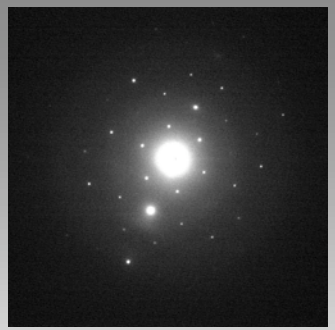
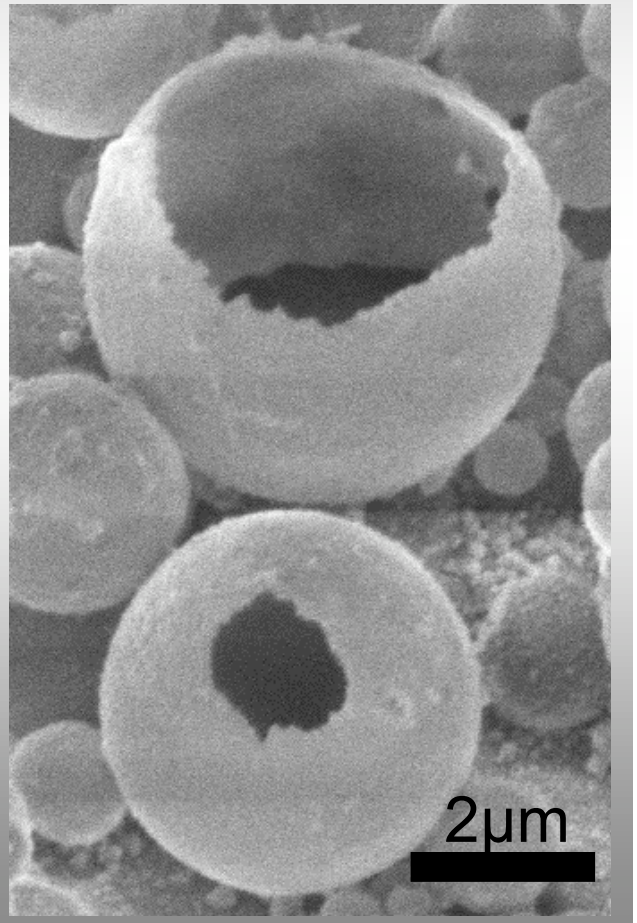
Could a template system of hydrophobic polymer amide on water have preceded an RNA-based code?



Julie E M McGeoch
Department of Molecular and Cellular Biology,
Harvard University,
7 Divinity Avenue, Cambridge,
MA 02138, USA.



Acknowledgements: NIH,
Center for Nanoscale Systems Harvard
Guido Guidotti
David C. Bell (TEM)
Richard Schalek (SEM)
Malcolm W McGeoch (physics)

Harvard daVinci Group meeting of 05.12.08

BASIC FIELD OF INTEREST:

We are interested in candidates for topologies that might have separated charge (a force) in the early Universe, while also providing the material basis for microenvironments to assist complex chemistry.

CHEMICAL STUDIED:

One such candidate is hydrophobic **Proteolipid,** the most conserved amide bonded polymer in all organisms (in terms of amino acid sequence) - nucleotide code-based evidence indicates it first arose 4Gya in the Pre-cells, the Archaea and the Eubacteria (McGeoch & McGeoch Entrapment of water by subunit c of ATP synthase. J.R.Soc.Interface (2008) 5, 311-318).

DETAILS:

In organic solvent when presented with water, proteolipid assembles into tubes that pinch off into vesicles - these observations were made via light microscopy in the course of crystallizing proteolipids for X-ray structure work.

A gene duplication of proteolipid, subunit c of V-ATPase, is in large amounts in the ER of Eukaria.

The ER in some Mammalian cells forms a 3D scaffold of tubes/conduits on the nanometer scale (McGeoch 2007 J Microscopy 227, 172-184).

QUESTION:

As we had a chemical that spanned the 3.8 Gyrs of cell based life on Earth our question was: If we examined pure proteolipid and pure water on the nanometer scale would we see any evidence of scaffolds that could suggest they were candidates that formed early topologies

- & just HOW EARLY in the UNIVERSE could this system have arisen?

Conundrum: Life (CO_2 fixation) arose on Earth after 70 million yrs, while the constituent elements of polymer amide (H,C,N,O) had been around in the Universe for 9 billion yrs (100 times longer).

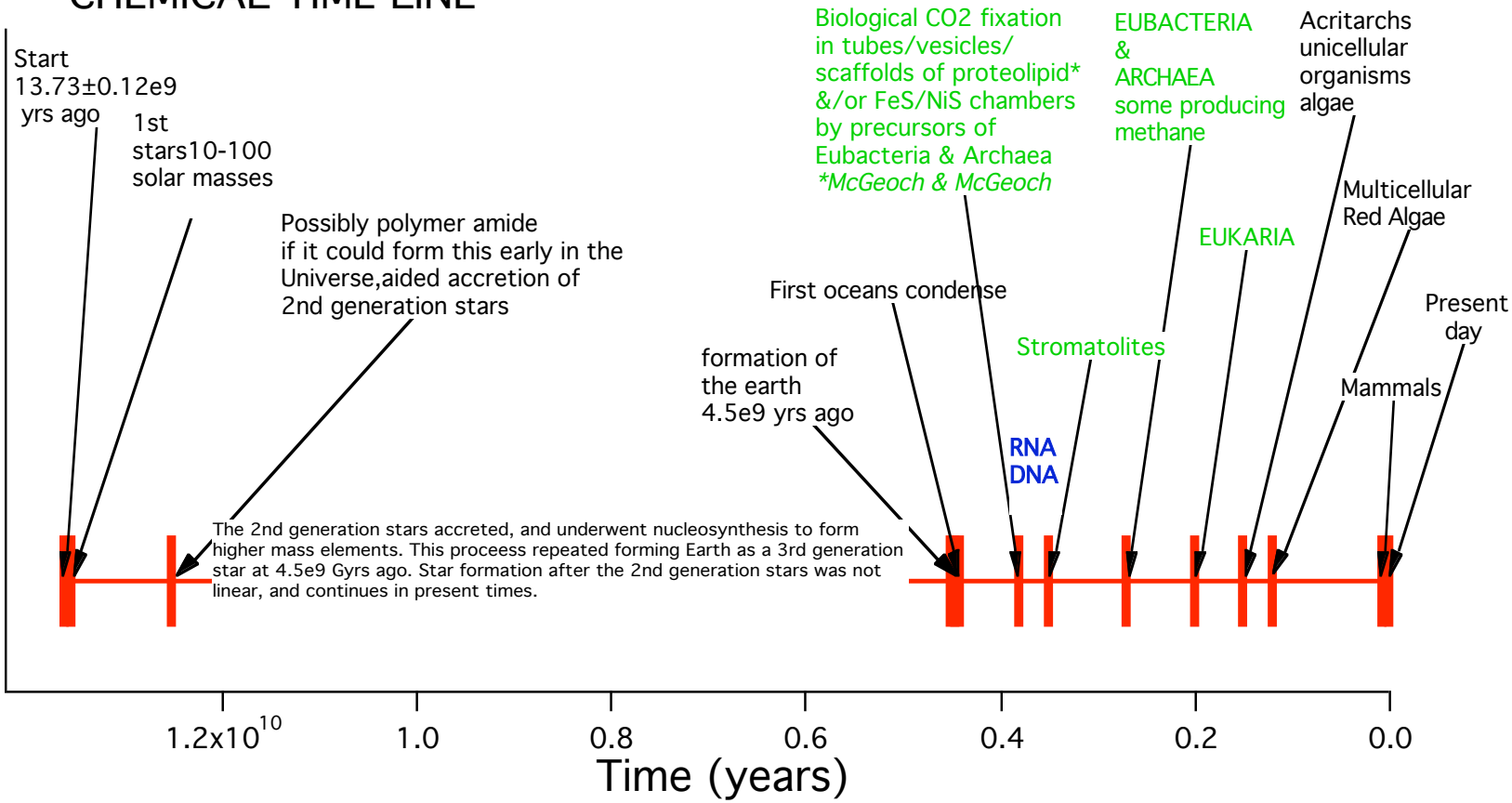
	YEARS Ago
Start (BB)	13.73±0.12e9
Low mass elements mainly hydrogen, carbon, nitrogen & oxygen ***NB no halogens yet***	13.2e9
Some higher mass elements	12e9
All Elements	7.36e9
Earth	4.5e9
Oceans	4.4e9
CO_2 fixation	3.8e9

9e9yrs

7e8yrs

Komatsu et al (2008) Five -Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations:Cosmological interpretation *Astrophys J* eprint arXiv:0803.0547, **Bibliographic Code:**2008arXiv0803.0547K

CHEMICAL TIME LINE



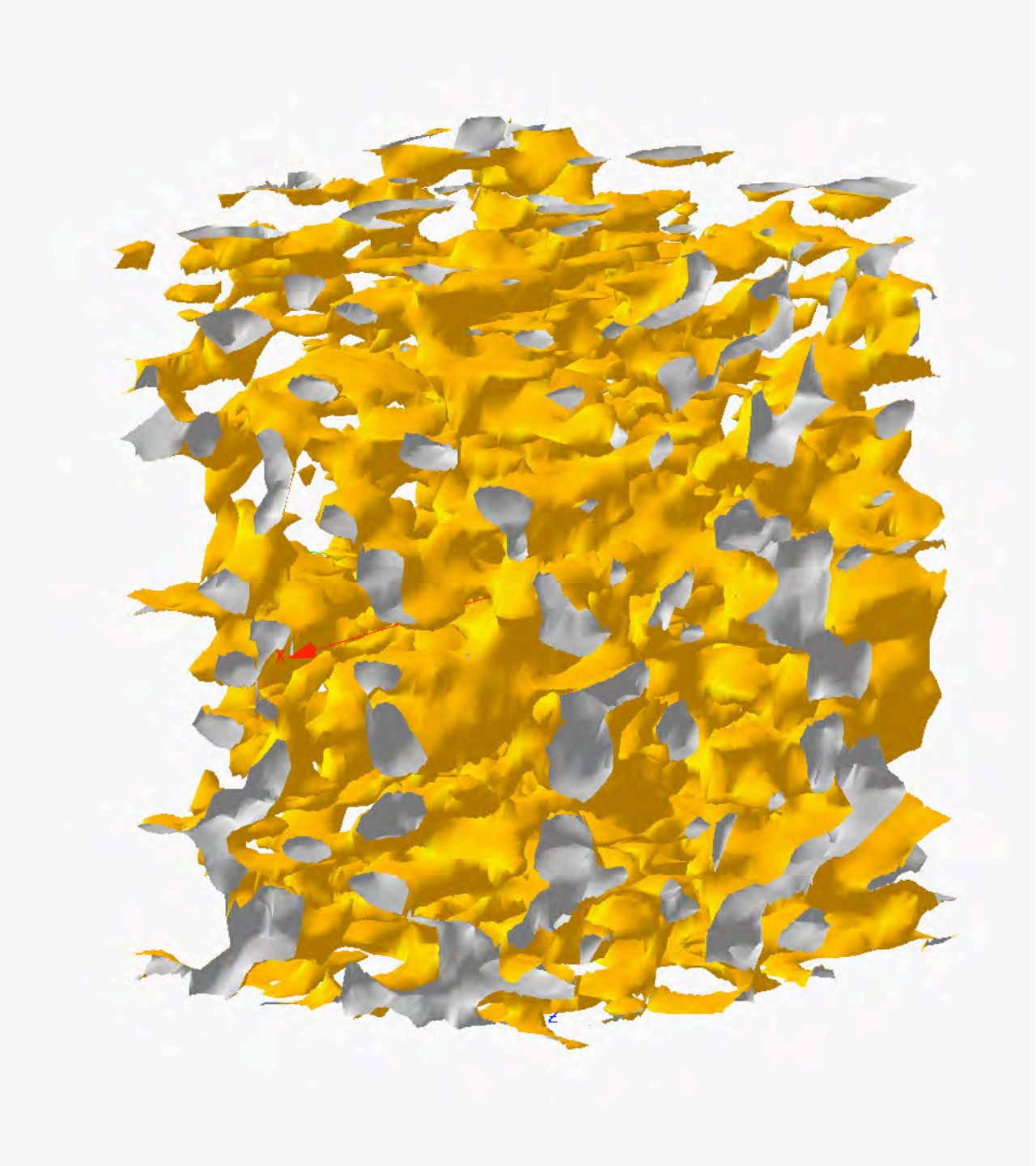
13.73±0.12Gyrs ago	**Start of the universe** - Quarks & Gluons
Minus 10ms	Baryons, Dark Matter, Dark Energy
Minus 100ms	1st Nucleosynthesis: nuclei H,D,He,Li then H_2
Minus 5e8	Accretion of first huge stars. 2nd nucleosynthesis forms low mass elements from H1-O16, mainly H,C,N,O
Minus 8e8	● Possibly Polymer Amide aids accretion of 2nd generation of stars
1.2e10 yrs	The 2nd generation stars accreted, and underwent nucleosynthesis to form higher mass elements. This proceess repeated forming Earth as a 3rd generation star at 4.5e9 Gyrs ago. Star formation after the 2nd generation stars was not linear, and continues in present times.
4.5e9 years	Formation of the earth
4.4e9 years	First oceans condense.
3.8e9 years	CO_2 fixation by precursors to Archaea and Eubacteria
RNA/DNA based code forms	
3.5e9 years	Stromatolites
2.7e9 years	Archaea and Eubacteria
2e9 years	Eukaria
1.5e9 years	Acritarchs, unicellular algea
1.2e9 years	Multicellular red algea
6.8e7 years	Mammals
0 years	Present day

Later Time line dates derived from William Martin and Michael Russell Phil Trans R.Soc. Lond B (2003) 358, 59-85

NB. That only 7e8 yrs elapsed between the Earths formation & CO2 fixation, implies fast life evolution or same time/longer evolution with it arriving on Earth via extra-terrestrial in-fall.

A 3D iso-surface through 29 planes of the ER of a buccal epithelial cell of dimensions: 600x600x770nm. It was Cryo-FIB etched with a 10pAmp current, a vacuum of 1e-4Pa, and at a temperature of -150^{0}C. There is complete continuity of the ER cisternae. Their high intensity grey iso-surface faces inward, and their low intensity yellow iso-surface faces outward into the cytoplasmic space. The branch pattern of the ER cisternae is mainly 3-way.

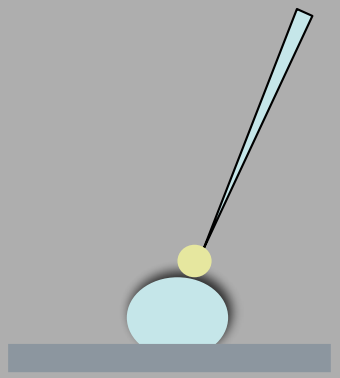
600x600x770nm.



