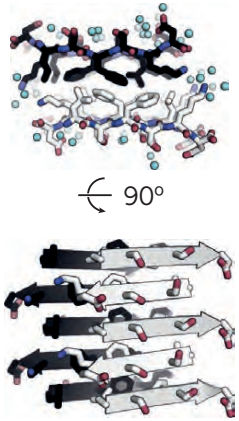
HSSNNF, IAPP (3FPO)





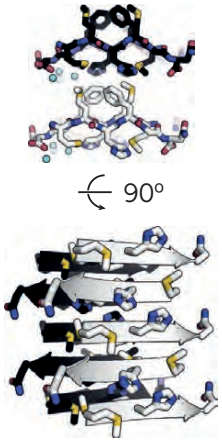
Class 5

LSFSKD, β_2 microglobulin (3LOZ)



Class 6

MMHFGN, Syrian hamster prion (3NVE)

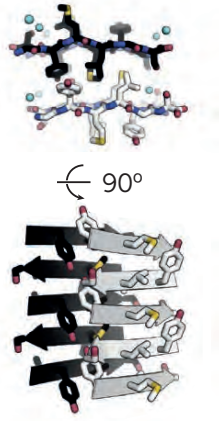


Class 7

MVGGVVIA, amyloid-β peptide (2Y3K)

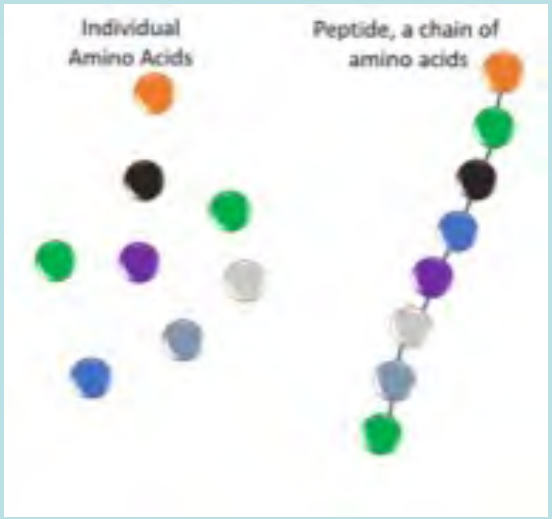
Class 8

GYMLGS, human prion (3NHC)



Eisenberg & Riek, Nature 2016

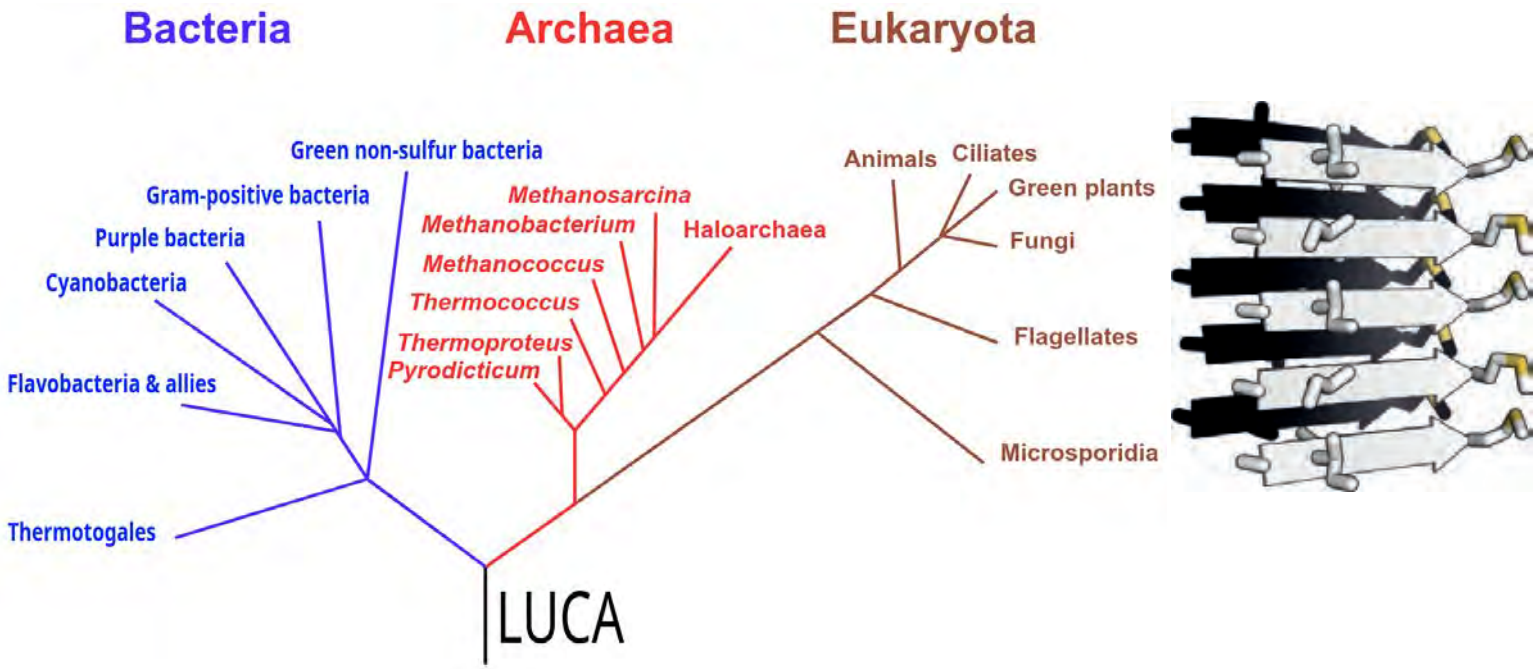
Screening the Protein Sequence Universe for amyloid presence (top down)



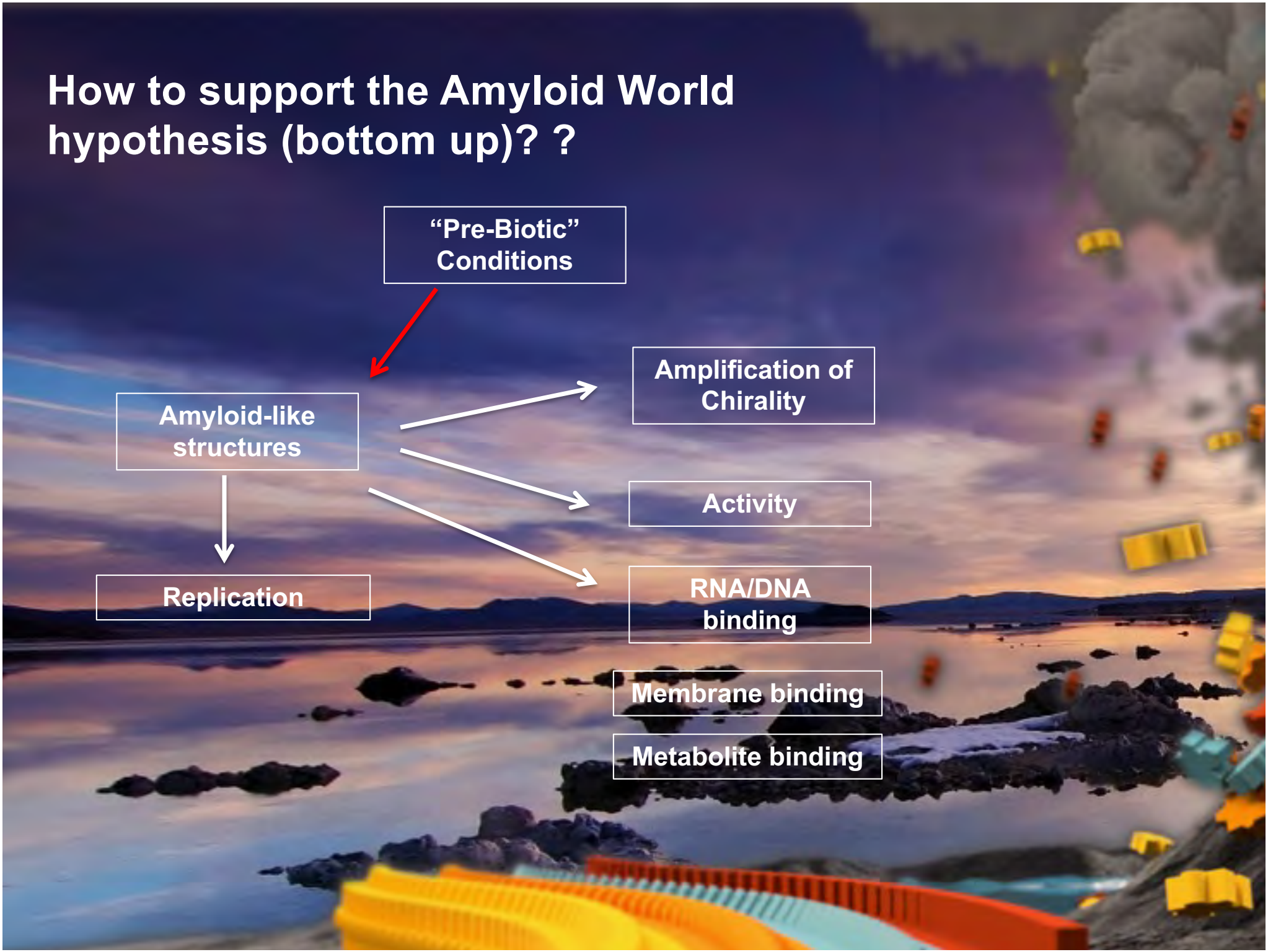
Amyloids

Take all proteomes known and study the amyloidogenecity of peptides

- Mehr amyloid Sequenzen als statistisch arbiträr
- Bakterien/Archaea haben mehr amyloidogene Sequenzen als Eukaryoten



How to support the Amyloid World hypothesis (bottom up)? ?



Our Scratch: Amino Acids

Presence of Amino acids in Meteorites -> Selecting: Gly, Ala, Asp, Val

- Composition

	M1	*M2*	*M3*	*I1*	*A1*
G - Gly	1.00	1.0	1.000	1.000	1.000
A - Ala	0.34	0.4	0.380	0.293	0.540
D - Asp	0.19	0.5	0.035	0.022	0.006
E - Glu	0.40	0.5	0.110	0.000	0.010
V - Val	0.19	0.3	0.100	0.012	0.000
S - Ser	0.00	0.0	0.003	0.072	0.000
I - Ile	0.13	0.0	0.060	0.000	0.000
L - Leu	0.04	0.0	0.035	0.000	0.000
P - Pro	0.29	0.1	0.000	0.001	0.000
T - Thr	0.00	0.0	0.003	0.000	0.000
K - Lys	0.00	0.0	0.000	0.000	0.000
F - Phe	0.00	0.0	0.000	0.000	0.000
R - Arg	0.00	0.0	0.000	0.000	0.000
H - His	0.00	0.0	0.000	0.000	0.000
NQCYMW	0.00	0.0	0.000	0.000	0.000

