

# Structural Co-Evolution of Viruses and Cells in the Primordial World

Matti Jalasvuori · Jaana K. H. Bamford

Received: 16 August 2007 / Accepted: 7 January 2008 /  
Published online: 29 January 2008  
© Springer Science + Business Media B.V. 2008

**Abstract** Viruses and cells co-evolve due to the parasitic nature of viruses. Yet there are no models suggesting how the unicellular organisms and their viruses might co-evolve structurally. Here, in this study, we plunge into this unexplored field from a wide perspective and try to describe some of the intriguing ways in which viruses may have shaped the cellular life forms on the ancient Earth. At first we propose a scenario where viruses act as a driving force in the emergence of bacterial cell walls by providing favorable intermediates for the otherwise improbable steps in the cell wall generation. We also discuss the role of viruses in the evolution of cell surface components such as receptors and second membranes. Finally we focus on hypothetical proto-viruses, the selfish abusers of the RNA-world, in explaining some of the very early stages in the origin and evolution of life. Proto-viruses may be responsible for creating the first true cells in order to support their selfish needs. In this model we also suggest a logical pathway to explaining the emergence of modern viruses.

**Keywords** Cell wall · Evolution of viruses · Origin of cells · Origin of viruses · RNA world

## Introduction

A cell is the primary unit when life on Earth is being defined. Yet, there are also other systematically reproducing elements existing on our planet which outnumber the cells by orders of magnitude (Bergh et al. 1989; Wommack et al. 1992). Current evidence suggests that at least some viruses existed before the last universal common ancestor (LUCA) of

---

M. Jalasvuori · J. K. H. Bamford  
Department of Biological and Environmental Science and Nanoscience Center, University of Jyväskylä,  
P.O. Box 35, FIN-410014 Jyväskylä, Finland

M. Jalasvuori (✉)  
Department of Biological and Environmental Science and Nanoscience Center, University of Jyväskylä,  
P.O. Box 35, 40014 Jyväskylä, Finland  
e-mail: matti.jalasvuori@jyu.fi

cells (Benson et al. 1999; Bamford 2003; Koonin et al. 2006). Therefore these abundant parasites must have affected the evolution of cellular life in numerous ways throughout natural history.

Viruses are a major factor in the horizontal gene transfer (HGT) since they can acquire genes from the host genomes and again deliver the genes to other organisms (Kidambi et al. 1994; Canchaya et al. 2003; Comeau and Krisch 2005). There are several examples in modern cells of HGT where genes do not show vertical but lateral relationships to one another (Iyer et al. 2006). It is possible that viruses can carry out major evolutionary changes in single evolutionary events due to their genome-editing competence (Witzany 2006) as they can integrate into host genomes (Holzel and Sokol 1974; Doerfler 1975; Brussow et al. 2004; Krupovic and Bamford 2007). Viruses are also thought to sustain the diversity of microbial communities by eliminating the most successful organism from the ecosystem i.e. “killing the winner” (Weinbauer and Rassoulzadegan 2004; Weinbauer 2004). Therefore viruses have been suggested to be the driving force of microbial evolution. The only structural approach of exploring the co-evolution of viruses and prokaryotes concerns the virus host recognition proteins and their cellular counterparts. In these studies it has been shown that viruses quickly adapt to bind host cell structures which have mutated beyond recognition (Buckling and Rainey 2002). However, we believe that the ways described above are not sufficient in recognizing the true role of viruses in the evolution of life. Therefore we propose a wide perspective and a more structure-oriented model to expand the current views.

