UNIVERSITÄT OSNABRÜCK

Lomonosov University,
Moscow

Anoxic geothermal fields and the early evolution of first cellular organisms

Armen Mulkidjanian

Опарин

International Conference
THE PROBLEM OF THE ORIGIN OF LIFE
Moscow – September 22-26, 2014

Merging phylogenomics with geochemistry:

Bioinformatics:

Daria Dibrova

University of Osnabrück

Geochemistry:

Andrew Bychkov

Moscow State University

Biochemistry:

Michael Galperin

Genomics:

Eugene Koonin

National Center for Biotechnology Information,
National Library of Medicine,
National Institutes of Health, USA

Support: DFG, DAAD, NIH, RSF

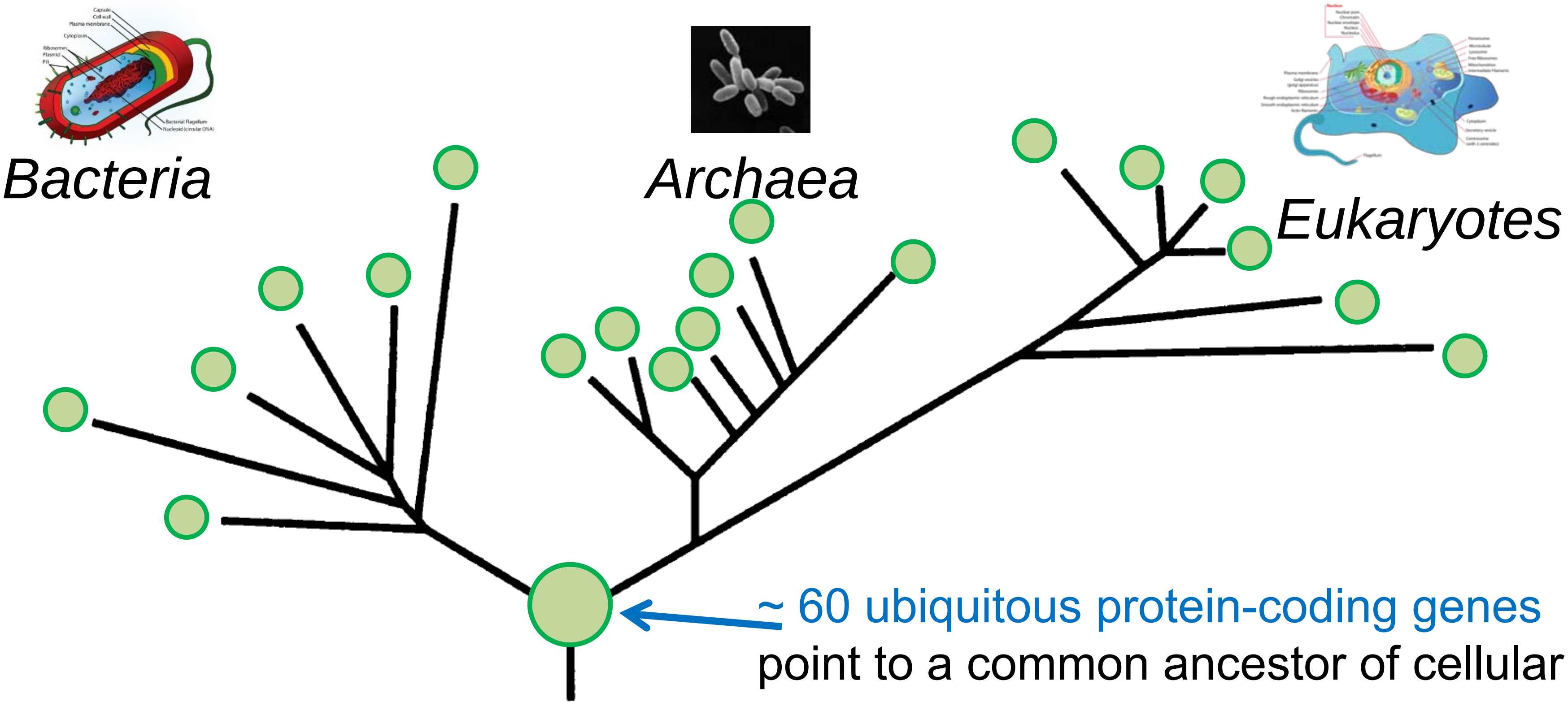
- **Introduction: Ubiquitous protein set and the Last Universal Cellular Ancestor**
- **Emergence of the first cells in K^+-rich environments**
- **Invasion of Na^+-rich habitats**
- **On the co-evolution of membranes and enzyme machinery**

It is possible to get into the past by descending along the branches of the tree of life

Image: Boris Kulikov (www.boriskulikov.com)

By finding those traits that are shared by modern organisms, it is possible to reconstruct their ancient forms

Last Universal Cellular Ancestor (**LUCA**)



~ 60 ubiquitous protein-coding genes point to a common ancestor of cellular organisms, the LUCA, that existed before the separation of bacteria and archaea.

(Koonin, *Annu Rev Genomics Hum Genet*, **2000**; Charlebois *et al.*, *Genome Res*, **2004**)

Based on 16S rRNA similarity, all cellular life forms could be separated into **three domains**.

(Woese *et al.*, *PNAS*, **1977**;
Woese *et al.*, *PNAS*, **1990**)

The figures of life forms are taken from the Wikipedia

Ubiquitous proteins should have been present in the LUCA

Translation (about 40 genes):
- ribosomal proteins
- aminoacyl-tRNA-synthetases
- translation factors

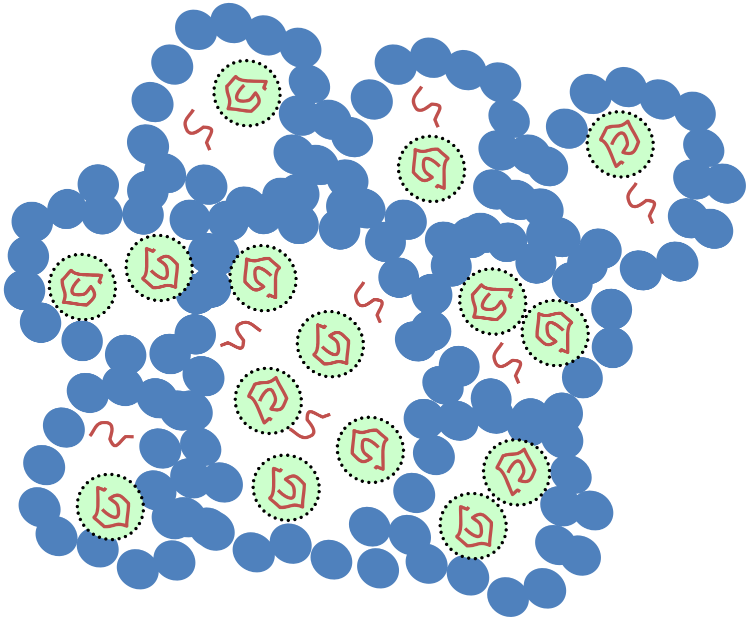
Membrane enzymes:
- Proteolipid subunit of the rotary ATPase
- CDP-diglyceride synthetase
- SecY translocase subunit

Procession of (poly)nucleotides
- recombinase RecA/RadA
- transcription factors NusG and NusA
- 5’-3’-exonuclease
- topoisomerase IA,
- α- and β-subunits of rotary membrane ATPase
- pseudouridylate synthase
- thymidylate kinase
- chaperone GroEL

Redox enzymes:
- thioredoxin
- thioredoxin reductase

Based on (Koonin, *Annu Rev Genomics Hum Genet* **2000**; Charlebois *et al.*, *Genome Res*, **2004**)

The LUCA could have been a consortium of mutually dependent organisms that could synthesize proteins and should have shared their enzymes and metabolites (Woese, 1998, Koonin & Martin, 2005, Szathmary et al. 2007).



- ❖ Introduction: Ubiquitous protein set and the Last Universal Cellular Ancestor
- ❖ **Emergence of the first cells in K^+-rich environments**
- ❖ Invasion of Na^+-rich habitats
- ❖ On the co-evolution of membranes and enzyme machinery

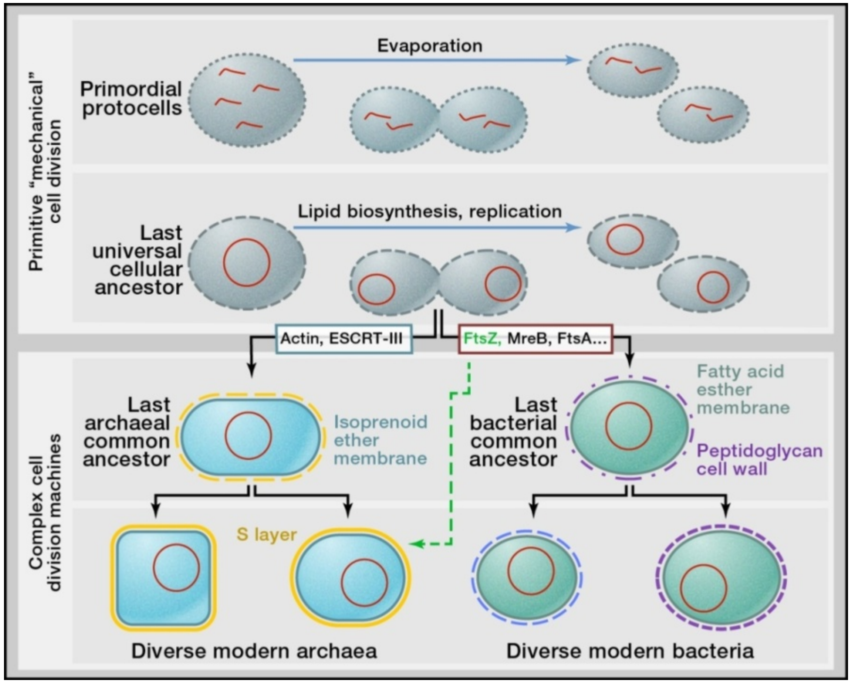
Origin of first cells at terrestrial, anoxic geothermal fields

Armen Y. Mulkidjanian[a,b,1], Andrew Yu. Bychkov[c], Daria V. Dibrova[a,d], Michael Y. Galperin[e], and Eugene V. Koonin[e,1]

[a]School of Physics, University of Osnabrück, D-49069 Osnabrück, Germany; [b]A. N. Belozersky Institute of Physico-Chemical Biology and Schools of [c]Geology and [d]Bioengineering and Bioinformatics, Moscow State University, Moscow 119992, Russia; and [e]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894

Edited* by Norman H. Sleep, Stanford University, Stanford, CA, and approved January 17, 2012 (received for review October 28, 2011)

Semipermeable membranes of the first cells

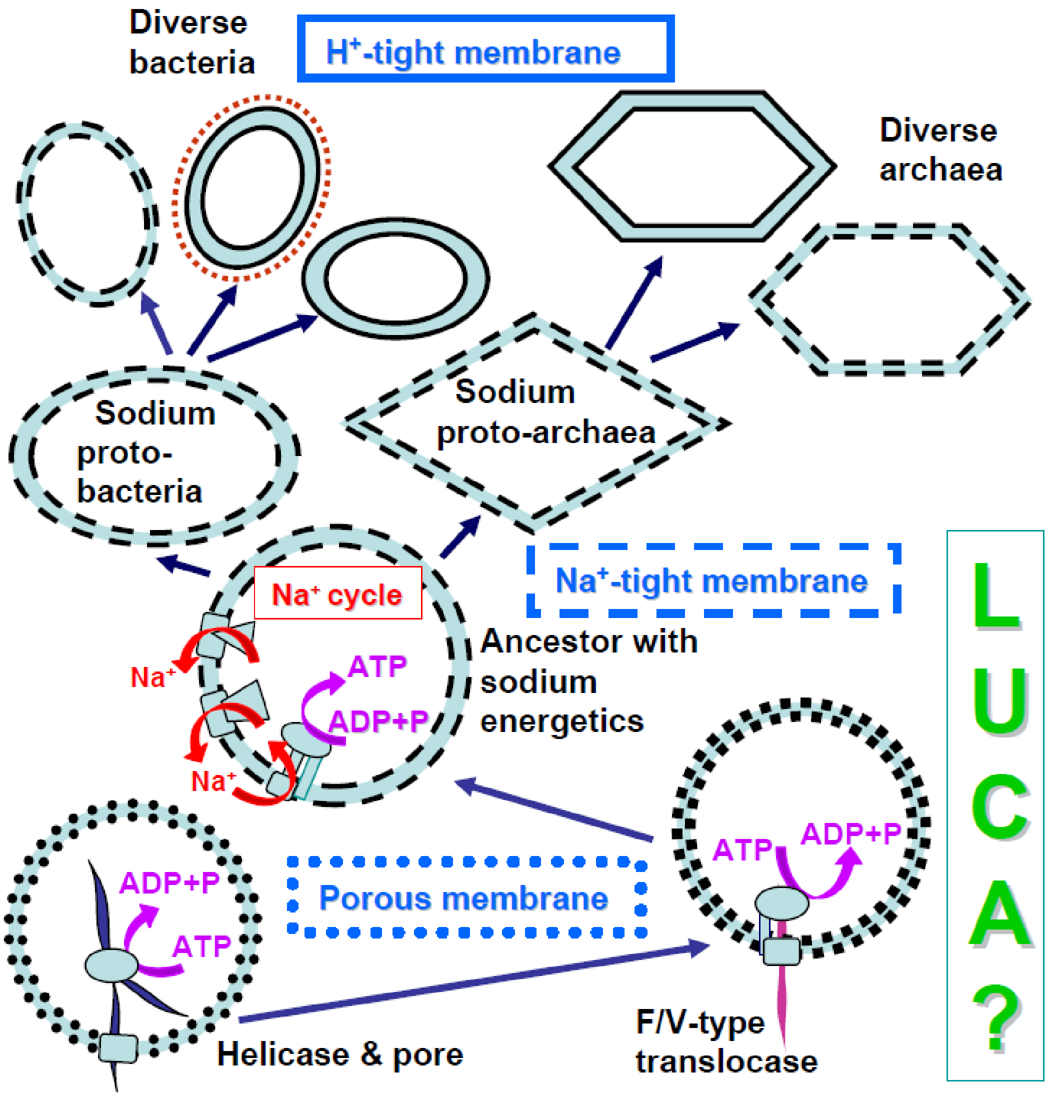


The tightness of cell envelopes increased with time; this evolution could be reconstructed by means of comparative genomics.

The first cells should have been in equilibrium with environment as it concerns small molecules and ions.

From: Koonin and Mulkidjanian, 2013, *Cell*

Co-evolution of the lipid bilayer and membrane bioenergetics



Transition from membranes that were permeable to small ions, but not to RNA and proteins, to the membranes that became tight first to Na$^+$ and then to H$^+$ ions.

From: Mulkidjanian et al., 2009, *TIBS*

"...the very earliest organisms ... had as yet no nuclei and an enclosing membrane could have been only of the most elementary character.

From the first there must have been an adjustment in the composition of very simple organisms to that of their medium,.... which diffusing into each minute multi-micellar mass brought into it the inorganic elements in the proportions in which they obtained in the external medium."

Archibald Byron Macallum, 1926
The paleochemistry of the body fluids and tissues,
Physiological Reviews

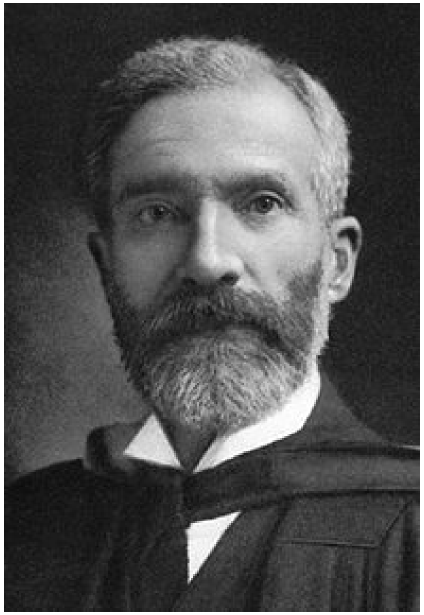


Photo: Wikipedia

The ionic composition of the cytoplasm differs dramatically from the chemistry of sea water.

The K^+/Na^+ ratio in all cells is > 1;
cells contain high levels of transition metals such as Zn, Fe, Mn.

How old are these traits?

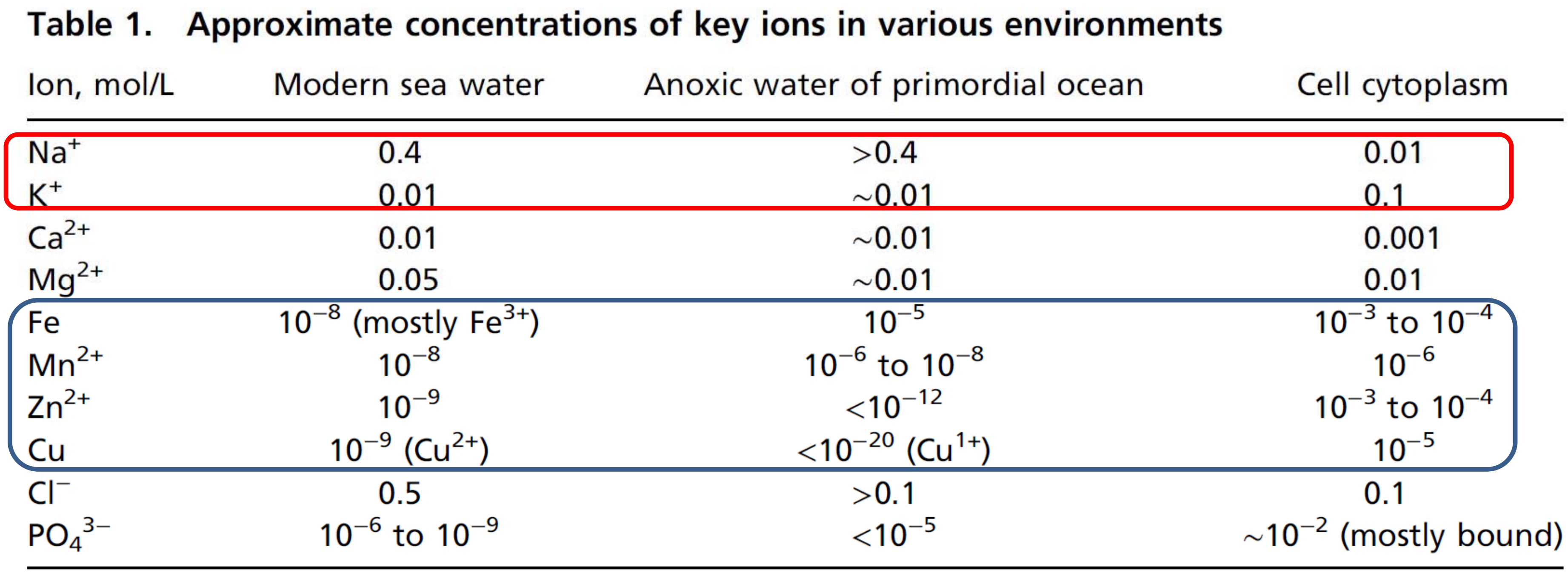


Table 1. Approximate concentrations of key ions in various environments

Ion, mol/L	Modern sea water	Anoxic water of primordial ocean	Cell cytoplasm
Na^+	0.4	>0.4	0.01
K^+	0.01	∼0.01	0.1
Ca^{2+}	0.01	∼0.01	0.001
Mg^{2+}	0.05	∼0.01	0.01
Fe	10^{-8} (mostly Fe^{3+})	10^{-5}	10^{-3} to 10^{-4}
Mn^{2+}	10^{-8}	10^{-6} to 10^{-8}	10^{-6}
Zn^{2+}	10^{-9}	$<10^{-12}$	10^{-3} to 10^{-4}
Cu	10^{-9} (Cu^{2+})	$<10^{-20}$ (Cu^{1+})	10^{-5}
Cl^-	0.5	>0.1	0.1
PO_4^{3-}	10^{-6} to 10^{-9}	$<10^{-5}$	$\sim10^{-2}$ (mostly bound)

From: Mulkidjanian et al., 2012, *PNAS*

The inorganic composition of cytoplasm seems to be an old trait because it is similar in Bacteria and Archaea, which separated more than 3.4 Ga ago