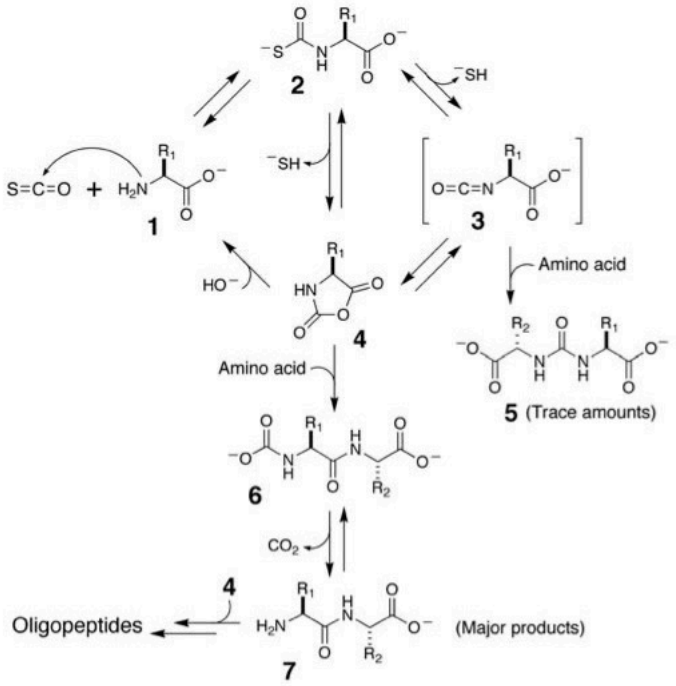
Higgs and Pudritz, 2009

Second requirement is the peptide synthesis

Carbonyl Sulfide–Mediated Prebiotic Formation of Peptides

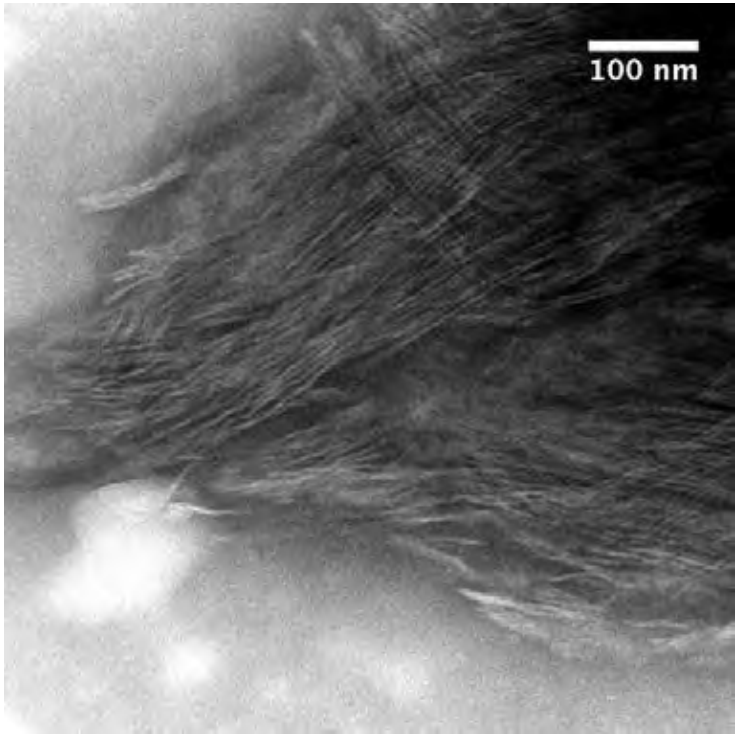
Luke Leman,[1] Leslie Orgel,[2] M. Reza Ghadiri[1]*

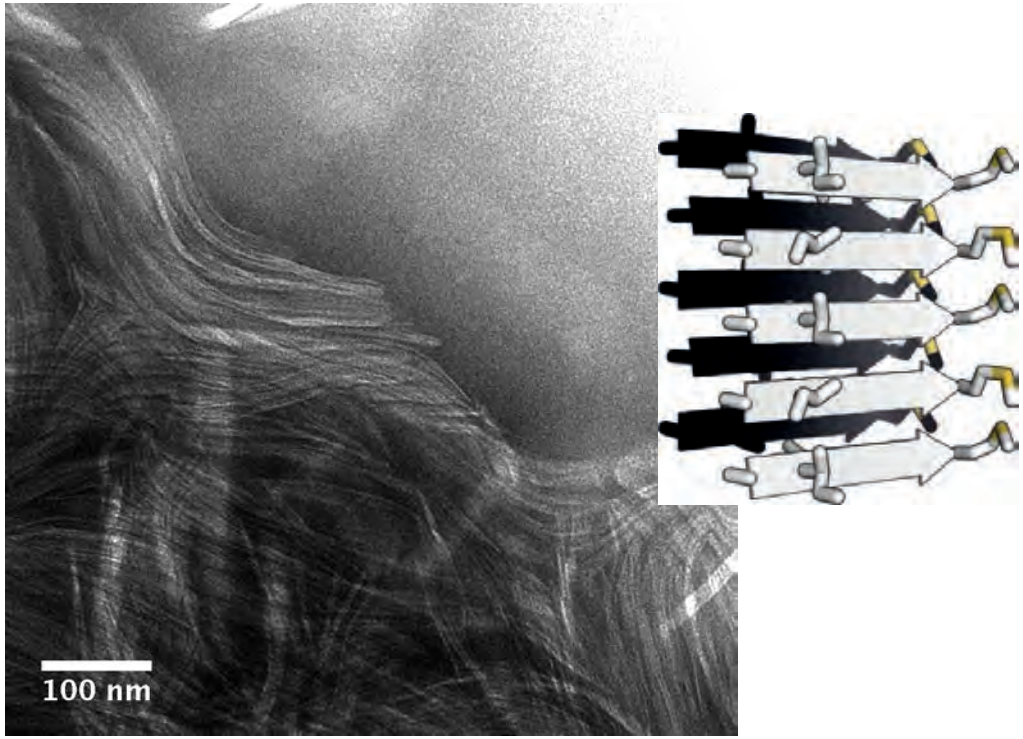
Science **306 283-286**



Poly Ala and Val from precipitate of COS mediated condensation (from 10 mM Ala/Val)

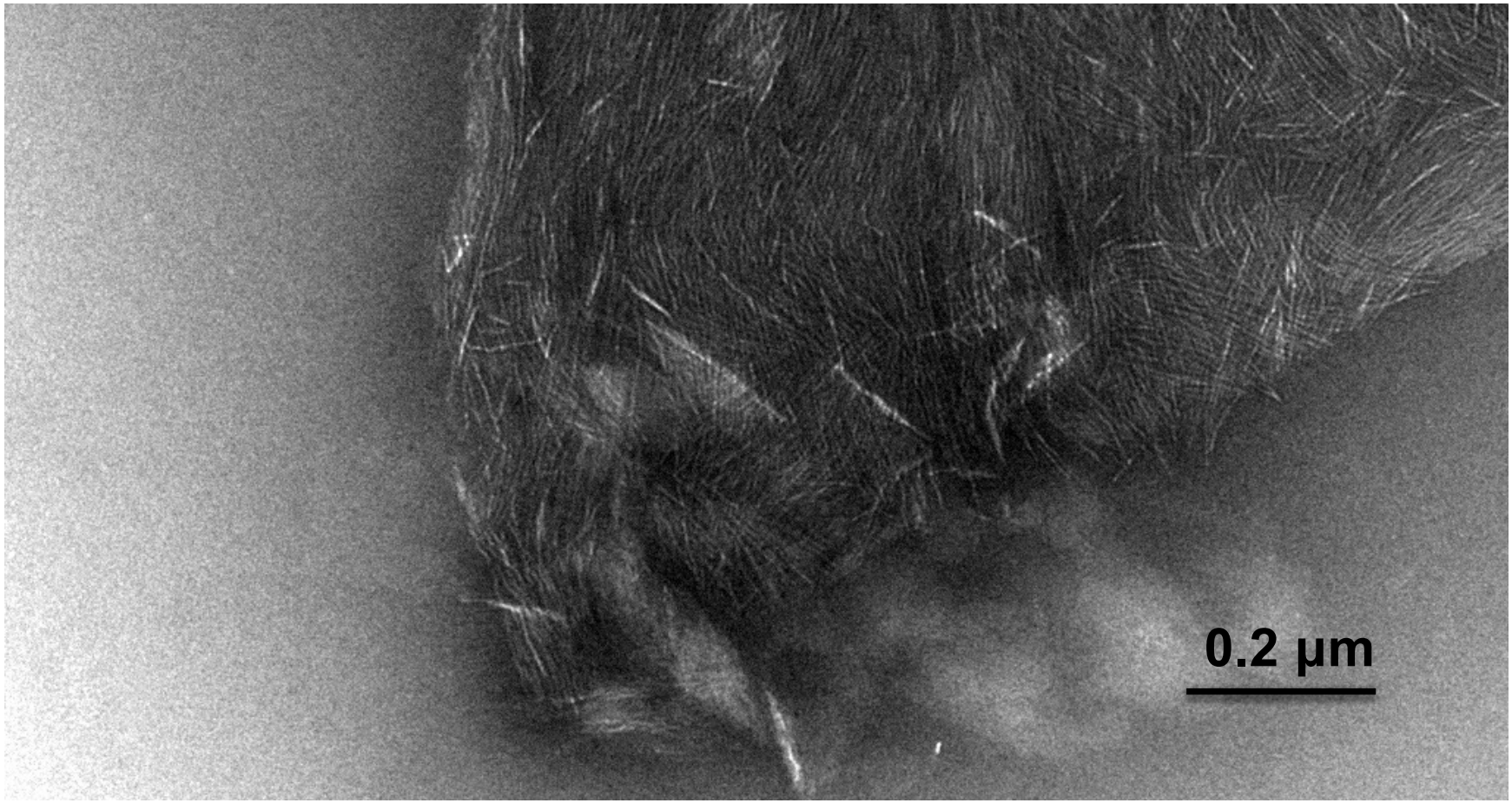
(Greenwald, Riek et al., Angewandte Chemie 2015)





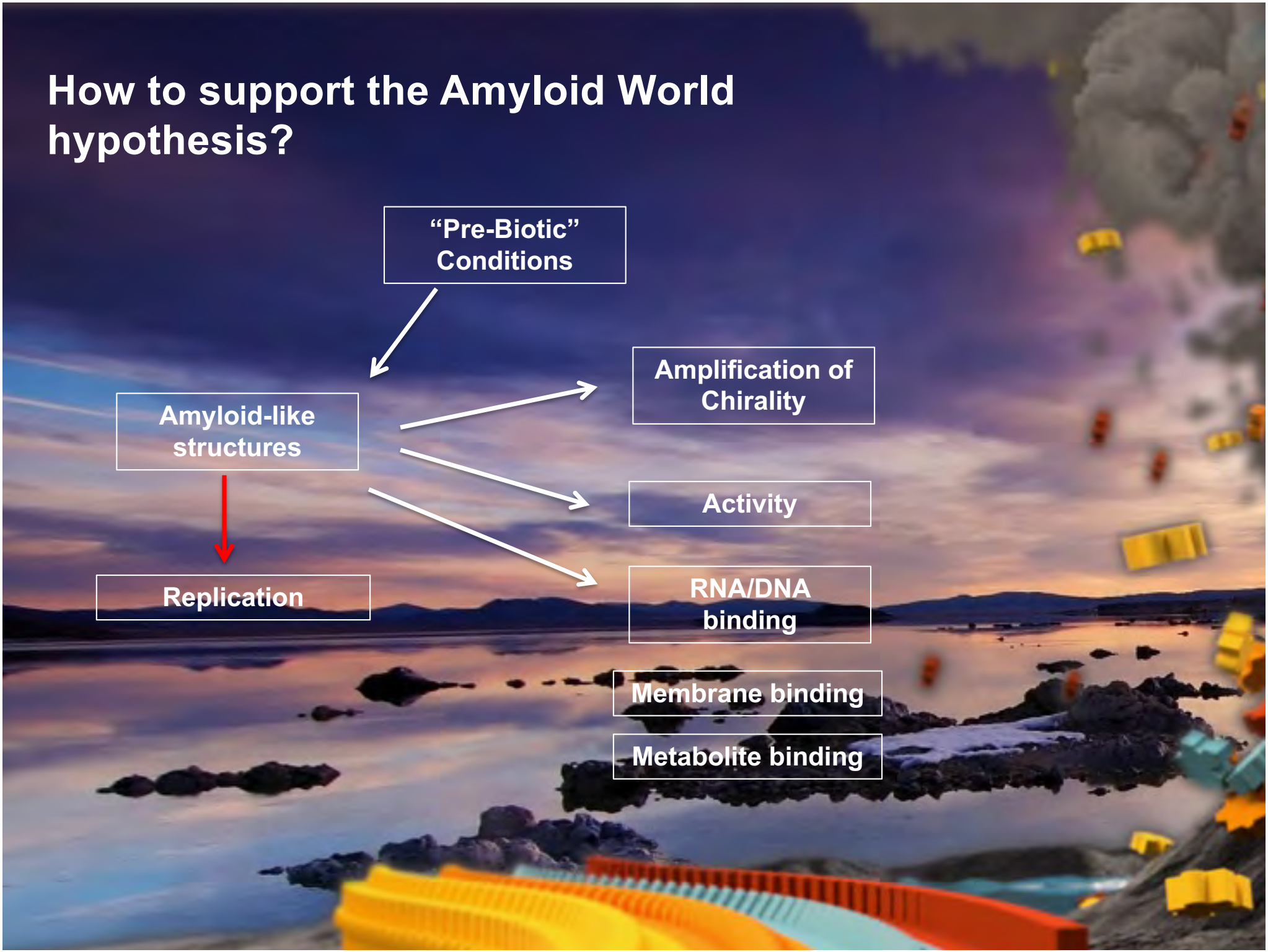
4-component co-polymerization
10 mM L-alanine, 5 mM L-valine, 2 mM L-glycine, 1 mM L-aspartate

CDI (Carbonyldimidazole) activated amino acids in H_2O



~5 nm wide fibrils

How to support the Amyloid World hypothesis?