Most of the microbial cells are surrounded by a rigid layer known as the cell wall which allows the cell to maintain a high inner osmotic pressure (Koch 2006). In the absence of the cell wall, the osmotic pressure causes the cell to lyse, which is a common way for viruses to release virions into the environment during the final stages of the infection (Young 1992). Cell wall chemistries differ tremendously between the domains of archaea and bacteria (Boucher et al. 2004), which indicates that cell walls might have emerged several times on the ancient Earth. Eukaryotic organisms also have a variety of diverse cell wall structures, but currently it is unclear whether Eukarya arose early in the past or whether it is a rather recent fusion of the domains of Archaea and Bacteria (Poole and Penny 2007). Therefore the eukaryal cell surface can be credibly used in the arguments only for the later stages of cell development, whereas archaeal and bacterial cell walls are most probably genuine relics of ancient evolution (Brown and Doolittle 1997).

The emergence of bacterial cell wall is suggested to be located in the very root of the tree of life since cell walls made it possible for the ancient organisms to maintain their genetic isolation effectively and thus made vertical evolution more efficient in comparison to the early genetic communities where lateral gene-movement was frequent (Koch 1994, 2003; Woese 2002). The formation of the bacterial cell wall has been suggested to be the point when the domain Bacteria originated (Koch 1998; Zorzopulos 2003). Bacteria have two basic types of cell structure: the Gram-negative cells have two membranes on both sides of a thin cell wall and the Gram-positive cells have only an inner membrane beneath a thick cell wall. The bacterial cell wall is made of peptidoglycan and its mechanisms of synthesis are well studied. The cell wall of Gram-positive bacteria is synthesized by a formation of peptidoglycan sub-units that are transferred through the membrane by a bactoprenol carrier (Anderson et al. 1965; Higashi et al. 1967). Individual sub-units are then covalently bond to each other. In this hypothesis we focus on Gram-positive like bacteria with only inner membrane as they might have been logical intermediates in forming modern bacteria and some studies also suggest Gram-positives to be close to the early forms of this domain (Gupta 2001; Koch 2003). The independent and detailed examination of Archaea does not

seem to broaden our idea in any significant way as our proposal does not concern events on enzymatic level of cell wall formation.

What were the driving forces that caused the first cell wall to emerge? The answer to this question provides details to our current knowledge of the origin of the three domains of life and therefore a better understanding in this matter is critical for building satisfying models for the early stages of life on Earth. One of the main focuses of this paper is to propose a logical model for what has promoted the emergence of cell walls. We also discuss the evolution of single membrane proteins in the presence of viruses.

Secondly, we reach out to the hypothesized initial stages of all evolution on Earth: into the RNA-world. Life's early biochemistry was possibly carried out with a RNA (or RNA-like) molecule (Orgel 2004) which was able to serve as genetic material and to perform enzymatic tasks. Probably no translated proteins or DNA existed in the beginning and this pre-protein and pre-DNA period is commonly known as the RNA-world (for a review see Dworkin et al. 2003). The pathway from the first self-replicating (or auto-replicating) molecule (or set of molecules) to the first functional cell is an unsolved mystery in the current models of life's origin. We combined the structural approach used in the cell wall model with the selfishness of viruses to come up with an idea which seems to provide some potential answers to the first steps in life's history. In our scenario the emergence of virus-like exploiters was placed in the beginning of life. This hypothesis revealed intriguing and possible solutions to explaining the origin of actual cells and modern viruses.

The goal of this hypothesis is to provide more understanding to the role of viruses in the evolution of cells (and viruses). The scenarios presented here, while being hypothetical, may possess valuable information for future studies, ideas and models concerning for example *in vivo* evolution of artificial life, computational simulations of early Earth biology and astrobiology research performed on extraterrestrial locations. To make this topic interesting and provoking, in the end of each of the next chapter's sections we suggest some highly hypothetical scenarios of how the possible absence of viruses could have altered the currently known pathway of the evolution of life.

## The Emergence of Cell Walls

### Virus Driven Evolution Could Explain Why Cell Walls Emerged

The Darwinian principle states that any two stages in the evolution of specie must have favorable intermediates. While the true origin of the first species might have occurred only after the first cell wall was reliably established as the cell wall provided the genetic privacy needed for vertical evolution (Zorzopulos 2003), the Darwinian principle had to be true in a somewhat different form before the formation of cell walls (Woese 2002). The machinery for building bacterial murein (or peptidoglycan) is rather complex and by no means could it have appeared in a single step. Here we have developed a rough model to explain some of the intermediates of the cell wall evolution starting from the early cell without a cell wall and ending in a cell with a wall.