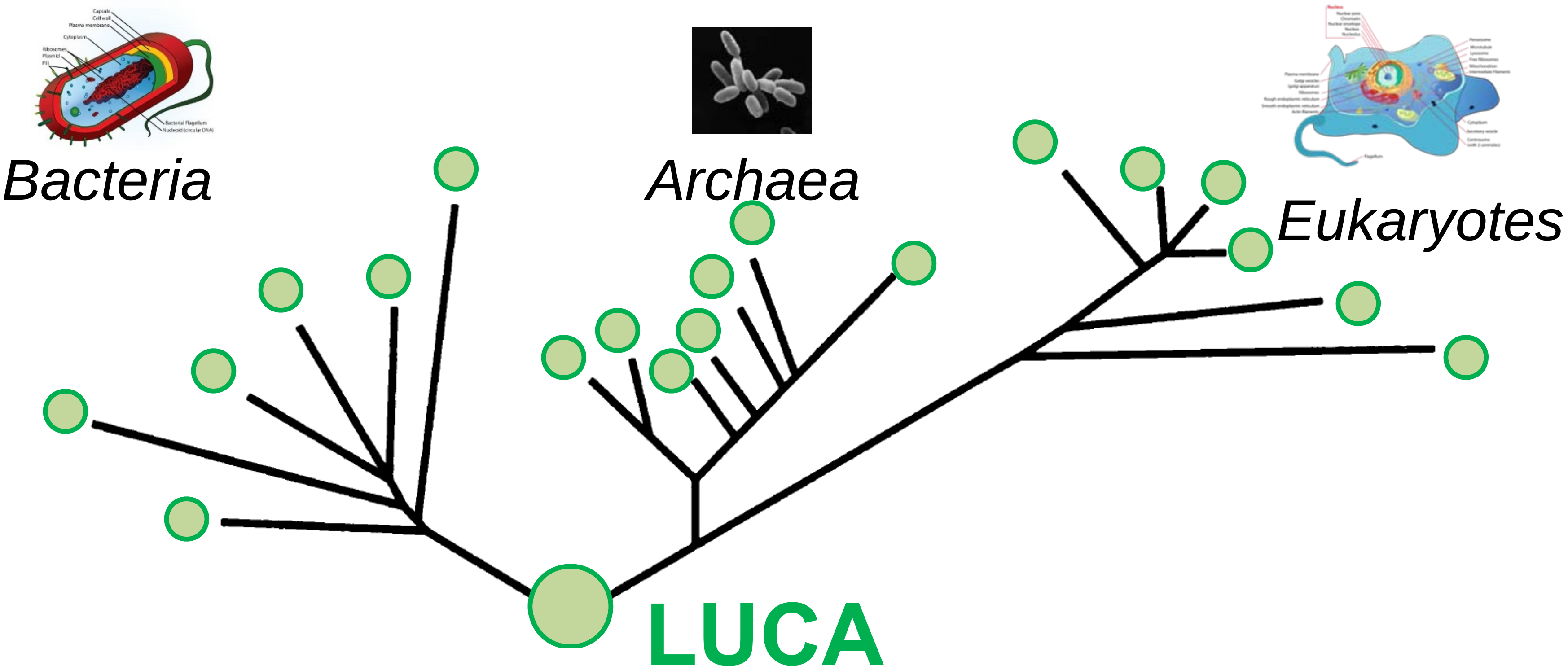
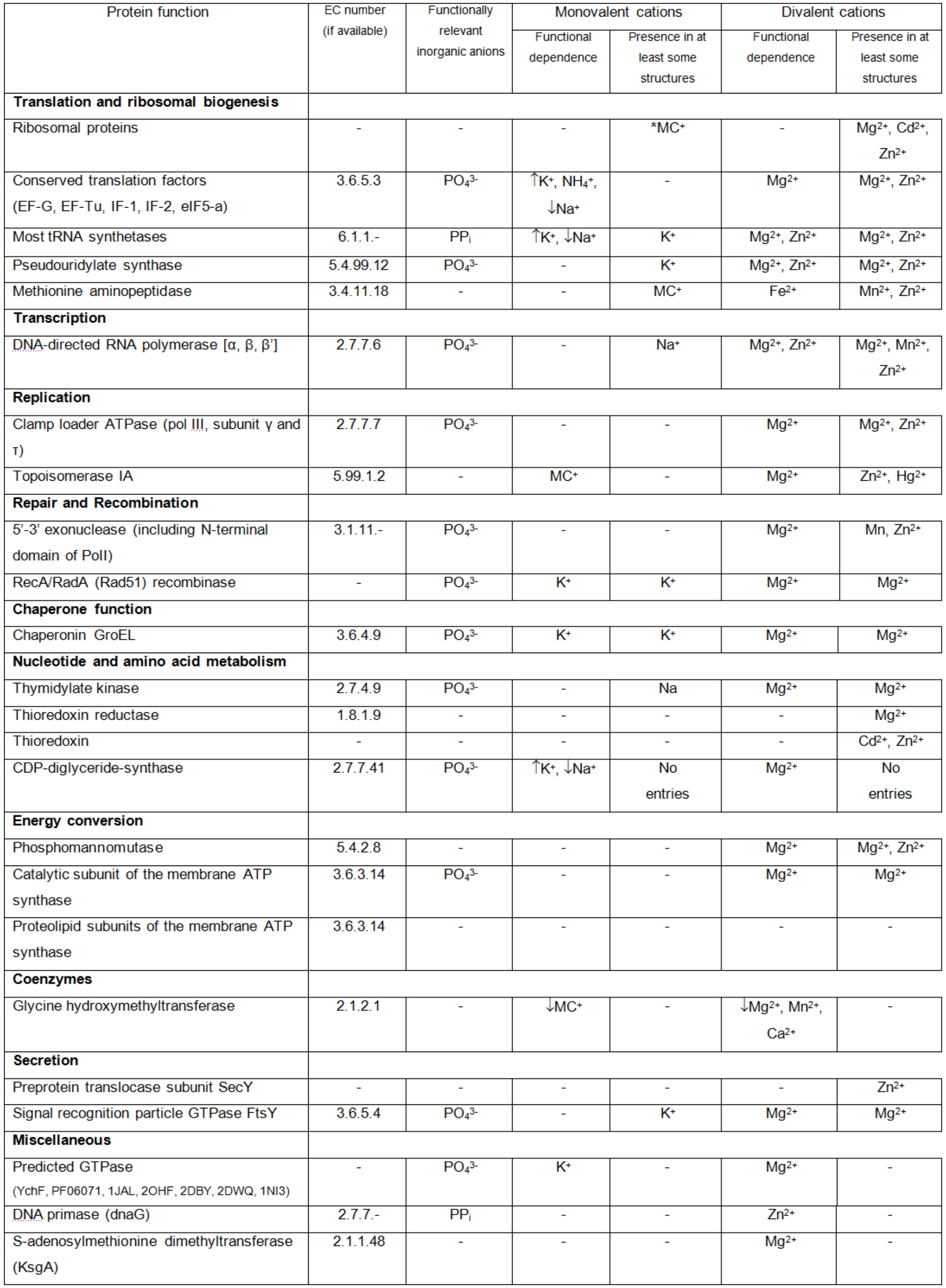


Figure: D. Dibrova

Ubiquitous proteins and inorganic ions

We checked the structural and functional association of ubiquitous proteins with inorganic ions

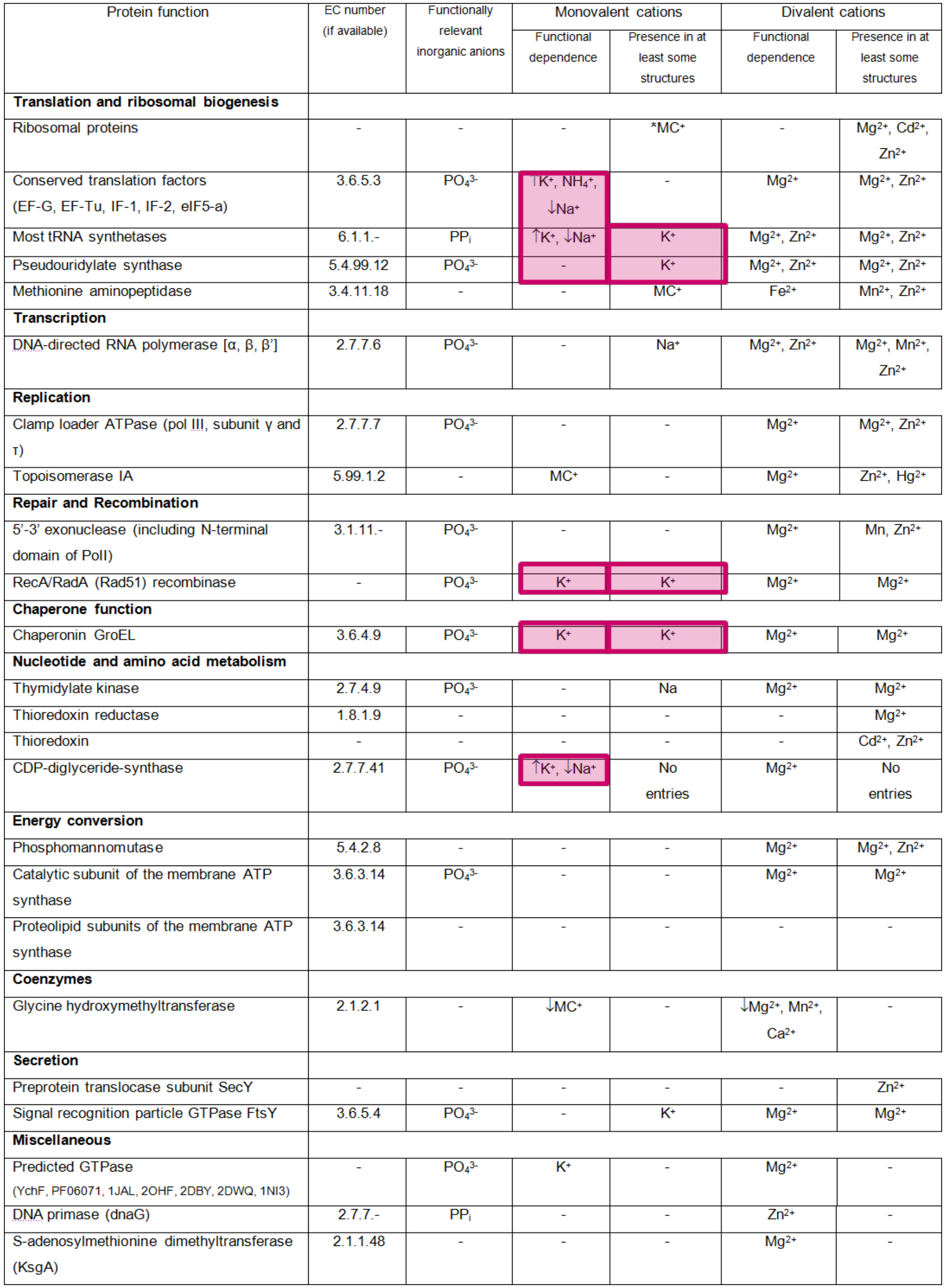


Protein function	EC number (if available)	Functionally relevant inorganic anions	Monovalent cations		Divalent cations	
			Functional dependence	Presence in at least some structures	Functional dependence	Presence in at least some structures
Translation and ribosomal biogenesis						
Ribosomal proteins	-	-	-	*MC$^+$	-	Mg^{2+}, Cd^{2+}, Zn^{2+}
Conserved translation factors (EF-G, EF-Tu, IF-1, IF-2, eIF5-a)	3.6.5.3	PO_4^{3-}	↑K^+, NH_4^+, ↓Na^+	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Most tRNA synthetases	6.1.1.-	PP_i	↑K^+, ↓Na^+	K^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Zn^{2+}
Pseudouridylate synthase	5.4.99.12	PO_4^{3-}	-	K^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Zn^{2+}
Methionine aminopeptidase	3.4.11.18	-	-	MC$^+$	Fe^{2+}	Mn^{2+}, Zn^{2+}
Transcription						
DNA-directed RNA polymerase [α, β, β’]	2.7.7.6	PO_4^{3-}	-	Na^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Mn^{2+}, Zn^{2+}
Replication						
Clamp loader ATPase (pol III, subunit γ and τ)	2.7.7.7	PO_4^{3-}	-	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Topoisomerase IA	5.99.1.2	-	MC$^+$	-	Mg^{2+}	Zn^{2+}, Hg^{2+}
Repair and Recombination						
5’-3’ exonuclease (including N-terminal domain of PolI)	3.1.11.-	PO_4^{3-}	-	-	Mg^{2+}	Mn, Zn^{2+}
RecA/RadA (Rad51) recombinase	-	PO_4^{3-}	K^+	K^+	Mg^{2+}	Mg^{2+}
Chaperone function						
Chaperonin GroEL	3.6.4.9	PO_4^{3-}	K^+	K^+	Mg^{2+}	Mg^{2+}
Nucleotide and amino acid metabolism						
Thymidylate kinase	2.7.4.9	PO_4^{3-}	-	Na	Mg^{2+}	Mg^{2+}
Thioredoxin reductase	1.8.1.9	-	-	-	-	Mg^{2+}
Thioredoxin	-	-	-	-	-	Cd^{2+}, Zn^{2+}
CDP-diglyceride-synthase	2.7.7.41	PO_4^{3-}	↑K^+, ↓Na^+	No entries	Mg^{2+}	No entries
Energy conversion						
Phosphomannomutase	5.4.2.8	-	-	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Catalytic subunit of the membrane ATP synthase	3.6.3.14	PO_4^{3-}	-	-	Mg^{2+}	Mg^{2+}
Proteolipid subunits of the membrane ATP synthase	3.6.3.14	-	-	-	-	-
Coenzymes						
Glycine hydroxymethyltransferase	2.1.2.1	-	↓MC$^+$	-	↓Mg^{2+}, Mn^{2+}, Ca^{2+}	-
Secretion						
Preprotein translocase subunit SecY	-	-	-	-	-	Zn^{2+}
Signal recognition particle GTPase FtsY	3.6.5.4	PO_4^{3-}	-	K^+	Mg^{2+}	Mg^{2+}
Miscellaneous						
Predicted GTPase (YchF, PF06071, 1JAL, 2OHF, 2DBY, 2DWQ, 1NI3)	-	PO_4^{3-}	K^+	-	Mg^{2+}	-
DNA primase (dnaG)	2.7.7.-	PP_i	-	-	Zn^{2+}	-
S-adenosylmethionine dimethyltransferase (KsgA)	2.1.1.48	-	-	-	Mg^{2+}	-

K^+ requirements of ubiquitous cellular systems

Set of ubiquitous proteins is enriched with K^+-dependent enzymes.

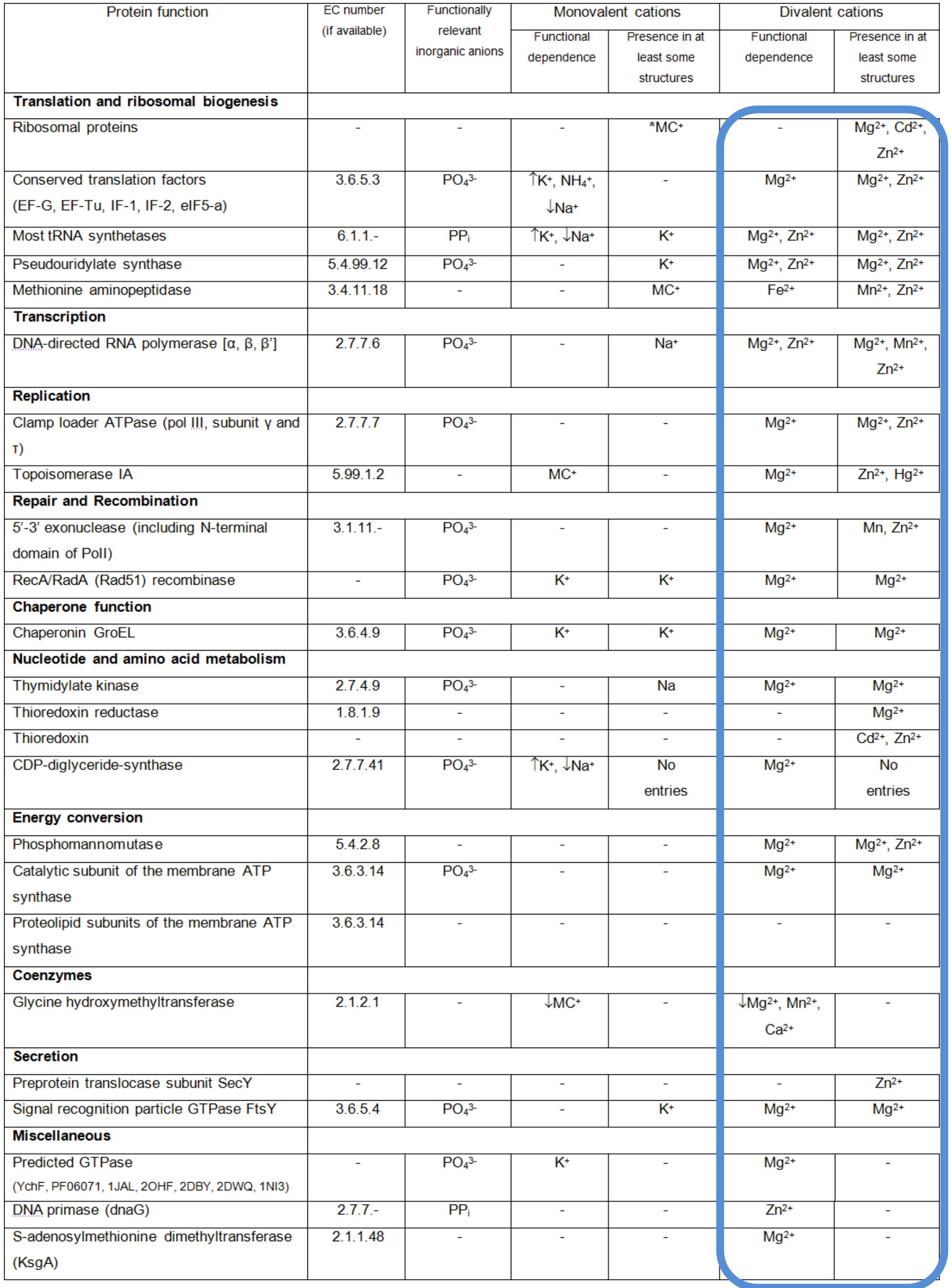
Specifically, protein synthesis requires high K^+/Na^+ ratio in the cytoplasm



Protein function	EC number (if available)	Functionally relevant inorganic anions	Monovalent cations		Divalent cations	
			Functional dependence	Presence in at least some structures	Functional dependence	Presence in at least some structures
Translation and ribosomal biogenesis						
Ribosomal proteins	-	-	-	*MC^+	-	Mg^{2+}, Cd^{2+}, Zn^{2+}
Conserved translation factors (EF-G, EF-Tu, IF-1, IF-2, eIF5-a)	3.6.5.3	PO_4^{3-}	↑K^+, NH_4^+, ↓Na^+	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Most tRNA synthetases	6.1.1.-	PP_i	↑K^+, ↓Na^+	K^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Zn^{2+}
Pseudouridylate synthase	5.4.99.12	PO_4^{3-}	-	K^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Zn^{2+}
Methionine aminopeptidase	3.4.11.18	-	-	MC^+	Fe^{2+}	Mn^{2+}, Zn^{2+}
Transcription						
DNA-directed RNA polymerase [α, β, β']	2.7.7.6	PO_4^{3-}	-	Na^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Mn^{2+}, Zn^{2+}
Replication						
Clamp loader ATPase (pol III, subunit γ and τ)	2.7.7.7	PO_4^{3-}	-	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Topoisomerase IA	5.99.1.2	-	MC^+	-	Mg^{2+}	Zn^{2+}, Hg^{2+}
Repair and Recombination						
5'-3' exonuclease (including N-terminal domain of PolI)	3.1.11.-	PO_4^{3-}	-	-	Mg^{2+}	Mn, Zn^{2+}
RecA/RadA (Rad51) recombinase	-	PO_4^{3-}	K^+	K^+	Mg^{2+}	Mg^{2+}
Chaperone function						
Chaperonin GroEL	3.6.4.9	PO_4^{3-}	K^+	K^+	Mg^{2+}	Mg^{2+}
Nucleotide and amino acid metabolism						
Thymidylate kinase	2.7.4.9	PO_4^{3-}	-	Na	Mg^{2+}	Mg^{2+}
Thioredoxin reductase	1.8.1.9	-	-	-	-	Mg^{2+}
Thioredoxin	-	-	-	-	-	Cd^{2+}, Zn^{2+}
CDP-diglyceride-synthase	2.7.7.41	PO_4^{3-}	↑K^+, ↓Na^+	No entries	Mg^{2+}	No entries
Energy conversion						
Phosphomannomutase	5.4.2.8	-	-	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Catalytic subunit of the membrane ATP synthase	3.6.3.14	PO_4^{3-}	-	-	Mg^{2+}	Mg^{2+}
Proteolipid subunits of the membrane ATP synthase	3.6.3.14	-	-	-	-	-
Coenzymes						
Glycine hydroxymethyltransferase	2.1.2.1	-	↓MC^+	-	↓Mg^{2+}, Mn^{2+}, Ca^{2+}	-
Secretion						
Preprotein translocase subunit SecY	-	-	-	-	-	Zn^{2+}
Signal recognition particle GTPase FtsY	3.6.5.4	PO_4^{3-}	-	K^+	Mg^{2+}	Mg^{2+}
Miscellaneous						
Predicted GTPase (YchF, PF06071, 1JAL, 2OHF, 2DBY, 2DWQ, 1NI3)	-	PO_4^{3-}	K^+	-	Mg^{2+}	-
DNA primase (dnaG)	2.7.7.-	PP_i	-	-	Zn^{2+}	-
S-adenosylmethionine dimethyltransferase (KsgA)	2.1.1.48	-	-	-	Mg^{2+}	-

Transition metals and ubiquitous cellular systems

Many ubiquitous proteins depend on such divalent cations as Mg^{2+}, Zn^{2+}, and Mn^{2+}.

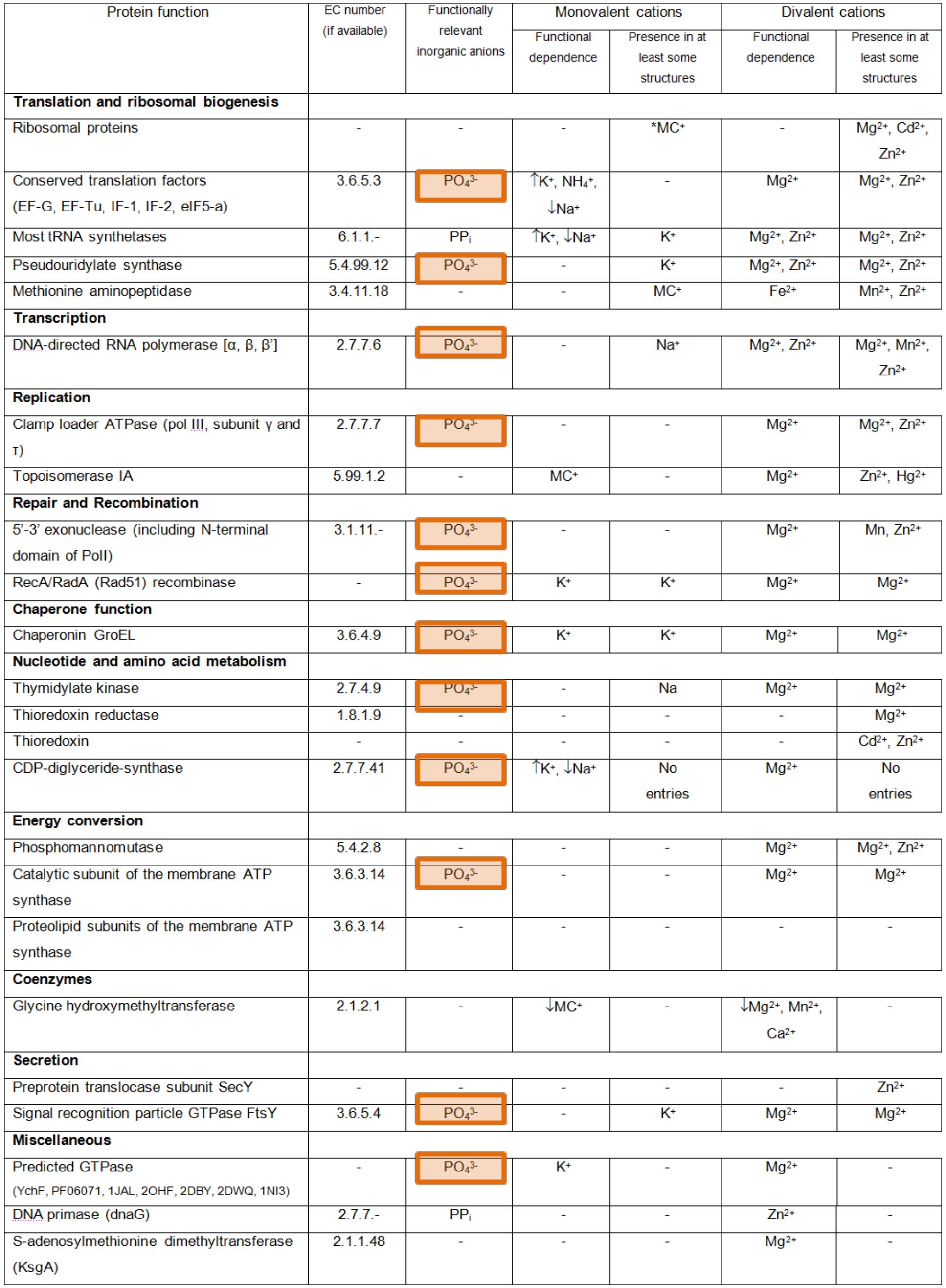


Protein function	EC number (if available)	Functionally relevant inorganic anions	Monovalent cations		Divalent cations	
			Functional dependence	Presence in at least some structures	Functional dependence	Presence in at least some structures
Translation and ribosomal biogenesis						
Ribosomal proteins	-	-	-	*MC^+	-	Mg^{2+}, Cd^{2+}, Zn^{2+}
Conserved translation factors (EF-G, EF-Tu, IF-1, IF-2, eIF5-a)	3.6.5.3	PO_4^{3-}	↑K^+, NH_4^+, ↓Na^+	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Most tRNA synthetases	6.1.1.-	PP_i	↑K^+, ↓Na^+	K^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Zn^{2+}
Pseudouridylate synthase	5.4.99.12	PO_4^{3-}	-	K^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Zn^{2+}
Methionine aminopeptidase	3.4.11.18	-	-	MC^+	Fe^{2+}	Mn^{2+}, Zn^{2+}
Transcription						
DNA-directed RNA polymerase [α, β, β']	2.7.7.6	PO_4^{3-}	-	Na^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Mn^{2+}, Zn^{2+}
Replication						
Clamp loader ATPase (pol III, subunit γ and τ)	2.7.7.7	PO_4^{3-}	-	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Topoisomerase IA	5.99.1.2	-	MC^+	-	Mg^{2+}	Zn^{2+}, Hg^{2+}
Repair and Recombination						
5'-3' exonuclease (including N-terminal domain of PolI)	3.1.11.-	PO_4^{3-}	-	-	Mg^{2+}	Mn, Zn^{2+}
RecA/RadA (Rad51) recombinase	-	PO_4^{3-}	K^+	K^+	Mg^{2+}	Mg^{2+}
Chaperone function						
Chaperonin GroEL	3.6.4.9	PO_4^{3-}	K^+	K^+	Mg^{2+}	Mg^{2+}
Nucleotide and amino acid metabolism						
Thymidylate kinase	2.7.4.9	PO_4^{3-}	-	Na	Mg^{2+}	Mg^{2+}
Thioredoxin reductase	1.8.1.9	-	-	-	-	Mg^{2+}
Thioredoxin	-	-	-	-	-	Cd^{2+}, Zn^{2+}
CDP-diglyceride-synthase	2.7.7.41	PO_4^{3-}	↑K^+, ↓Na^+	No entries	Mg^{2+}	No entries
Energy conversion						
Phosphomannomutase	5.4.2.8	-	-	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Catalytic subunit of the membrane ATP synthase	3.6.3.14	PO_4^{3-}	-	-	Mg^{2+}	Mg^{2+}
Proteolipid subunits of the membrane ATP synthase	3.6.3.14	-	-	-	-	-
Coenzymes						
Glycine hydroxymethyltransferase	2.1.2.1	-	↓MC^+	-	↓Mg^{2+}, Mn^{2+}, Ca^{2+}	-
Secretion						
Preprotein translocase subunit SecY	-	-	-	-	-	Zn^{2+}
Signal recognition particle GTPase FtsY	3.6.5.4	PO_4^{3-}	-	K^+	Mg^{2+}	Mg^{2+}
Miscellaneous						
Predicted GTPase (YchF, PF06071, 1JAL, 2OHF, 2DBY, 2DWQ, 1NI3)	-	PO_4^{3-}	K^+	-	Mg^{2+}	-
DNA primase (dnaG)	2.7.7.-	PP_i	-	-	Zn^{2+}	-
S-adenosylmethionine dimethyltransferase (KsgA)	2.1.1.48	-	-	-	Mg^{2+}	-

Phosphorous compounds in ubiquitous proteins

Set of ubiquitous proteins is enriched with enzymes that process phosphate groups of ATP, GTP and PPi.