Basic Methodology

Combining in various ways:

1) Water de-ionized to 19 Mega Ohms
2) Pure proteolipid1-10mg/ml



Imaging of the forms via:

1. Phase contrast
   light microscopy

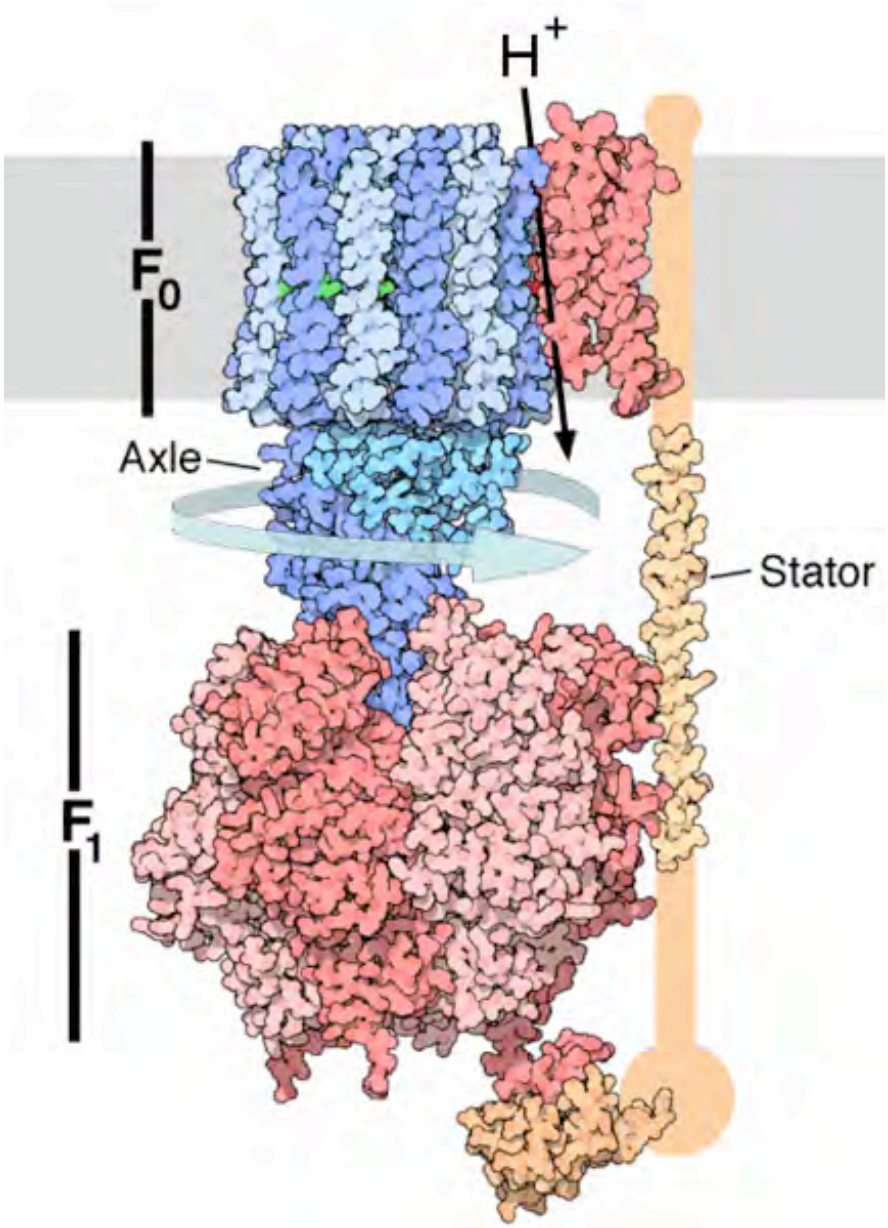
2. Cryo SEM FEI 235 dual-beam/Gatan Alto2500 cold chamber & stage

   Temp: 93-123K, electron beam source: field emission gun 5kV, vacuum: 1e-4Pa.
   Electron beam minimum spot size 2nm on a silicon substrate.
   Images formed from secondary electrons entering instrument collectors/detectors.

3. Cryo TEM/electron diffraction

   Jeol JEM-2100F
   Temp 83K, Electron beam energy: 200keV, Vacuum: 1e-6Pa
   Electron wavelength 2.5pm
   Calibrated just prior to experiment with nickel oxide

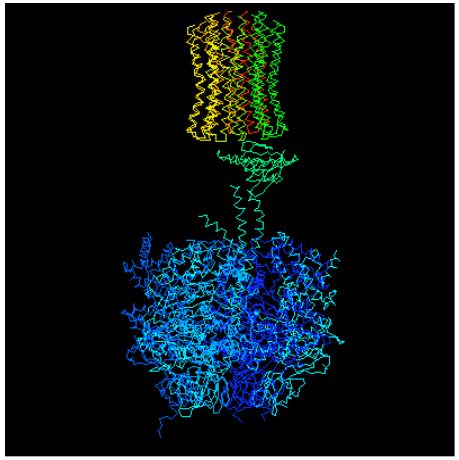
Structure of ATP synthase, the F_0 proton channel and rotating stalk are shown in blue, the F_1 synthase domain in red and the membrane in grey.



Daniela Stock, Andrew Leslie and John Walker 1999

Yeast ATP Synthase Complex
F1: α,β,γ,δ,ε; F0: c (-F0: a, b + f, d, h, OSCP, ATP8)

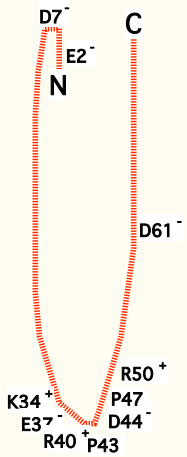
RasMol Backbone Colour by chain



Proteolipid

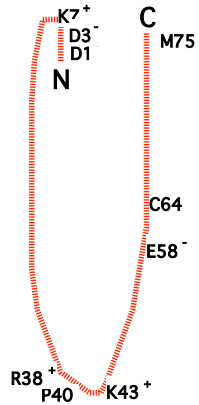
E coli

Subunit c of
ATP synthase F0
7.9kD, 79aa

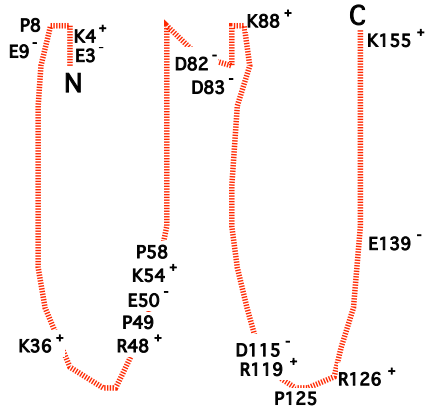


Mammal

Subunit c of
ATP synthase F0
7.6kD, 75 aa

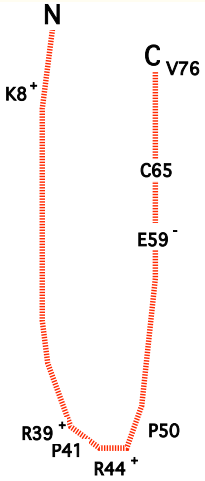


V0 of
Vacuolar ATPase
16kD, 155aa

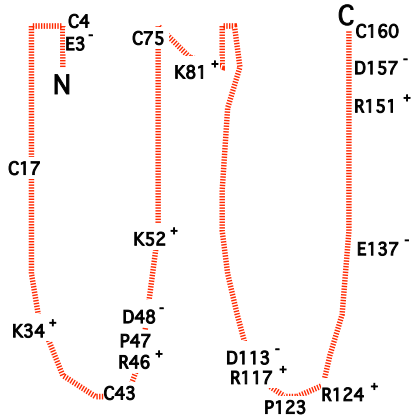


Yeast

Subunit c of
ATP synthase F0
7.7kD, 76 aa



V0 of
Vacuolar ATPase
16kD, 160aa

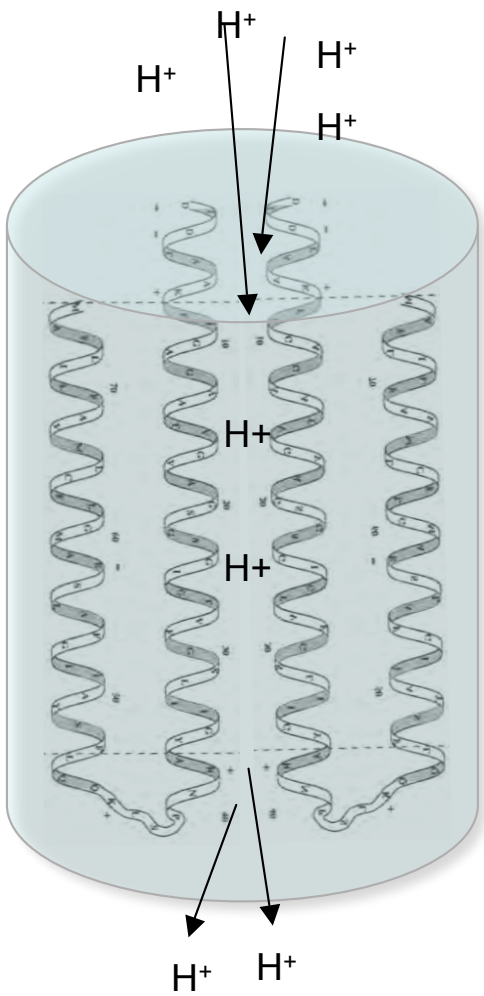


Model diagram of polymer amide pore - to emphasize the cation change with time. The model is based on the patch clamp data: McGeoch & McGeoch 1994, McGeoch & Guidotti 1997, NMR structure data Girvin, Rastogi, Abildgaars, Markley and Fillingame 1998, 2D Crystal data McGeoch & Guidotti 2001 and yeast structure of Stock, Leslie & Walker 1999.

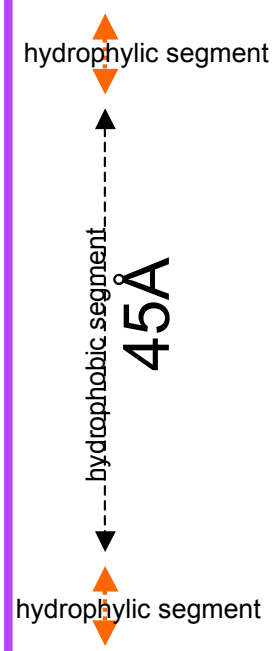
"13Gyrs ago"

Only elements from H1 to O8 have formed

No lipid yet, therefore the if an α-helix structure could form to act as a pore it would not be within a lipid bilayer.

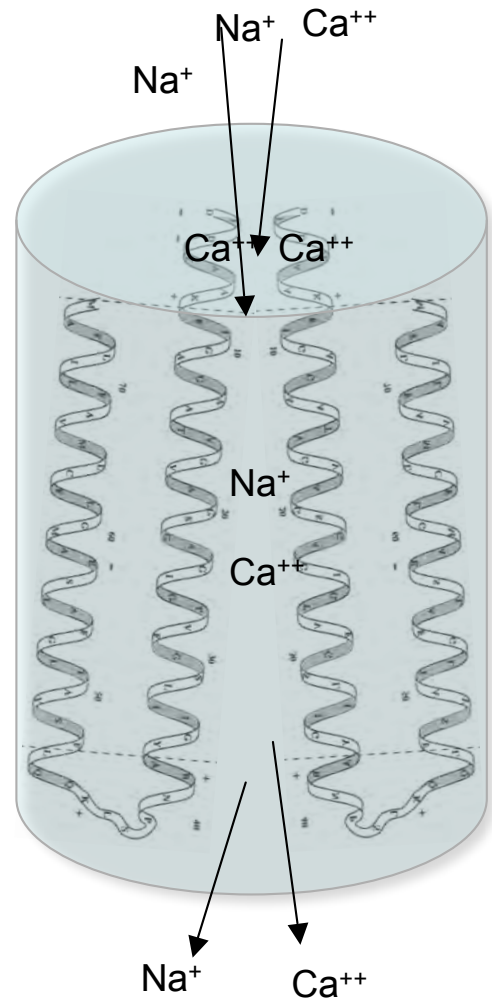


Polymer amide conduit interior



"11Gyrs ago or later"

After higher mass elements have formed



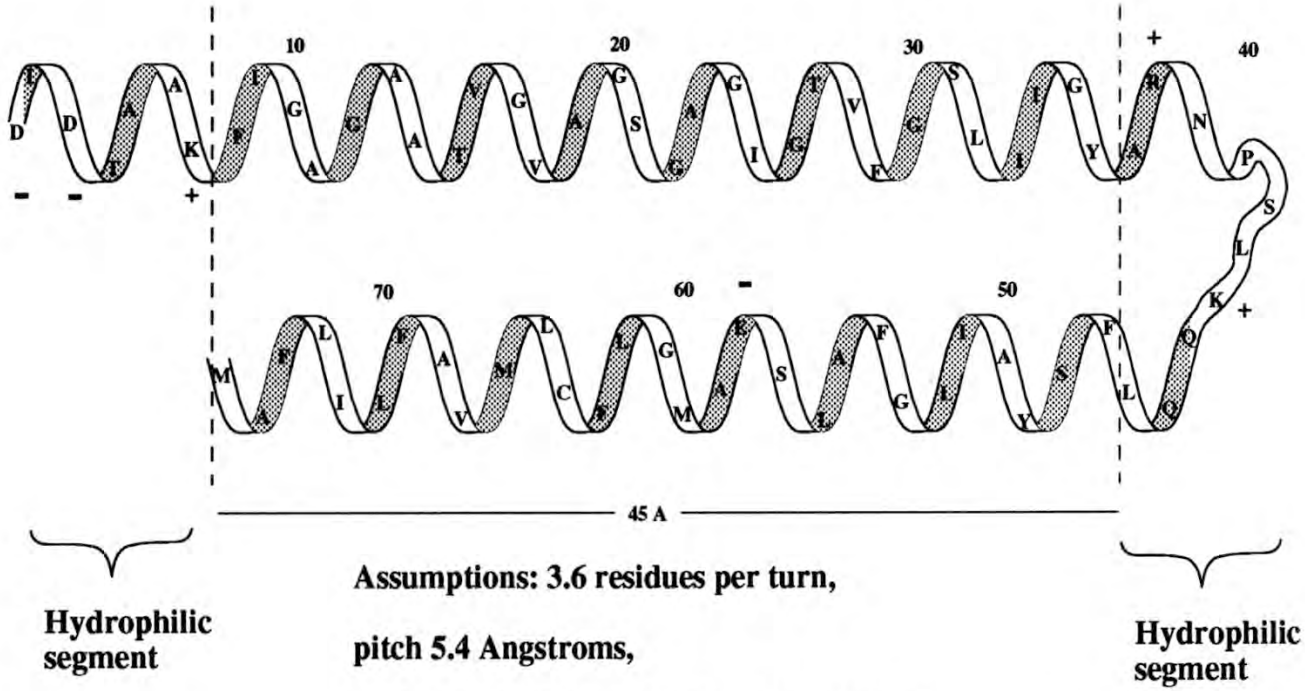
NB: Ca^{++} allows oscillatory closure of the pore by blocking the charge repulsion of the aspartic acid residues at the N-terminus of each subunit. A 10mer pore could therefore operate with 20 aspartate sites being variably regulated by calcium.