Template-Assisted Chemical Replication (using CDI or COS activated amino acids)

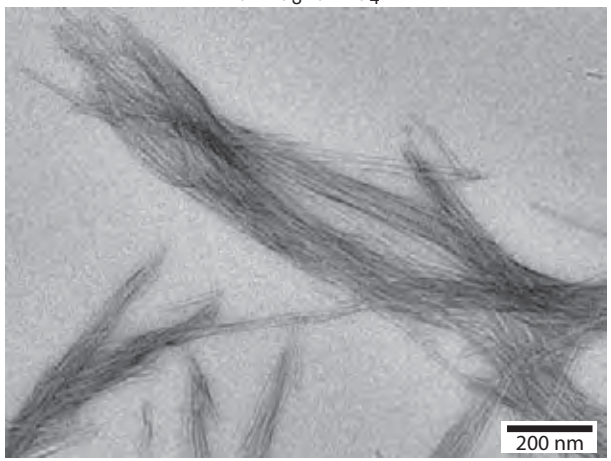
(Rout, Riek, Greenwald et al., Nat. Comm. 2018
Bomba, Riek, Greenwald.et al., Orig. of Life Evol Biosph 2019)

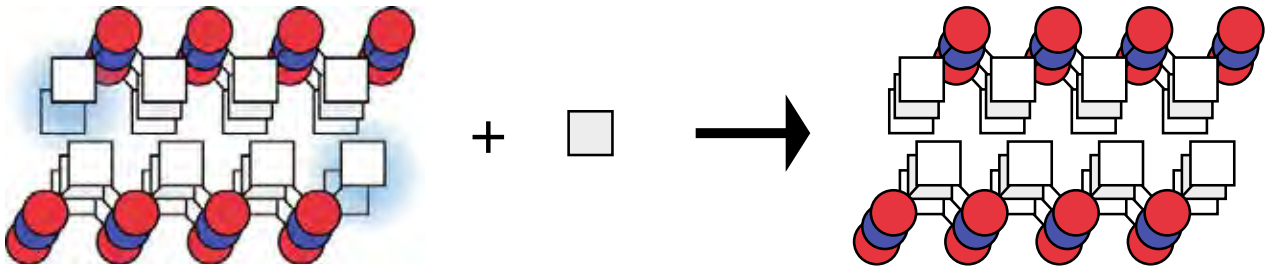
Ac-FEFEFEFE-NH_2 (Template)

RFRFRFR-NH_2

RFRFRFR-NH_2 (Substrate)

$R(FR)_3/(FE)_4$





(CDI carbonyl diimidazole amino acids)

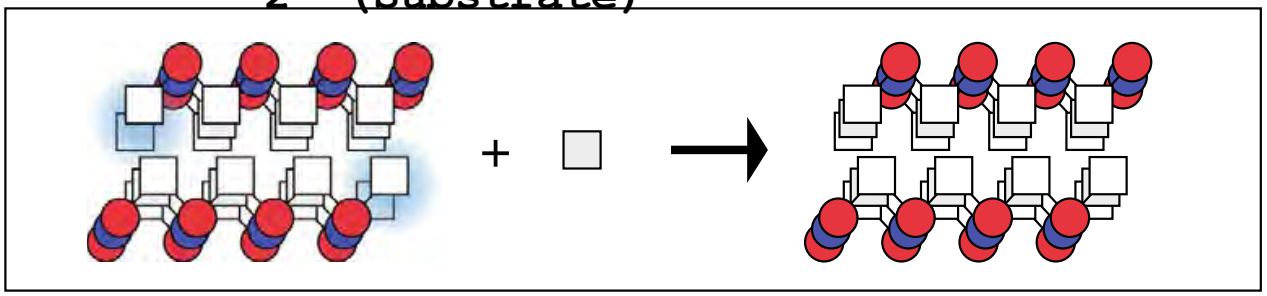
Template-Assisted Chemical Replication (using CDI or COS activated amino acids)

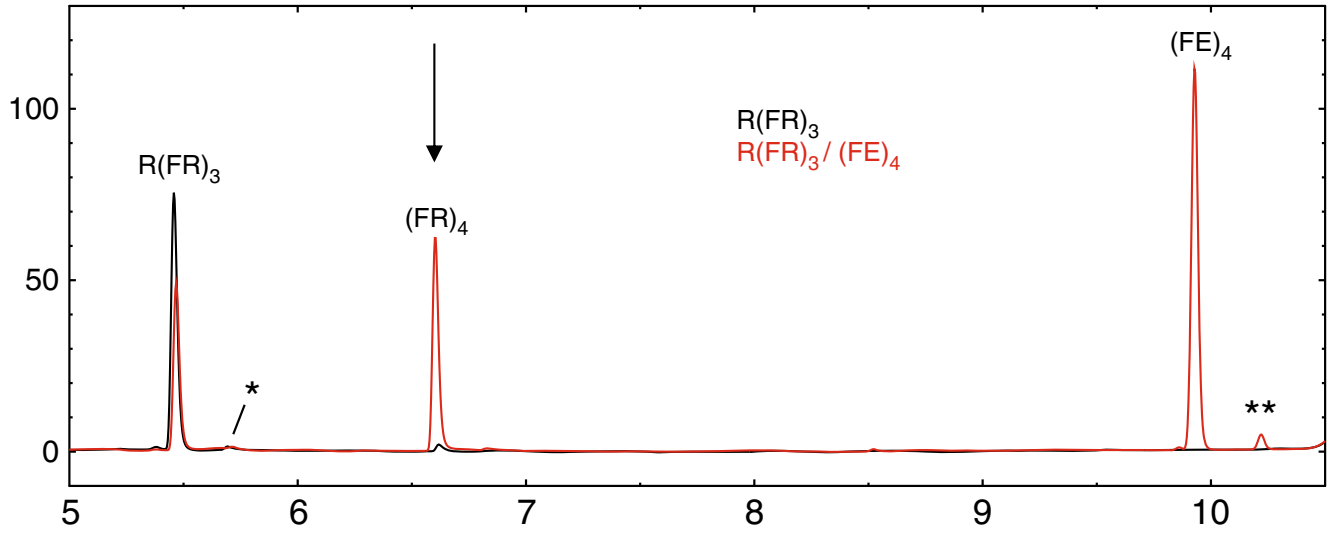
(Rout, Riek, Greenwald et al., Nat. Comm. 2018
Bomba, Riek, Greenwald.et al., Orig. of Life Evol Biosph 2019)

$Ac\text{-}FEFEFEFE\text{-}NH_2$ (Template)

$RFRFRFR\text{-}NH_2$

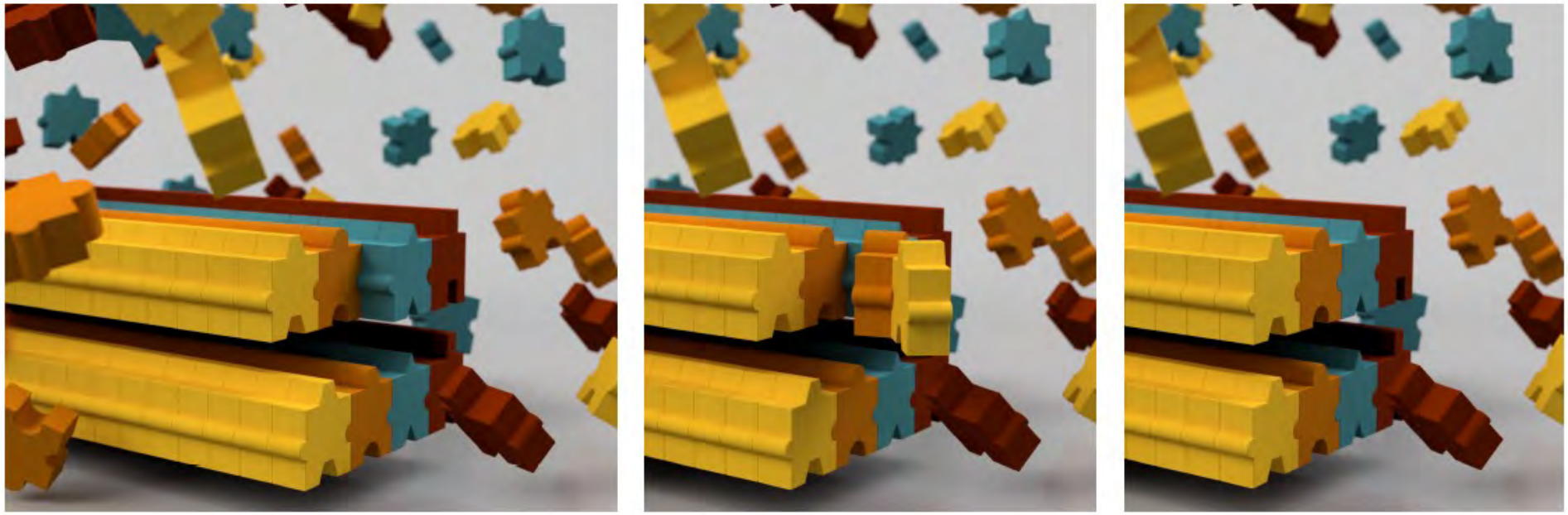
$RFRFRFR\text{-}NH_2$ (Substrate)



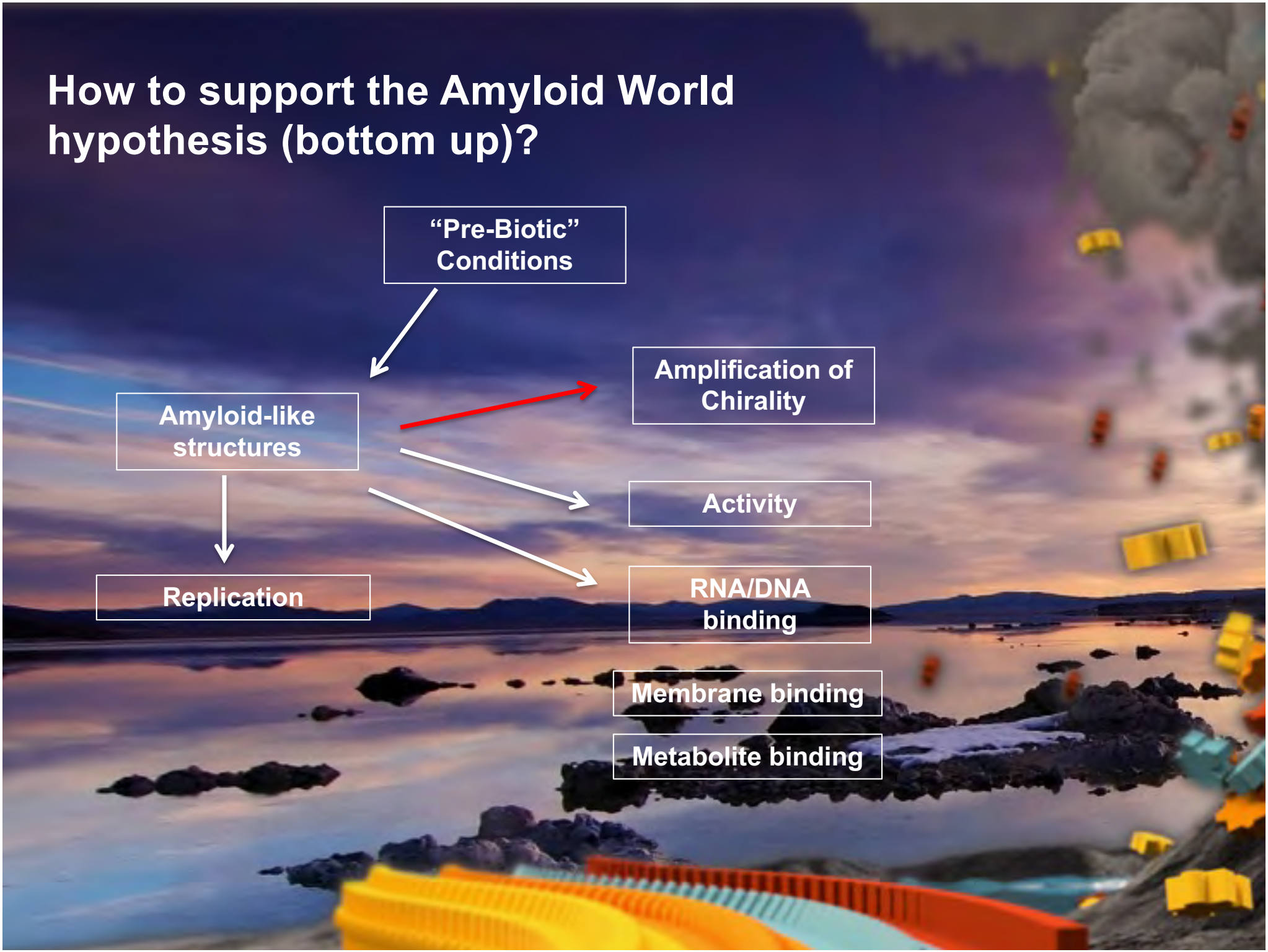


(CDI carbonyl diimidazole amino acids)

Peptide Replication within an Amyloid



How to support the Amyloid World hypothesis (bottom up)?