The energetics of the first cells could be based on the transfer of phosphate groups.



Protein function	EC number (if available)	Functionally relevant inorganic anions	Monovalent cations		Divalent cations	
			Functional dependence	Presence in at least some structures	Functional dependence	Presence in at least some structures
Translation and ribosomal biogenesis						
Ribosomal proteins	-	-	-	$^*MC^+$	-	Mg^{2+}, Cd^{2+}, Zn^{2+}
Conserved translation factors (EF-G, EF-Tu, IF-1, IF-2, eIF5-a)	3.6.5.3	PO_4^{3-}	↑K^+, NH_4^+, ↓Na^+	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Most tRNA synthetases	6.1.1.-	PP_i	↑K^+, ↓Na^+	K^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Zn^{2+}
Pseudouridylate synthase	5.4.99.12	PO_4^{3-}	-	K^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Zn^{2+}
Methionine aminopeptidase	3.4.11.18	-	-	MC^+	Fe^{2+}	Mn^{2+}, Zn^{2+}
Transcription						
DNA-directed RNA polymerase [α, β, β’]	2.7.7.6	PO_4^{3-}	-	Na^+	Mg^{2+}, Zn^{2+}	Mg^{2+}, Mn^{2+}, Zn^{2+}
Replication						
Clamp loader ATPase (pol III, subunit γ and τ)	2.7.7.7	PO_4^{3-}	-	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Topoisomerase IA	5.99.1.2	-	MC^+	-	Mg^{2+}	Zn^{2+}, Hg^{2+}
Repair and Recombination						
5’-3’ exonuclease (including N-terminal domain of PolI)	3.1.11.-	PO_4^{3-}	-	-	Mg^{2+}	Mn, Zn^{2+}
RecA/RadA (Rad51) recombinase	-	PO_4^{3-}	K^+	K^+	Mg^{2+}	Mg^{2+}
Chaperone function						
Chaperonin GroEL	3.6.4.9	PO_4^{3-}	K^+	K^+	Mg^{2+}	Mg^{2+}
Nucleotide and amino acid metabolism						
Thymidylate kinase	2.7.4.9	PO_4^{3-}	-	Na	Mg^{2+}	Mg^{2+}
Thioredoxin reductase	1.8.1.9	-	-	-	-	Mg^{2+}
Thioredoxin	-	-	-	-	-	Cd^{2+}, Zn^{2+}
CDP-diglyceride-synthase	2.7.7.41	PO_4^{3-}	↑K^+, ↓Na^+	No entries	Mg^{2+}	No entries
Energy conversion						
Phosphomannomutase	5.4.2.8	-	-	-	Mg^{2+}	Mg^{2+}, Zn^{2+}
Catalytic subunit of the membrane ATP synthase	3.6.3.14	PO_4^{3-}	-	-	Mg^{2+}	Mg^{2+}
Proteolipid subunits of the membrane ATP synthase	3.6.3.14	-	-	-	-	-
Coenzymes						
Glycine hydroxymethyltransferase	2.1.2.1	-	↓MC^+	-	↓Mg^{2+}, Mn^{2+}, Ca^{2+}	-
Secretion						
Preprotein translocase subunit SecY	-	-	-	-	-	Zn^{2+}
Signal recognition particle GTPase FtsY	3.6.5.4	PO_4^{3-}	-	K^+	Mg^{2+}	Mg^{2+}
Miscellaneous						
Predicted GTPase (YchF, PF06071, 1JAL, 2OHF, 2DBY, 2DWQ, 1NI3)	-	PO_4^{3-}	K^+	-	Mg^{2+}	-
DNA primase (dnaG)	2.7.7.-	PP_i	-	-	Zn^{2+}	-
S-adenosylmethionine dimethyltransferase (KsgA)	2.1.1.48	-	-	-	Mg^{2+}	-

From Mulkidjanian et al, 2012, PNAS

The ubiquitous, and, by inference, primordial proteins and functional systems show affinity to and functional requirement for Mg^{2+}, K^+, Zn^{2+}, Mn^{2+} and phosphate.

Thus, protocells must have evolved in habitats with a high K^+/Na^+ ratio and relatively high concentrations of transition metals and phosphorous compounds.

Macallum: Since cells contain more K^+ than Na^+, the first cells may have emerged in K^+-rich habitats.

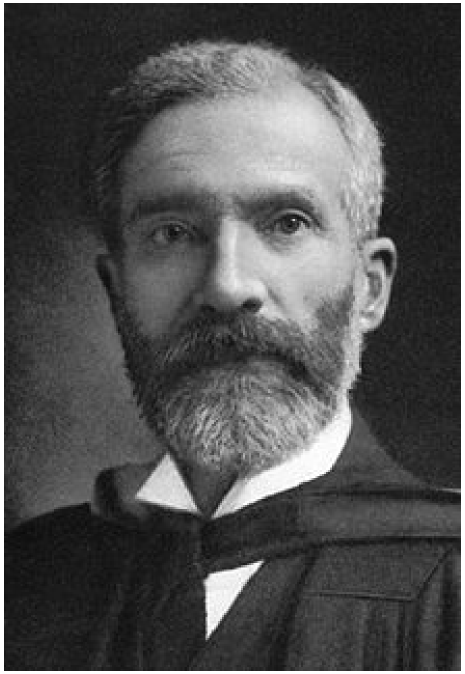


Photo: Wikipedia

The sea water always contained more Na^+ than K^+, which seems to rule out any marine environment as a cradle of the first cells.

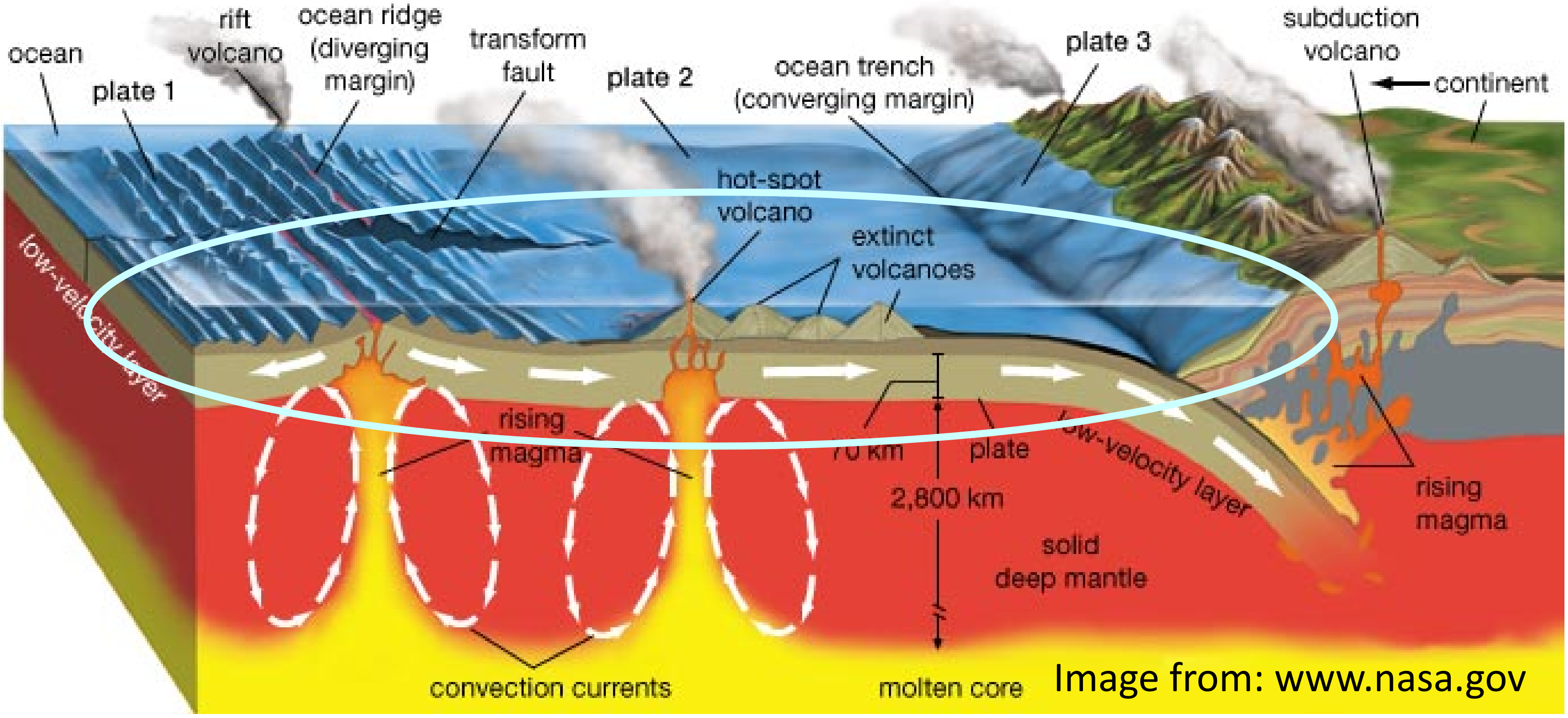
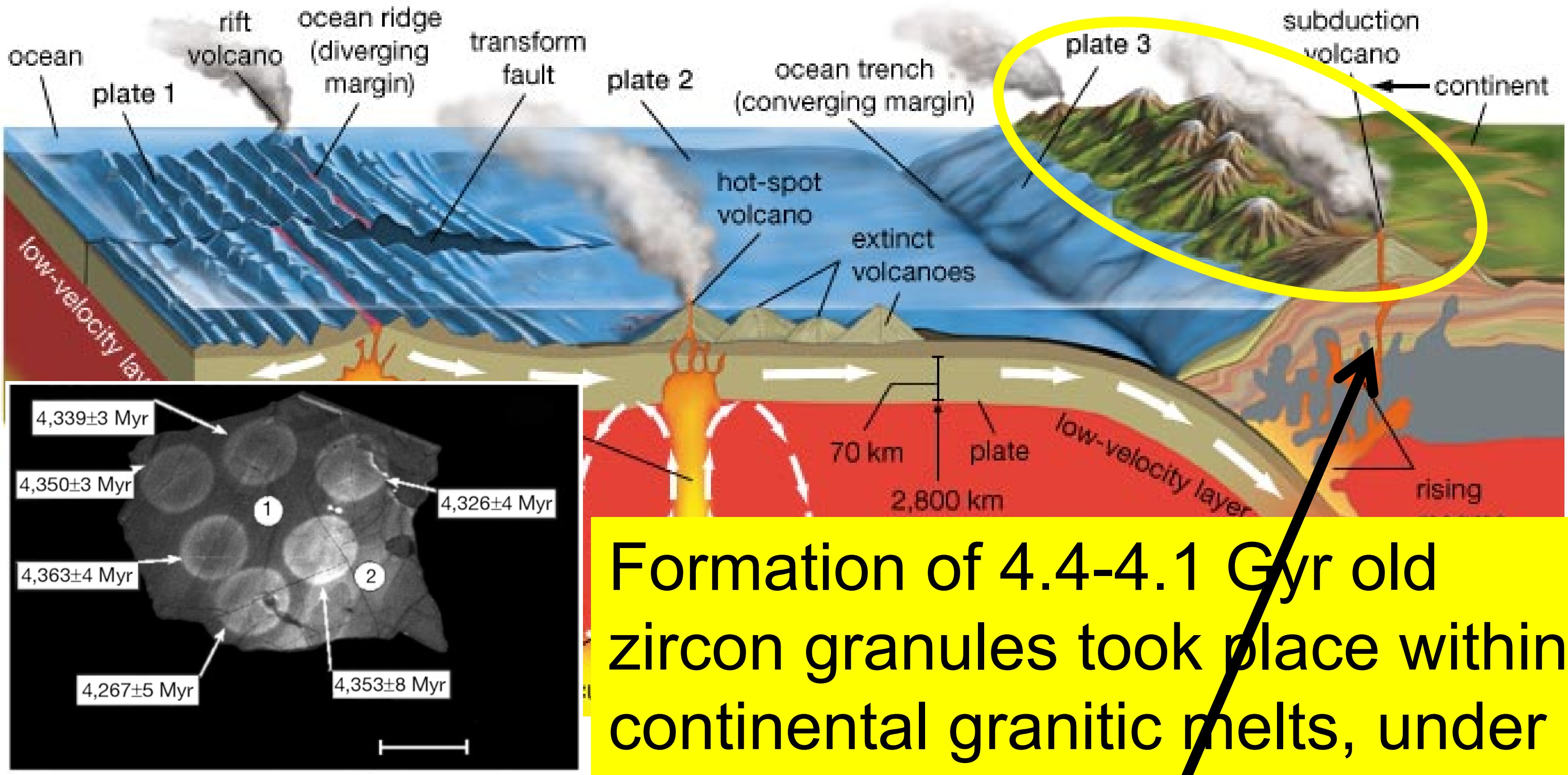


Image from: www.nasa.gov

What about continental geothermal systems?

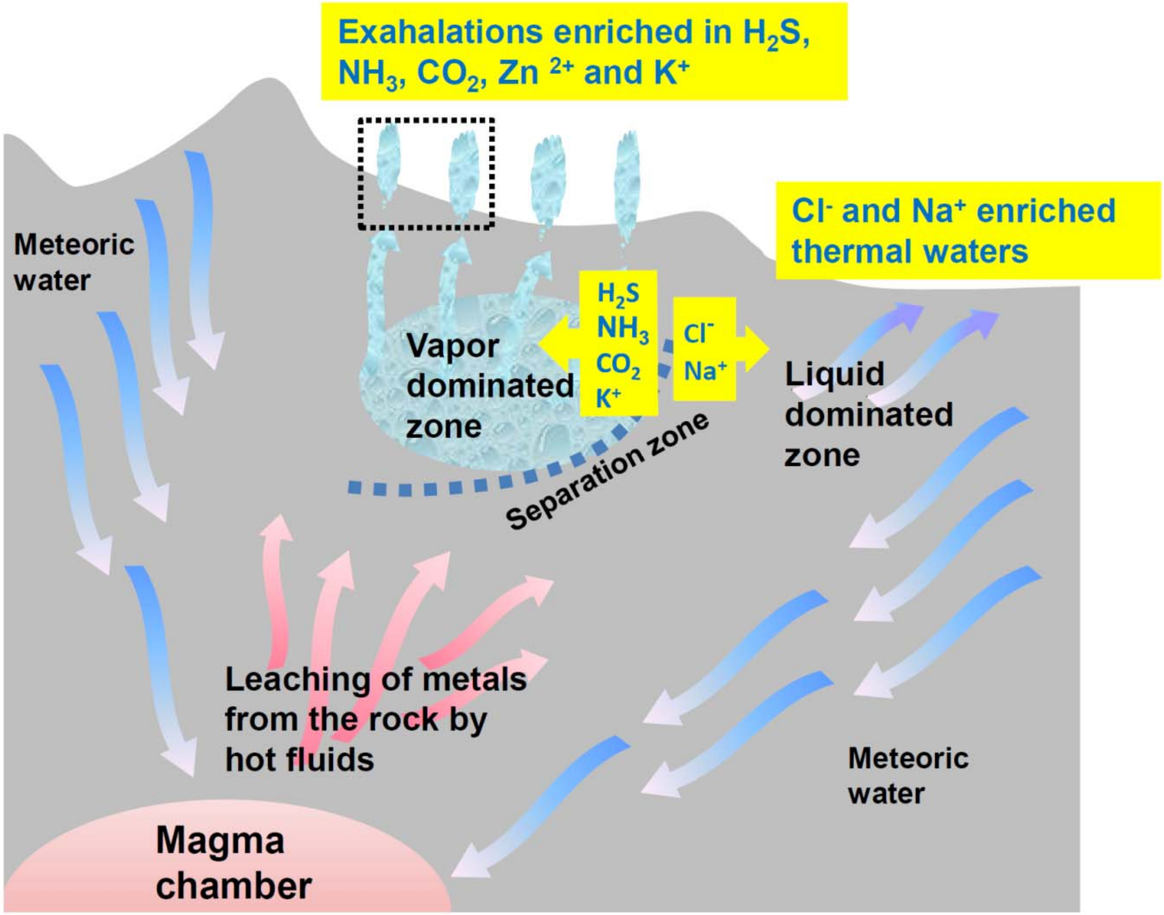


Formation of 4.4-4.1 Gyr old zircon granules took place within continental granitic melts, under conditions resembling the modern subduction zones

From: Wilde et al. Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago (2001) *Nature*

Terrestrial geothermal systems are intrinsically heterogeneous because, upon the boiling of the ascending hot fluid, the vapor separates from the liquid.

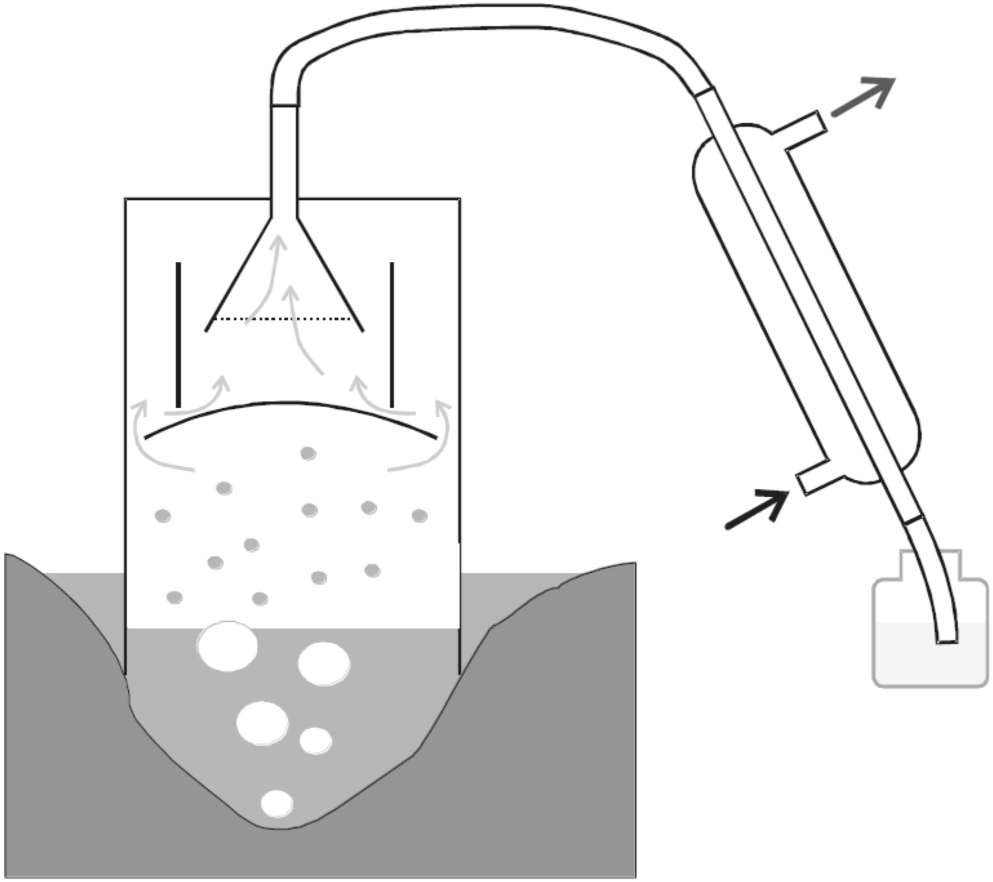
The vapor phase becomes enriched in volatile compounds, such as CO_2, H_2S, and NH_3.



From: Mulkidjanian et al., 2012, *PNAS*

The areas where geothermal vapor discharges via boiling silica pots and fumaroles are called geothermal fields.