Bacterial cell wall makes it possible for the bacterial cells to maintain high osmotic pressure, which in turn provides the cell better means to obtain energy. This, however, is an unlikely reason for the emergence of the cell walls as the cell wall has to fully surround the cell and the peptidoglycan units have to be connected to each other before it provides the needed structural support. We suggest that viruses made the formation of even a few peptidoglycan units favorable for the cell as they stuck out from the membrane and

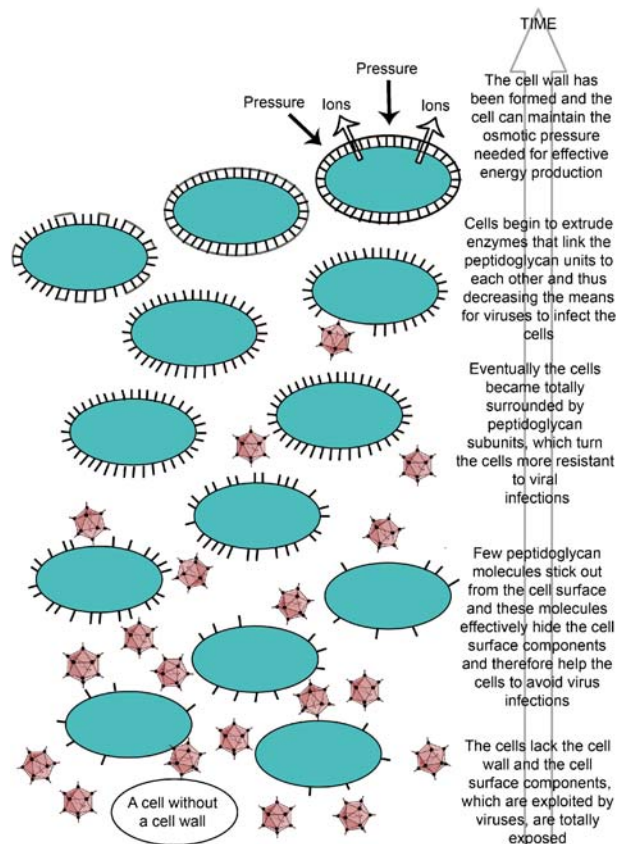
therefore reduced the suitable area where viruses can infect the cell. The less membrane area there was for primitive viruses to fuse with, the less likely it was that the cells got abused by viruses. The basic idea is presented in Fig. 1.

Membrane-bound cells probably lived in a hydrothermal vent system (Monnard and Deamer 2002). In this environment viruses might have already been exploiting the resources of cells (Koonin and Martin 2005; Koonin et al. 2006). Due to the lack of cell walls, the original viruses probably had lipid envelopes which fused with the cell membrane during the infection. Thus, the more membrane area there was available for the viruses to bind the easier it was for a cell to get infected.

Let us hypothesize that one cell produces a peptidoglycan unit which is transferred outside the cell membrane. Now the virus seeking for a suitable host in order to initiate an infection has a slightly smaller chance to infect a cell which has a region camouflaged from viruses. Thus even a single peptidoglycan unit can be of advantage to a cell in an environment where there are viruses present. However, if no viruses were around, this peptidoglycan unit might only be a waste of resources (although it can provide some protection from other external threats, such as preventing direct contact with rock surface as rock surface can be inhibiting agent for many biomolecules).