Assumptions:
3.6 residues/turn
Pitch=5.4Å
Relative orientation of helices
Is not known

Conduit interior/cell cytosol

2 subunits of each polymer amide pore shown, 8-12 subunits per pore

Transmembrane Helices of Bovine Fo subunit c modeled from possible structure of E. Coli Fo subunit c, from R. H Fillingame, Biochim. Biophys. Acta, 1101 (1992) pp240-243



Assumptions: 3.6 residues per turn,

pitch 5.4 Angstroms,

relative orientation of helices is not known

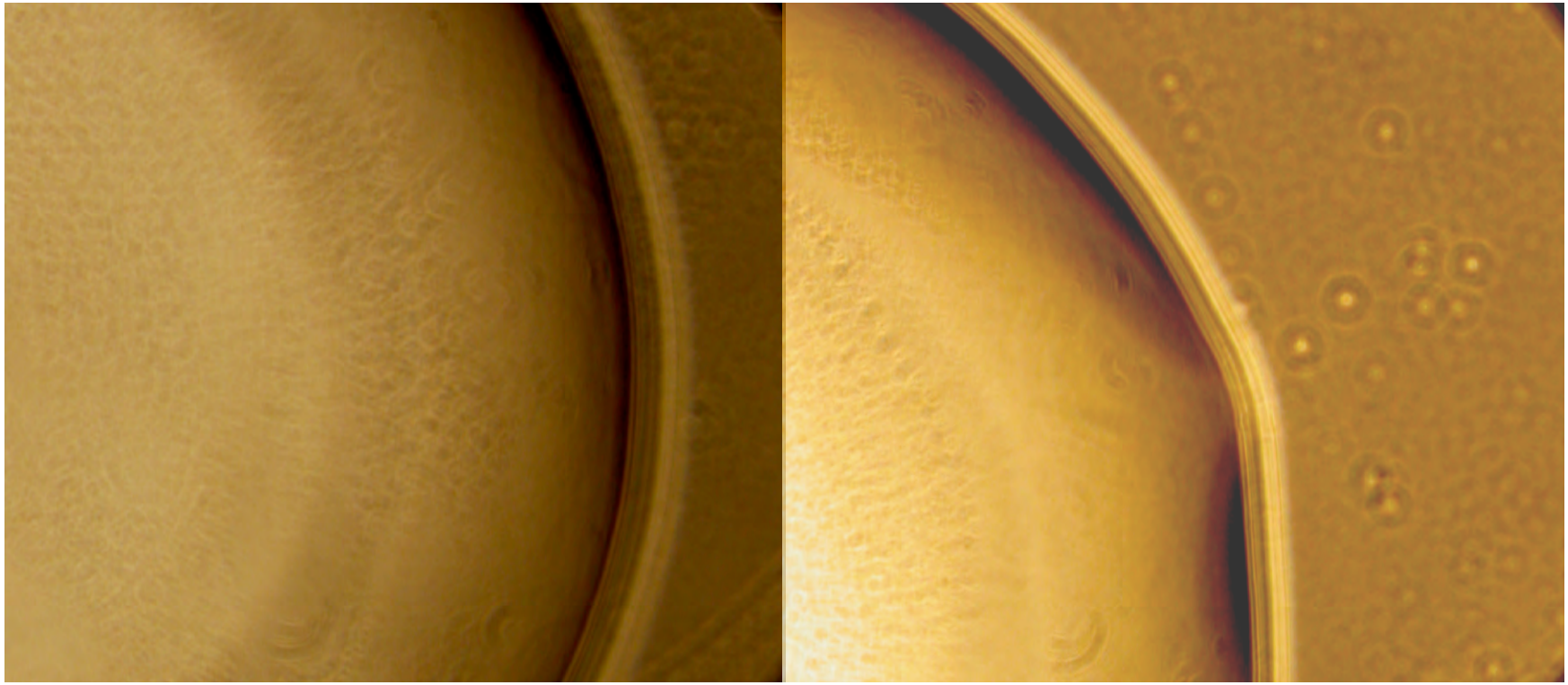
Location of subunit c of ATP synthase coding & non-coding regions (*Homo sapiens*)

CHROMOSOME	CODING SEQUENCE	NON-CODING SEQUENCE	ASSOCIATED but "OTHER"
1		PSEUDOGENE of ATP5G1 VEGA RP5-1198020.5 0TTHUMG0000008302 PSEUDOGENE of ATP5G2 VEGA RP11-385M4.4 0TTHUMG00000036167	
2	HGNC: ATP5G3 P3 ENSG00000154518		
3			
4			
5			
6		PSEUDOGENE of ATP5G1 VEGA: 0TTHUMG00000014203	
7		PSEUDOGENE of ATP5G2 VEGA: AC000362.1 0TTHUMG00000034029	
8			
9			
10		PSEUDOGENE of ATPRG VEGA: RP11-404C6.3 0TTHUMG00000018379 PSEUDOGENE of ATP5G1 VEGA: RP11-443A13.4 0TTHUMG00000018533	
11			
12	HGNC: ATP5G2 P2 ENSG00000135390		
13		PSEUDOGENE of VEGA: RP11-29716.2 0TTHUMG00000017322	
14			HGNC: ATP5S Coupling factor B
15			
16			
17	HGNC: ATP5G1 P1 ENSG00000159199		
18		PSEUDOGENE of ATP5G1 VEGA: AC036176.2 0TTHUMG00000141308	
19			
20			
21			HGNC: GABPA α-sub respir factor 2
22			
X		PSEUDOGENE of ATP5G1 VEGA: RP1-50A13.2 0TTHUMG00000021381	
Y			

Subunit c on the surface of water induces ordering of the surface

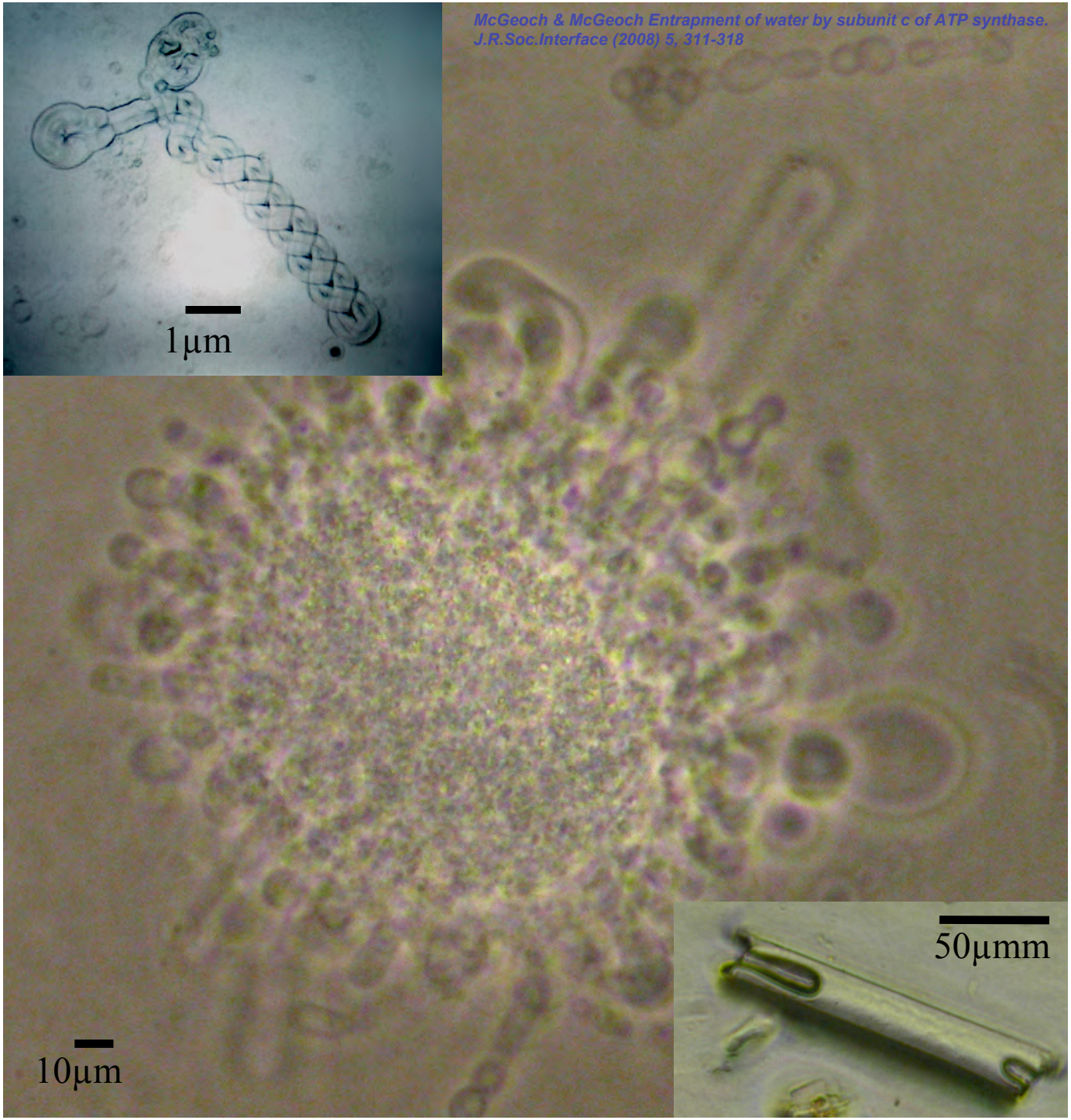
Glass slide with 1µl water

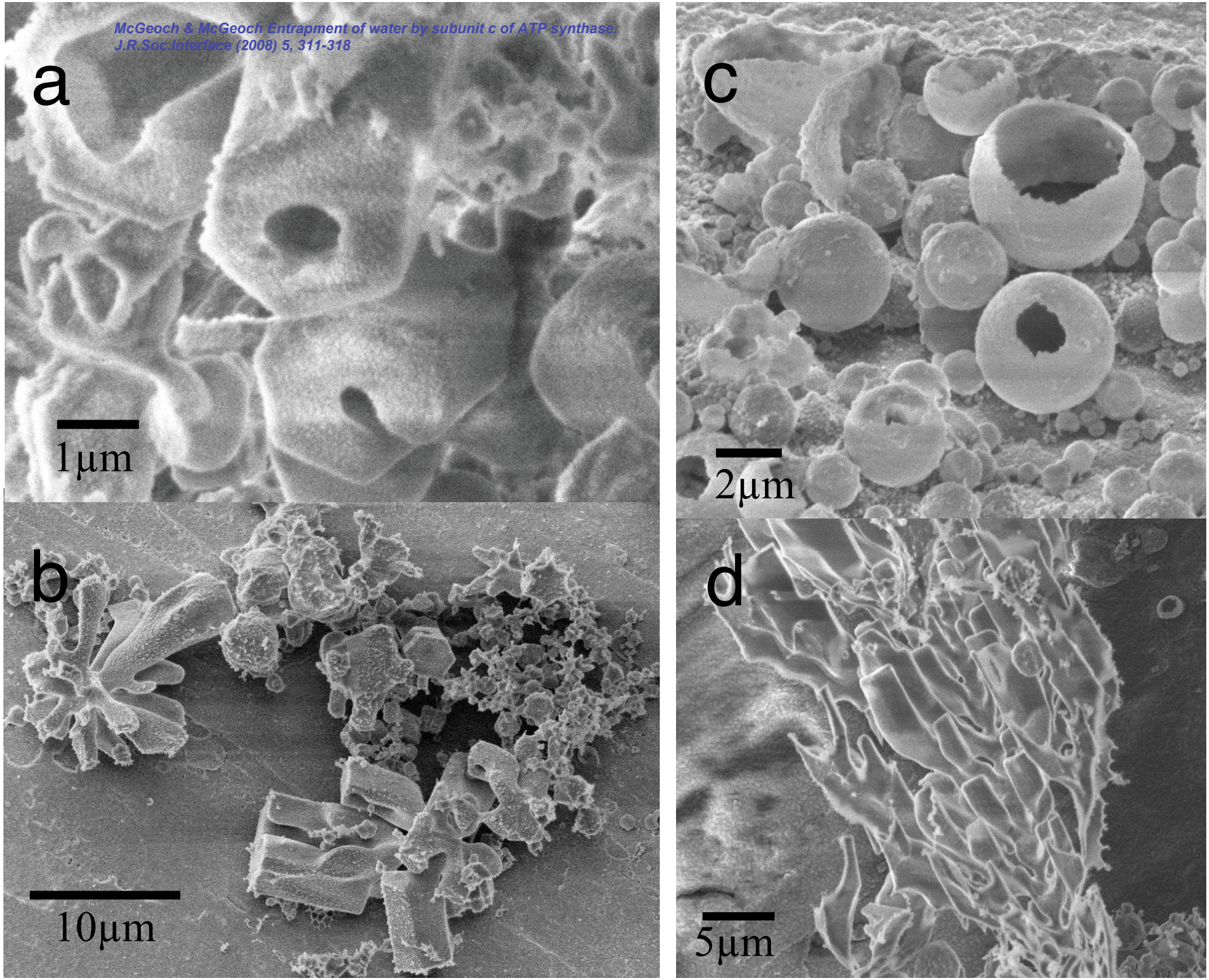
Same + 1ng subunit c. Note the edges of the droplet are becoming straight rather than circular due to the ordering via the protein