Mutnovsky geothermal field, Kamchatka (photos by Dr. Anna S. Karyagina)



Liquids of thermal springs and the vapor condensate of the same springs could be sampled separately

Image from: Mulkidjanian et al., 2012, *PNAS*

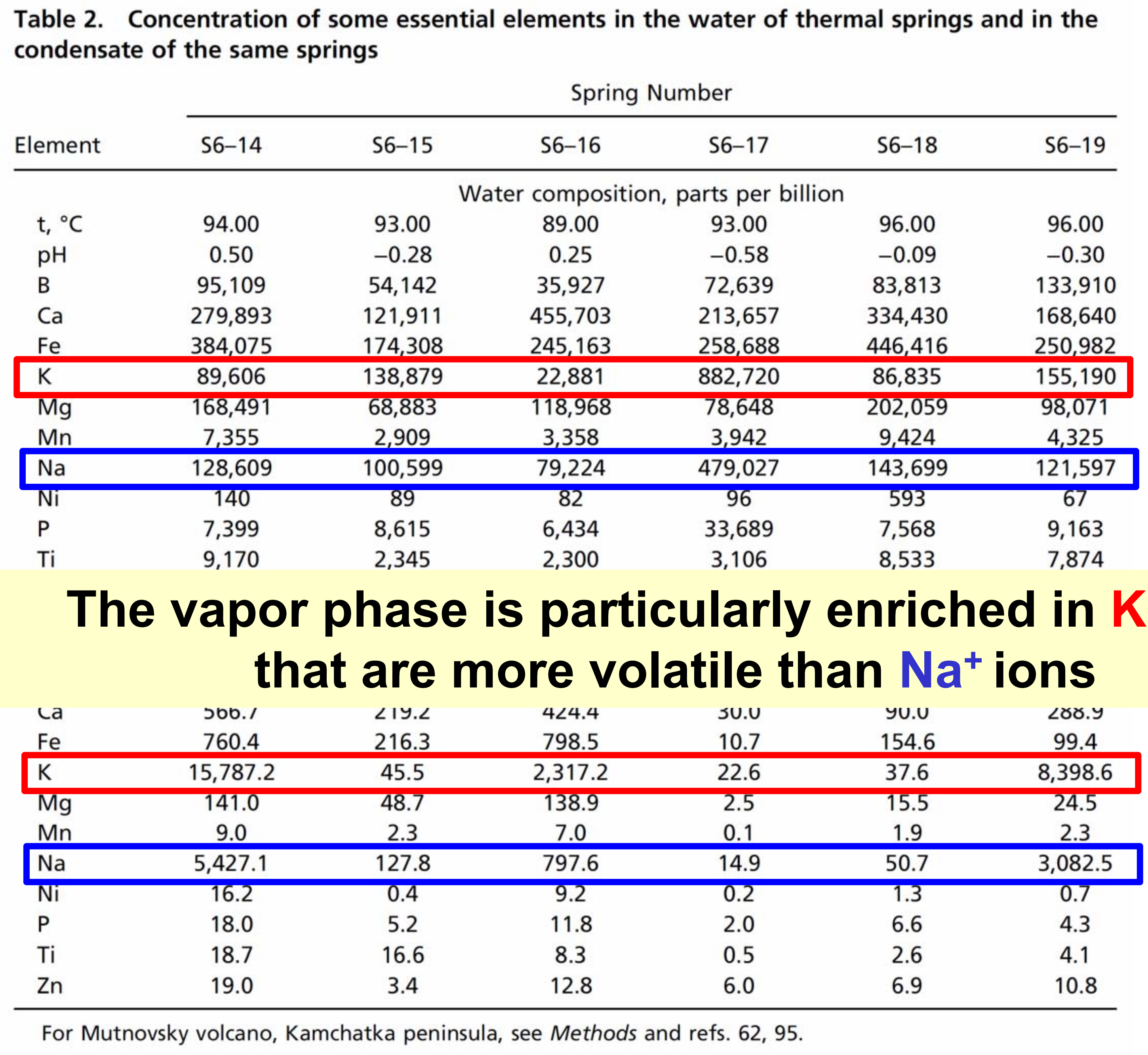


Table 2. Concentration of some essential elements in the water of thermal springs and in the condensate of the same springs

Element	S6–14	S6–15	S6–16	S6–17	S6–18	S6–19
			Water composition, parts per billion			
t, °C	94.00	93.00	89.00	93.00	96.00	96.00
pH	0.50	−0.28	0.25	−0.58	−0.09	−0.30
B	95,109	54,142	35,927	72,639	83,813	133,910
Ca	279,893	121,911	455,703	213,657	334,430	168,640
Fe	384,075	174,308	245,163	258,688	446,416	250,982
K	89,606	138,879	22,881	882,720	86,835	155,190
Mg	168,491	68,883	118,968	78,648	202,059	98,071
Mn	7,355	2,909	3,358	3,942	9,424	4,325
Na	128,609	100,599	79,224	479,027	143,699	121,597
Ni	140	89	82	96	593	67
P	7,399	8,615	6,434	33,689	7,568	9,163
Ti	9,170	2,345	2,300	3,106	8,533	7,874
Ca	566.7	219.2	424.4	30.0	90.0	288.9
Fe	760.4	216.3	798.5	10.7	154.6	99.4
K	15,787.2	45.5	2,317.2	22.6	37.6	8,398.6
Mg	141.0	48.7	138.9	2.5	15.5	24.5
Mn	9.0	2.3	7.0	0.1	1.9	2.3
Na	5,427.1	127.8	797.6	14.9	50.7	3,082.5
Ni	16.2	0.4	9.2	0.2	1.3	0.7
P	18.0	5.2	11.8	2.0	6.6	4.3
Ti	18.7	16.6	8.3	0.5	2.6	4.1
Zn	19.0	3.4	12.8	6.0	6.9	10.8

For Mutnovsky volcano, Kamchatka peninsula, see *Methods* and refs. 62, 95.

K^+/Na^+ ratio in geothermal fluids: up to 1.4

K^+/Na^+ ratio in the vapor: up to 2.9

The vapor phase is particularly enriched in K^+ ions that are more volatile than Na^+ ions

From: Mulkidjanian et al., 2012, *PNAS*

At modern geothermal fields the atmospheric oxygen promptly oxidizes the ascending H_2S to sulfuric acid

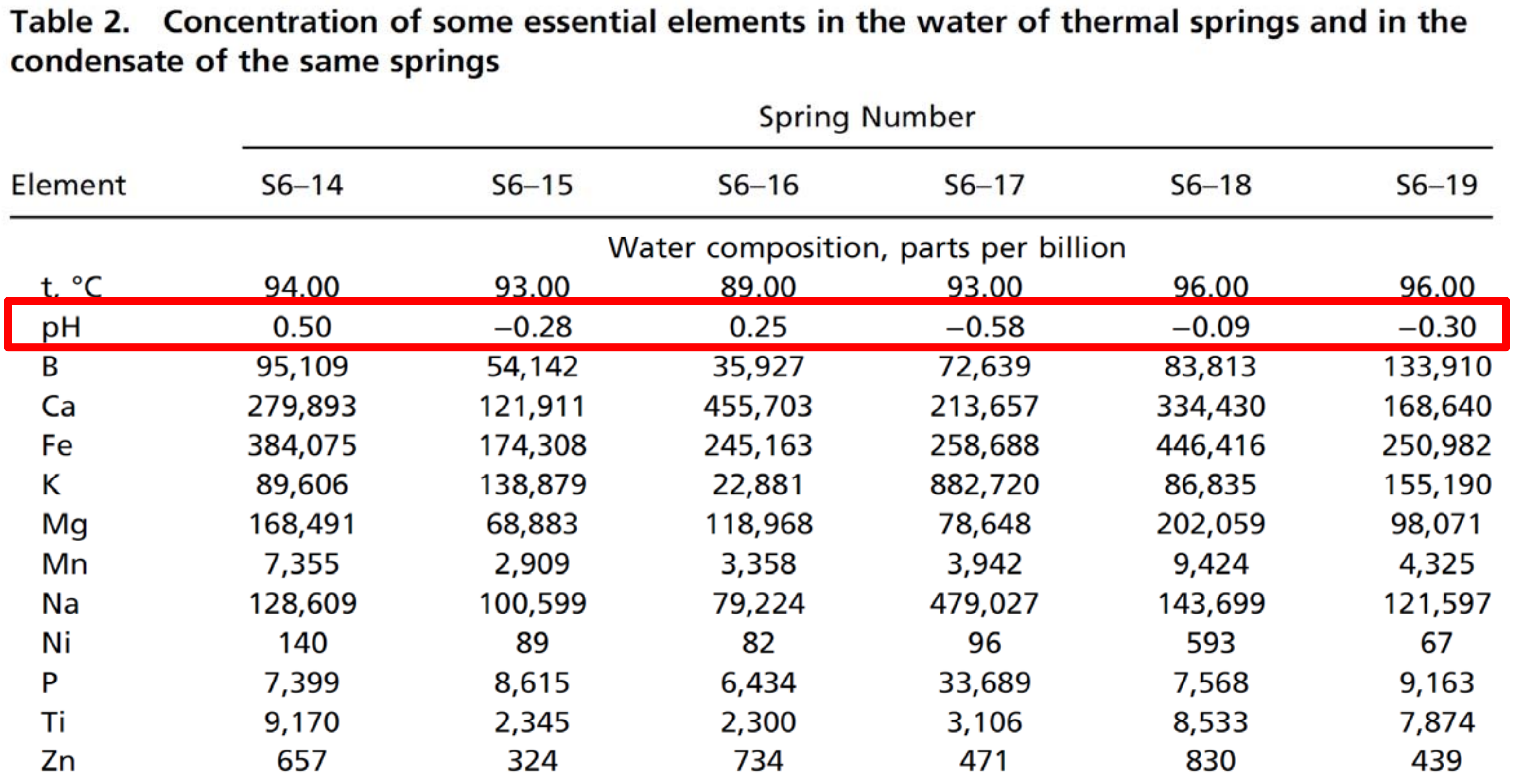


Table 2. Concentration of some essential elements in the water of thermal springs and in the condensate of the same springs

	Spring Number					
Element	S6–14	S6–15	S6–16	S6–17	S6–18	S6–19
	Water composition, parts per billion					
t, °C	94.00	93.00	89.00	93.00	96.00	96.00
pH	0.50	−0.28	0.25	−0.58	−0.09	−0.30
B	95,109	54,142	35,927	72,639	83,813	133,910
Ca	279,893	121,911	455,703	213,657	334,430	168,640
Fe	384,075	174,308	245,163	258,688	446,416	250,982
K	89,606	138,879	22,881	882,720	86,835	155,190
Mg	168,491	68,883	118,968	78,648	202,059	98,071
Mn	7,355	2,909	3,358	3,942	9,424	4,325
Na	128,609	100,599	79,224	479,027	143,699	121,597
Ni	140	89	82	96	593	67
P	7,399	8,615	6,434	33,689	7,568	9,163
Ti	9,170	2,345	2,300	3,106	8,533	7,874
Zn	657	324	734	471	830	439

From: Mulkidjanian et al., 2012, *PNAS*

On the primeval Earth, in the absence of oxygen:

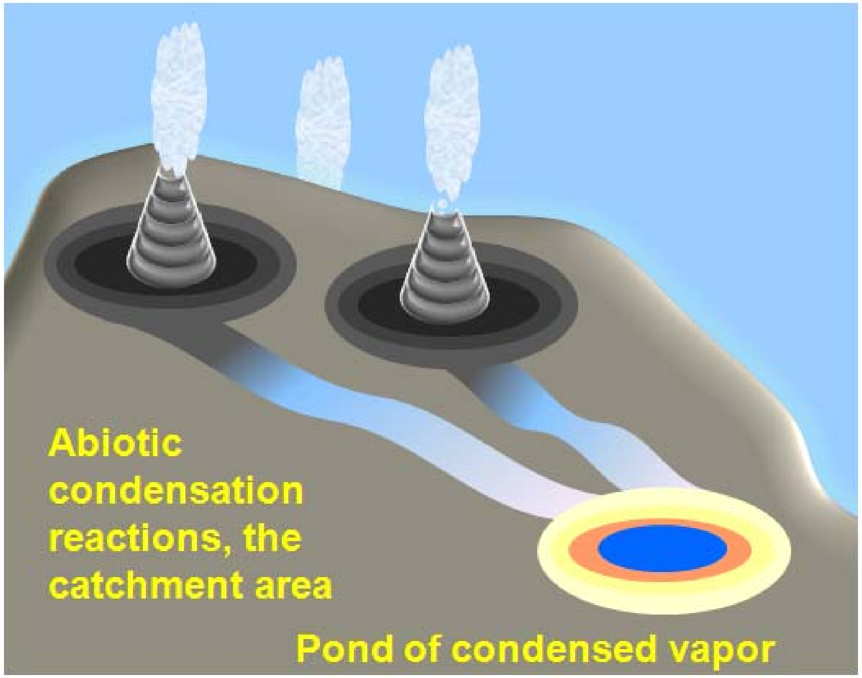
The pH of the discharges
from the vapor-dominated zones
should have been closer to neutral;

At neutral pH, silica would precipitate not as mud,
but as porous, ordered silicate minerals
- clays and zeolites;

Because of high sulfide levels,
formation of metal–sulfide–contaminated clays
and zeolites would be expected.

Two types of environments at anoxic geothermal fields

Periodically wetted and illuminated mineral surfaces could serve as templates and catalysts for diverse abiotic syntheses.



Remote pools and puddles of cool, condensed geothermal vapor, enriched in slowly precipitating ZnS and MnS, may have served as tentative hatcheries of the first cells.

From: Mulkidjanian et al., 2012, *PNAS*

Two types of environments at anoxic geothermal fields

Periodically wetted and illuminated mineral surfaces could serve as templates and catalysts for diverse abiotic syntheses.

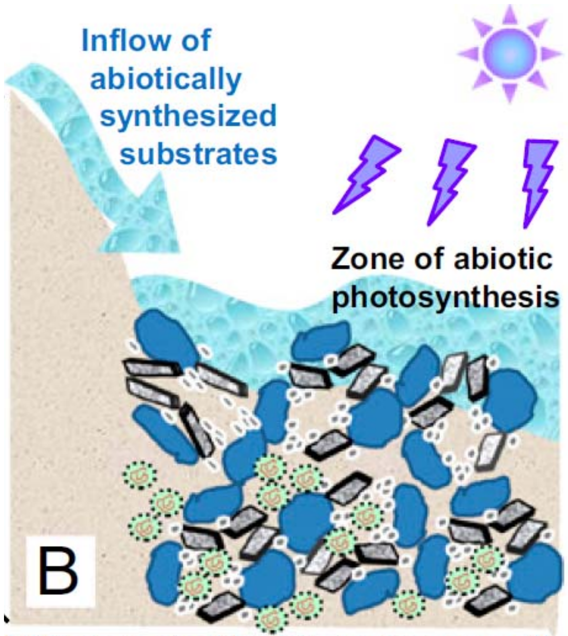
Remote pools and puddles of cool, condensed geothermal vapor, enriched in slowly precipitating ZnS and MnS, may have served as tentative hatcheries of the first cells.

Image: Dominique Decobecq

A stratified system of a primordial pond:

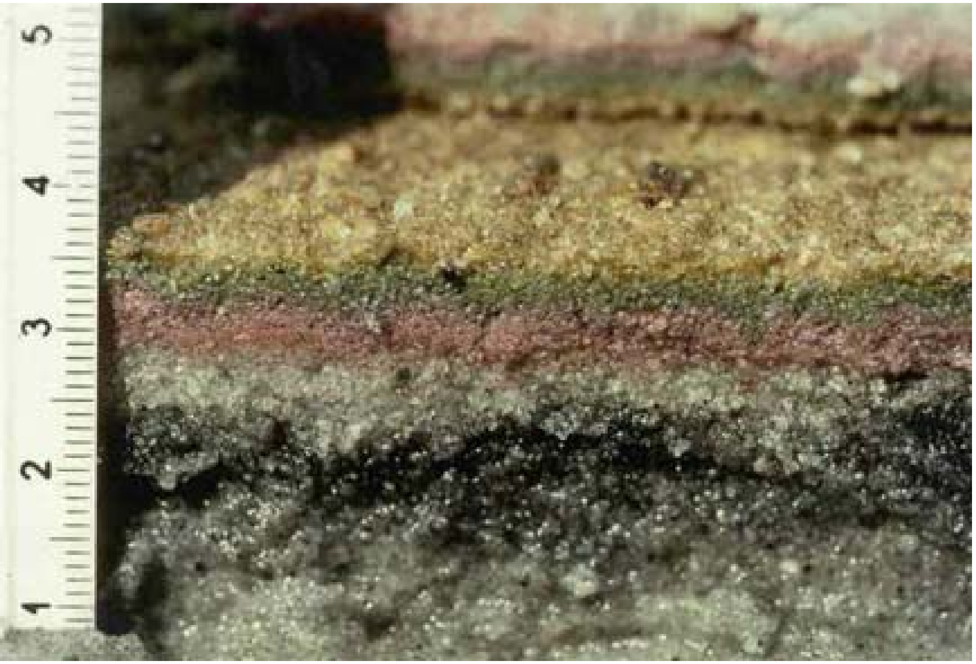
the illuminated upper layers could protect the protocells from the UV damage, “harvest” organic compounds and also produce them in photocatalytic reactions;

the deeper, less productive, but better protected layers could provide shelter for the protocells.



UV-protected habitat of protocells

5 μm layer of **ZnS** attenuates the UV light by a factor of 10^{10}
(that corresponds to > 100 m of water)



Both the light gradient and the interlayer metabolite exchange are typical of modern stratified phototrophic microbial communities.

Source: http://textbookofbacteriology.net/themicrobialworld/microbial_mat.xsecs.jpg

From the chemical viewpoint,
spontaneous formation of nucleobases, ribose, and (poly)nucleotides should have required evaporative basins rich in
formamide and other simple amides,
carboneous compounds,
phosphorous compounds,
borate,
oxidized molybdenum,
and UV light as a photocatalist
and a selective factor
(Schoffstall, Benner, Szostak, Di Mauro, Saladino, Powner, Sutherland, Hud and their co-workers).

Vapor condensate: up to 1 mM of phosphorus, up to 10 mM of borate, as well as ammonia, cyanide, hydrocarbons, molybdenum compoinds.

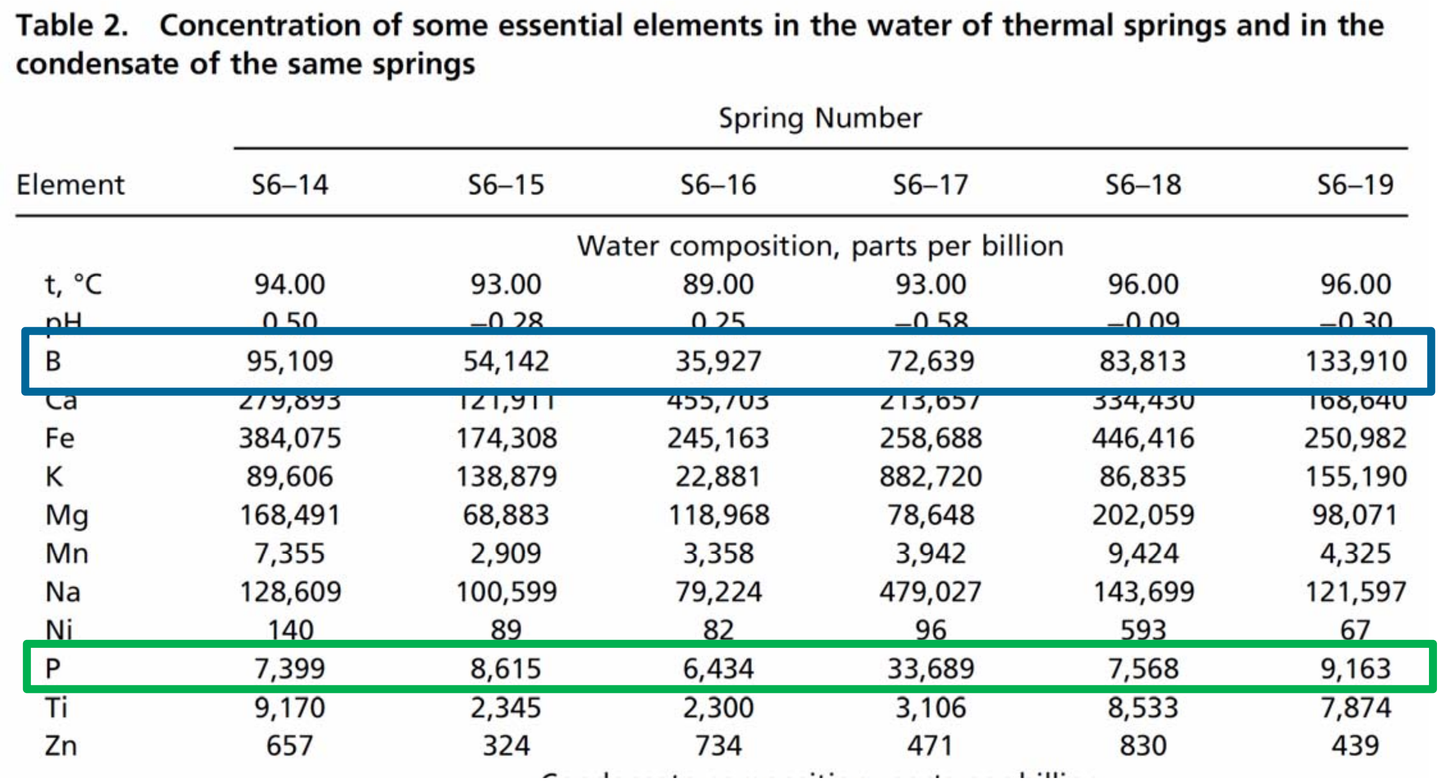


Table 2. Concentration of some essential elements in the water of thermal springs and in the condensate of the same springs

Element	S6–14	S6–15	S6–16	S6–17	S6–18	S6–19
			Water composition, parts per billion			
t, °C	94.00	93.00	89.00	93.00	96.00	96.00
pH	0.50	−0.28	0.25	−0.58	−0.09	−0.30
B	95,109	54,142	35,927	72,639	83,813	133,910
Ca	279,893	121,911	455,703	213,657	334,430	168,640
Fe	384,075	174,308	245,163	258,688	446,416	250,982
K	89,606	138,879	22,881	882,720	86,835	155,190
Mg	168,491	68,883	118,968	78,648	202,059	98,071
Mn	7,355	2,909	3,358	3,942	9,424	4,325
Na	128,609	100,599	79,224	479,027	143,699	121,597
Ni	140	89	82	96	593	67
P	7,399	8,615	6,434	33,689	7,568	9,163
Ti	9,170	2,345	2,300	3,106	8,533	7,874
Zn	657	324	734	471	830	439

From: Mulkidjanian et al., 2012, *PNAS*

Anoxic geothermal fields, identified as tentative cradles of the first cells by using the top-down approach and phylogenomic analysis,

could provide exactly those geochemical conditions that were suggested as most conducive for the emergence of life by the chemists who pursued the complementary bottom-up strategy.

Geothermal fields *versus* deep sea vents:

Anoxic geothermal fields

share all of the advantages of the deep sea hydrothermal vents, including the presence of inorganic compartments, versatile catalysts, and sources of organic matter.

In contrast to deep-sea vents,
the geothermal fields:

- are conducive to synthetic reactions and

- enable the involvement of solar light as an energy source and a selective factor.

Anoxic geothermal fields could habour not only the first cells, but also the earlier life forms by providung:

1) Basins that chemically resembled the cell cytoplasm,

2) Sources of organic and phosphorous compounds, as well as simple amides,

3) Porous silicates lined with powerful (photo)catalysts, such as ZnS and MnS,

4) Vast opportunities for condensation of reagents upon evaporation.