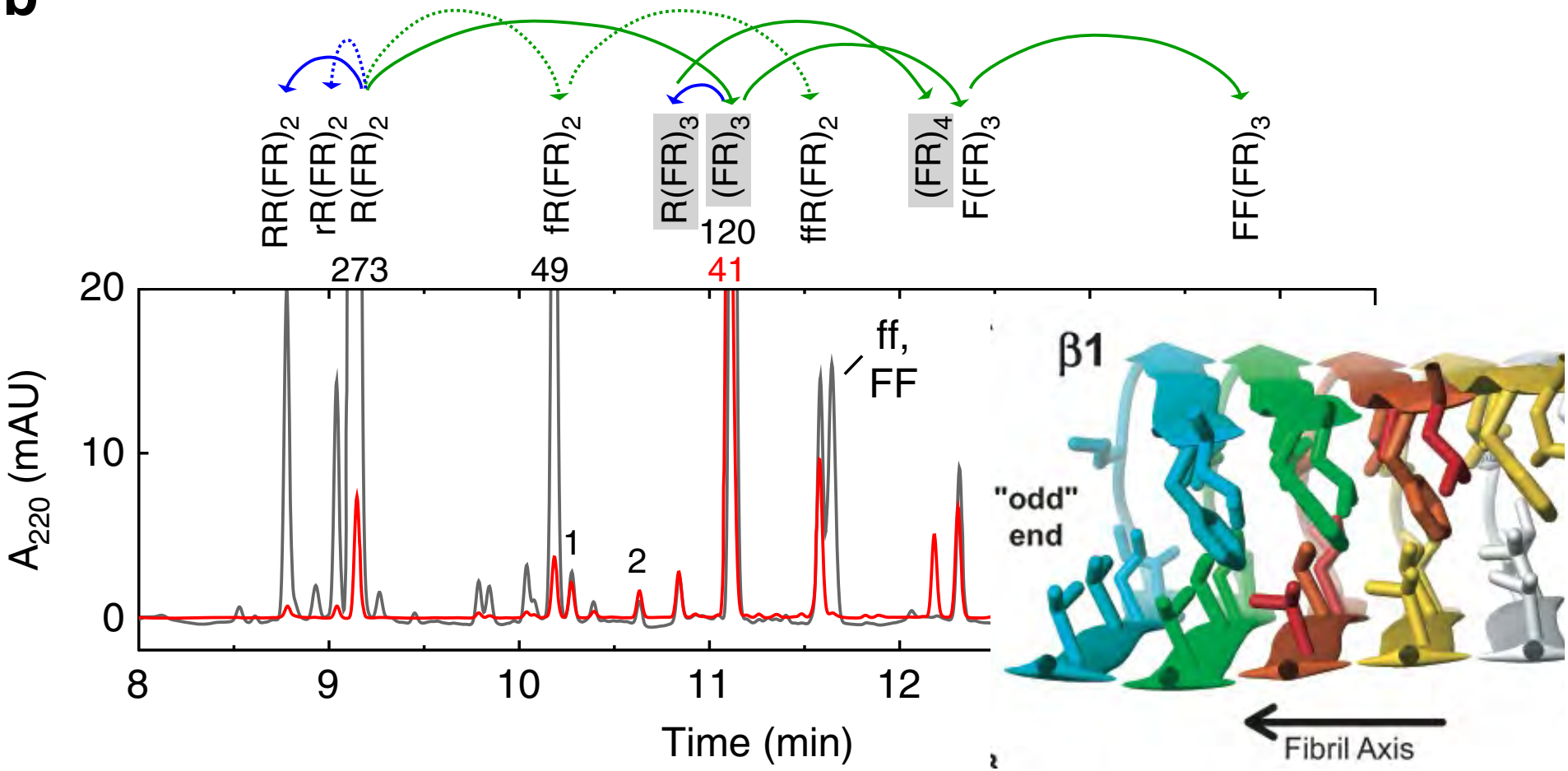
HPLC of with and without template of the reaction of CDI-activated amino acids L-Arg (R) and D-Arg(r) and L-Phe (F) and D-Phe (f)

Red: XXXRFRFR-NH_2

Ac-FEFEFEFE-NH_2

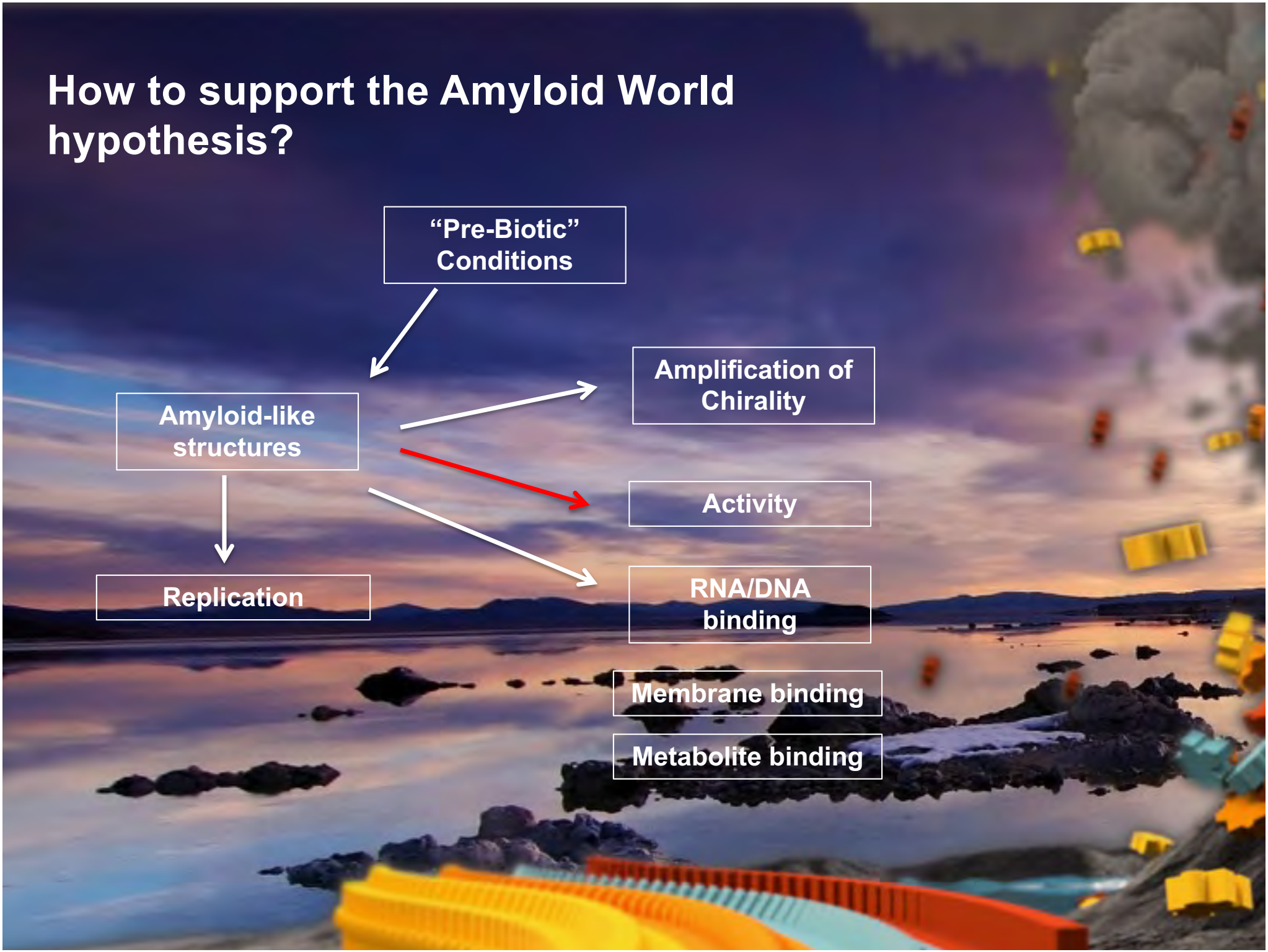
Black: XXXRFRFR-NH_2

b



100 microM substrate and template 200 microM amino acids

How to support the Amyloid World hypothesis?

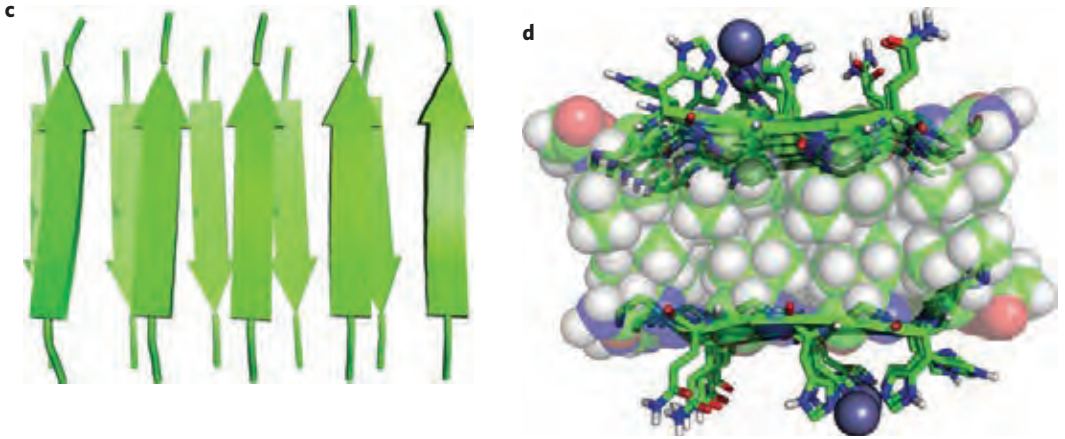


Can Amyloids by enzymatic active: Peptide of length 7 plus Y for detection Amino acid composition: V, A, S, D, and H

Table S1 and associated data: Biophysical characterization of library peptides.