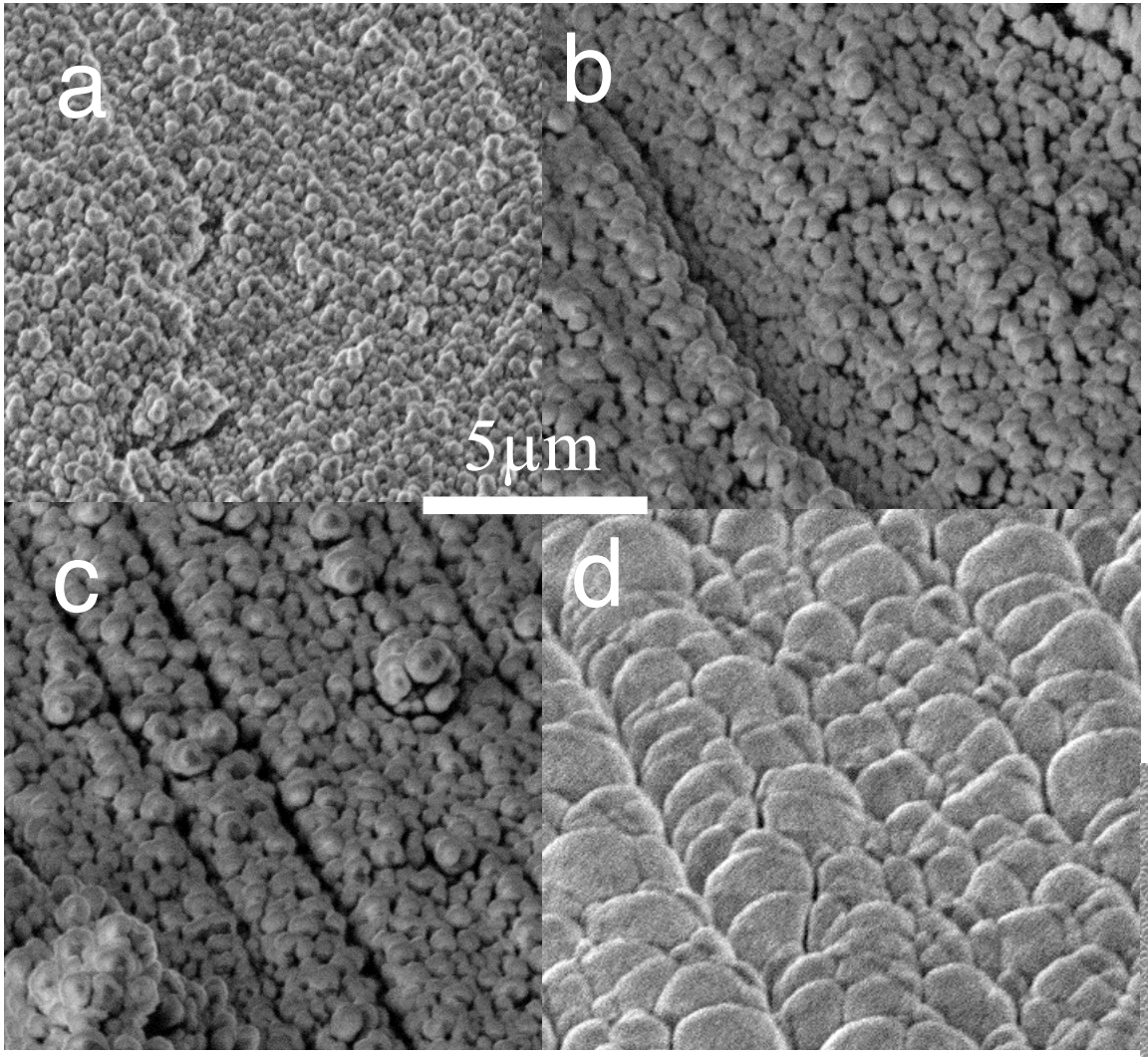


Light microscopy phase contrast image

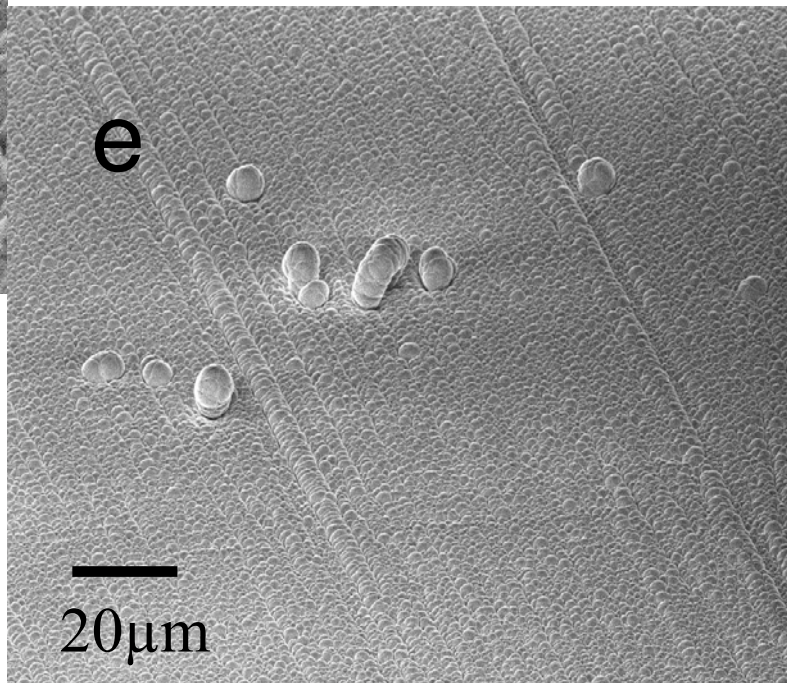
Phase contrast images of tubes of proteolipid. When the volume of solvent above the initial lipid/protein droplet is sufficient to allow the tubes to extend without reaching an air/solvent interface a mass of tubes grow out and up from the droplet to form a medusa-like structure. Proteolipid tubes at the outer surface of the water droplet (top right) round off and form vesicles. The range of diameter of the tubes is 5-40µm. If evaporation of the water is considerably slowed then crystals grow that are sometimes hollow (bottom right inset). Long thin tubes fold and twist back (top left inset, inverted phase contrast image to reveal twist detail).




McGeoch & McGeoch Entrapment of water by subunit c of ATP synthase,
J.R.Soc.Interface (2008) 5, 311-318
a
1μm
b
10μm
c
2μm
d
5μm



SEM images showing the effect of raising the temperature of proteolipid vesicles from 83 to 233K (-190 to -40C) to sublime the central water. **(a)** Vesicles are initially <1µm diameter on the Al stub at 83K (-190C). **(b)** They expand on warming as their central ice reconfigures. **(c)** At 213K (-60C) each vesicle has expanded further and has a very thin relatively dark wall area. *In all experiments once the temperature reaches 233K (-40C) the central water in the vesicles suddenly escapes from each vesicle resulting in the instrument abruptly registering a significant "out-gassing or sublimation event". The instrument vacuum of 1e-4Pa however remains intact but images cannot be captured until the temperature is lowered to halt further sublimation.*
(d) Here the temperature is being lowered again and has reached about 123K (-150C). The vesicles are now multiple consisting of several flattened vesicles collapsed around each other to form spheres of about >4 µm diameter. They are aligned completely across the stub, having nucleated to an arbitrary sub-micron straight groove on the stub. **(e)** a larger area than D showing the extent of the alignment of the vesicles. Some grow up above the monolayer and collapse back.



The proteolipid skin retained the entrapped central water over a temperature rise from -190 to -50^{0}C.

Free water should have started to sublime at -120^{0}C.

McGeoch & McGeoch Entrapment of water by subunit c of ATP synthase. J.R.Soc.Interface (2008) 5, 311-318

TEM images of transitional forms of proteolipid encapsulating water.

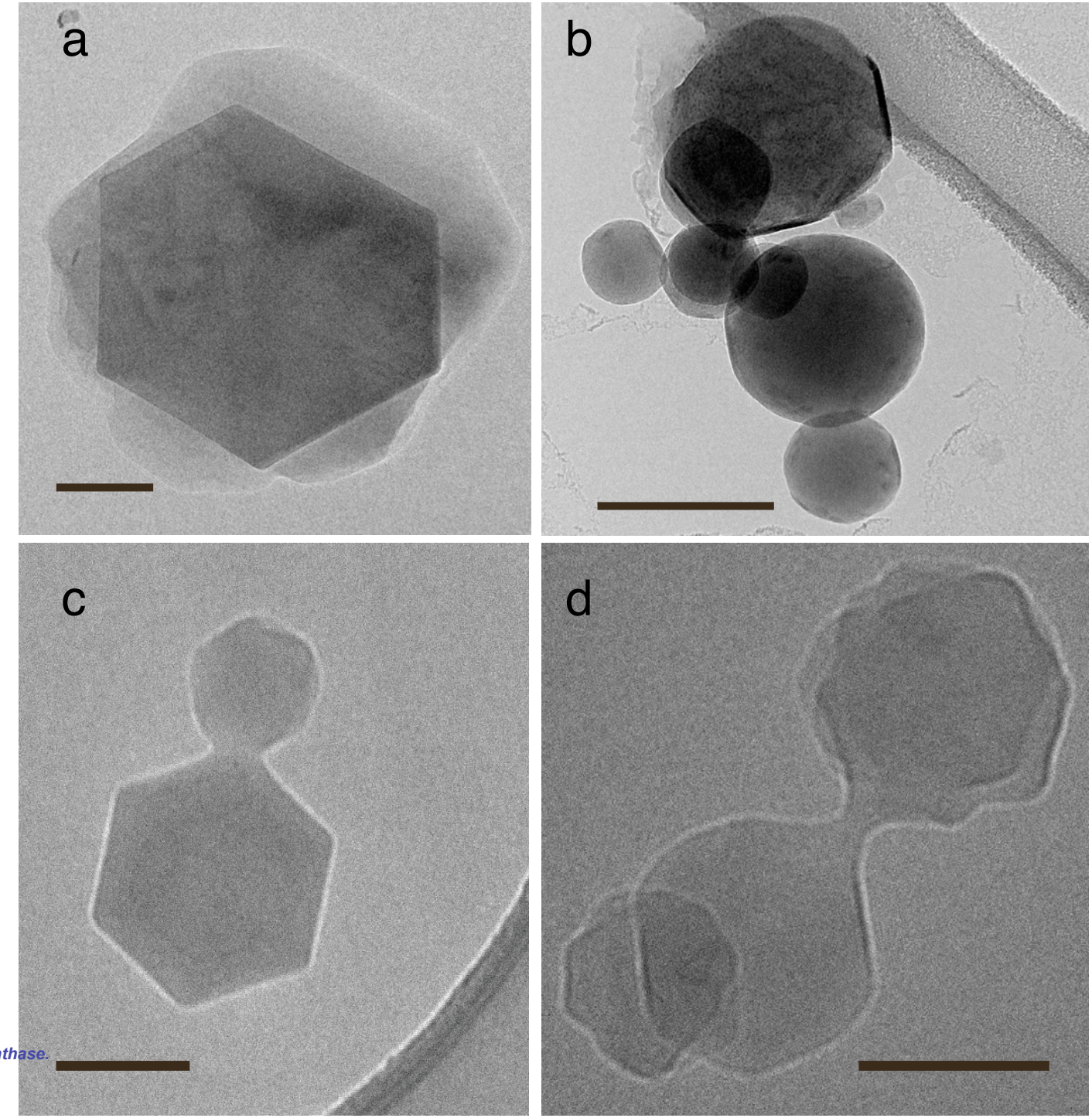
(a) spacious vesicles with central crystals.

(b) vesicles whose sides are beginning to straighten.

(c) spherical vesicles with attached hexagon crystals.

(d) vesicles and crystals attached to one another where neither one or the other is obviously present.

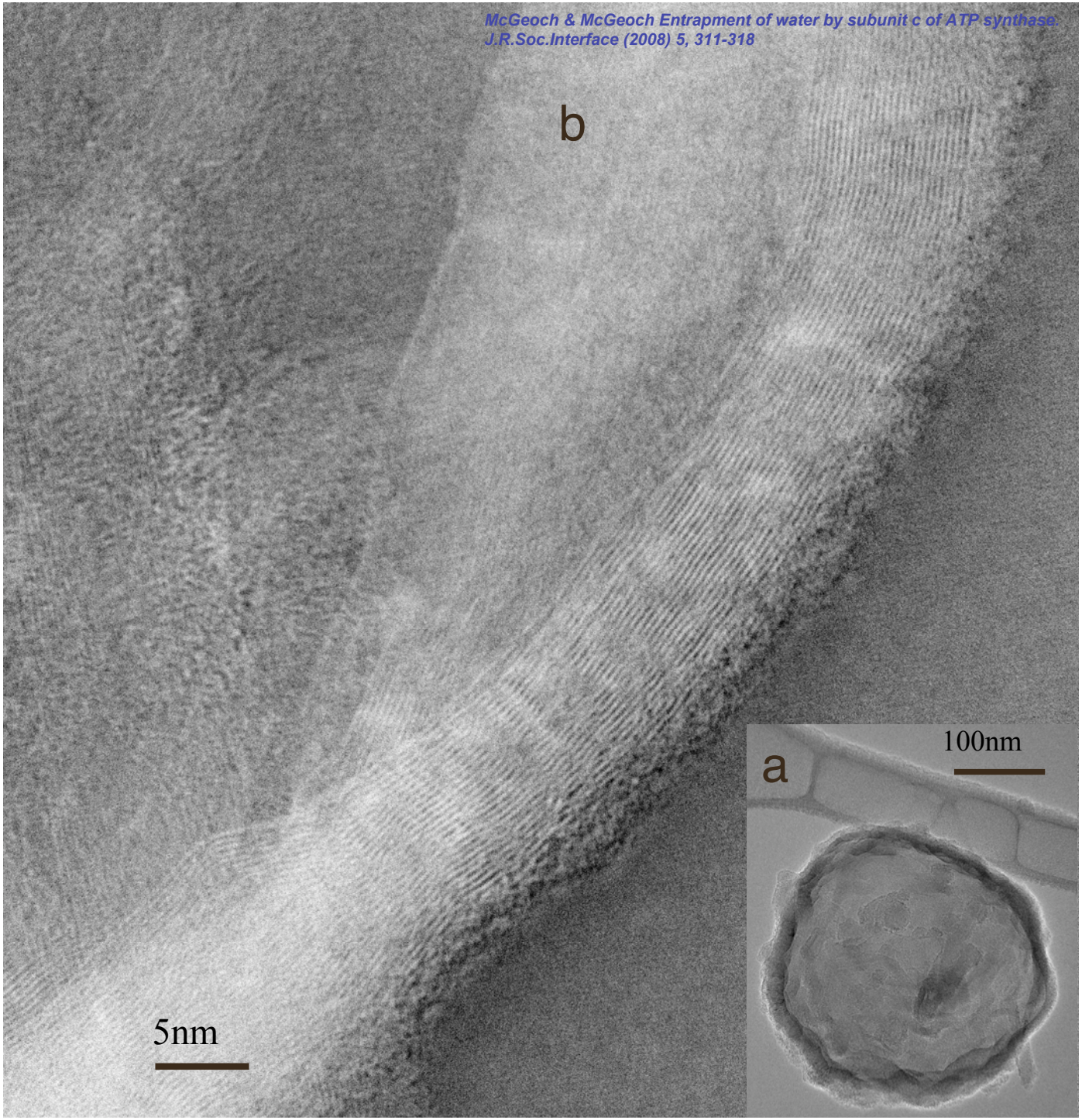
All scale bars = 100nm.



McGeoch & McGeoch
Entrapment of water by subunit c of ATP synthase.
J.R.Soc.Interface (2008) 5, 311-318

(a) TEM image of a vesicle consisting of ribbons of proteolipid winding around water.

(b) TEM *inverted* image of ribbons of proteolipid. A ribbon has β-sheets of strands of proteolipid laterally spaced 3.7Å. The edges of the ribbon away from the vesicle are α-helices. Lengthways the strands/carbon backbones of the protein can be followed in the ribbons as continuous strands of length >134nm diagonally across the entire image. The nearest ribbon has 24 strands with inter-molecular β-sheet type bonding and α-helices at its right edge. TEM 200kV, 1e-6Pa.

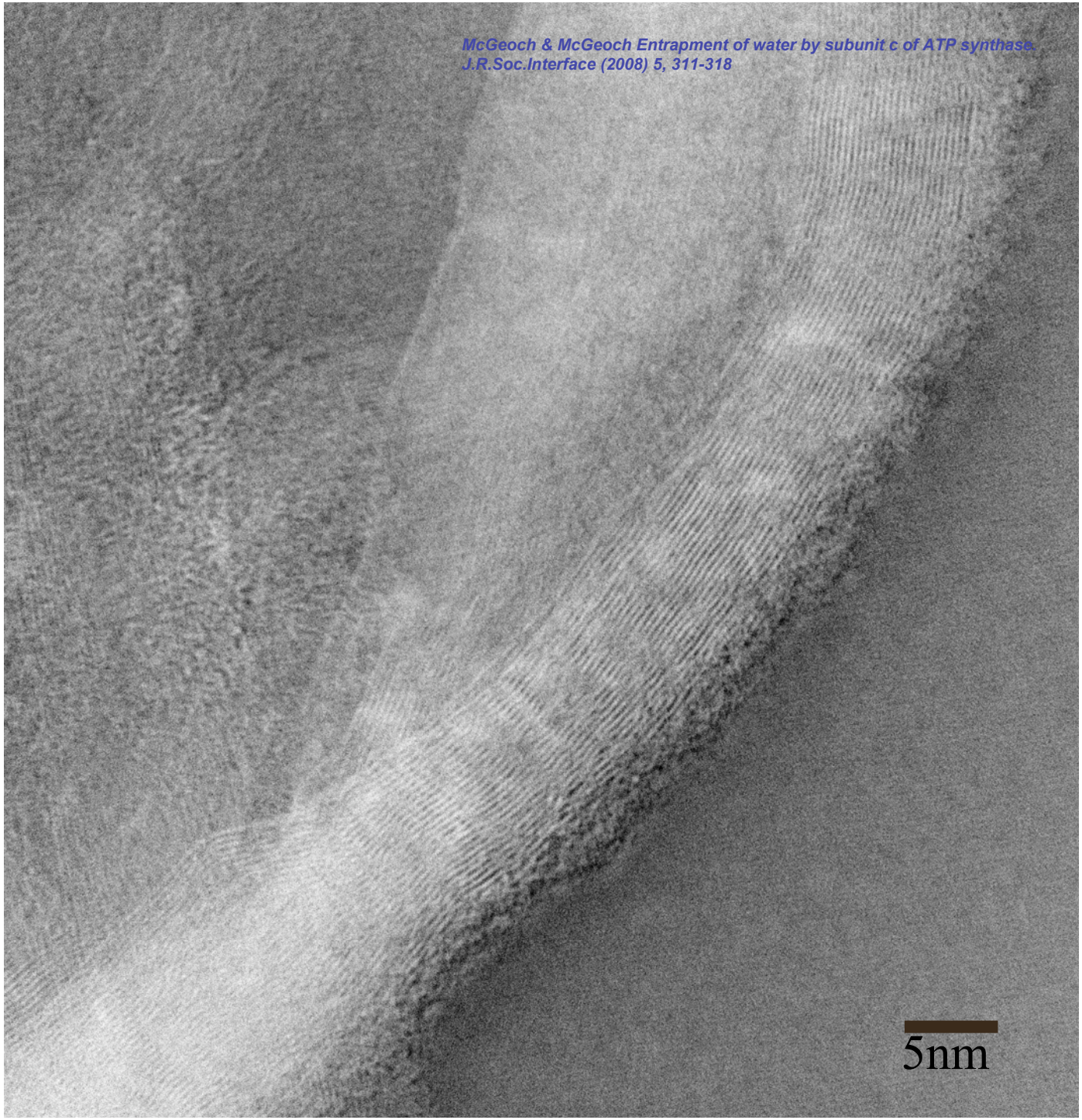


STACKING ISSUE

The ribbons of proteolipid stack

Proteolipid aa sequences contain alternating hydrophobic/hydrophilic aa in groups of **2-3 aa's**

This allows each layer to contain both hydrophobic & hydrophilic R groups in the same orientation. If the peptide sequence had single alternating hydrophobic & hydrophlic aa's, the top R groups away from the water would all be hydrophobic & the bottom towards the water, all hydrophilic. A degree of repulsion between successive layers would ensue, which would inhibit the stacking and alignment of the layers.