The cradle of the first organisms according to:

Oparin - „nutrient mixture“

Haldane: - „thin hot soup“

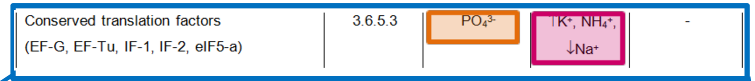
A soup-like, dense solution of organic molecules implies only a limited amount of water

Image credit: Bill Sanderson / Science Photo Library

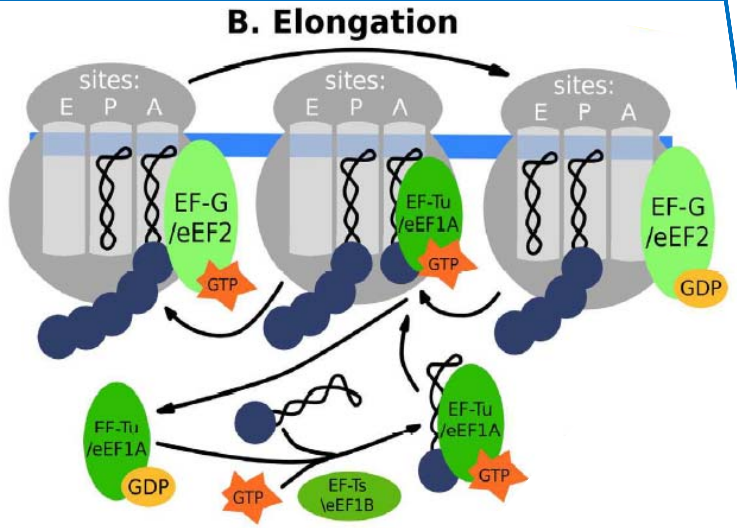
- Introduction: Ubiquitous protein set and the Last Universal Cellular Ancestor
- Emergence of the first cells in K^+-rich environments
- **Invasion of Na^+-rich habitats**
- On the co-evolution of membranes and enzyme machinery

K^+-dependence of translation

Ubiquitous translation factors EF-Tu and EF-G are GTPases that provide energy for protein synthesis.



Conserved translation factors (EF-G, EF-Tu, IF-1, IF-2, eIF5-a)	3.6.5.3	PO_4^{3-}	↑K^+, NH_4^+, ↓Na^+	-



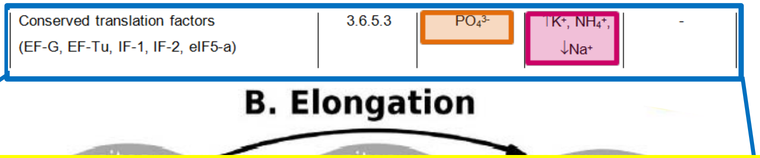
B. Elongation

Protein function	EC number (if available)	Functionally relevant inorganic anions	Monovalent cations	
			Functional dependence	Presence in at least some structures
Conserved translation factors (EF-G, EF-Tu, IF-1, IF-2, eIF5-a)	3.6.5.3	PO_4^{3-}	↑K^+, NH_4^+, ↓Na^+	-

Figure from (Zhouravleva *et al.*, “Gene Duplication”, **2011**)

K^+-dependence of translation

Ubiquitous translation factors EF-Tu and EF-G are GTPases that provide



Conserved translation factors (EF-G, EF-Tu, IF-1, IF-2, eIF5-a)	3.6.5.3	PO_4^{3-}	↑K^+, NH_4^+, ↓Na^+	-

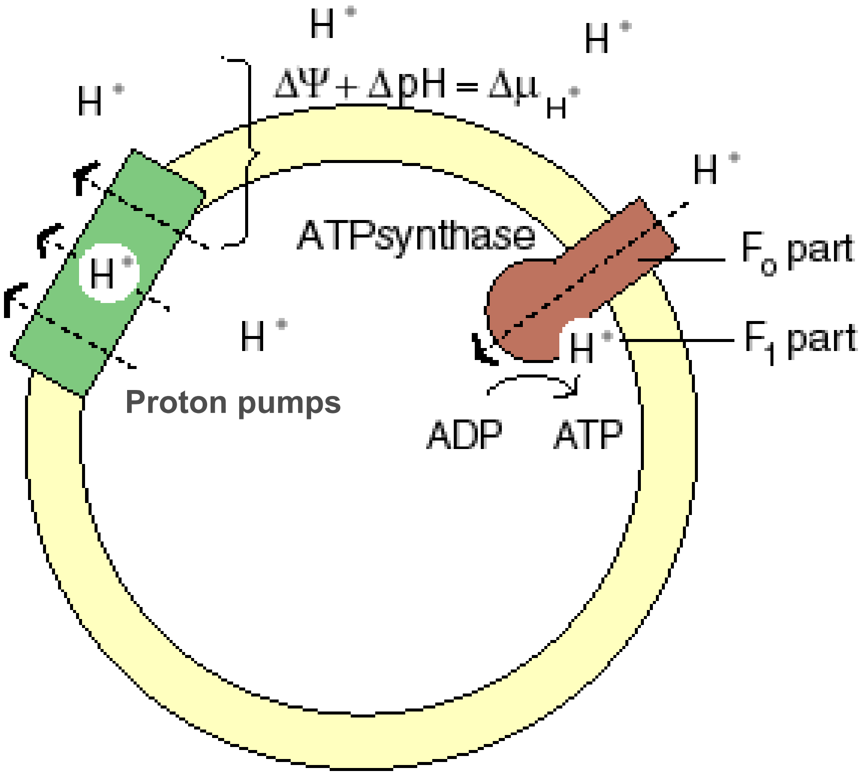
Protein synthesis by ribosomes demands K^+/Na^+ ratio > 1.

The K^+/Na^+ ratio is < 1 in all known environments, except geothermal fields.

How could the first cells survive in Na^+-rich habitats?

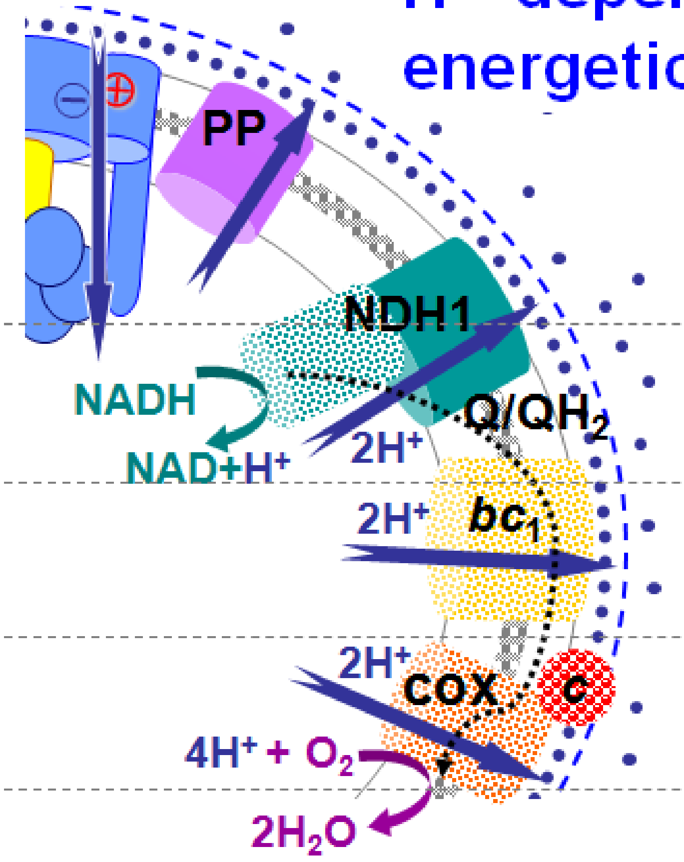
Figure from (Zhouravleva *et al.*, "Gene Duplication", **2011**)

Textbook bioenergetics: Membrane pumps generate H^+ gradient that can be used for synthesis of ATP



F/V-ATPases

H⁺- dependent energetics

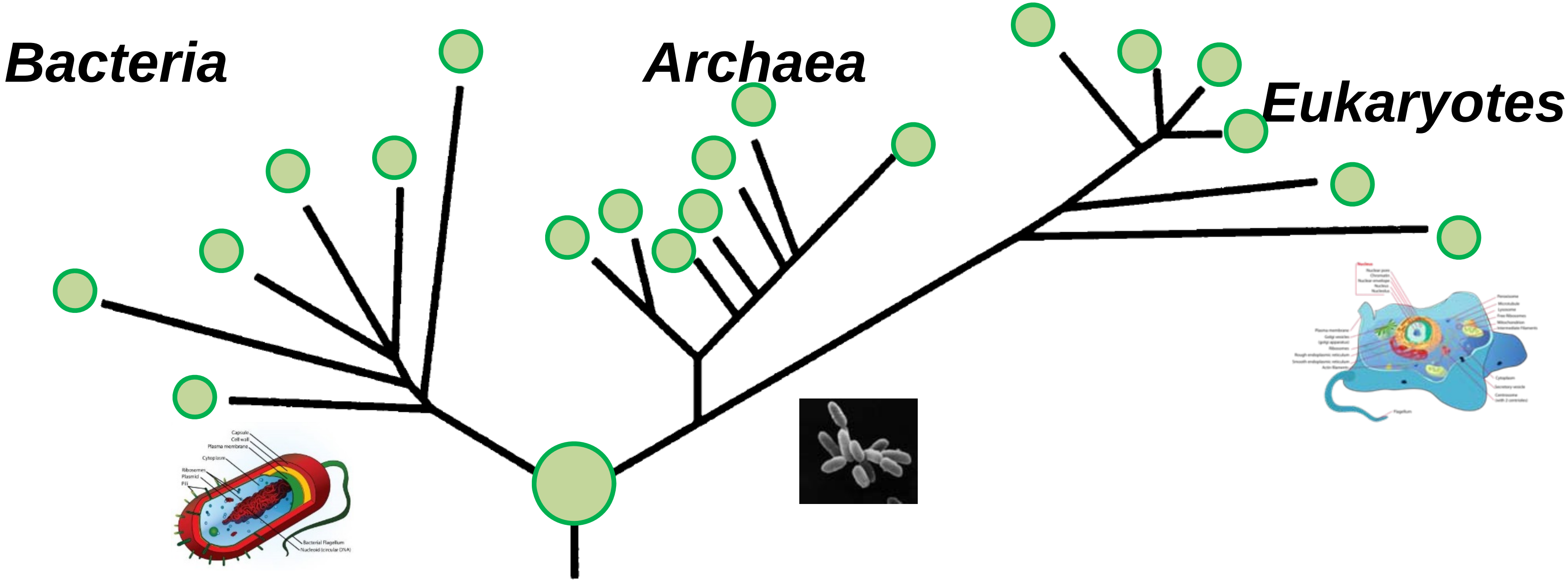


“The use of ion gradients over membranes for energy conservation, as in chemiosmotic coupling, is as universal as the genetic code itself“

Lane, Martin, 2012, *Cell*

Is it indeed so?

	Bacteria	Archaea	Eukaryotes
Energy conservation at the cellular membrane	**+**	**+**	**-**



Based on 16S rRNA similarity, all cellular life forms could be separated into three domains (Woese *et al.*, *PNAS*, 1977; Woese *et al.*, *PNAS*, 1990).

Figures of the life forms are taken from the Wikipedia

	Bacteria	Archaea	Eukaryotes
Energy conservation at the cellular membrane	+	+	-

The use of ion gradients over cell membranes for energy conservation is absent in eukaryotes and therefore

is not

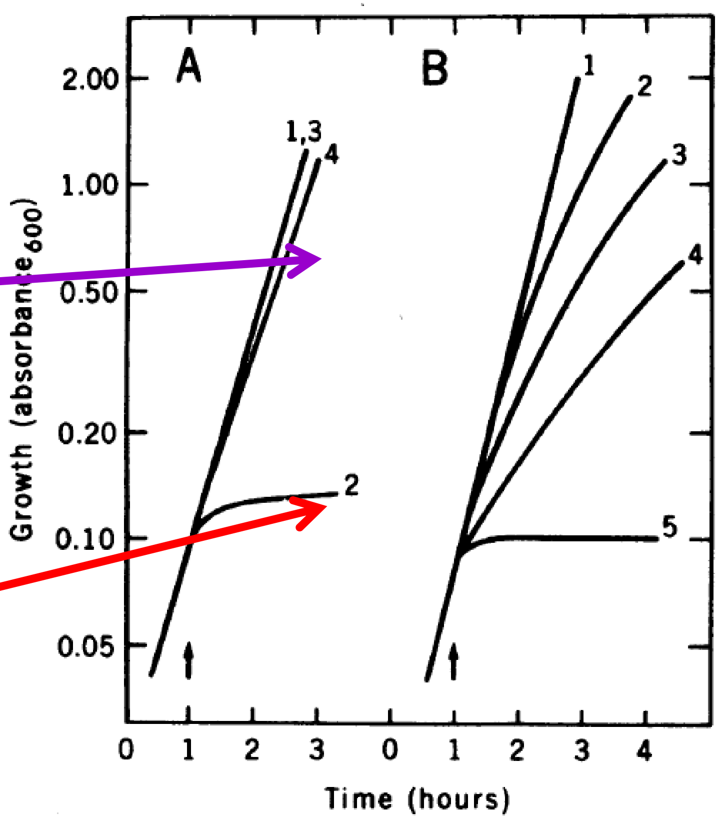
as universal as the genetic code itself.

If the membrane potential is not obligatory, is it possible then to switch it off?

Membrane potential can be switched off without affecting bacterial growth

When the membrane potential was collapsed by gramicidin and other ionophores, bacteria grew happily in a K^+-containing, rich medium (curve 4).

In a Na^+-containing, rich medium, collapsing of the potential stopped the growth (curve 2).



Circulation of H^+ and K^+ Across the Plasma Membrane Is Not Obligatory for Bacterial Growth

Abstract. Streptococcus faecalis *grows normally in the presence of gramicidin and other ionophores under conditions such that there is no gradient of* pH *or of electrical potential across the plasma membrane and that currents of* H^+, K^+, *and* Na^+ *are short-circuited. Growth requires a rich medium, a slightly alkaline* pH, *and a high concentration of external* K^+. *The proton circulation maintains the cytoplasmic* pH *and pools of ions and other metabolites but is not obligatory for biosynthetic functions including DNA replication, cell division, or assembly of the structural framework of the cell.*

FRANKLIN M. HAROLD
JENNIFER VAN BRUNT

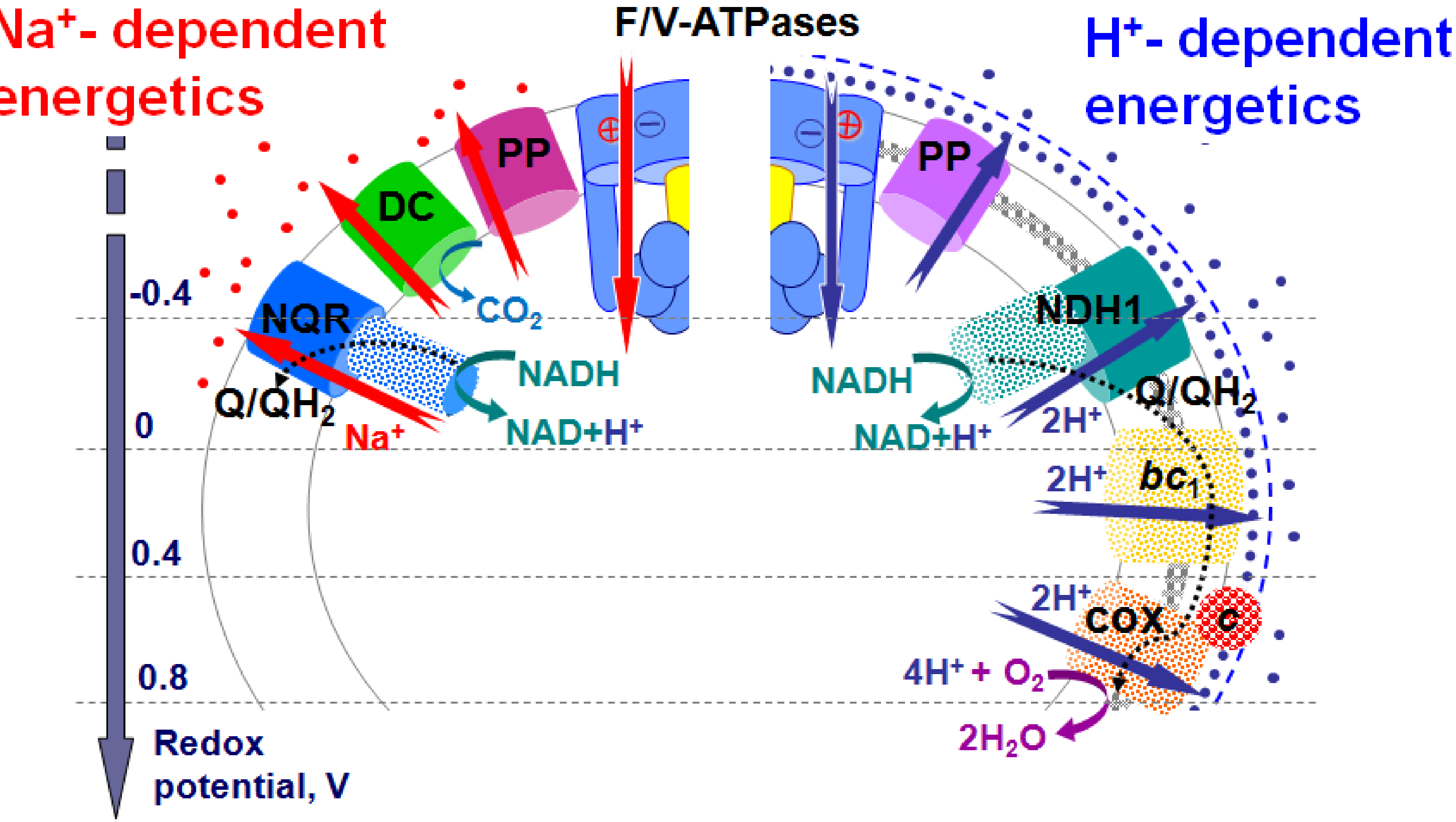
SCIENCE, VOL. 197, 1977

	Bacteria	Archaea	Eukaryotes
Energy conservation at the cellular membrane	+	+	-
$[K^+]/[Na^+]$ gradient at the cellular membrane	+	+	+

The primary function of membrane bioenergetics is to keep the cytoplasmic K^+/Na^+ ratio > 1 in natural environments, which are usually Na^+-rich.

What is the relation of the K^+/Na^+ disequilibrium to the membrane bioenergetics?

Some prokaryotes, among them many human pathogenes, use Na^+ ions instead of protons

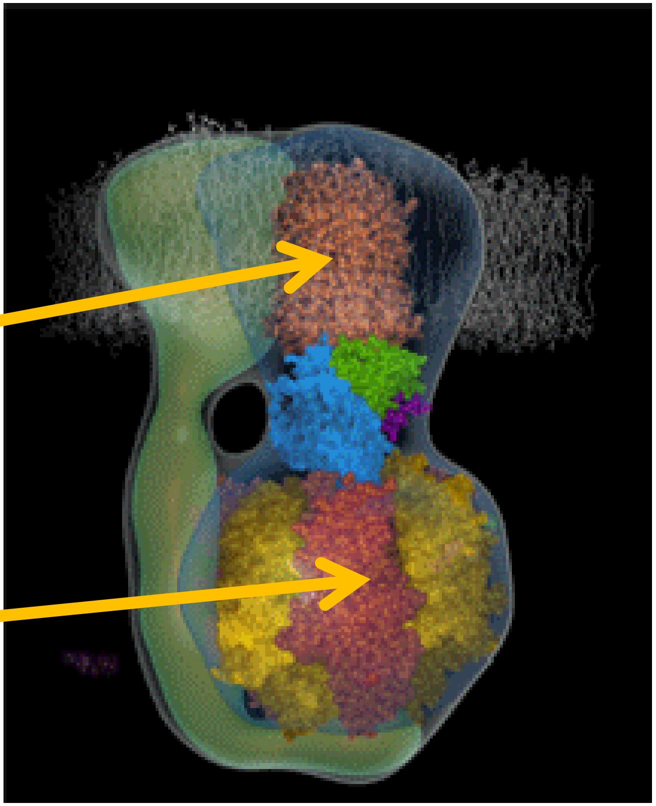


From: Mulkidjanian et al., 2009, *TiBS*

Components of the rotary ATPase belong to the ubiquitous proteins

The only energy-converting proteins that should have been present in the LUCA are:

- the proteolipid subunit and
- the catalytic subunits of the rotary ATPase



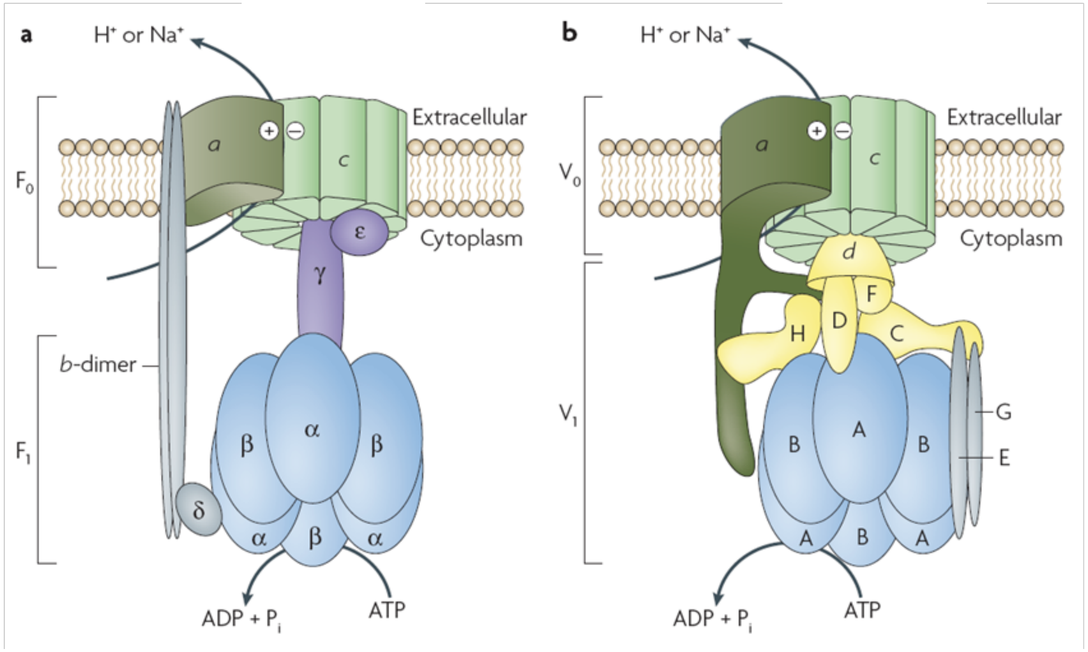
From the web-page of John Walker's lab

Based on (Koonin, *Annu Rev Genomics Hum Genet* **2000**; Charlebois *et al.*, *Genome Res*, **2004**)

Rotary ATPases are found in all organisms with distinct forms in Bacteria and Archaea

F-type

Bacteria, mitochondria, chloroplasts



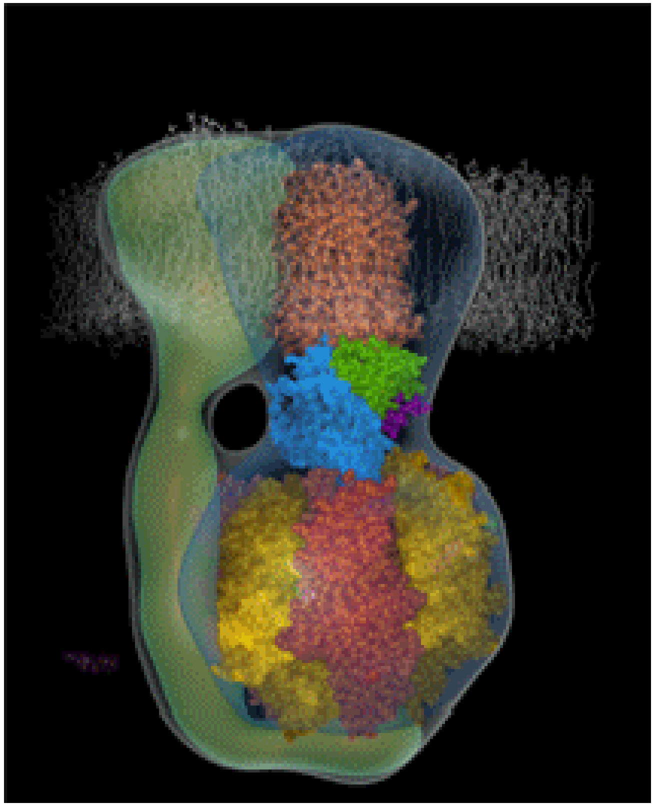
A/V-type

Archaea, some Bacteria, eukaryotic membranes

From Mulkidjanian et al., 2007, *Nature Rev. Microb.*

Only rotary ATP synthases can couple the proton-motive or sodium-motive force to the ATP synthesis.