	Sequence	pH 7.3 ThT[1]	CD[2]	FTIR[3]	EM[4]	pH 4.0 ThT	CD	FTIR	EM	X-ray[5]
1	Ac-YV**D**V**D**V**D**V-CONH$_2$	1	–	–		89	β^{++}	1625	+	
2	Ac-YV**S**V**D**V**D**V-CONH$_2$	24	–	1624		322	β^{++}	1624	+	cross-β
3	Ac-YV**D**V**S**V**D**V-CONH$_2$	2	–	–		120	β^{++}	1623	+	
4	Ac-YV**D**V**D**V**S**V-CONH$_2$	3	–	–		160	β^{++}	1618	+	
5	Ac-YV**S**V**S**V**D**V-CONH$_2$	322	β^{+}	1619		322	β^{++}	1618	+	
6	Ac-YV**S**V**D**V**S**V-CONH$_2$	309	β^{+}	1618		322	β^{++}	1622	+	
7	Ac-YV**D**V**S**V**S**V-CONH$_2$	208	–	–		128	β^{++}	1619	+	
8	Ac-YV**H**V**H**V**H**V-CONH$_2$	126	β^{+}	1627	+	1	–	–	–	cross-β
9	Ac-YV**S**V**H**V**H**V-CONH$_2$	82	β^{++}	1624	+	1	–	–		
10	Ac-YV**H**V**S**V**H**V-CONH$_2$	29	β^{+}	1623	+	1	–	–		
11	Ac-YV**H**V**H**V**S**V-CONH$_2$	18	β^{++}	1624	+	1	–	–		
12	Ac-YV**S**V**S**V**H**V-CONH$_2$	183	β^{++}	1623		5	?	1621	+	cross-β
13	Ac-YV**S**V**H**V**S**V-CONH$_2$	0	β^{+}	1622	+	11	β^{++}	1621		
14	Ac-YV**H**V**S**V**S**V-CONH$_2$	322	β^{+}	1624	+	2	β^{+}	1623		cross-β
15	Ac-YV**D**V**H**V**S**V-CONH$_2$	114	β^{+}	1623	+	13	β^{+}	1622	+	
16	Ac-YV**D**V**S**V**H**V-CONH$_2$	177	β^{++}	1618	+	7	β^{++}	1619	+	
17	Ac-YV**H**V**D**V**S**V-CONH$_2$	163	β^{+}	1622	+	5	β^{++}	1622		
18	Ac-YV**S**V**D**V**H**V-CONH$_2$	31	β^{++}	1622	+	1	β^{++}	1622		
19	Ac-YV**S**V**H**V**D**V-CONH$_2$	142	β^{++}	1624	+	9	β^{++}	1624		
20	Ac-YV**H**V**S**V**D**V-CONH$_2$	98	β^{+}	–	+	23	β^{++}	1622	+	
21	Ac-YV**H**V**H**V**D**V-CONH$_2$	133	β^{+}	1626	+	1	?	–	+	
22	Ac-YV**H**V**D**V**H**V-CONH$_2$	44	β^{+}	1624	+	1	?	–	+	
23	Ac-YV**D**V**H**V**H**V-CONH$_2$	21	β^{+}	1624	+	1	–	1626	+	
24	Ac-YV**D**V**D**V**H**V-CONH$_2$	2	–	–	+	1	β^{++}	1626	+	
25	Ac-YV**D**V**H**V**D**V-CONH$_2$	2	–	–		9	β^{++}	1623	+	
26	Ac-YV**H**V**D**V**D**V-CONH$_2$	5	–	–		5	β^{++}	1624	+	
27	Ac-YV**A**V**H**V**H**V-CONH$_2$	129	β^{+}	1625	+	1	–	–		cross-β
28	Ac-YV**H**V**A**V**H**V-CONH$_2$	114	β^{+}	1624	+	1	–	–		
29	Ac-YV**D**V**H**V**A**V-CONH$_2$	183	β^{+}	1622	+	25	?	1618		
30	Ac-YV**A**V**D**V**H**V-CONH$_2$	42	–	1623	+	10	β^{++}	1624		
31	Ac-YV**H**V**D**V**A**V-CONH$_2$	259	β^{+}	1622		32	β^{++}	1621	+	
32	Ac-YV**H**V**A**V**D**V-CONH$_2$	322	β^{+}	1621		111	β^{++}	1621	+	
33	Ac-YV**A**V**H**V**D**V-CONH$_2$	67	β^{+}	1624		2	β^{++}	1623	+	

Ac-YV**A**V**H**V**H**V-CONH2



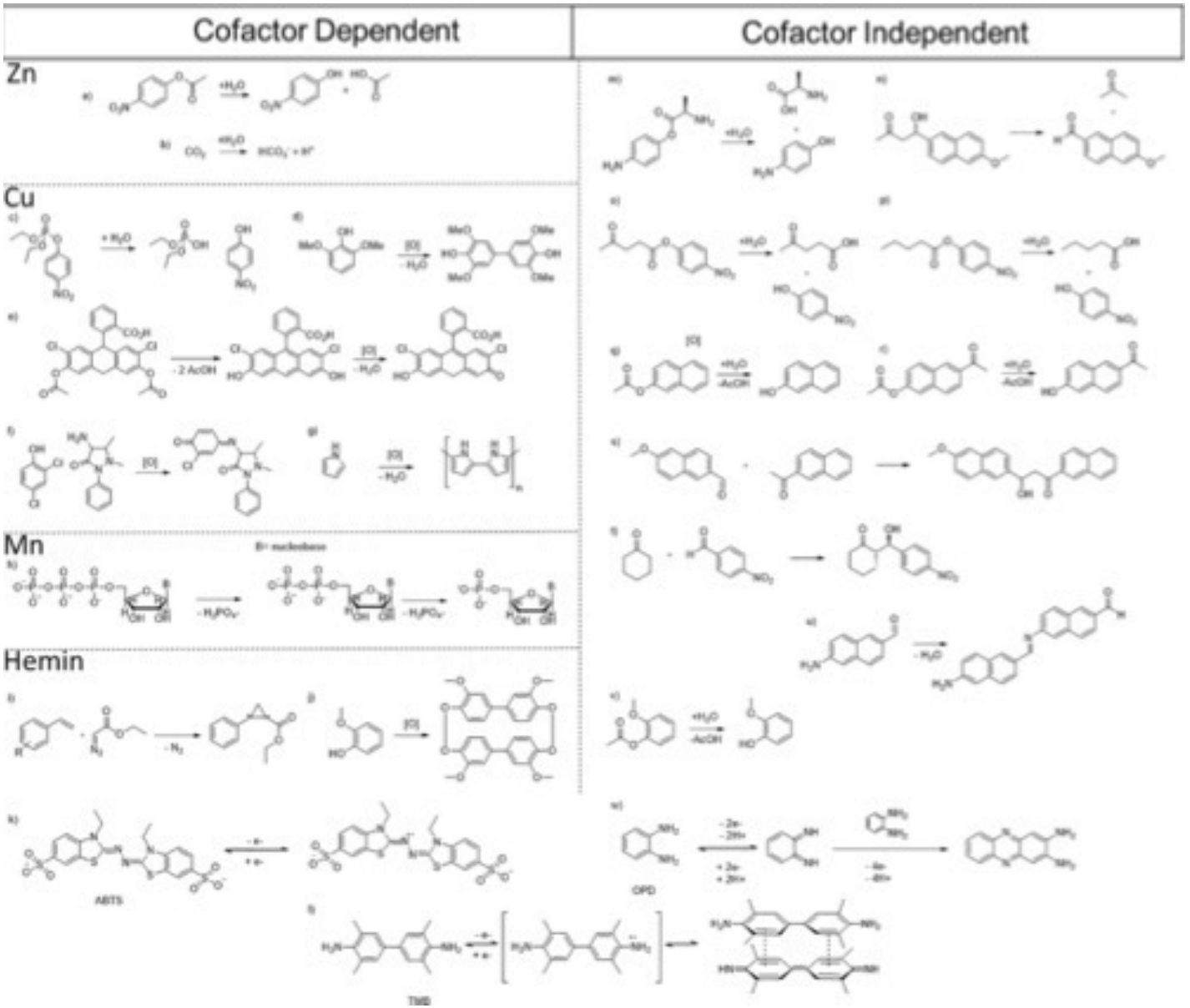
Short peptides self-assemble to produce catalytic amyloids

Caroline M. Rufo[1†], Yurii S. Moroz[1†], Olesia V. Moroz[1†], Jan Stöhr[2], Tyler A. Smith[1], Xiaozhen H
William F. DeGrado[3*] and Ivan V. Korendovych[1*]

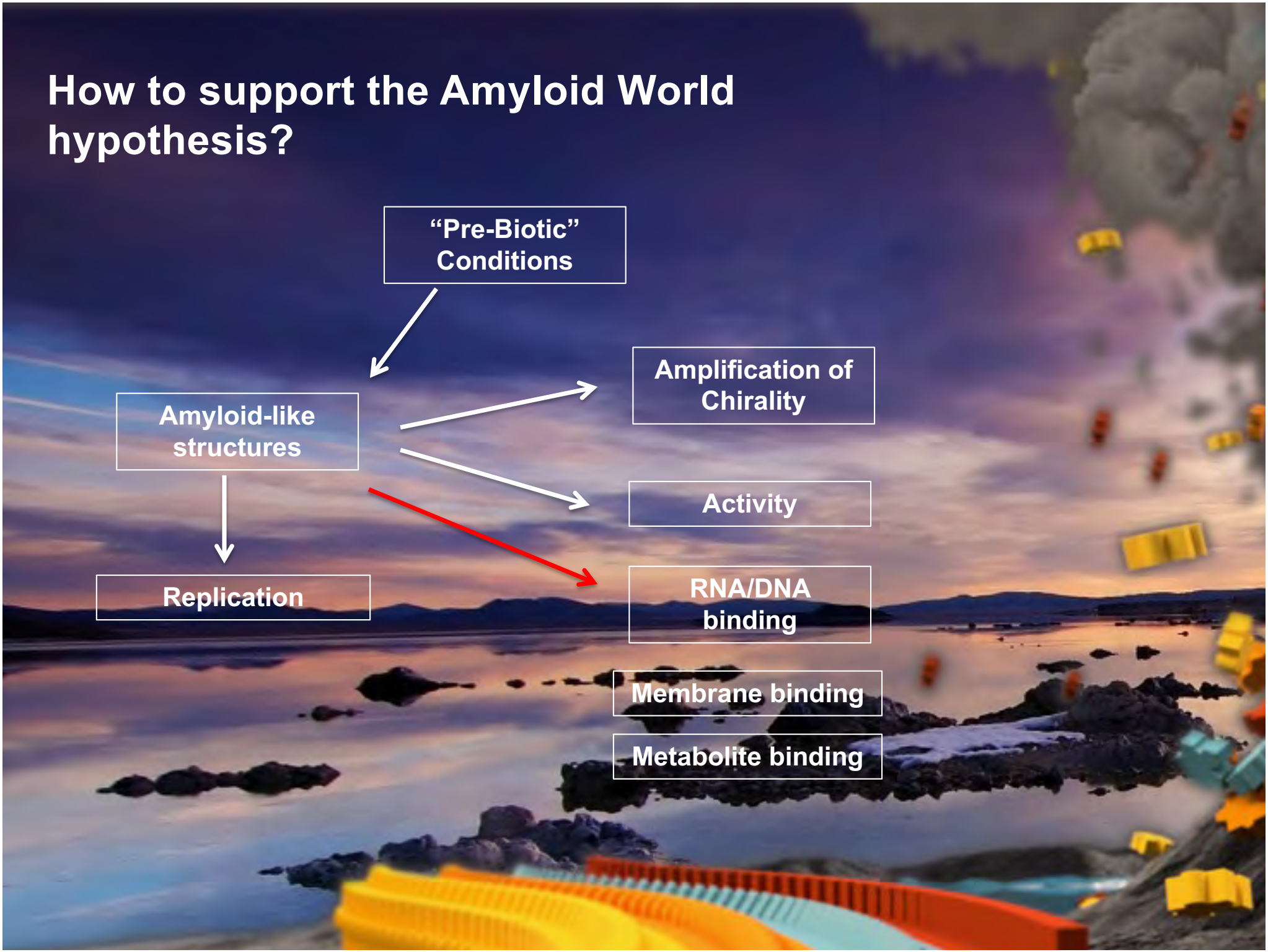
Nat. Chem. 2014

(Friedmann, Riek et al. PLoS ONE, 2015)

Many catalytic Activities of Peptide Amyloids (Marshall and Korendovych, 2021)

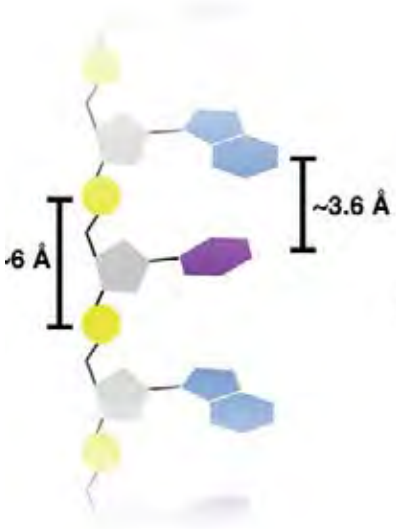


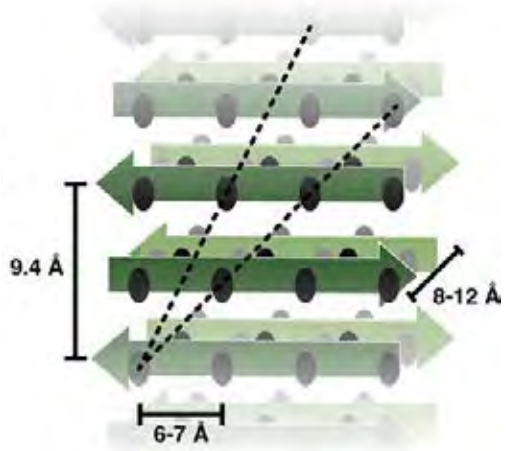
How to support the Amyloid World hypothesis?