PROTEOLIPID SEQUENCES OF MAMMAL, YEAST & E. coli showing how the amino acids might distribute at a hydrophobic/hydrophilic interface - the mammal is seen to have the most alternating sequence, that goes above and below the interface every 2-3 amino acids and only at the C-terminus remains in the hydrophobic phase for 14 amino acids.

R groups should alternate above and below the air/water interface based on hydrophathy.
That the sequences alternate in 2's and 3's poses a topological problem. Phenylalanine (F) induced helical twists in the backbone could solve this.*

Red=hydrophobic amino acids
Blue/grey =hydrophilic amino acids

Mammalian subunit c **NP_788822 Bov**

DIDTAAKFIG AGAATVGVAG SGAGIGTVFG SLIIGYARNP SLKQQLFSYA ILGFALSEAM GLFCLMVAFL ILFAM

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
		I				A	A		F	I		A		A	A		V		V	A				A		I			V	F			L	I	I		Y	A			P
	D		D	T				K			G		G			T		G			G	S	G		G		G	T			G	S				G			R	N

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
		L					L	F		Y	A	I	L		F	A	L			A	M		L	F	C	L	M	V	A	F	L	I	L	F	A	M
	S		K	Q	Q				S					G				S	E			G														

Yeast subunit c **NP_009319**

MQLVLAAKYI GAAIATIGLL GAGIGIAIVF AALINGTSRN PSLRNTLFPA ILGFALSEAT GLFCLMISFL LLYGV

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	M		L	V	L	A	A		Y	I		A	A	I	A		I		L	L		A		I		I	A	I	V	F	A	A	L	I						
		Q						K			G					T		G			G		G		G										N	G	T	S	R	N

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
	P		L				L	F	P	A	I	L		F	A	L			A			L	F	C	L	M	I		F	L	L	L	Y		V
		S		R	N	T							G			S	E		T	G								S						G	

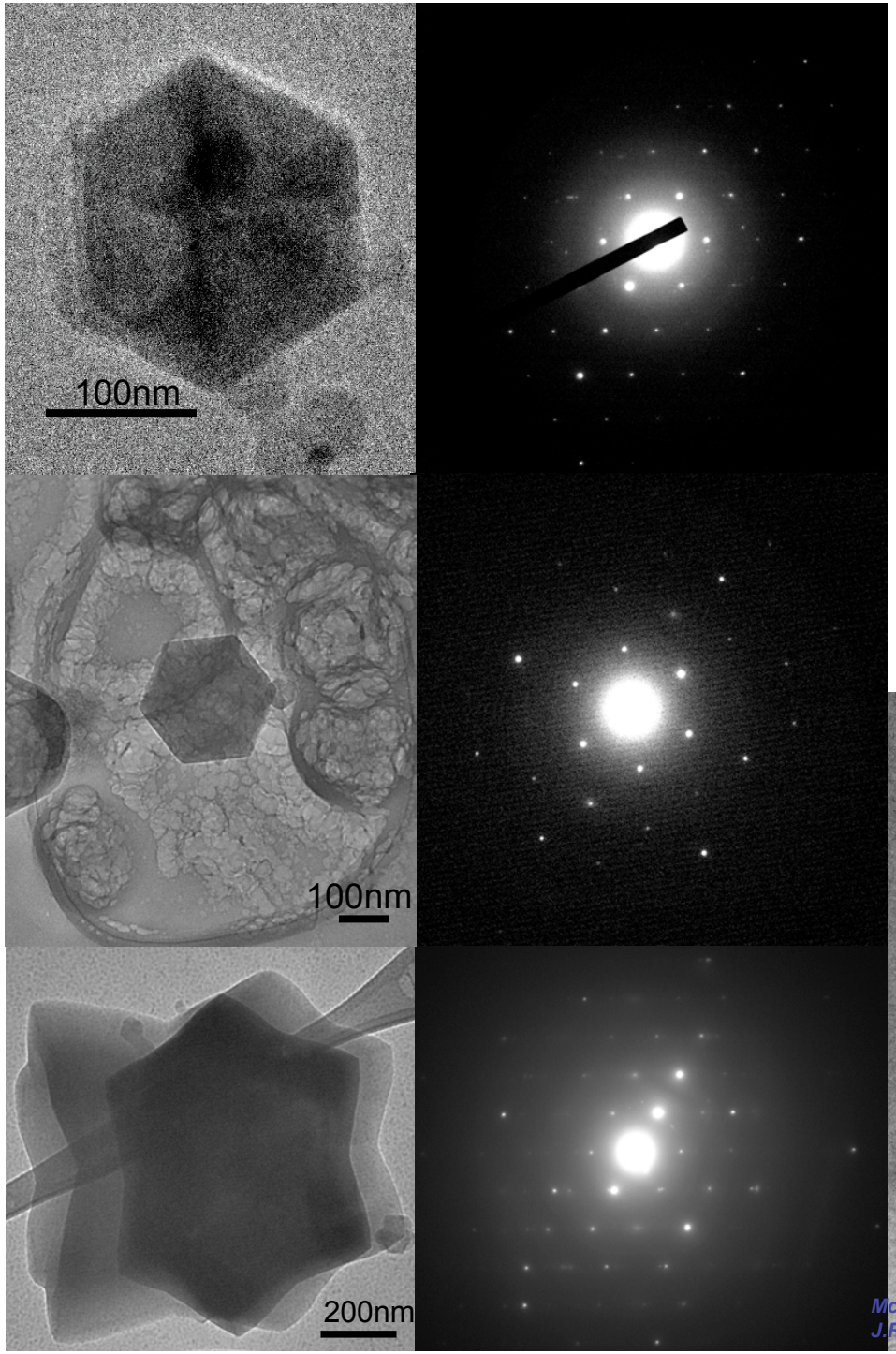
E.coli subunit c **AAC76760**

MENLNMDLLY MAAAVMMGLA AIGAAIGIGI LGGKFLEGAA RQPDLIPLLR TQFFIVMGLV DAIPMIAVGL GLYVMFAVA

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	M			L		M		L	L	Y	A	M	A	A	V	M	M		L	A	A	I		A	A	I		I		I	L				F	L			A	A
		E	N		N		D											G					G				G		G			G	G	K			E	G		

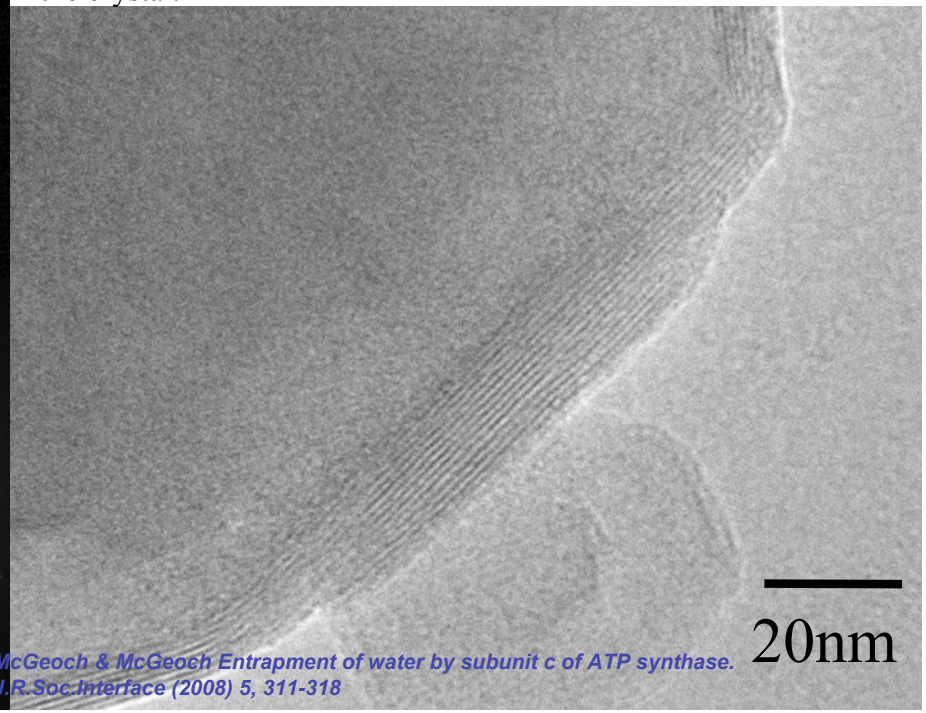
	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
		P		L	I	P	L	L			L	F	F	I	L	V	M		L		A	I	P	M	I	A	V		L		L	Y	V	M	F	A	V	A	
	R	Q		D						R	T	Q							G		D							G		G									

*Matsummura, Uemura, Mihara 2004 Chem Eur J 10, 2789-2794.



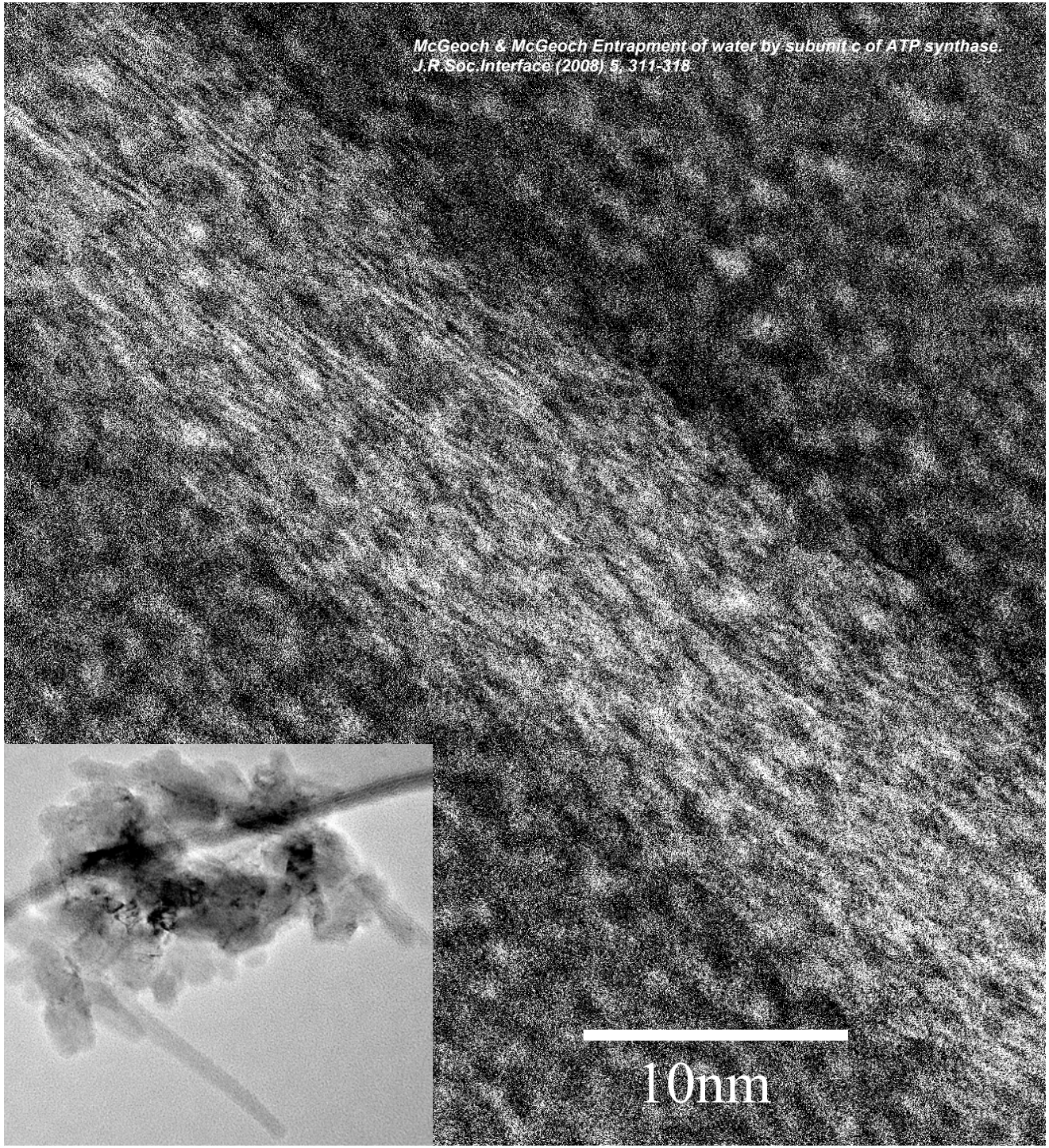
TEM of hexagonal crystals of water encapsulated by proteolipid and their diffraction pattern to the right. The average spacing from the center of the pattern to the first hexagonal water spot corresponds to 3.73±0.06Å in 3 separate crystals measured from 1-3 directions. These sharp diffraction patterns are seen with a electron wavelength of 2.5pm. It implies the existence of a regular repeating structure in which the atoms (eg Oxygen of water) are aligned to their ideal positions with a displacement of somewhat less than the wavelength e.g. 1pm.

In the middle crystal image electron dense spherical vesicles encapsulated by protein can be seen together with pale masses of lipid in the background. The bottom R image is another hexagon water crystal that shows detail of the protein inter-sheet spacing as it encapsulates the crystal. At different points on the edge of this crystal there are 7-13 separate subunit c sheets spaced 6.0±0.1Å. The sheets bend over the 60^0 turns of the crystal.



McGeoch & McGeoch Entrapment of water by subunit c of ATP synthase. J.R.Soc.Interface (2008) 5, 311-318

TEM (*inverted image*) of a tube whose surface is covered by strands of proteolipid subunit c with a lateral spacing of 3.7Å amidst circular structures that could be assembled channels of subunit c of 3nm diameter. The strands adopt a woven path over the tube surface. The inset image shows a lower magnification of the entire pile of collapsed tubes on the TEM grid and in their center are crystals.



OBSERVATION

In a given proteolipid peptide strand:
X direction spacing i.e, β-sheet hydrogen bonding = 3.7Å
Y direction spacing i.e, distance between peptide bonds = 3.5Å
Z direction spacing i.e. stacking of the layers - 6Å

Entrapped water lattice spacing = 3.73±0.06Å*
It has the clear 6 fold symmetry of hexagonal ice but this state existed at a RT of 17^0C. The state was conserved by flash freezing in milliseconds to -195^0C. Therefore the entrapped water is ORDERED and compressed.

Further the amide sequence of proteolipid would seem to "fit" onto a water surface.