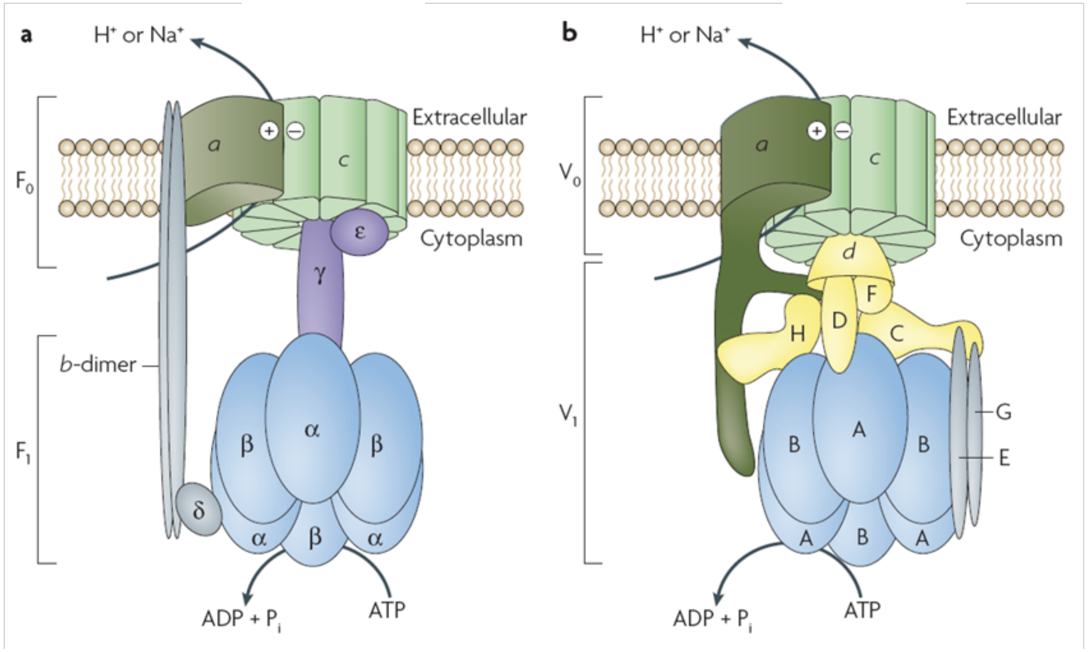
Therefore the coupling ion specificity of the ATP-synthase defines whether the organism relies on a Na^+- or H^+-type membrane energetics.



From the web-page of John Walker‘s lab

Both F-type and A/V-type rotary ATPases are found in Na^+-translocating and H^+-translocating versions

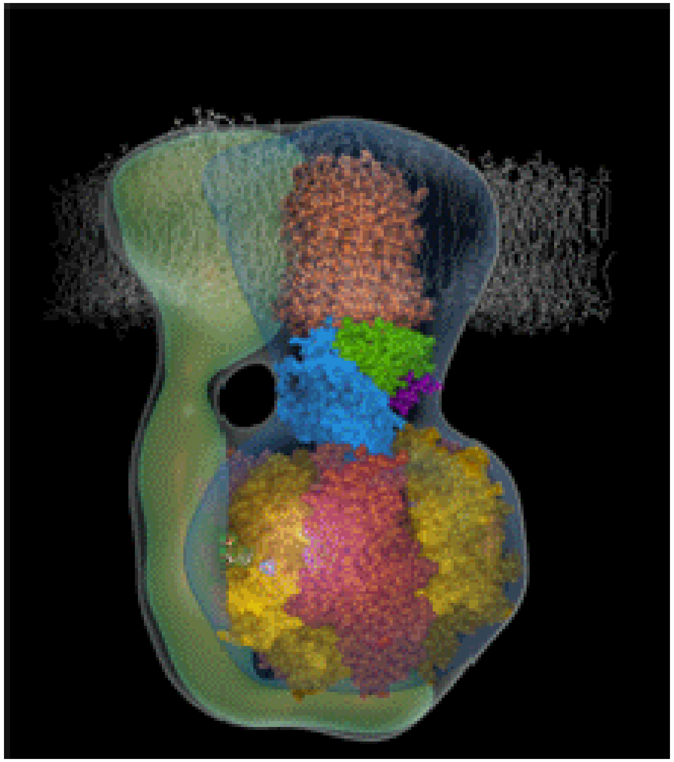
F-type



A/V-type

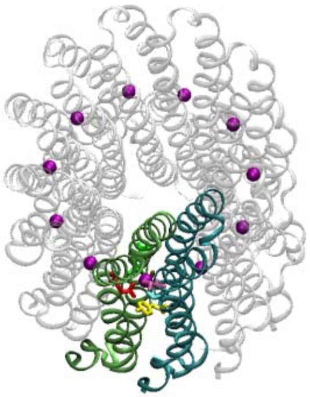
From Mulkidjanian et al., 2007, *Nature Rev. Microb.*

The Na$^+$-binding ligands are homologous in the F- and V-type ATP synthases



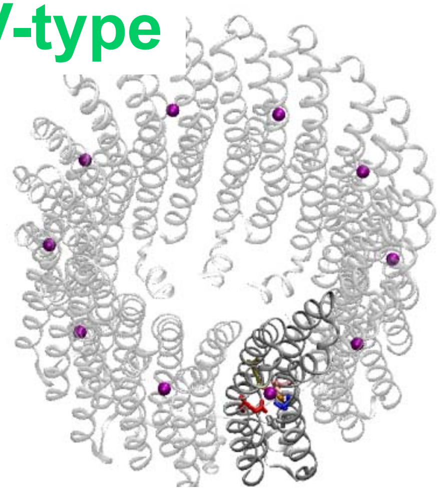
From the web-page of John Walker‘s lab

F-type

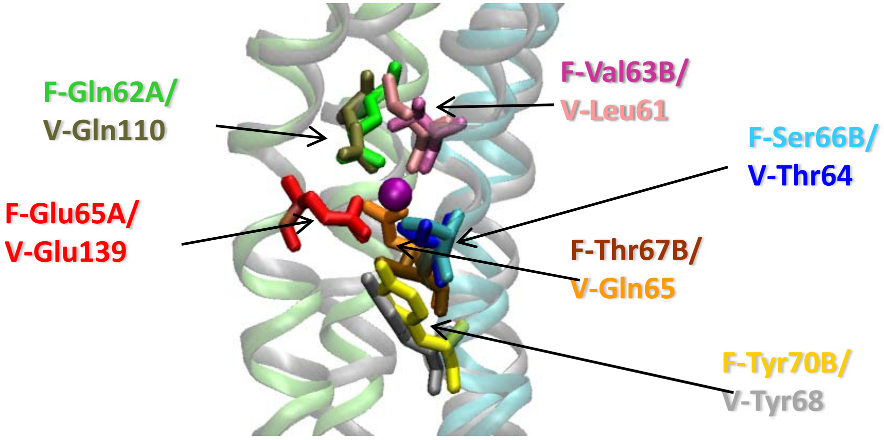


Ilyobacter tartaricus,
Meier et al., 2005, *Science*

A/V-type

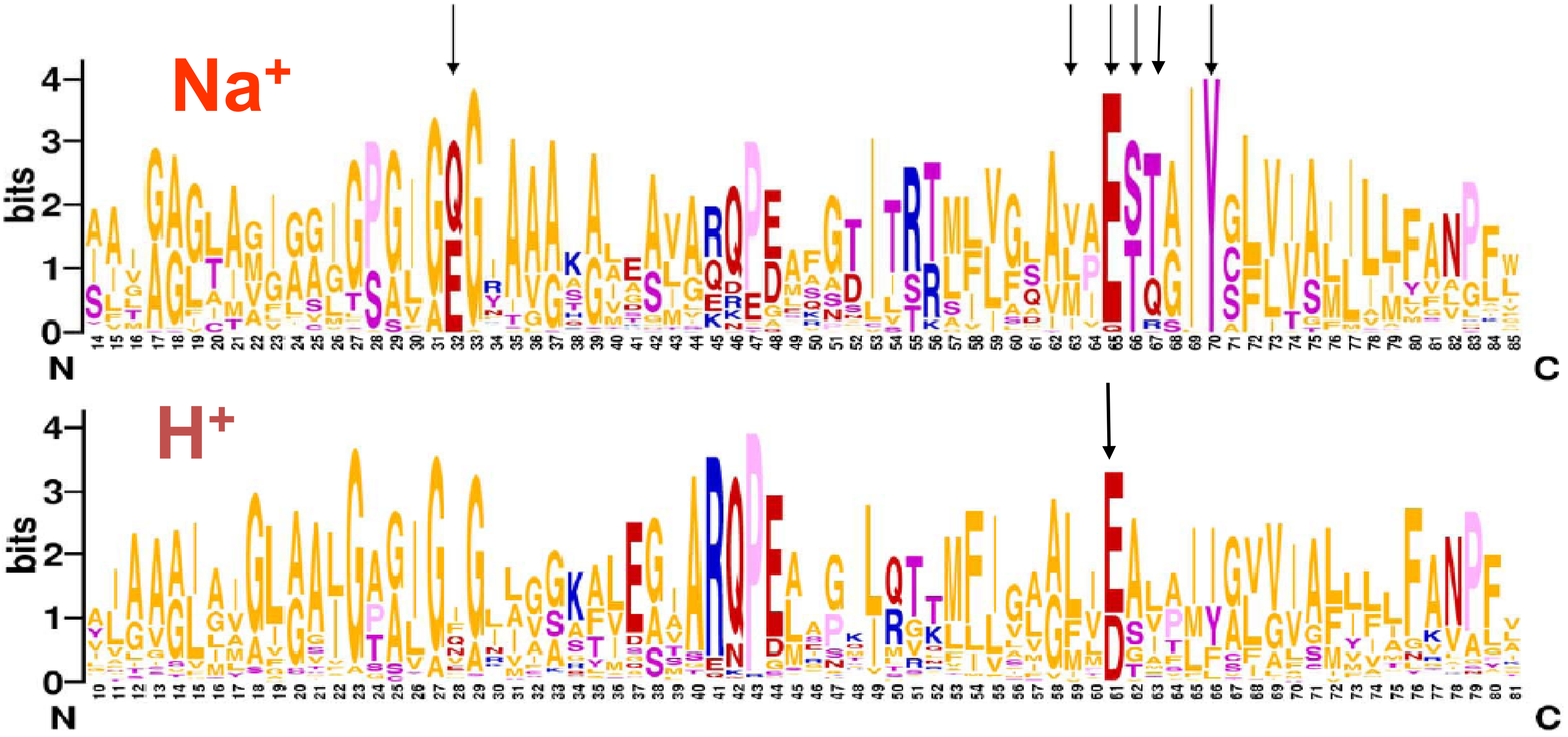


Enterococcus hirae,
Murata et al., 2005, *Science*

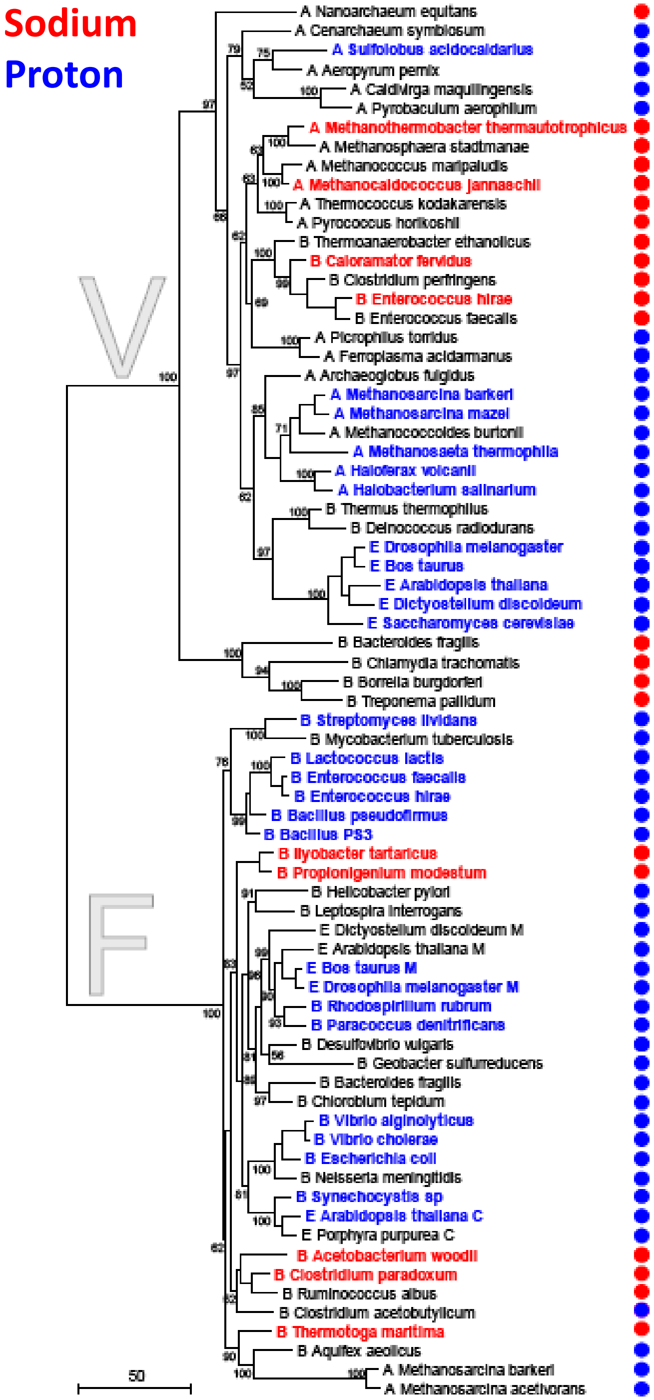


From Mulkidjanian et al., 2008, *Biology Direct*

Translocation of Na$^+$ ions requires a complete set of ligands, while H$^+$ can be translocated by a single charged amino acid residue



From Mulkidjanian et al., 2008, *Biology Direct*



The Na^+-dependent ATPases are spread over the phylogenetic tree. Independent emergence of same sets of Na^+ ligands in several lineages does not look plausible.

The common ancestor of the rotary ATPases should have had a Na^+-binding site.

The Na^+ energetics comes out as the ancestral one.

Most likely, the Na^+-tight membranes preceded the H^+-tight membranes in evolution

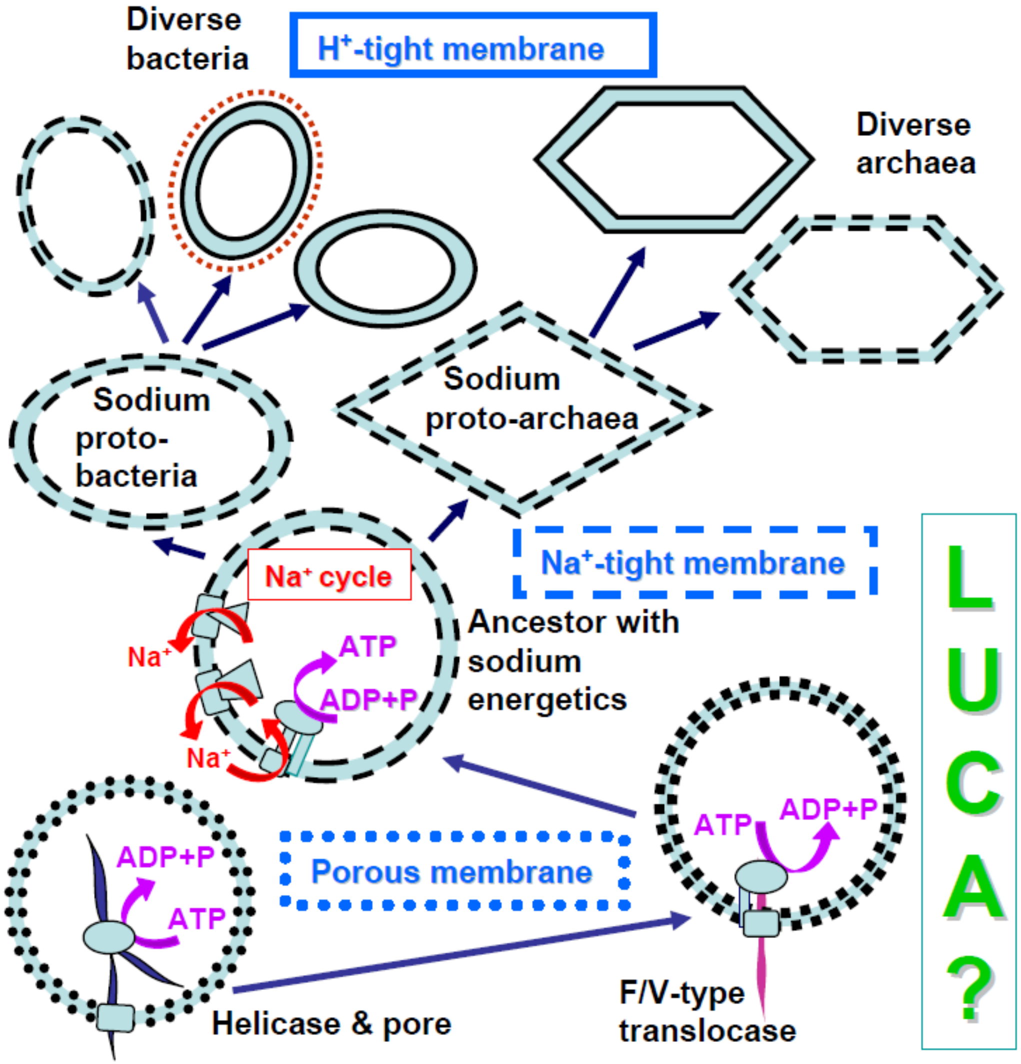
From Mulkidjanian et al., 2008, *Biology Direct*

The membrane permeability increases dramatically in the presence of voltage that is needed to drive ATP synthesis
(the field strength across a prokaryotic membrane is about 300000 V/cm)

In mitochondrial membranes proton leaks account for the loss of 10-50% of the whole energy.

Primodial cells may have used sodium as a less penetrative cation.

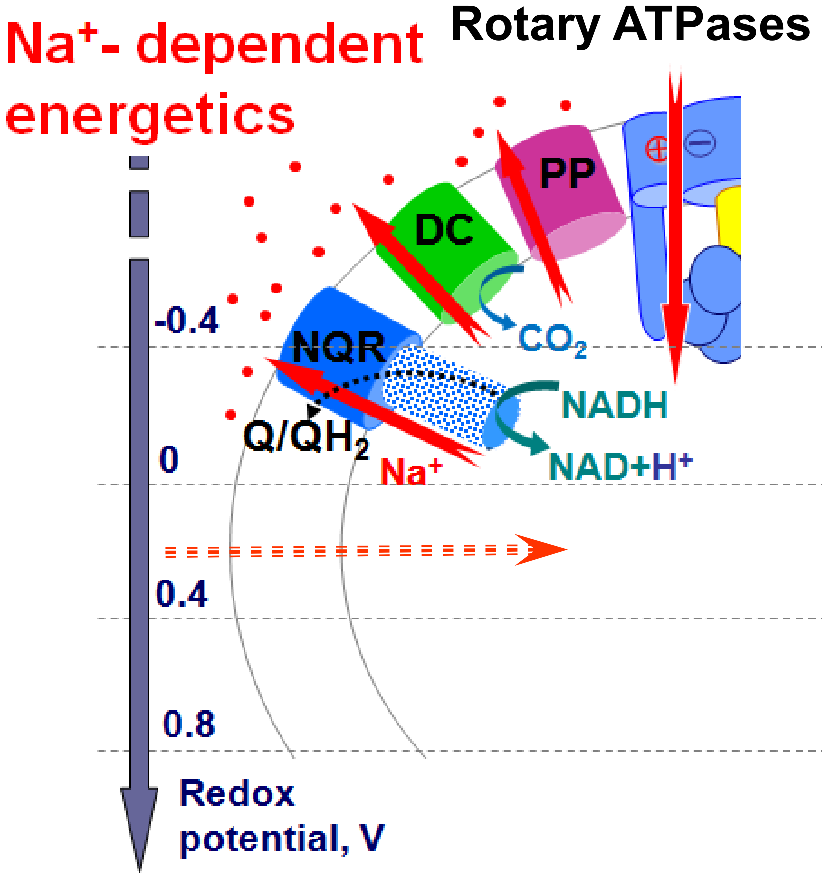
Separate transition to proton bioenergetics in Bacteria and Archaea



Since the common ancestor of the rotary ATPase was a Na$^+$-binding enzyme, the transition to the proton bioenergetics may have proceeded separately in Bacteria and Archaea.

From: Mulkidjanian et al., 2009, *TIBS*

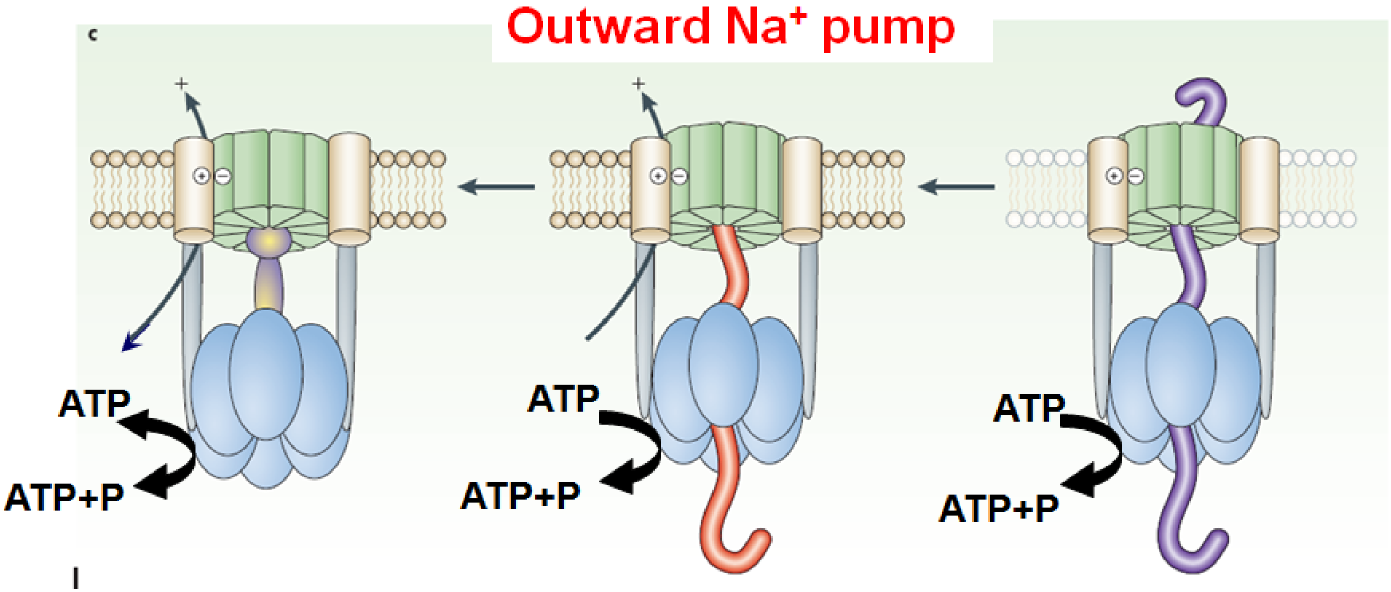
Where the Na^+-dependent, energy-converting enzymes came from?