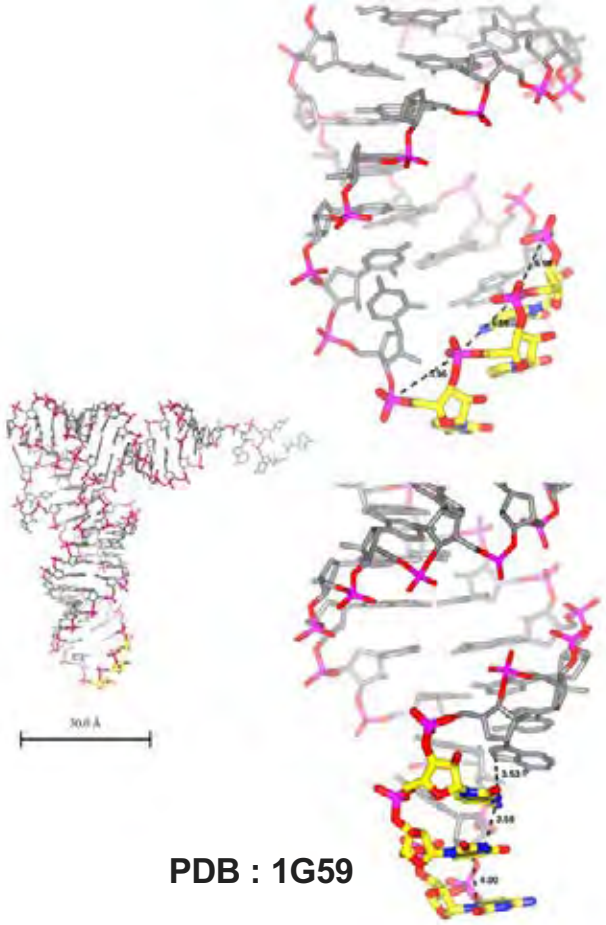


(III) Amyloid and RNA

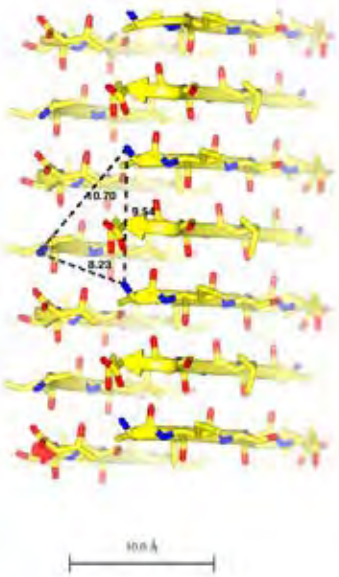
- **Both Amyloids and RNA are repetitive structures.**







PDB : 1G59

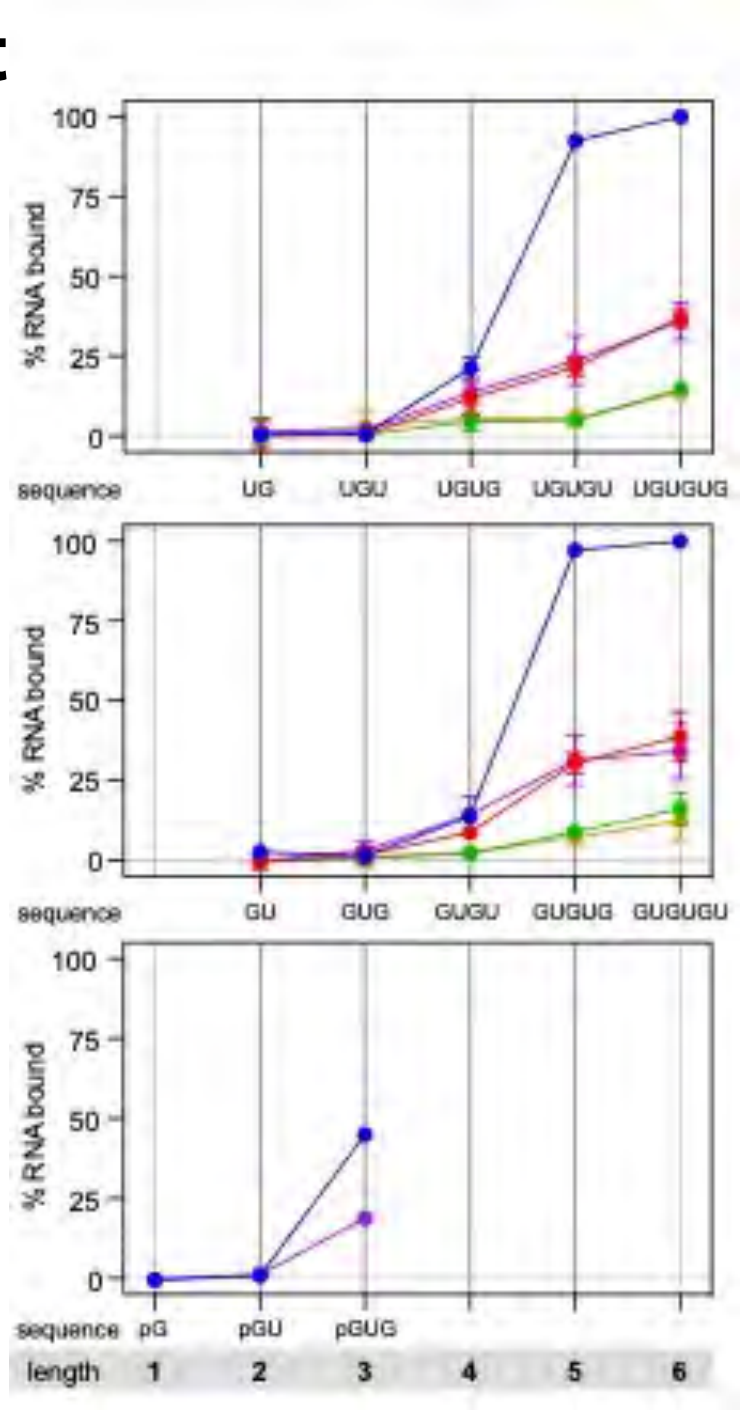


PDB : 3FOD

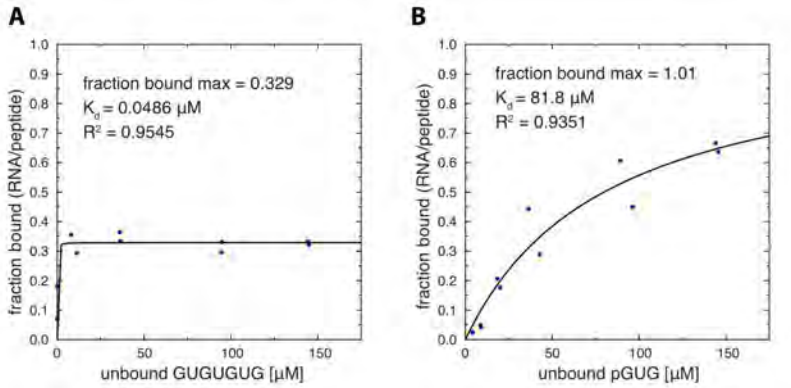
Hypothesis:

Amyloid-RNA interactions can shed light on the origin of the genetic code.