Could there be a co-operative association here in terms of:
1. Early protein condensation
2. And more speculative ~ even a degree of dependence for long term stability

*the strongest lattice diffraction of hexagonal ice is 3.89 at -180C. The spacing of the first peaks from the centre of the diffraction pattern in 3 separate crystals from 1-3 directions n=36. This measurement is smaller than expected for hexagonal water ice at 3.9Å which indicates compression.

If an early system arose in the Universe why select for:

1. Amino acids particularly?
2. What was the source of the amino acids i.e. synthesized or delivered?
3. Why was homochirality adopted (sugars D and aa's L)?
4. Energy consideration for polymerization of amino acids in/on water?

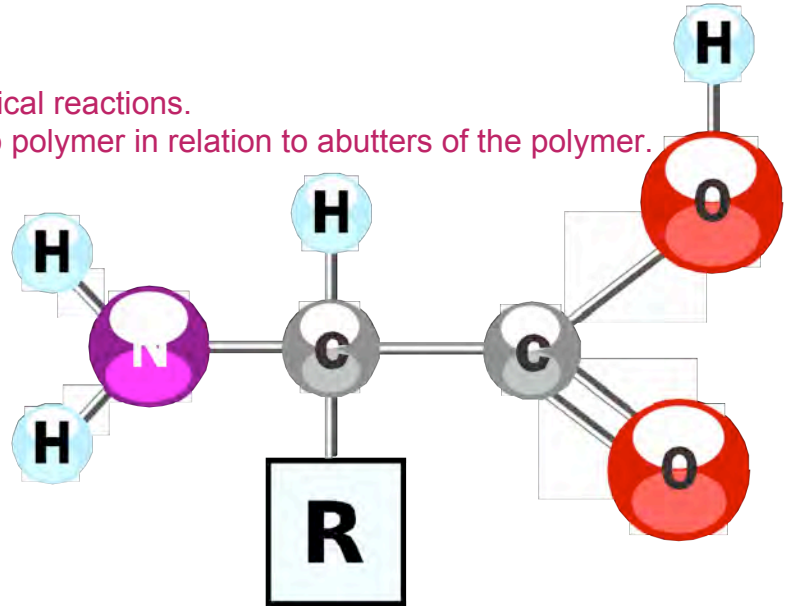
No halogens at the point H,C,N,O first arise
-hydrophobic polymer amide remains stable

Once H, C, N or O formed, amino acids arising from these elements offered high value to a system

An α-AMINO ACID (aa)

Has a primary amine grp & a primary carboxyl grp
~ each aa has a specific isoelectric point pH when: $NH3^+/COO^-$ termed a Zwitterion

Reactions:
Nucleophilic addition.
Amide bond formation - polymers.
Amine group: Imine formation & esterification.
Carboxyl grp: Amide bond formation and decarboxylation.
The multiple side chains of amino acids can undergo chemical reactions.
*aa side chains confer hydrophobic/hydrophilic character to polymer in relation to abutters of the polymer.



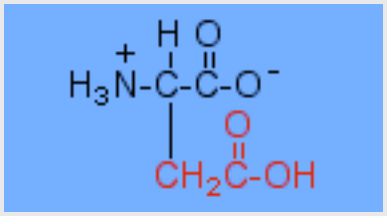
Amino acid amides to polymers

Converting an amino acid to its amide gives it sufficient energy to drive the subsequent reaction with an amino acid to form a peptide bond.

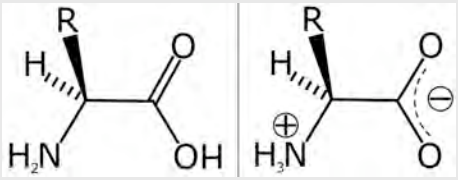
Heating amino acid amides (amides of glycine, alanine, valine, aspartic acid) to dryness with kaolin (clay mineral) - 10 hydration/dehydration cycles produces polymers linked by peptide bonds:
6% α-helix
42% β-sheet
52% Random

Amino acid

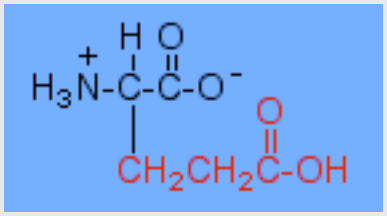
Aspartic acid Asp



Zwitterion

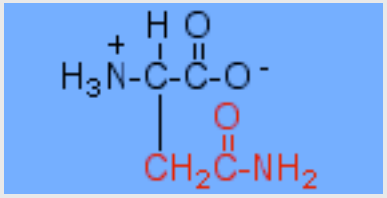


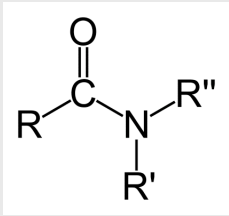
Glutamic acid Glu



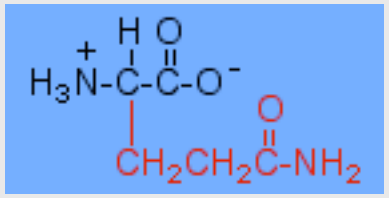
Amide - terminal hydroxyl grp replaced with an amine

Asparagine Asn





Glutamine Gln



PREBIOTIC CONDENSATION OF AMINO ACIDS TO PEPTIDES

From Andre Brack's 2007 review

EXPERIMENTS that synthesize polymer

Simulations equivalent to those of thermal vents- oscillating high/low temp & constant high pressure with reducing conditions: glycine + $CuCl_2$➔10mer G

Simulations of high salt lagoons - Amino acids + 3M Nacl + Cu + high temp➔simple amino acids + dipeptides

Amino acids + *N,N'*-carbonyldiimidazole + H_2O➔condensate enriched for protein amino acids

Activated amino acids (aminoacylphosphate derivatives) + mineral clays (montmorillonite)➔56mer peptides

Amino acid thioesters + CdS/ZnS➔8mer peptides

Acidic Asp/Glu + hydroxylapatite $Ca_5(OH)(PO_4)_3$ or clay illite➔polymers of Asp/Glu (*NB Ca^{++} binding essential to the mechanism)*

Basic Arg + kaolinite/illite/montmorillonite clays + *N,N'*-carbonyldiimmidazole➔polymers after 50 cycles

β–sheet/α-helix formation

Strictly alternating homochiral poly (val-lys) + NaCl (screens the lys $NH3^+$ charges) in water➔β-sheets
Rule: Strict alternation of hydrophobic/hydrophilic aa's in water➔bilayer thermally stable β-sheets - also need to be homochiral all L or all Daa's.
Tetrapeptide periodicity produces thermally labile α-helices - the aa's in an α-helix don't have to be homochiral e.g conotoxins

If alcohol is added to a pre-formed β-sheet it transitions to α-helix - alcohol reverses the H_2O induced constraint on the hydrophobic grps
NB Lipid must have a similar effect on proteolipid

Peptide Nucleic acids (PNA) were discussed by the Brack review - backbone of achiral poly(*N*-(2-aminoethylglycine) with bases attached through an acetate linker. PNA cyclizes easily making oligomers hard to form. Can get PNA-PNA.

BUT a direct pathway from peptide to Nucleotide-based code is not necessary if a peptide/lipid topology (formed much earlier than RNA) can host the development of RNA/DNA alone.

1. Why Amino acids particularly?

DATE REPORTED	REFERENCE	CHEMICAL FORMED	DETAILS
1850, 1854	A. Strecker	AMINO ACIDS	Amino acids formed from: KCN, aldehyde s and ketones
1953 1959 2004	S l Miller & H C Urey A Lazcano & J L Bada	AMINO ACIDS	Amino acids 7 othe r organic monomers from reduced mixtures of methane,ammonia & hydrogen.
1967	C N Matthews & R E Moser	PROTEIN	Polymerization of HCN in the presence of water
1960 reviewed in 1994	S W Fox., P R Bahn., K Dose., K Harada ., L Hsu ., Y Ishima ., J Jungck. , J Kendrick., G Kram pitz., N Tadayoshi., A Pappelis., A Po., D L Rohlfing., A Vegotsky., T V Waehneldt., H Wx & B Yu.	PROTEIN	Protein was fo rmed from ammonia, methane, water and a discharge.
1960 to present time	Large number of competent chemists	PROTEIN	Protein will readily form from HCN and other H, C, N and O based chemicals.
2000	G D Cody., N Z Boctor., T R Filley., R M Hazen., J H Scott., A Sharma. & H S Yoder Jr.	Basic Metabolites like pyruvate	Experimental time is the only limit to producing almost any chemical provided its constituent elements are present.
2007	M W Powner., C Anastasi ., M A Crowe., A L Parkes., J Raftery. & J D Sutherland.	RNA-like molecule	RNA-like molecules can arise from photoanomerisation of cysteine nucleotides

TABLE 1. CHEMICALS from HYDROGEN, CARBON, NITROGEN & OXYGEN.

NB: Chemicals requiring constituent higher mass elements like RNA are in blue backgrou nds – these would have first formed later than 13Gyrs ago.

1. Why Amino acids particularly?

Voet & Voet 1995
Metabolic breakdown
of individual amino acids

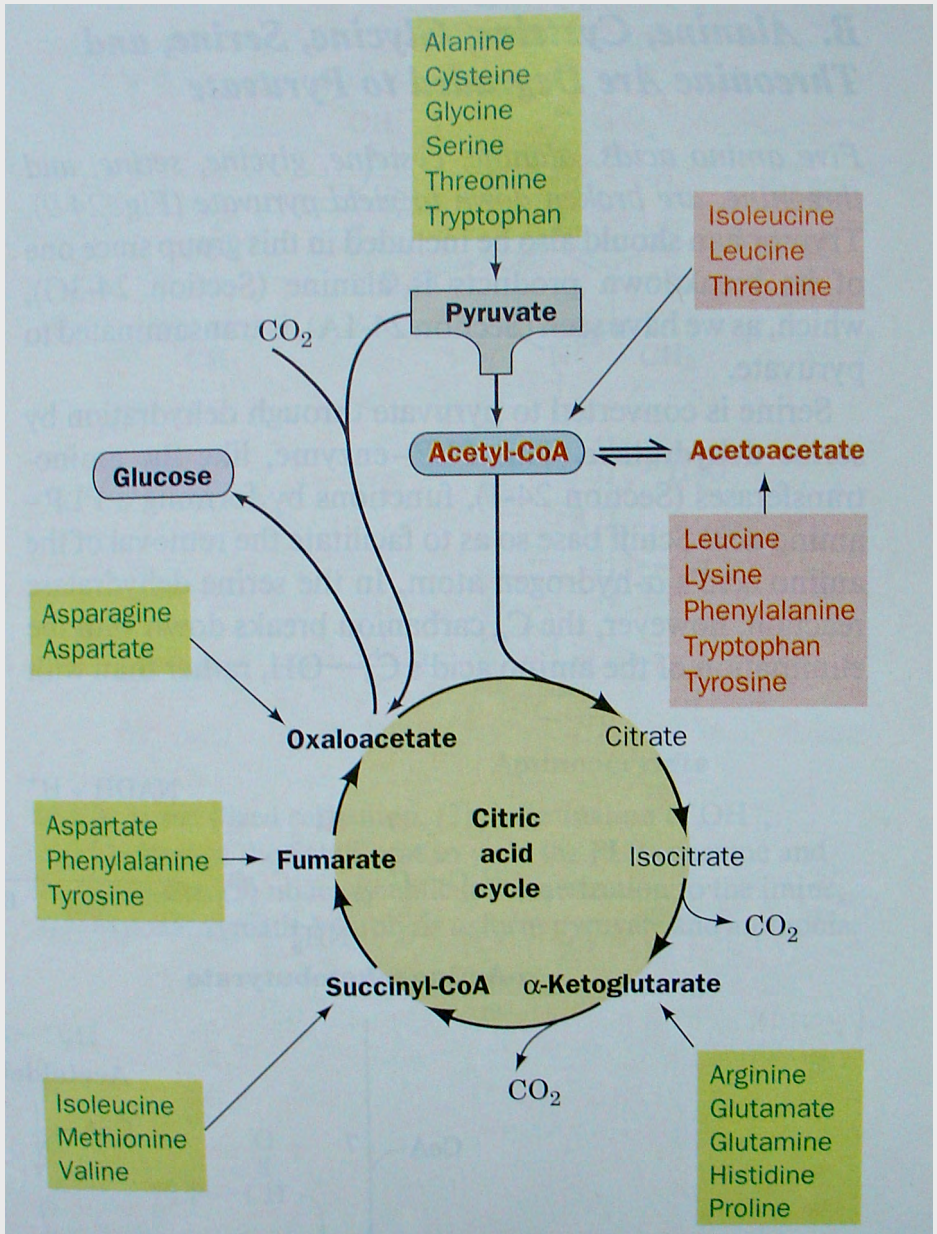


FIGURE 24-8. Amino acids are degraded to one of seven common metabolic intermediates. Glucogenic and ketogenic degradations are indicated in green and red, respectively.

Source of the amino acids?

MISSION	SAMPLE SOURCE	AMINO ACIDS	LIPID	SUGAR	OTHER
	Murchison Carbonaceous chondrite Australia1969 Engel & Nagy 1982	70-90% aa with L bias Non-earth aa's with L-aa bias at incidence of 2-9%			
	Murchison Carbonaceous chondrite Australia1969 Deamer & Pashley 1989		Fractions (chloroform/methanol extraction) act like membranes		
	Murray Deamer & Pashley 1989		Fractions (ibid) act like membranes		
	Mighei Deamer & Pashley 1989		Fractions (ibid) act like membranes		
	Murchison Carbonaceous chondrite Australia1969 Mautner, Leonard & Deamer1995	glycine	Nonanoic acid-forms vesicles**		pyrene
	Deamer 1997 METIORITES: Murchison AntarcticALH84044 (origin through to be Mars-ejected from Mars 16 million yrs ago into Earth crossing orbit-landed in Antarctica 13,000 yrs ago) AntarcticLEW 8534		Many examples lipid from meteorite samples Very good vesicles		
	Tagish Lake Cananda Pizzarello et al 2001				Dicarboxylic acids Pyridine carboxylic acids Sulphonic acid Aliphatic and aromatic hydrocarbons The insoluble carbon has deuterium enrichment Fullerines containing planetary He and Ar
	Murchison Carbonaceous chondrite Australia1969 Cooper et al 2001			POLYOLS polyhydroxylated cpds sugars, sugar alcohols, sugar acids	
	Murray metiorite Cooper et al 2001			POLYOLS polyhydroxylated cpds sugars, sugar alcohols, sugar acids	
	Murchison Carbonaceous chondrite Australia1969 Shimoyama & Ogasawara 2002	Dipeptides: Gly-Gly 11pmo23/g Cyclo(Gly-Gly) 18pmol/g No L or LL stereoisomers of protein aa's Monomer glycine 3 orders of mag [higher] than dipeptides			
	Yamato-791198 Carbonaceous chondrite Shimoyama & Ogasawara 2002	Dipeptides: Gly-Gly 11pmol/g Cyclo(Gly-Gly) 18pmol/g No L or LL stereoisomers of protein aa's Monomer glycine 4 orders of mag [higher] than dipeptides			
	Murchison Carbonaceous chondrite Australia1969 Meierhenrich et al 2004	aa's and Di peptides with no chiral bias			
Rosetta					
Stardust					
Mars					
Voyager 1 to Titan					Thick pink haze surrounding Titan. Tholins Indicative of organic cpds
Voyager 1 to dark hemisphere of Saturus-satellite Inpetus					tholins
Voyager 1 to Centaur 5145 Pholus					tholins
	Comet 81P/wild2 Sandford S A 2006				Organics rich in oxygen and nitrogen compared to meteorites Deuterium and N15 content suggest interstellar heritage

Table 1 Comet & Meteorite chemicals

Enlarge to read the table

Providing potential source of chemicals to earth from extraterrestrial in-fall

Table 1 References in the order that they appear in the table:

Engel & Nagy (1982) *Nature* **296**, 838.