The enzymes of Na^+-dependent bioenergetics are outward Na^+-pumps

The early cells could invade sodium-rich, e.g. marine environments by „inventing“ special enzymes for expelling Na^+ ions

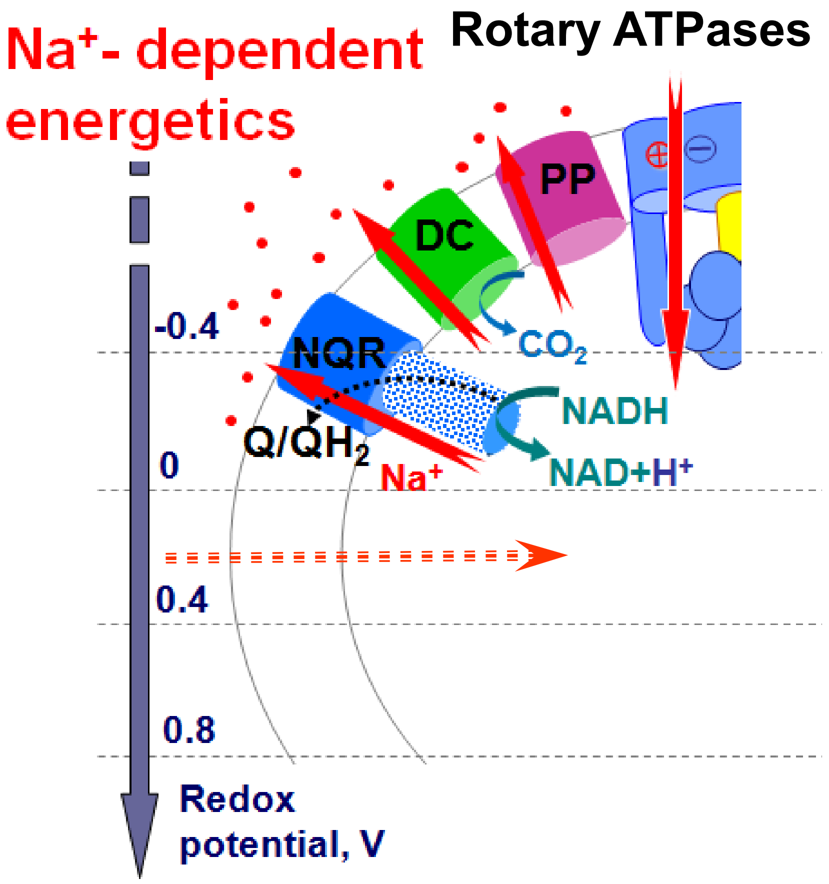
From: Mulkidjanian et al., 2009, *TIBS*



From Mulkidjanian et al., 2007, *Nature Rev. Microb.*

The rotary Na-translocating ATP synthase may have evolved from a ATP driven protein translocase via an intermediate step of a Na$^+$-expelling ATPase

Some recently described sodium out-pumps



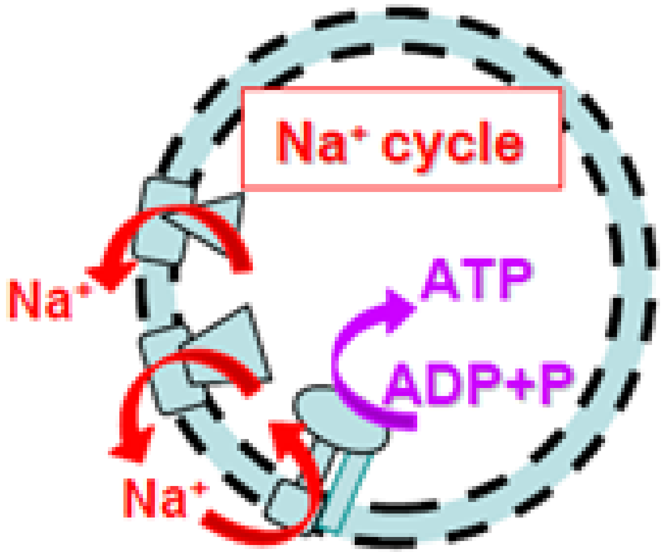
N-ATPase (Dibrova et al. 2010, Soontharapirakkul eat al. 2011)

Ferredoxin:NAD$^+$oxidoreductase (Biegel and Mueller, 2010)

Na$^+$-proteorhodopsin (Inoue et al, 2013)

From: Mulkidjanian et al., 2009, *TIBS*

Interplay between initially independent Na^+ outpumps could yield the first bioenergetic cycle.

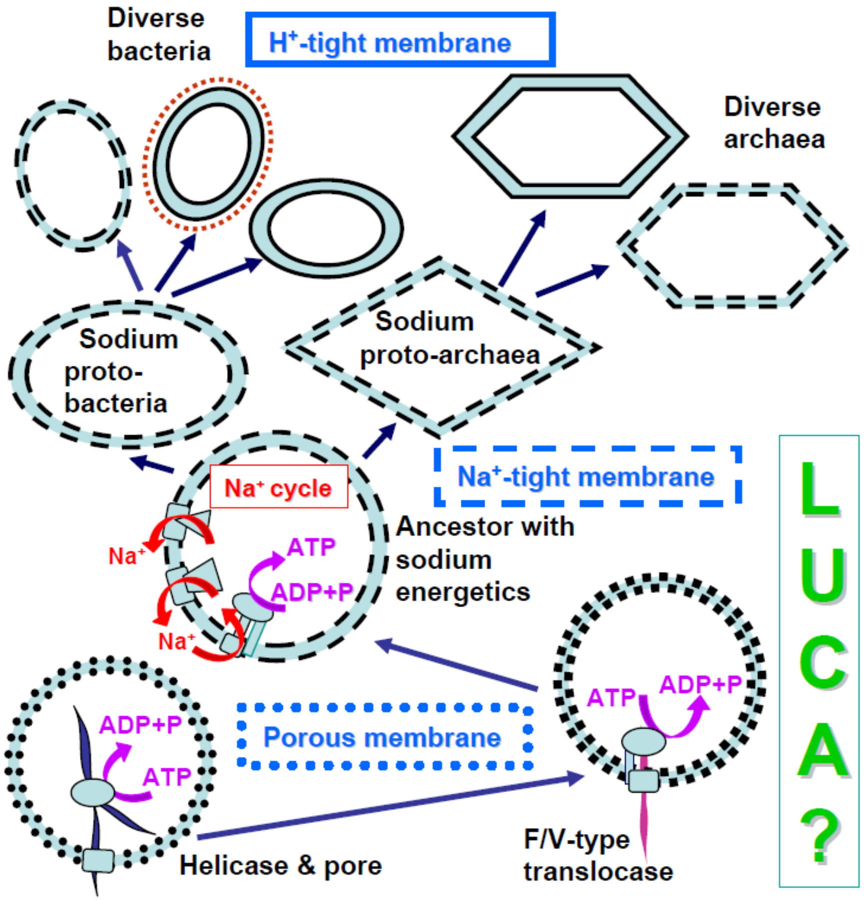


The ancient rotary ATPase, owing to its rotating scaffold, would be potentially able to translocate Na^+ ions in both directions.

The concentration of Na^+ in the Archean ocean was about 1 M; the Na^+ gradient could be powerful enough to cause the switch of the rotary machine from the hydrolysis to the synthesis of ATP <u>provided that the membranes were already tight enough</u>.

- Introduction: Ubiquitous protein set and the Last Universal Cellular Ancestor
- Emergence of the first cells in K^+-rich environments
- Invasion of Na^+-rich habitats
- **On the co-evolution of membranes and enzyme machinery**

Co-evolution of membranes and bioenergetics



From: Mulkidjanian et al., 2009, *TIBS*

Proton-tight membranes, proton-dependent bioenergetics

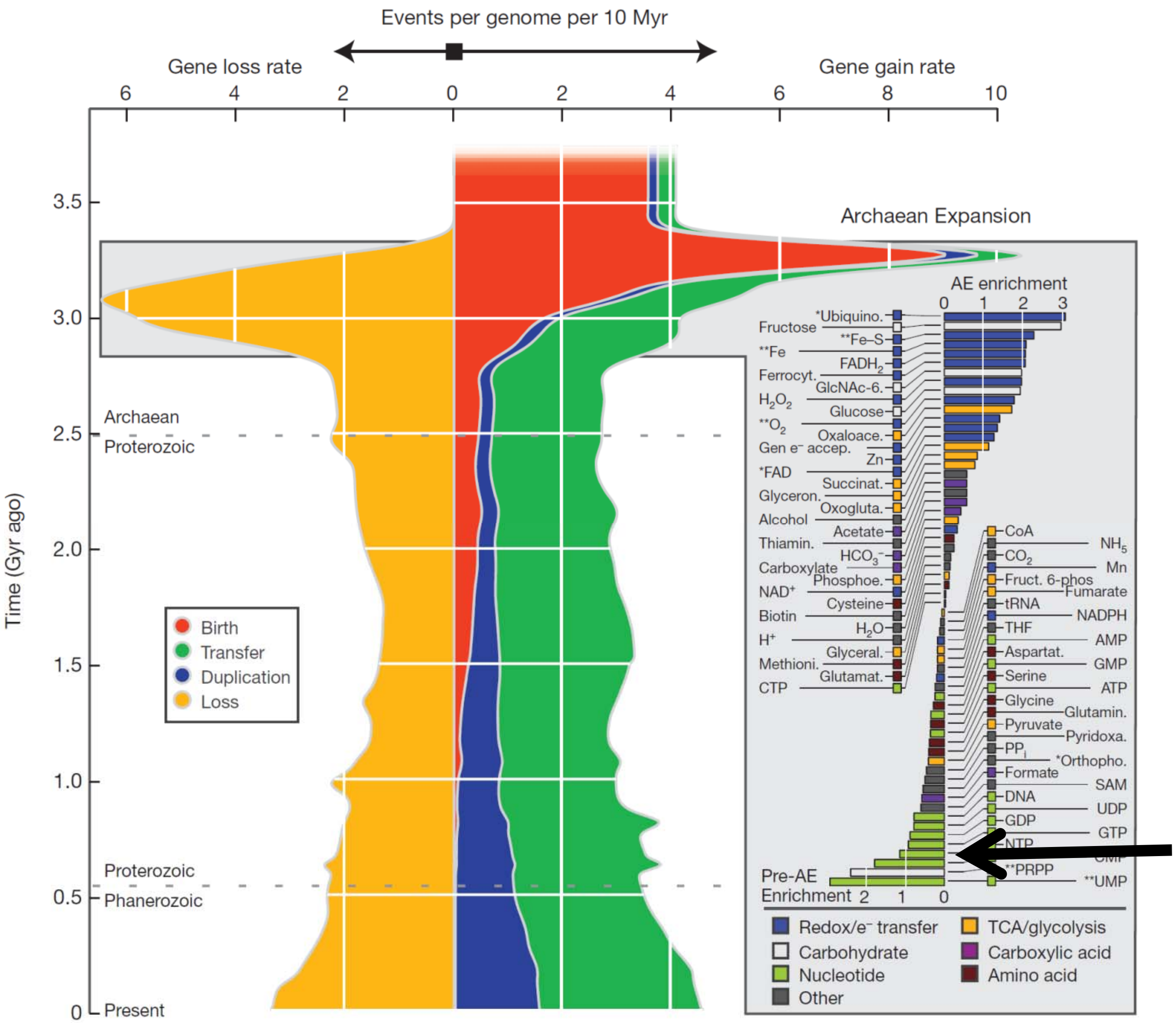
primitive, sodium-dependent bioenergetics in marine environments (?),

Ion-tight membranes, sodium out-pumping systems

K^+-rich environments, leaky membranes, phosphate transfer reactions as a source of energy

time

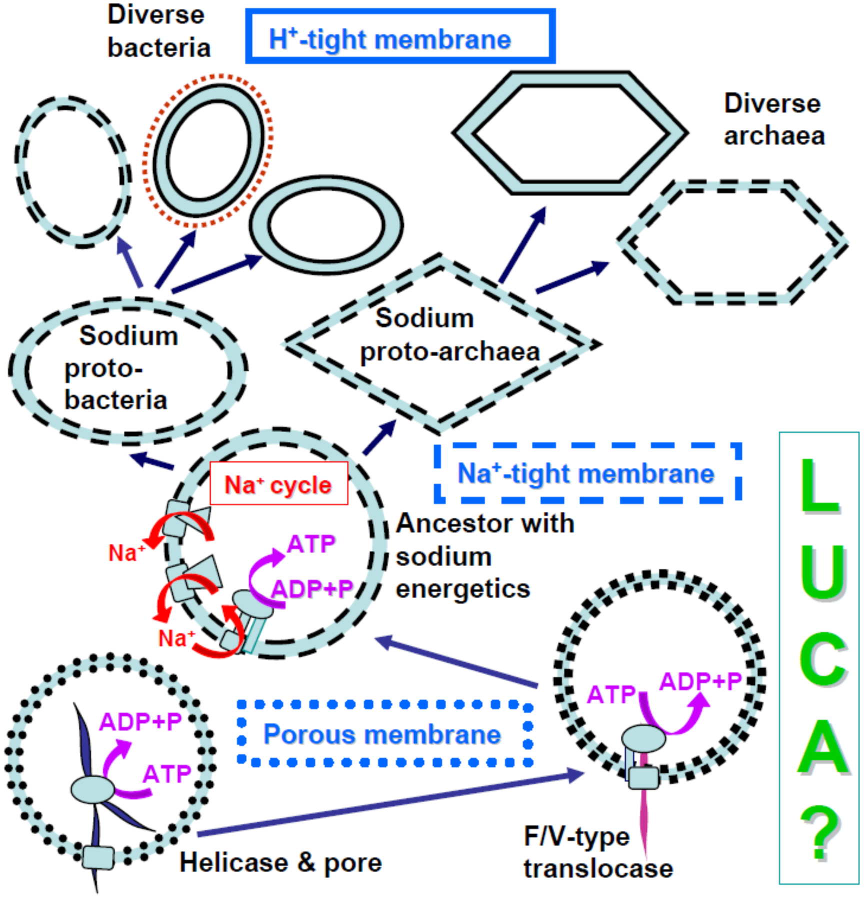
The primordial RNA organisms and the first cells could have used phosphate transfer reactions as an energy source



From: David & Alm, 2011, Nature

The earliest emergence of phosphate-transferring enzymes

Co-evolution of membranes and bioenergetics



From: Mulkidjanian et al., 2009, *TIBS*

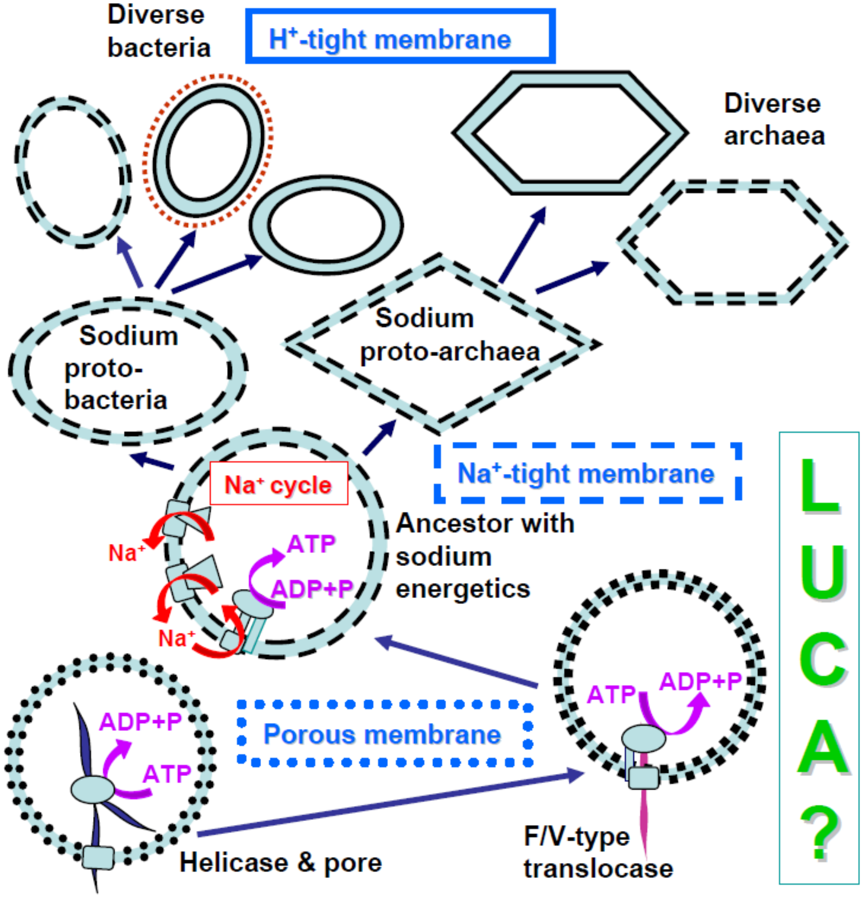
Proton-tight membranes, proton-dependent bioenergetics

primitive, sodium-dependent bioenergetics in marine environments (?),

Ion-tight membranes, sodium out-pumping systems

K$^+$-rich environments, leaky membranes, phosphate transfer reactions as a source of energy

Co-evolution of membranes and bioenergetics



From: Mulkidjanian et al., 2009, *TIBS*

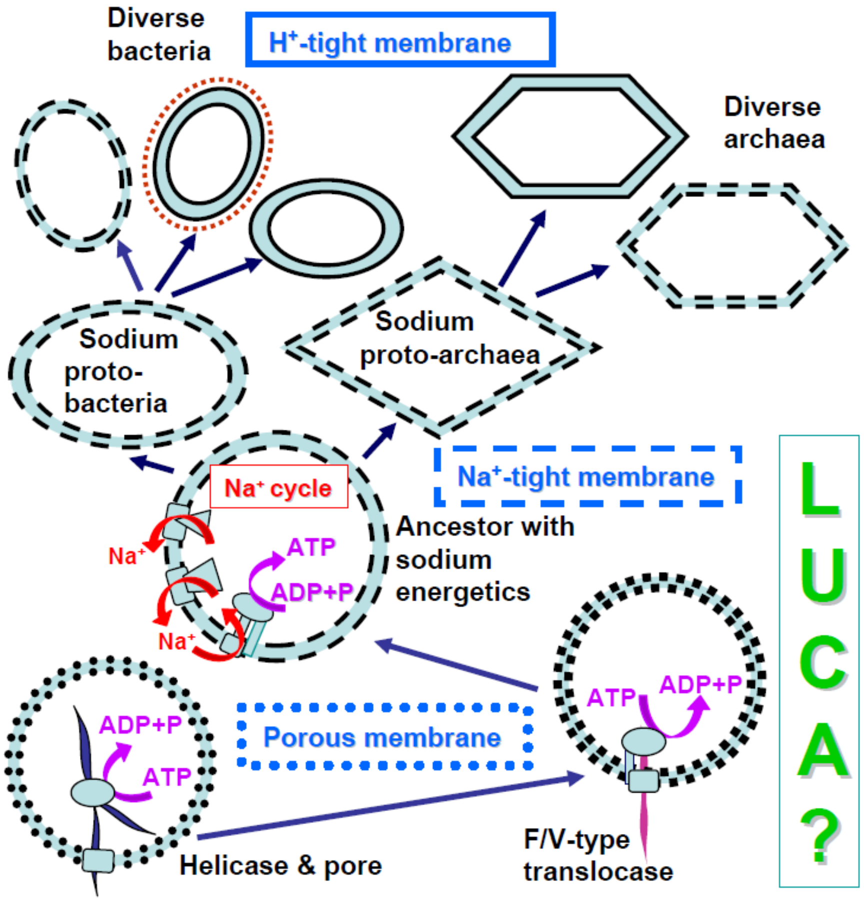
Proton-tight membranes, proton-dependent bioenergetics

primitive, sodium-dependent bioenergetics in marine environments (?),

Ion-tight membranes, sodium out-pumping systems

K⁺-rich environments, leaky membranes, phosphate transfer reactions as a source of energy

Co-evolution of membranes and bioenergetics



From: Mulkidjanian et al., 2009, *TIBS*

Proton-tight membranes, proton-dependent bioenergetics

primitive, sodium-dependent bioenergetics in marine environments (?),

Ion-tight membranes, sodium out-pumping systems

K^+-rich environments, leaky membranes, phosphate transfer reactions as a source of energy

Conclusions

The first cells may have emerged in K^+-rich environments, supposedly, at anoxic geothermal fields.