Nt binding of RNA

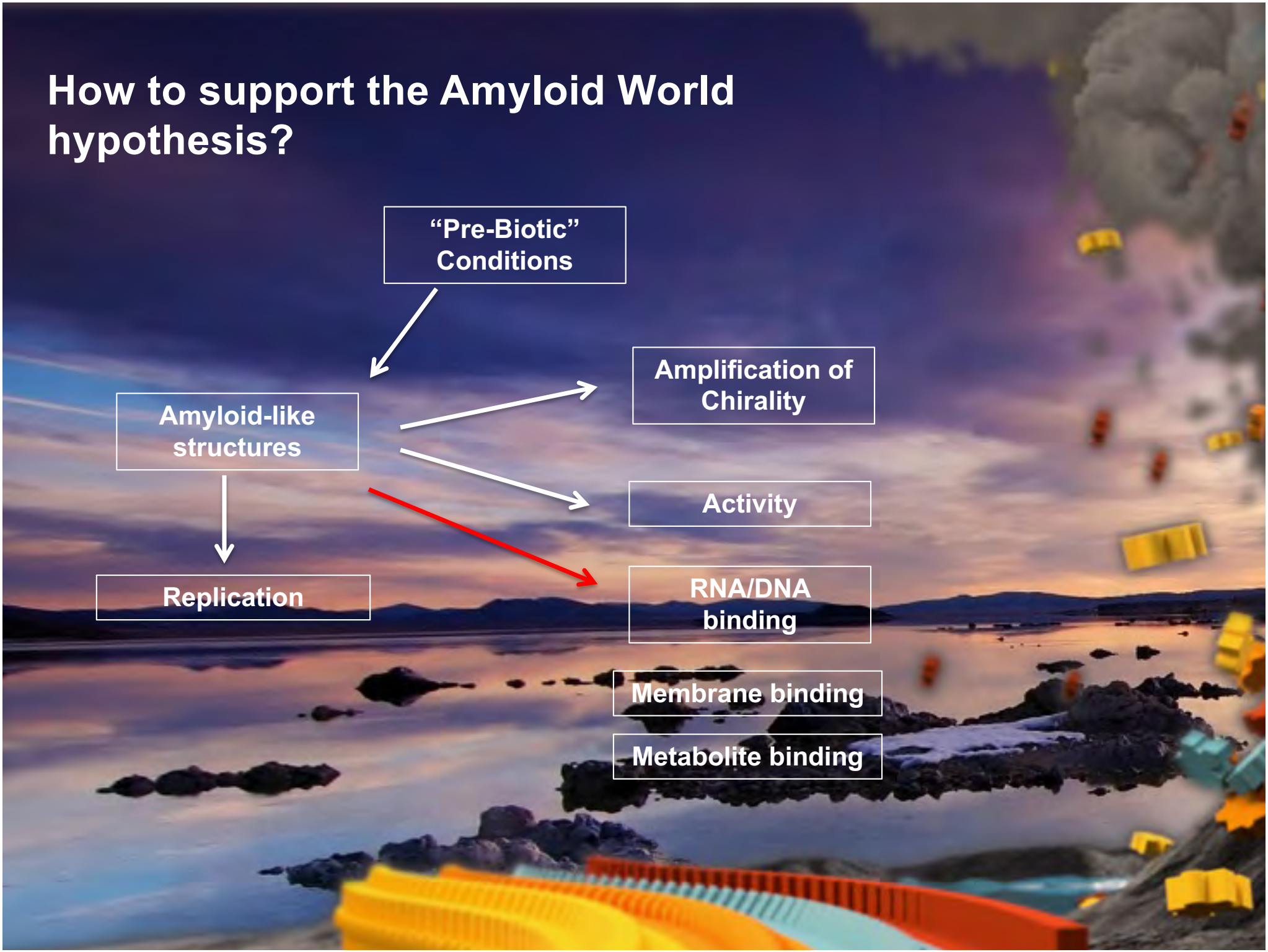


100 uM peptide
50 uM RNA Oligo
Citrate-Phosphate buffer



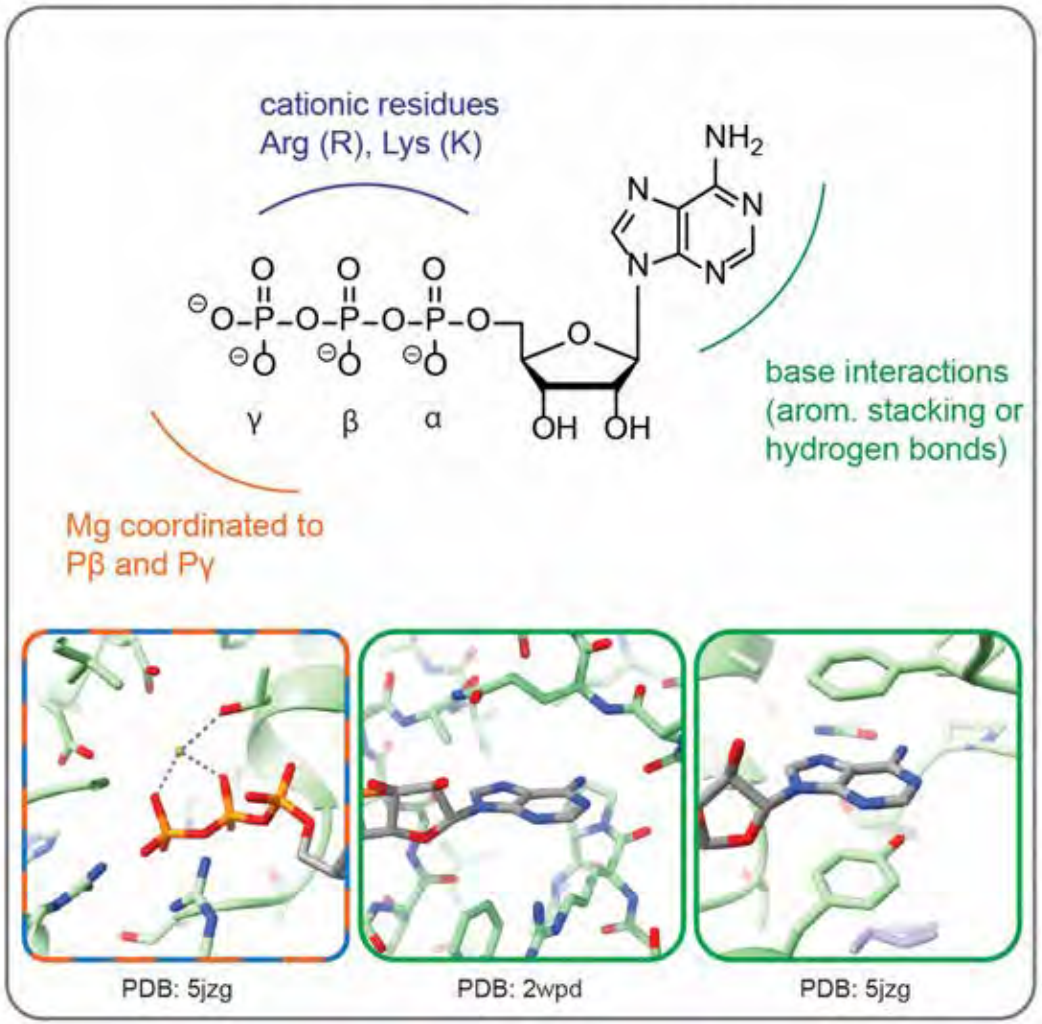
No.	Composition [a]	Ionization at neutral pH [b]			
		N	Sidechain	C	Sum
1	FVF**E**FQFQ	+	-	-	-1
2	FVFQFQFQ	+		-	0
3	FVFQFQFQ-NH_2	+			+1
4	F**E**F**E**F**K**F**K**	+	- - + +	-	0
5	F**E**F**E**FQFQ	+	- -	-	-2
6	Ac-F**E**F**E**FQFQ		- -	-	-3
7	VF**E**F**E**F**K**F**K**	+	- - + +	-	0
8	VF**E**F**E**FQFQ	+	- -	-	-2
9	FVF**E**FQFV	+	-	-	-1
10	**D**F**E**F**E**FQFQ	+	- - -	-	-3
11	F**D**F**E**FQF**D**	+	- - -	-	-3

How to support the Amyloid World hypothesis?