Deamer & Pashley (1989) *Orig Life Evol Biosph* **19**, 21-38.

Mautner, Leonard & Deamer (1995) *Plan Space Sci* **43**, 139-147.

Deamer (1997) Microbiol *Mol Biol Rev* **61**, 239-261.

Pizzarello et al (2001) *Science* **293**, 2236-2239.

Cooper et al (2001) *Nature* **414**,879-883.

Shimoyama & Ogasawara (2002) *Orig Life Evol Biosph* **32**, 165-179.

Meierhenrich et al (2004) *PNAS* **101**,9182-9186.

Sandford et al (2006) *Meteor Plan Sci* **36**, 1117-1133.

L D discussion

Relates to amino acid condensation to protein on a water surface

By having 2 entities in a system choice is possible
Choice conveys a "value" to each entity

Value synonymous with "information"

Why was only one amino acid chirality selected (Laa)?

A polymer-based system of *n* units (typically n=100's), using homochirality (on-handed monomers)) is simplified in terms of the sequence combinations by a factor of 2.

How were L amino acids possibly enriched?

i.e what stellar events might have produced a greater proportion of L aa's making it more statistically likely that L would be selected by a system?

Data from *Murchison* meteorite:
interstellar grains from a neutron star remnant of a supernova show circular polarized synchrotron radiation (Bonner 1991).

Data from *Orion OMC-1*:
In the star formation region dust shows strong infrared circular polarization (Bailey et al 1998).

Bailey in 2001 suggested that circular polarization at short wavelengths might induce chiral asymmetry in interstellar organic molecules that were subsequently delivered to Earth by extra-terrestrial in-fall.

NB. Earth collects 20,000 tons of interplanetary dust/micro-meteorites per year - 99% are 50-500µm diameter and earlier when the Earth had just formed at 4.5e9yrs ago, this influx was greater.

Condensation of aa's to polymer in Water

Energetically unfavorable : H-Ala-OH + H-Gly-OH ➔ H-Ala-Gly-OH + H_2O ΔG^0 = 4.13kcal pH 7
Equilibrium constant of the reaction is only *ca* 10^{-3}

Equilibrium concentration of the dipeptide for 1M solutions of free amino acids is only slightly above 10^{-5}M

Dixon and Webb 1958 estimated that 1M solutions in each of the 20 proteinaceous aa's would yield at equilibrium a 10^{-99}M concentration for a 12kDa protein & the volume of this solution would be **10^{50} times the volume of the Earth.**

Therefore Energy input needed (bucket chemistry)

But -

On a smooth **Water Surface** with "snug & equal" atomic spacing of 3.7Å for reactants, with >66% hydrophobic aa's, possibly not so unfavorable?

No halogens at the point H,C,N,O first arise
-hydrophobic polymer amide remains stable

In summary:

•*A HYPOTHESIS is presented that POLYMER AMIDE functioned in the VERY EARLY UNIVERSE*

We suggest that polymer amide first formed at the condensation stage of the 2nd generation stars *ca.* 13Gyrs ago.

Hydrophobic polymer amide in conjunction with lipid = proteolipid.

•*The EVIDENCE that POLYMER AMIDE functioned at 4Gyrs ago in UNIVERSE*
Polymer amide has all the properties to provide a topology for complex chemistry 4Gyrs ago ago *(McGeoch & McGeoch Entrapment of water by subunit c of ATP synthase. J.R.Soc.Interface (2008) 5, 311-318).*

Proteolipids are the most conserved proteins in all organisms in terms of their amino acid sequence-they first arose 4Gya in the Pre-cells, the Archae and the Eubacteria.

Proteolipid in organic solvent when presented with water, assembles into a 3D scaffold of tubes that pinch off into vesicles when the solvent starts to evaporate.

A gene duplication of proteolipid polymer amide, subunit c of V-ATPase, is in large amounts in the ER of Eukaria.

The ER in some Mammalian cells forms a 3D scaffold of tubes/conduits on the nanometer scale *(McGeoch 2007 Topology of the mammalian cell via cryo-FIB etching. J Microscopy 227, 172-184).*

Polymer amide can thus provide a topology when:

1) It is just a basic chemical in solvent.
or
2) When it is a component in a complex cell structure like the ER.

This demonstrates that it could in principal have functioned from the time its elemental components first arose 13Gyrs ago to the present day.

Applied considerations:

Present day diseases related to loss of control of ancient template systems

Applies also to the most fundamental aspect of polymer amide which is that it is the MATERIAL BASIS for SEPARATING CHARGE.

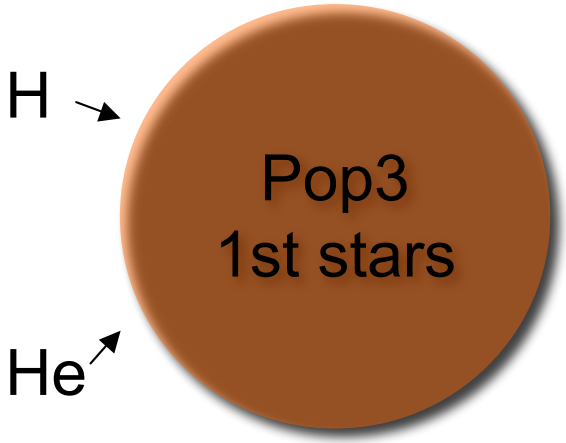
α-helix of polymer = the postman
β-sheet = the insulator
Solution with ions between the β-sheets = the force (charge) that has been confined.

Neurological diseases that could have as their basis, aberrant control (via nucleotide-based mutations in controlling proteins) of ancient template systems

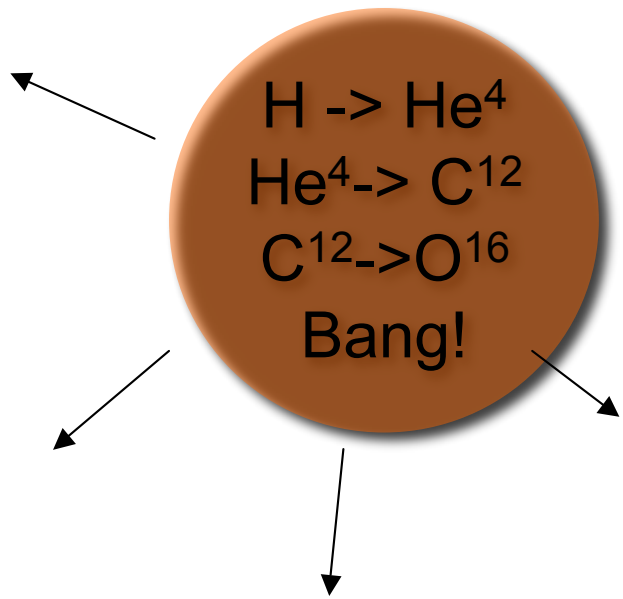
1)	2)	3)	4)
ANCIENT PROTEIN & how far back is their lineage relative to 4Gyrs ago when the first pre-cells started	FUNCTION	CODE-BASED SYSTEM that possibly suppresses/tightly regulates the ANCIENT PROTEIN's non-code based chemical reactions	DISEASE THAT COULD BE ASSOCIATED WITH LOSS OF **3**
PROTEOLIPID Back to 4e9yrs ago	1) ATP synthase rotor 2) Cation channel	**Enzymes associated with lipid synthesis, control and transport** Caspases Palmitoyl-protein-thioesterase Tripeptidyl-peptidase 7 transmembrane protein	**BATTENS DISEASE alias** **Neuronal Ceroid Lipofiuscinoses (CLN's)** CLN1 CLN2 CLN3
AMYLOID Back 2.7e9yrs a g o Amyloid fibres found in Ecoli biofilms *Otzen D, Nielsen PH Cell mol Life Sci 2008 65,910-27.*	Memory	**PRESILINS: PS1, PS2** Proteolytic action on amyloid precursor protein that results in beta amyloid production (plaques). When PS1 and PS2 are mutated the beta amyloid is miss-folded. Same for **Tau protein** whose function is to phosphorylate microtubules of the cytoskeleton. Over phosphorylated tau aggregates with other tau to form tangles. Abnormal **Caspase-6 activation** disrupts CNS cytoskeleton	**ALZHEIMER'S**
HUNTINGTIN Back to the evolution of insects 4e8yrs ago (Htt/mHtt (mutant)) DNA code acquires Polyglutamate (CAG) repeats Homologies in Drosophila	Vesicle trafficing involving clathrin.	**Caspases** degrade the N-terminus of the normal Htt and remove the glutamate repeats. In protein with extra repeats they are not all removed by the carpases. mHtt binds with $TAF_{II}130$, a coactivator to CREB dependent transcription.	**HUNTINGTON'S CHOREA**

Spare slides for discussion

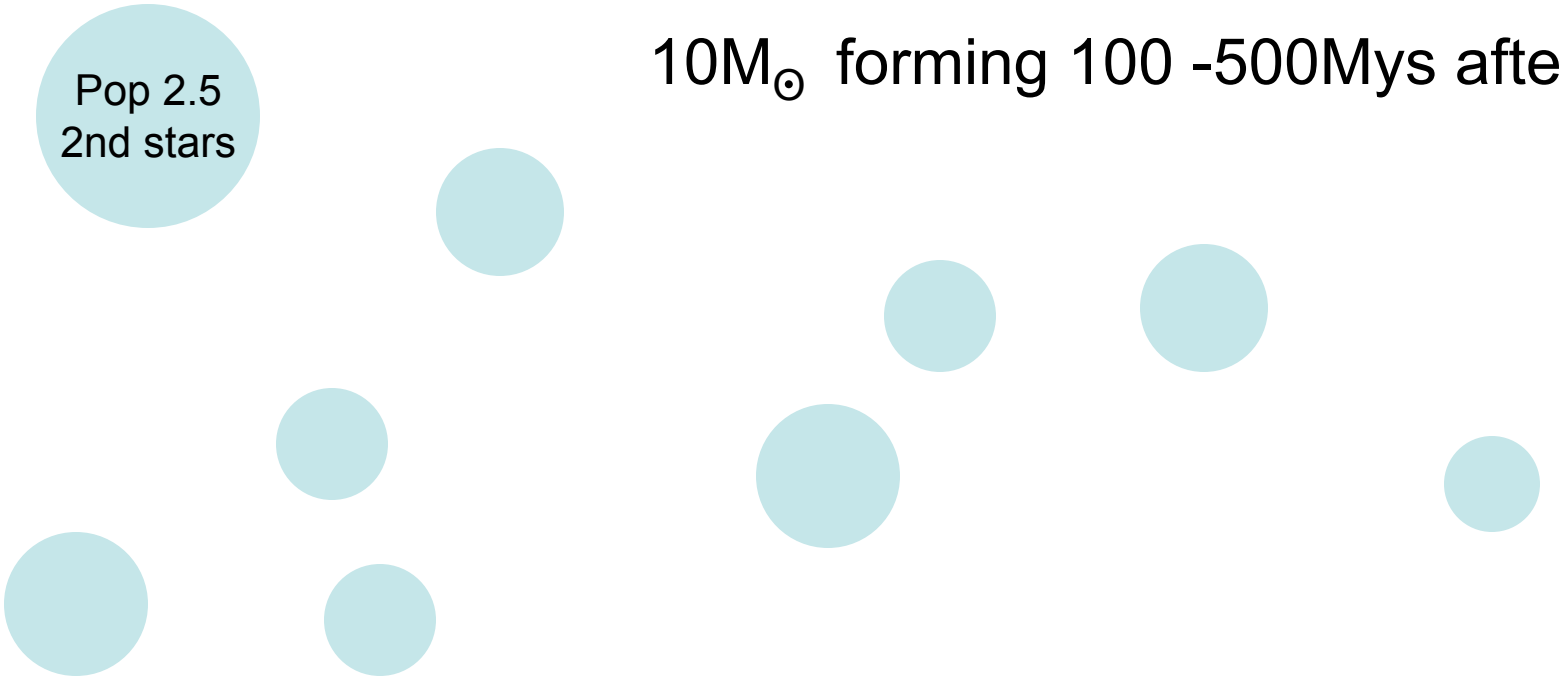
The 1st STARS



$100M_{\odot}$ forming 100M yrs after BB



After few 100M yrs: H, He, O, C, N and not much else

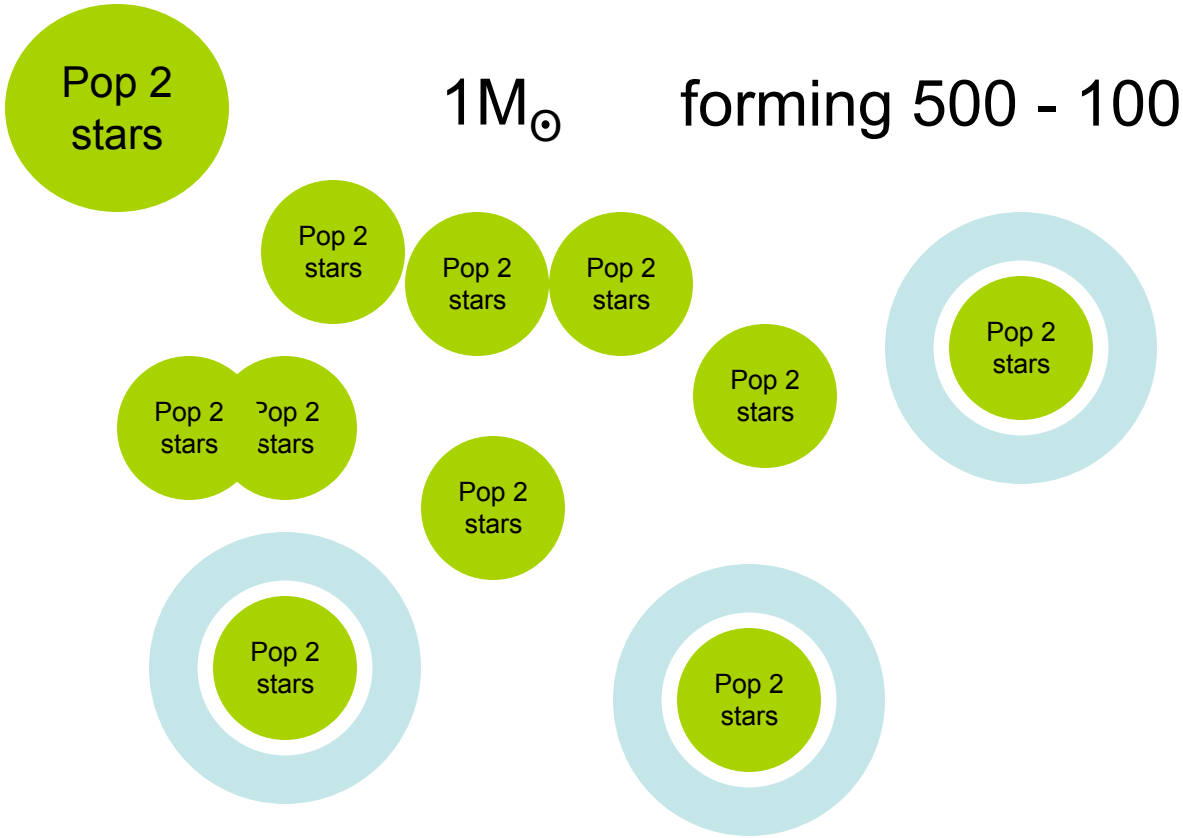


$10M_{\odot}$ forming 100 -500Mys after BB

Re-ionization (harsh environment) due to radiation from Pop3 stars

Oxygen could aid Pop 2.5 formation at smaller size than Pop3 (hyperfine transition radiation from ground state)

Pop 2.5 produces yet more O, C and N



$1M_{\odot}$ forming 500 - 1000M yrs after BB

Hypothesis: Water was dominant constituent of accretion discs

First chance of complex chemistry of proteins to occur

Acknowledgements: NIH, Center for Nanoscale Systems Harvard
Guido Guidotti
David C. Bell (TEM)
Richard Schalek (SEM)

North point of Harbour Island

North Eleuthera Bahamas

In calm warm ocean, alternate layers of cyanobacteria/biofilms of $CaCO_3$ form localized coral mounds termed stromatolites.
Date from 3.5e9yrs ago

Experimental Stromatolite layer depositions have been produced abiotically - via controlled nano scale sprayed layer deposition.