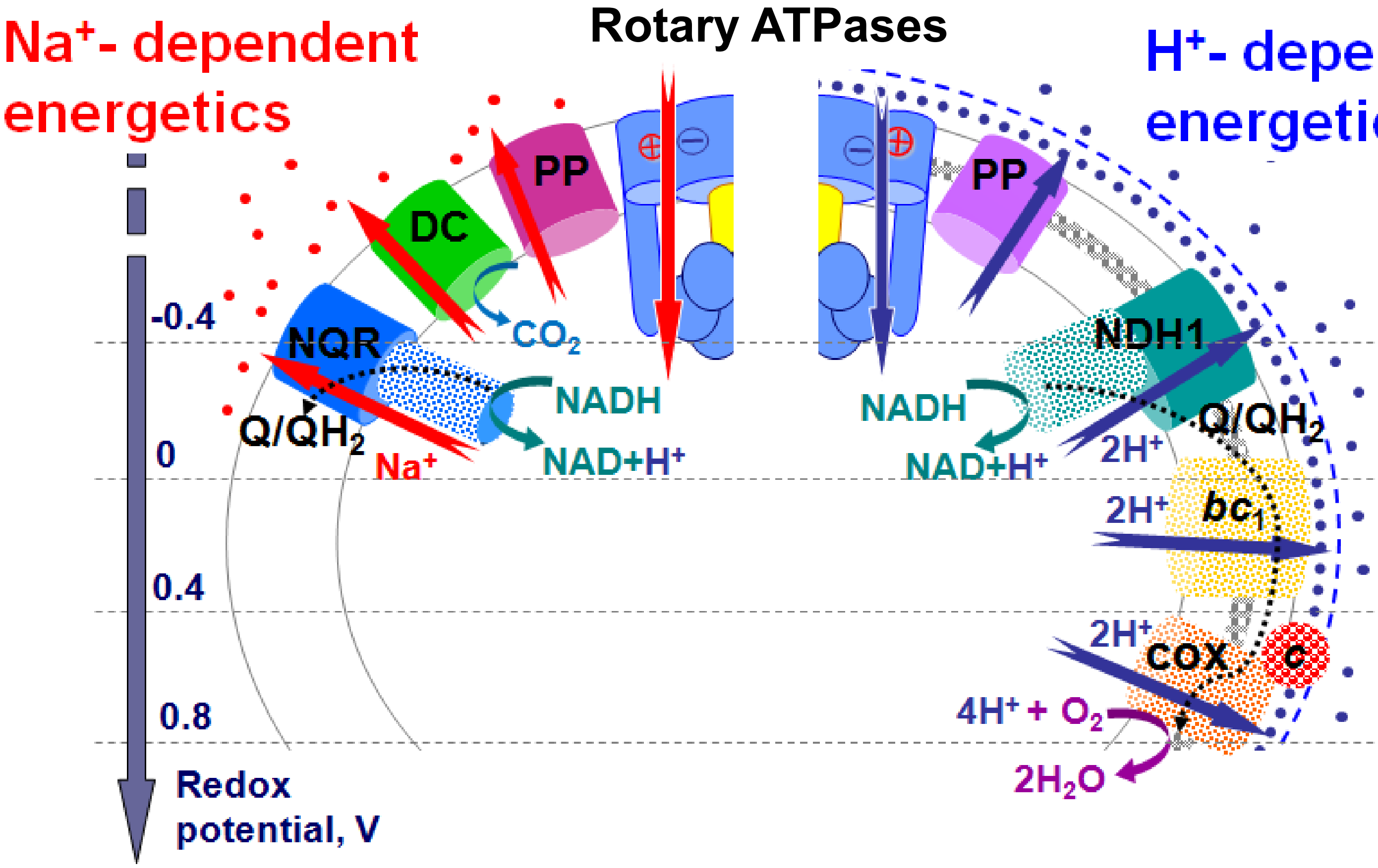
The foray into new, Na^+-rich habitats was the likely driving force for the emergence of diverse redox-, light, chemically- or osmotically-dependent sodium export pumps, as well as for the improvement of membrane tightness.

An interplay between several, initially independently acting sodium pumps may have yielded the Na^+-dependent membrane bioenergetics and the energy-harvesting rotary ATP synthase.

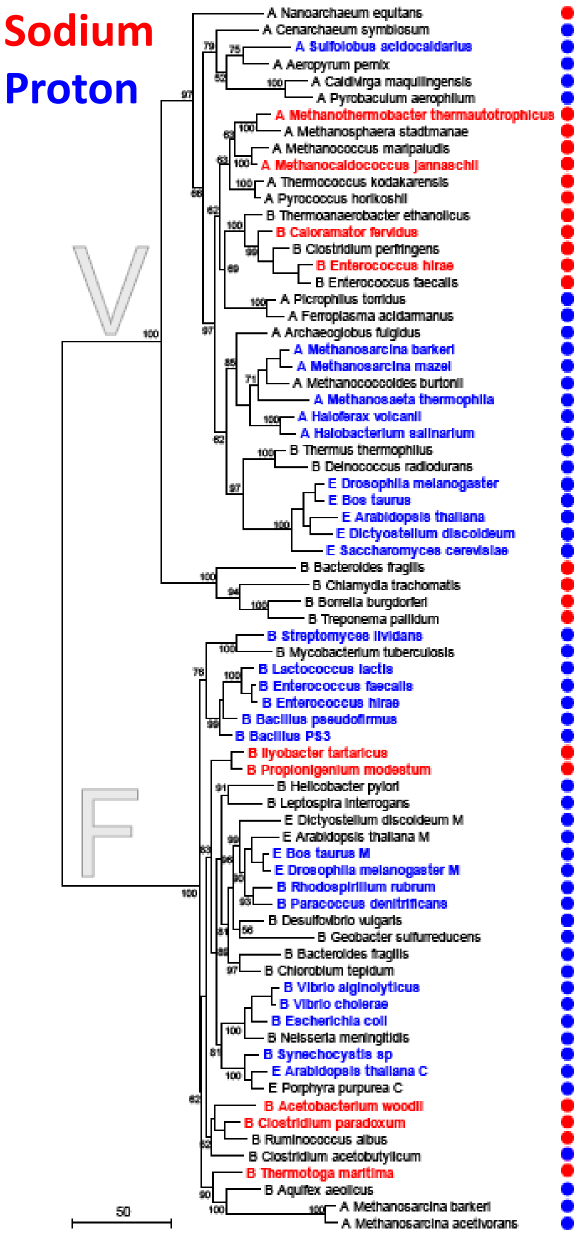
The key function of the membrane energy converting systems is to maintain internal elementary composition similar to those in the K^+-rich habitats of the very first cells.

Additional slides

What determines the choice between the Na^+ and H^+-dependent energetics in modern organisms?

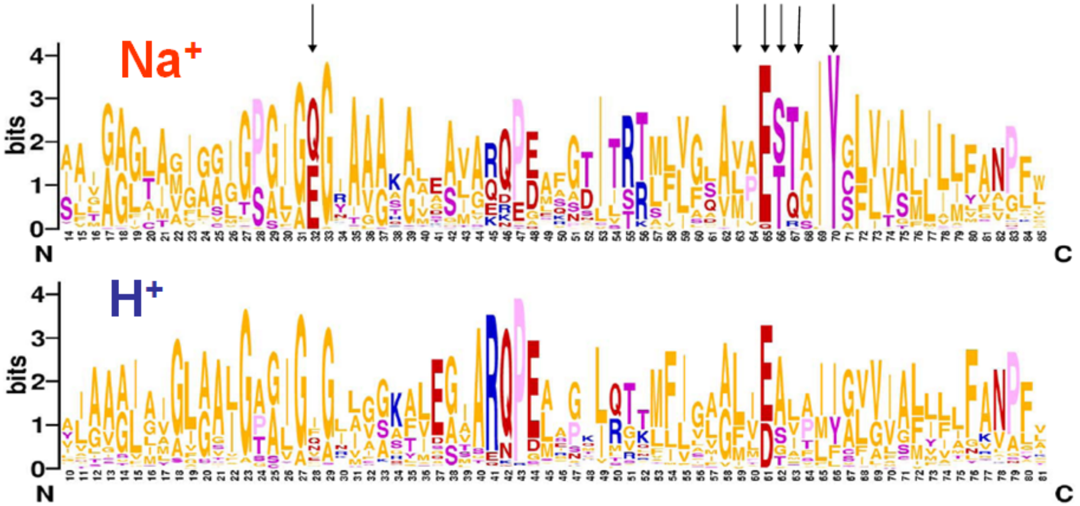


Is the Na^+-dependent energetics preferred:
by alkaliphiles? by thermophiles?

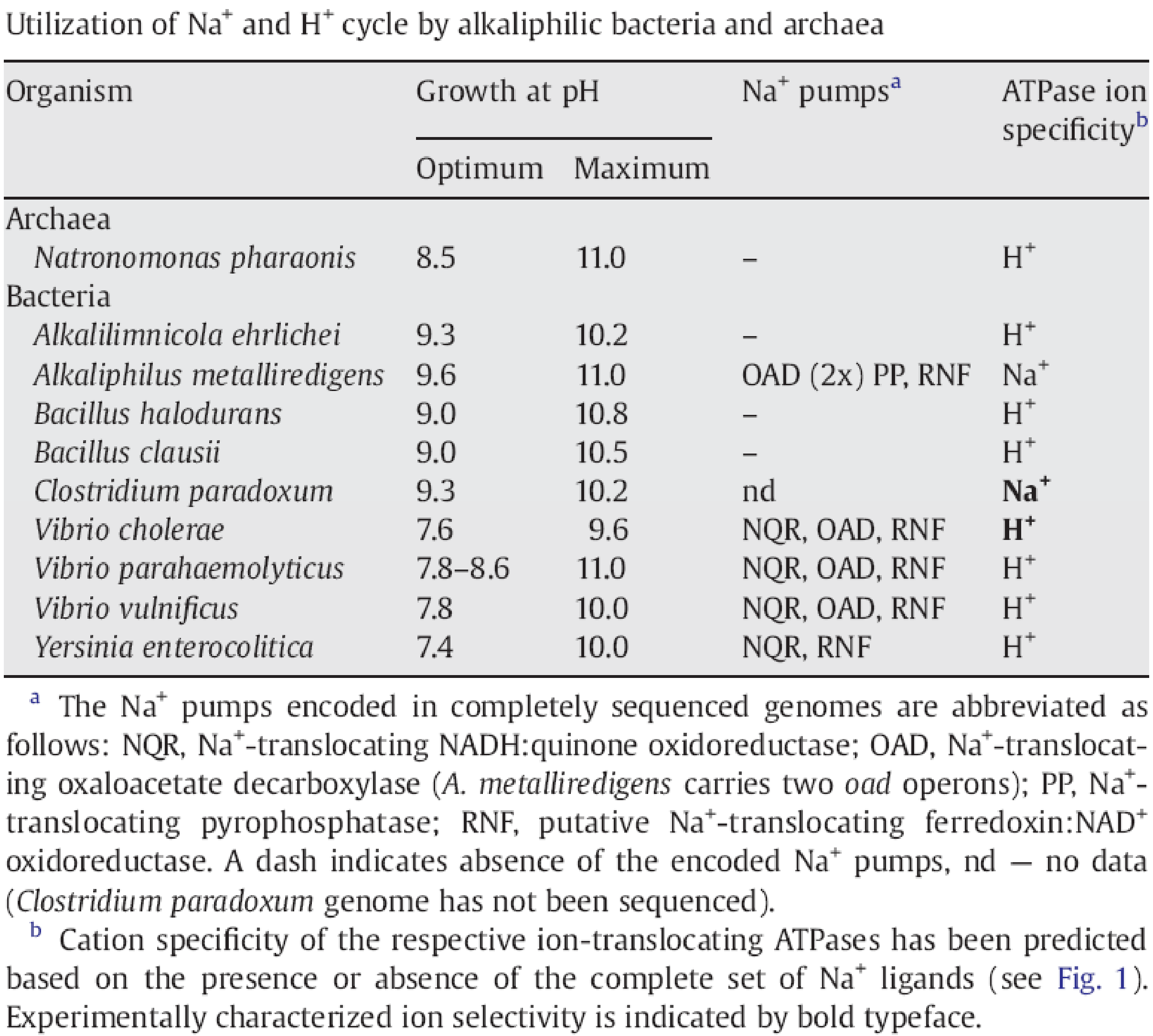


It is possible to determine the type of energetics in an organism just from the sequence of the proteoilipid subunit, without doing any experiments.

Those organisms that have a full complement of Na^+ ligands have also a Na^+-dependent bioenergetics.



From Mulkidjanian et al., 2008, *Biology Direct*



Utilization of Na^+ and H^+ cycle by alkaliphilic bacteria and archaea

Organism	Growth at pH		Na^+ pumps[a]	ATPase ion specificity[b]
	Optimum	Maximum		
Archaea				
Natronomonas pharaonis	8.5	11.0	–	H^+
Bacteria				
Alkalilimnicola ehrlichei	9.3	10.2	–	H^+
Alkaliphilus metalliredigens	9.6	11.0	OAD (2x) PP, RNF	Na^+
Bacillus halodurans	9.0	10.8	–	H^+
Bacillus clausii	9.0	10.5	–	H^+
Clostridium paradoxum	9.3	10.2	nd	**Na^+**
Vibrio cholerae	7.6	9.6	NQR, OAD, RNF	**H^+**
Vibrio parahaemolyticus	7.8–8.6	11.0	NQR, OAD, RNF	H^+
Vibrio vulnificus	7.8	10.0	NQR, OAD, RNF	H^+
Yersinia enterocolitica	7.4	10.0	NQR, RNF	H^+

[a] The Na^+ pumps encoded in completely sequenced genomes are abbreviated as follows: NQR, Na^+-translocating NADH:quinone oxidoreductase; OAD, Na^+-translocating oxaloacetate decarboxylase (*A. metalliredigens* carries two *oad* operons); PP, Na^+-translocating pyrophosphatase; RNF, putative Na^+-translocating ferredoxin:NAD^+ oxidoreductase. A dash indicates absence of the encoded Na^+ pumps, nd — no data (*Clostridium paradoxum* genome has not been sequenced).

[b] Cation specificity of the respective ion-translocating ATPases has been predicted based on the presence or absence of the complete set of Na^+ ligands (see Fig. 1). Experimentally characterized ion selectivity is indicated by bold typeface.

Genome analysis shows that Na^+ energetics does not correlate with alkaliphily

From: Mulkidjanian, Dibrov, Galperin, 2008, BBA

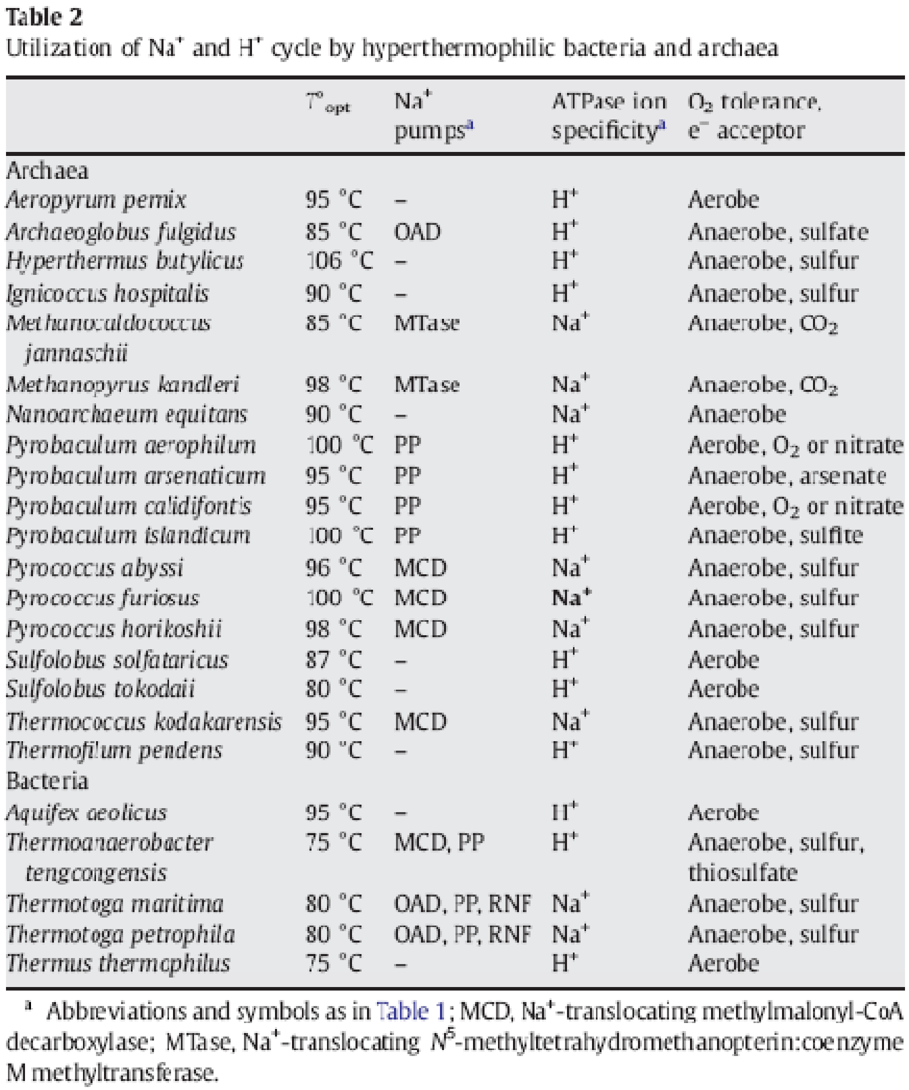


Table 2
Utilization of Na^+ and H^+ cycle by hyperthermophilic bacteria and archaea

	T_{opt}	Na^+ pumps[a]	ATPase ion specificity[a]	O_2 tolerance, e^- acceptor
Archaea				
Aeropyrum pernix	95 °C	–	H^+	Aerobe
Archaeoglobus fulgidus	85 °C	OAD	H^+	Anaerobe, sulfate
Hyperthermus butylicus	106 °C	–	H^+	Anaerobe, sulfur
Ignicoccus hospitalis	90 °C	–	H^+	Anaerobe, sulfur
Methanocaldococcus jannaschii	85 °C	MTase	Na^+	Anaerobe, CO_2
Methanopyrus kandleri	98 °C	MTase	Na^+	Anaerobe, CO_2
Nanoarchaeum equitans	90 °C	–	Na^+	Anaerobe
Pyrobaculum aerophilum	100 °C	PP	H^+	Aerobe, O_2 or nitrate
Pyrobaculum arsenaticum	95 °C	PP	H^+	Anaerobe, arsenate
Pyrobaculum calidifontis	95 °C	PP	H^+	Aerobe, O_2 or nitrate
Pyrobaculum islandicum	100 °C	PP	H^+	Anaerobe, sulfite
Pyrococcus abyssi	96 °C	MCD	Na^+	Anaerobe, sulfur
Pyrococcus furiosus	100 °C	MCD	Na^+	Anaerobe, sulfur
Pyrococcus horikoshii	98 °C	MCD	Na^+	Anaerobe, sulfur
Sulfolobus solfataricus	87 °C	–	H^+	Aerobe
Sulfolobus tokodaii	80 °C	–	H^+	Aerobe
Thermococcus kodakarensis	95 °C	MCD	Na^+	Anaerobe, sulfur
Thermofilum pendens	90 °C	–	H^+	Anaerobe, sulfur
Bacteria				
Aquifex aeolicus	95 °C	–	H^+	Aerobe
Thermoanaerobacter tengcongensis	75 °C	MCD, PP	H^+	Anaerobe, sulfur, thiosulfate
Thermotoga maritima	80 °C	OAD, PP, RNF	Na^+	Anaerobe, sulfur
Thermotoga petrophila	80 °C	OAD, PP, RNF	Na^+	Anaerobe, sulfur
Thermus thermophilus	75 °C	–	H^+	Aerobe

[a] Abbreviations and symbols as in Table 1; MCD, Na^+-translocating methylmalonyl-CoA decarboxylase; MTase, Na^+-translocating N^5-methyltetrahydromethanopterin:coenzyme M methyltransferase.

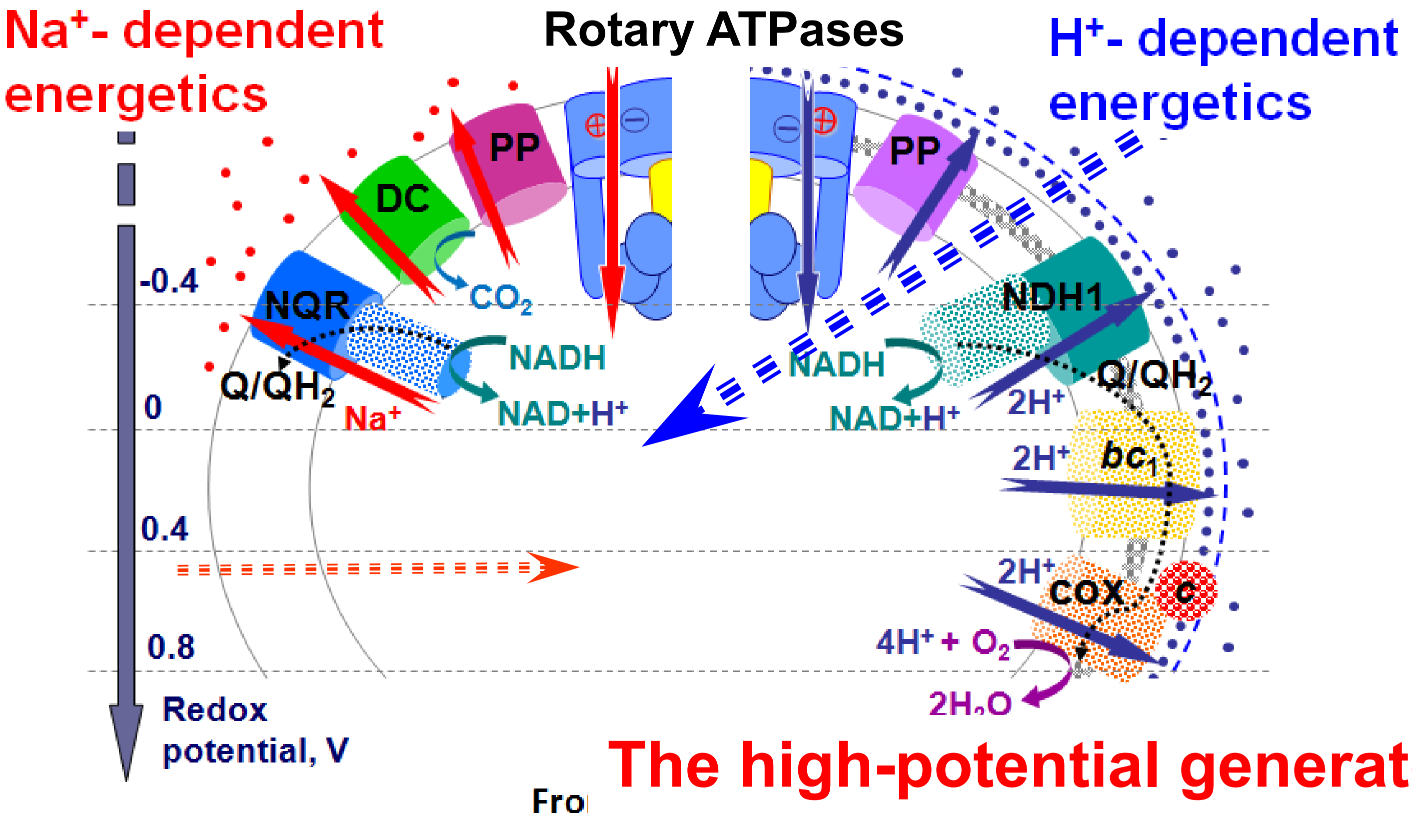
As well, the Na^+ energetics does not correlate with thermophily...

However, there is a clear correlation between the H^+ cycle and the ability to utilyze oxygen or other terminal electron acceptors such as nitrate, sulfate, or sulfite.

Elementary sulfur makes an intermediate case.

From: Mulkidjanian, Dibrov, Galperin, 2008, BBA

Whether an organism relies on Na^+ or H^+, seems to depend on a trade-off between the amount of potentially available free energy and the intensity of ion leakage across the coupling membrane.



Fro

The high-potential generators of membrane potential are obligatory proton-tranlocators.

Those extremophiles that have ample supply of electron acceptors, be it oxygen, nitrate or sulfate, are less energy-limited and can rely on proton energetics even while paying a heavy price of elevated proton leakage at high temperature or at high pH.

In some obligate anaerobes, whose energy budget is tight and cannot cover the losses caused by proton leaks, the Na^+ energetics may become the favored one.

Under the reducing conditions of the primordial Earth, the proton-dependent energetics could be hardly beneficial (Mårtin Wikström).

The Earth could be populated by organisms similar to the δ-proteobacterium *Desulfotignum phosphitoxidans* (Poehlein et al. 2013, BMC Genomics)