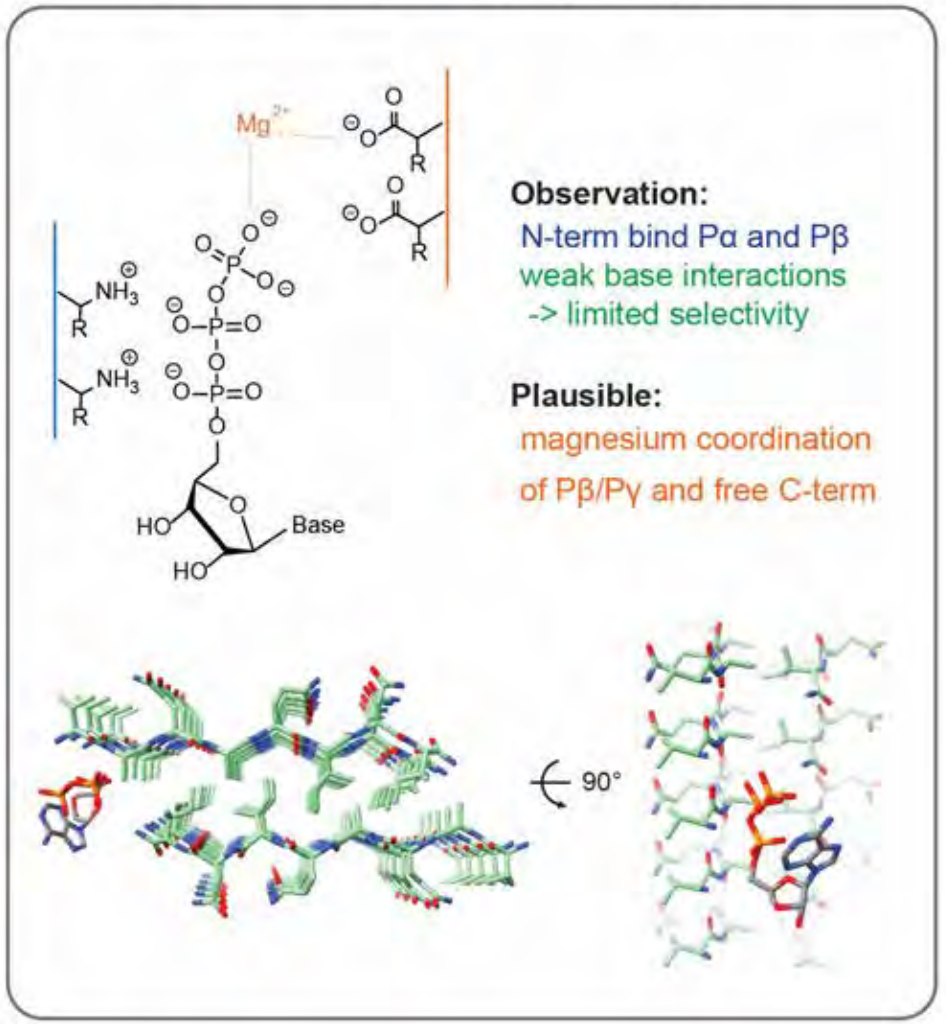


Ancient Canonical Motif

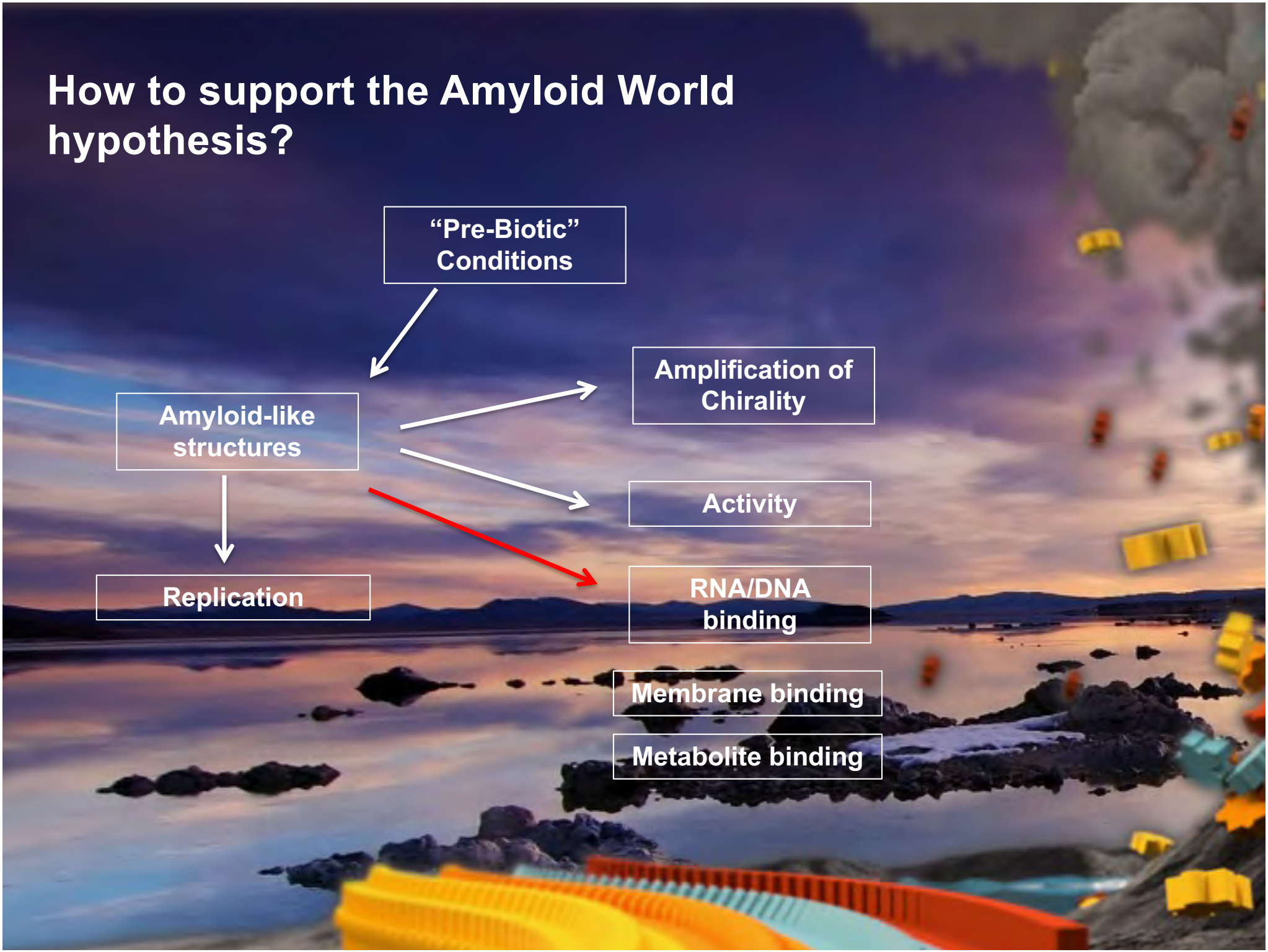
ATP binding to proteins



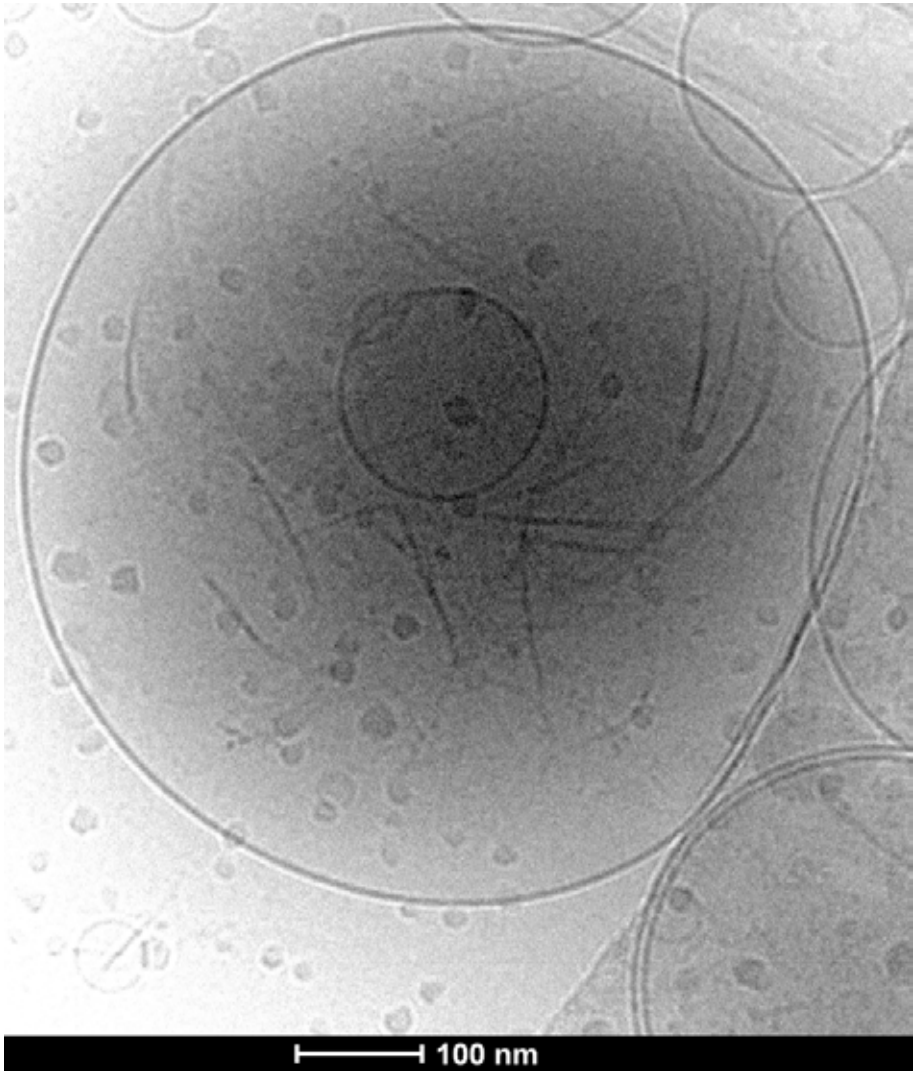
NMR structure: ATP-Amyloid

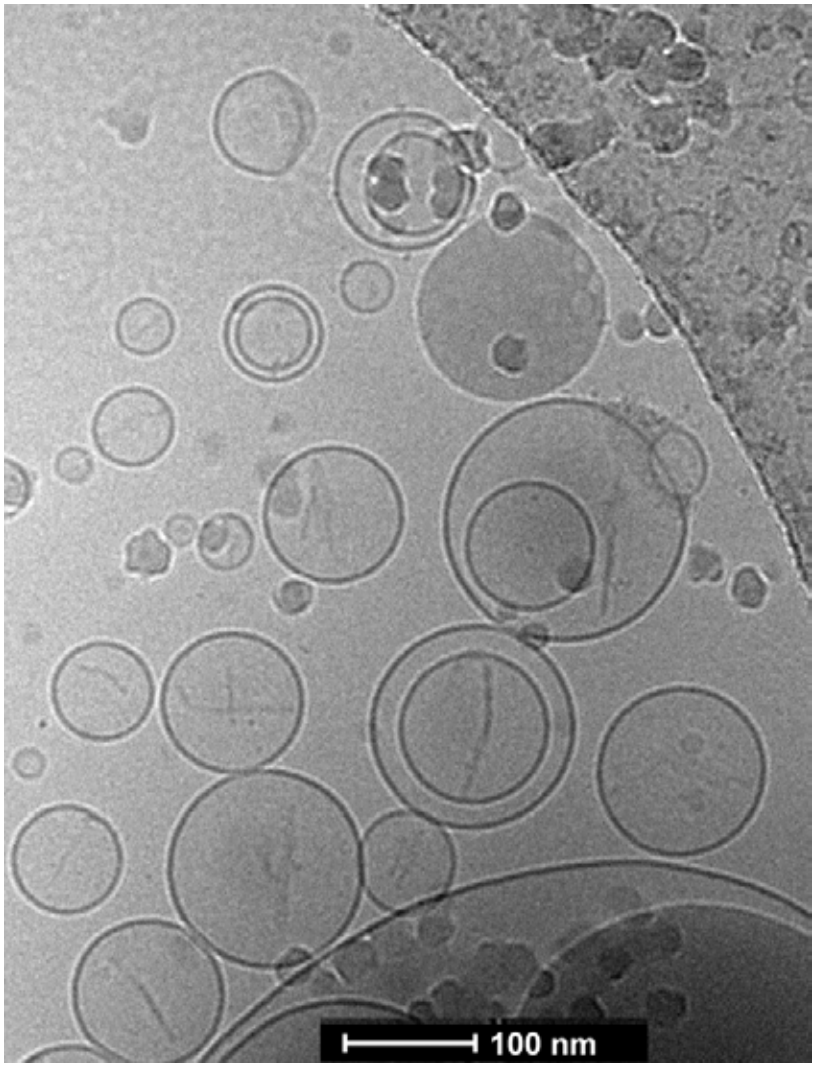


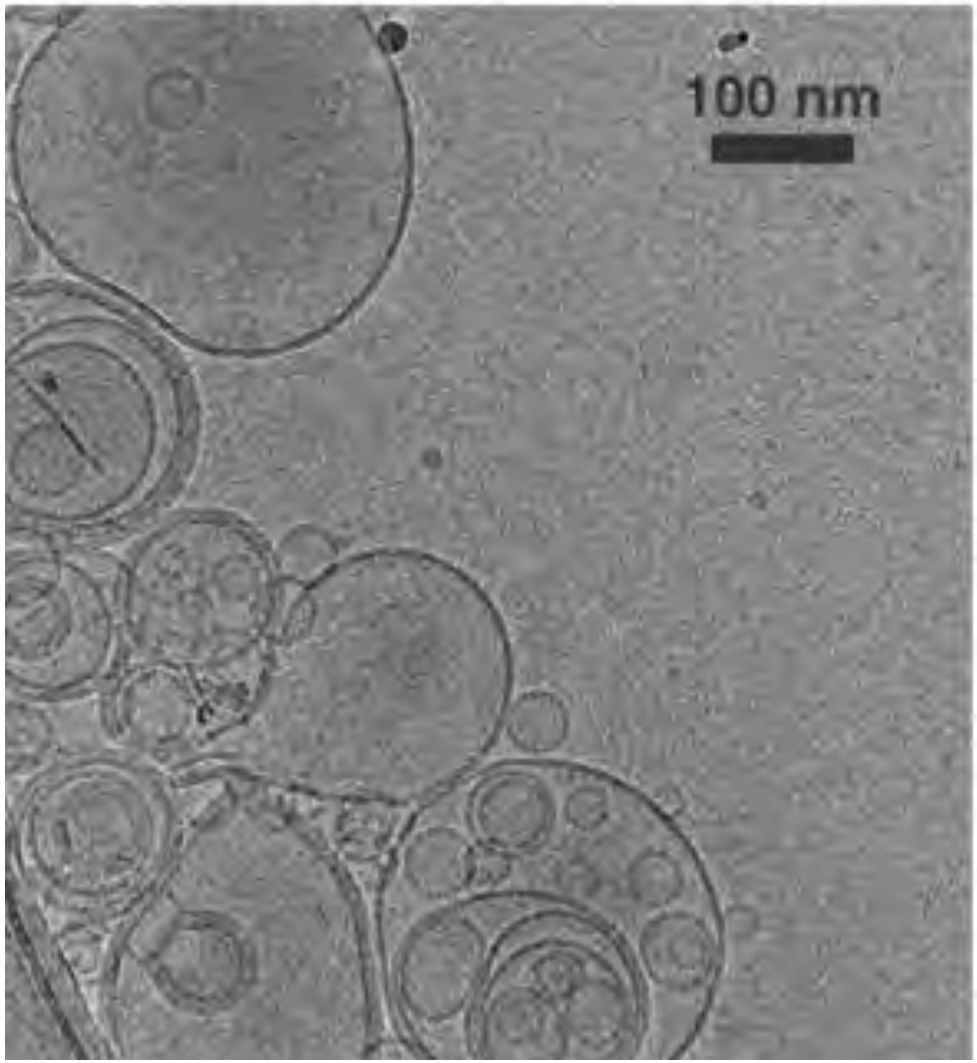
How to support the Amyloid World hypothesis?



Amyloid Fibrillen in einem Membran Container



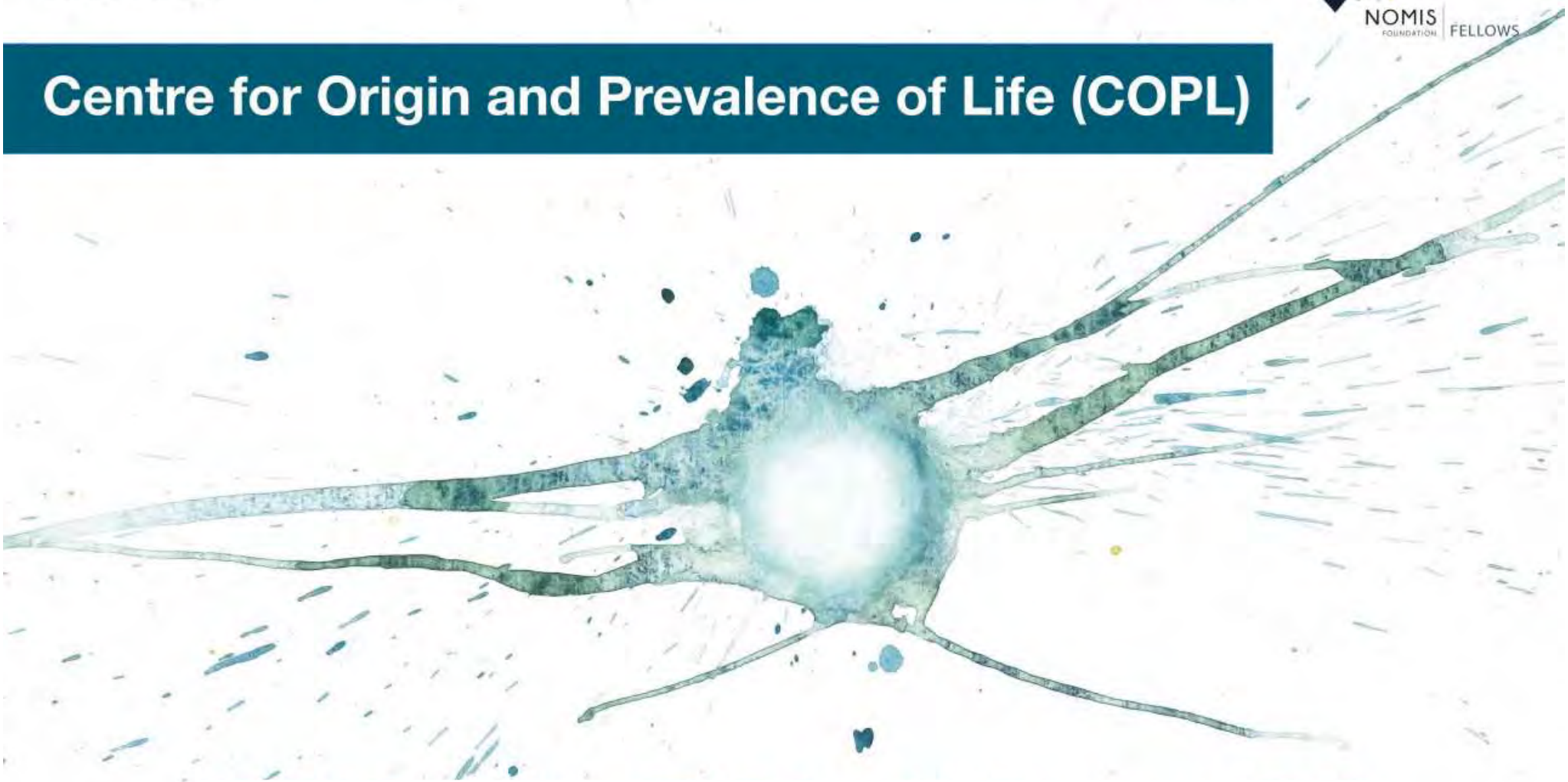



100 nm

ETH zürich

NOMIS FOUNDATION | FELLOWS

Centre for Origin and Prevalence of Life (COPL)



40 groups at ETH over 5 departments

Head of COPL: Didier Queloz

Vice-directorate:

Sascha Quanz, Cara Carmagnabosco, Roland Riek

copl.ethz.ch

“Progress in science depends on new techniques, new discoveries and new ideas, probably in that order”

Sydney Brenner


Beta-endorphin
Carolin Seuring
Nadja Nespovitaya
Beat Meier
Samir Maji
Orgin of life
Saroj Rout
Witek Kwiatkowski
Radek Bomba
Jason Greenwald
Cryo EM of a-synuclein
Henning Stahlberg (Biozentrum)
Ricardo Rodriguez (Biozentrum)
Pratibha Kumari
Secondary Nucleation of a-synuclein
Cedric Eichmann
Pratibha Kumari
Dhiman Ghosh
Yannick Fleischmann
A-Synuclein chaperone
Björn Burmann (Biozentrum Basel)
Sebastian Hiller (Biozentrum Basel)
Juan Gerez
Pratibha Kumari