A Stromatolite is a Biogeochemical Engine which translates energy into morphology

They could be candidates for containing a polymer amide template system along side the cyanobacterial nucleotide code?

Movie of Marine Foam - salt water/air bubbles encapsulated by amino acids

cf: Mopper & Zika 1987 Nature 325,246-249: Free amino acids in marine rains.

Universe diameter = 20billion light years = 1.1757e17miles

Jeffrey Weeks from Canton, New York diameter = 70 billion light years = 2.3e18miles

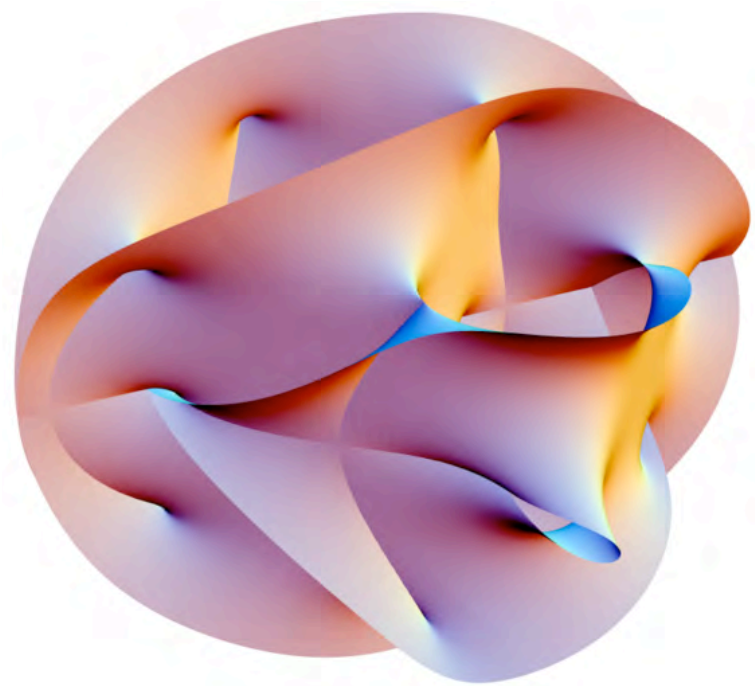
At the centre of the debate are observations by NASA's Wilkinson Microwave Anisotropy Probe (WMAP), which was launched in 2001. The probe measures temperature ripples in the "cosmic microwave background", the afterglow radiation from the big bang fireball. Weeks and his colleagues, a team of astrophysicists in France, say the WMAP results suggest that the Universe is not only small, but that space wraps back on itself in a bizarre way (***Nature*, vol 425, p 593 2003**).

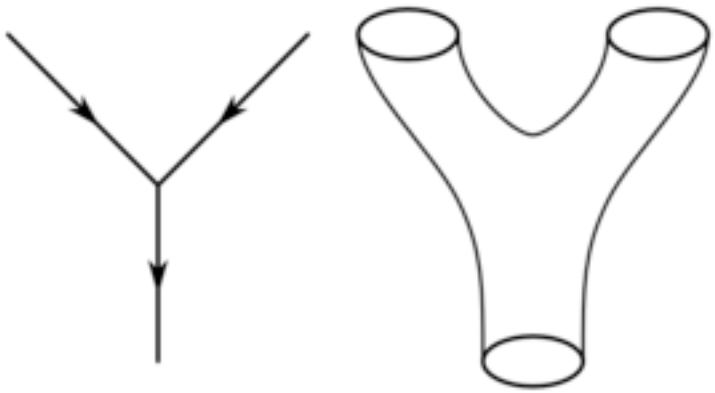
A **light-year** is equal to:

- 9,460,730,472,580.8 km (about 9.461 Pm)
- 5,878,625,373,183.61 statute miles
- about 63,241 astronomical units
- about 0.3066 parsecs

http://en.wikipedia.org/wiki/String_theory

String Theory





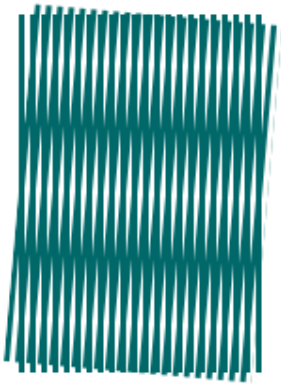
Interaction in the subatomic world: world lines of point-like particles in the Standard Model or a world sheet swept up by closed strings in string theory

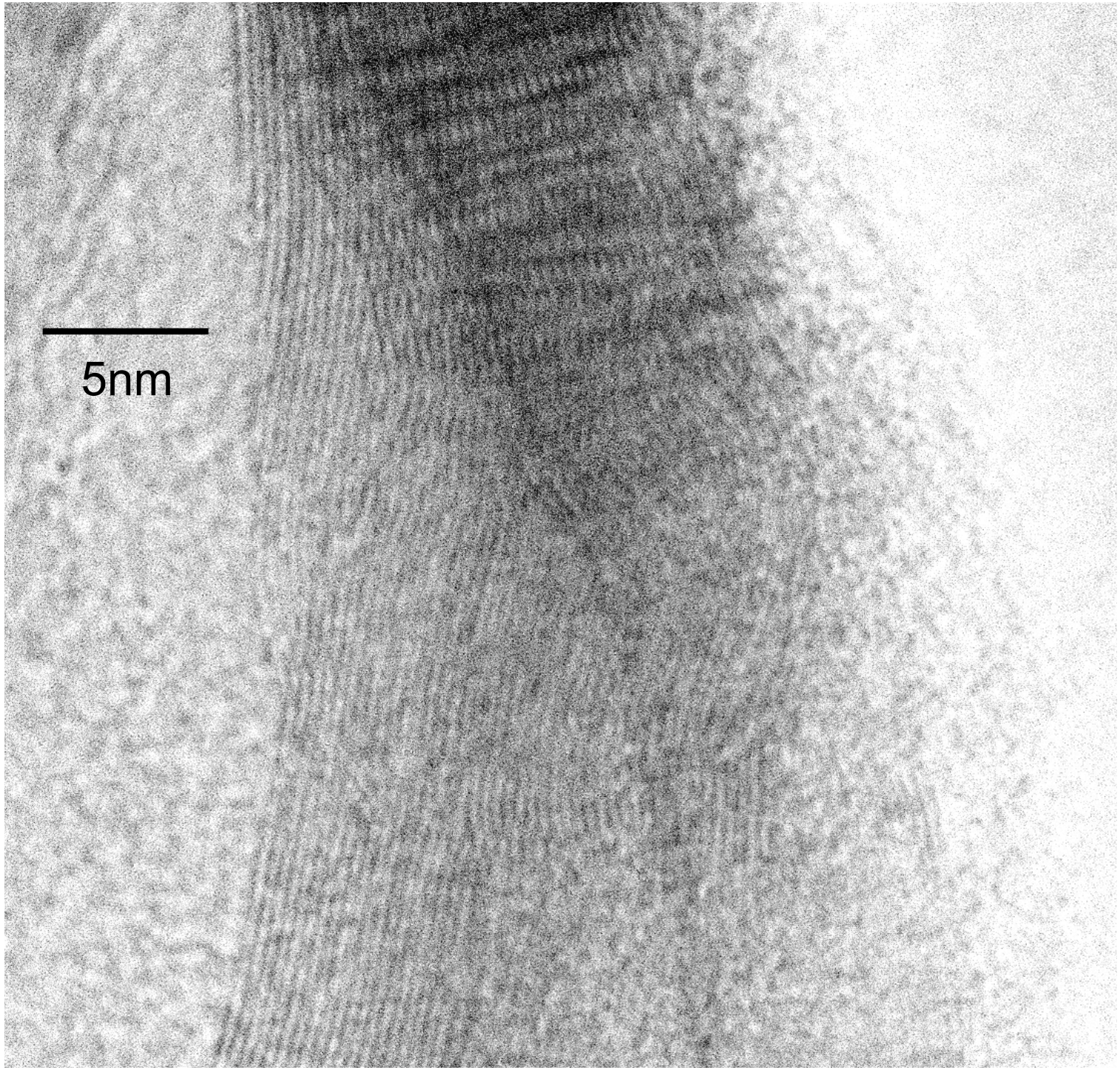
String theory is an incomplete mathematical approach to theoretical physics, whose building blocks are one-dimensional extended objects called strings, rather than the zero-dimensional point particles that form the basis for the standard model of particle physics. By replacing the point-like particles with strings, an apparently consistent quantum theory of gravity emerges, which has not been achievable under quantum field theory. Usually, the term string theory includes a group of related superstring theories and a few related frameworks such as M-theory, which seeks to unite them all.

TIME (years ago)	EVENT
13.73±0.12e9	Start
Minus 10ms	Baryons, Dark Matter, Dark Energy
Minus 100ms	1st Nucleosynthesis: nuclei H, D, He, Li then H_2
Minus 5e8yrs	Accretion of 1st generation stars 2nd Nucleosynthesis: low mass elements from H1 to O16, mainly H, C, N, O
Minus 8e8yrs	Start of accretion of 2nd generation stars Water (ice + polymers like polymer amide could have been involved in this accretion)
12e9	2nd generation stars 3rd Nucleosynthesis: higher mass elements *NB now halogens exist*
	FUSION reactions continue - star formation after the 2nd generation stars was not linear and continues in present times
4.5e9 4.4e9 3.8e9 7.	Formation of Earth – a 3rd generation star 1st oceans condense CO_2 fixation by pre-cells ***NB only 7e8yrs between Earths formation & CO_2 fixation*** RNA/DNA code forms in pre-cell systems
3.5e9	Stromatolites: alternate layers of bacteria and ppted Ca carbonate due to the bacteria lowering the CO_2 Are these structures 100% nucleotide code based? Protein synthesis inhibitors do not completely inhibit their nitrogen metabolism.
2.7e9	Archaea & Eubacteria
2e9	Eukaria
1.5e9	Acritachs, unicellular algae
1.2e9	Multicellular red algae
6.8e7 0	Mammals Present day

The 3 proteolipid ribbons at the top left overlap lengthwise and show “moire patterns” (term derived from *Bombyx mori)* due their overlapping protein backbones attenuating rather than transmitting some of the electrons.

A moire pattern, formed by two sets of parallel lines, one set inclined at an angle of 5^0 to the other



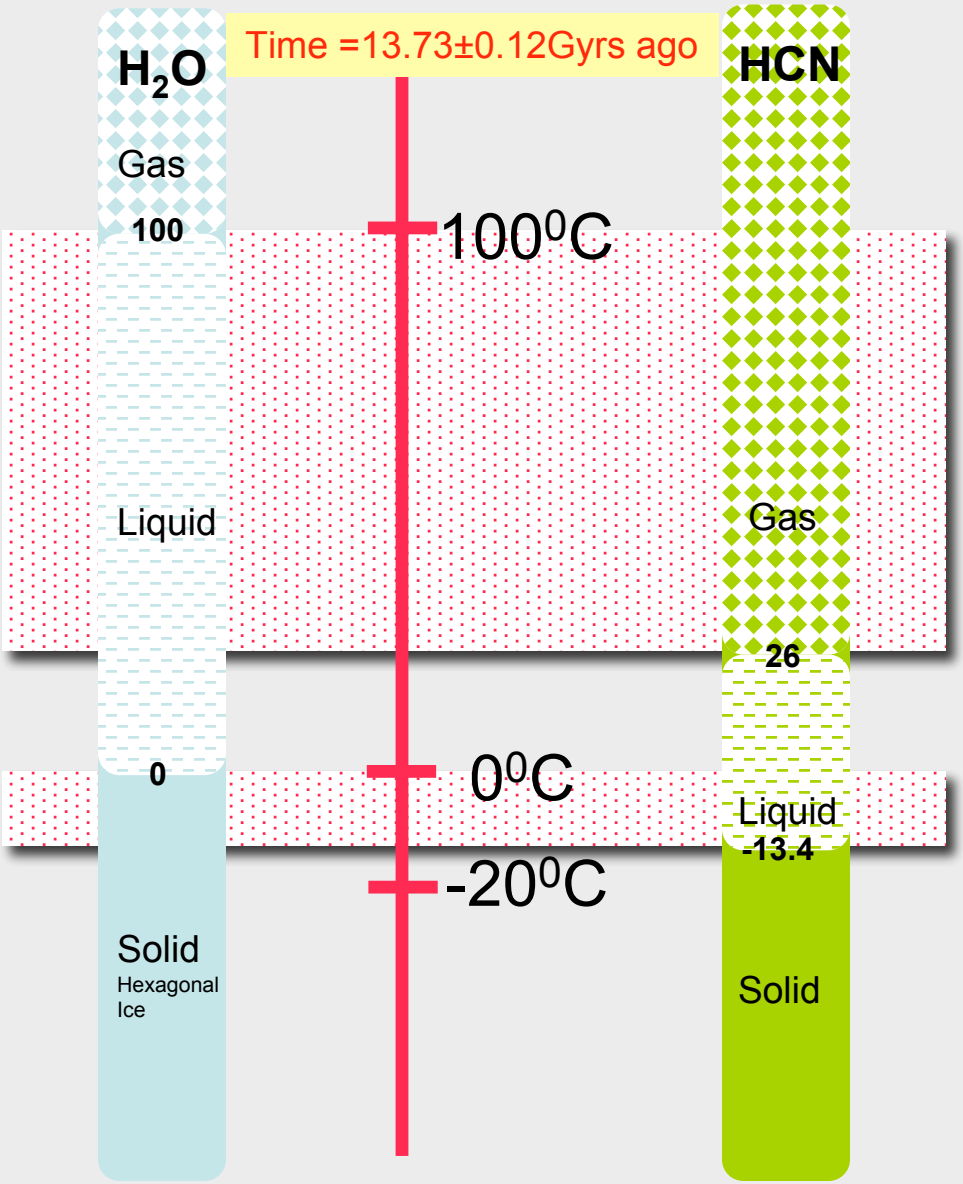


The Lotus leaf surface is of hydrophobic wax with upward projections of nano-hairs and micro-bumps
Water on this surface has a contact angle of 170^0
As a reference Human skin which is slightly hydrophobic, has a contact angle for water of 90^0

In strong sunlight when the leaf is undergoing photosynthesis the O_2 gas emerges through the open leaf somatae at such a rate, that it ejects micro-volumes of water outwards from the surface of the main droplet.

Photosynthesis: $6\ CO_2 + 12\ H_2O \rightarrow C_6H_{12}O_6 + 6\ O_2 + 6\ H_2O$

Small amount of CO_2 too, from leaf respiration



We start with the assumption that amide is synthesized in HCN from aldehydes & ketones

Shadowed Boxes represent temperature ranges that would support exchange of amide from gaseous HCN to liquid H_2O & liquid HCN to the surface of solid hexagonal water ice.

NB these are reactions at Earths Atmospheric Pressure