It is possible that this initial selective advantage provided by some peptidoglycan subunits favored some cells over their naked cousins. Subsequently, any cell that produced

**Fig. 1** The emergence of bacterial cell walls due to a selective pressure caused by virus infections



more peptidoglycan masked itself from viruses more effectively and was favored again. In the end, there would be a cell totally covered in murein which would be safe from the original viruses. Due to this advantage it would eventually become the dominant species in the hydrothermal world. If cells were not reproducing organisms at that time, then this ability could have spread with randomly budded vesicles that enclosed the genetic information of peptidoglycan synthesis within them. Viruses, of course, would later develop means to recognize peptidoglycans and introduce their genomes inside the cells. However, until such a transition occurs in viruses, the murein covered bacteria beats their murein-less relatives.

The support needed to maintain high osmotic pressure became relevant in the population where cells were surrounded by peptidoglycan subunits. The only thing the cells required in order to evolve was to produce enzymes which are able to catalyze the bonding of peptidoglycans and then transfer these enzymes outside the cell. Initially this covalent linking of subunits may have also provided additional safety against viruses as it might have prevented virus entry and, on the other hand, the release of virions to the environment (as the cell wall blocked the way). Nonetheless, a cell with a rigid surface was able to begin the bumping of ions outside the cell and thus provide more energy for the ATPases. This energy-efficiency is the next step in favorability for the cell. It should be noted that basic hydrogen ATPases are very ancient enzymes whose emergence occurred much before the formation of cell walls (Jekely 2006). Prior to cell walls the inner pressure in cells might have been much lower or hydrothermal vent compartment walls might have provided the support.

However, cells need to divide in order to reproduce and the cell wall needs to divide along with the cells. The virus induced selective pressure may have given the cells time to come up with solutions in order to produce such mechanism because it was better to maintain even a partly functional cell wall than lose it completely. Interestingly, if the virus pressure truly caused the emergence of cell walls and eventually the cell division apparatus, then it would be the internal pressure of the system (caused by viruses) which originally forced the cells to leave their primordial hydrothermal hatchery. The cell wall provided the cells mean to inhabit various ecological environments. Thus, wherever life emerges in the universe, viruses might play a crucial part in indirectly forcing the cellular organisms to abandon their original environments. In other words, if all parasites are lost early in the evolution of life, it might be less likely for cells to evolve into free living organisms as the emerging cells can safely rely on the benefits of hydrothermal vent-like systems.

### The Origin of the Second Membrane on Gram-Negative Bacteria

Gram-positive bacteria have only one membrane which is surrounded by a cell wall. In contrast to this, Gram-negative bacteria have two membranes with a cell wall in between. There is some evidence suggesting that Gram-positive bacteria are of older origin than their Gram-negative cousins (Gupta 2001; Koch 2003). However, phylogenetic studies also support the reversed order where Gram-negative *Aquifex* was the first to separate from other bacteria (Burggraf et al. 1992; Battistuzzi et al. 2004). Nevertheless, murein surrounded bacteria might have preceded the origin of Gram-negative bacteria. The outer membrane (OM) of Gram-negative bacteria is synthesized by transferring lipids across the inner membrane and delivering them to the OM (Doerfler 2006).

What caused bacteria to obtain the second layer over the cell wall? Similarly to the above, the selective pressure to avoid virus infections could have caused the emergence of the OM. When viruses adapted to recognize and penetrate the cell wall structures, the wall no longer provided protection from infections. As the cellular wall also provided the

structural support needed for a high inner osmotic pressure, viruses learnt to exploit the blow up after cell wall digestion in order to release a vast number of viruses into the environment. This blowing up was not possible before the murein (as there was not enough internal pressure to cause it) and thus the earlier viruses were not that effective cell-killers, only effective resource exploiters. Therefore, in a sense, the initially protective cell wall turned viruses into more hazardous parasites later on.

An additional lipid layer over the murein effectively masked all the receptors that viruses used for attaching to the cell. Therefore the OM freed the first Gram-negative bacteria from viral infections, which in turn provided outstanding selective advantage over the Gram-positive like bacteria. Therefore the OM of the first Gram-negative bacteria did not need to be too effective or very functional as long as it prevented viruses from binding the cell. However, the cell wall digesting enzymes produced by other organisms might also have been a major factor in this process as these enzymes could not reach the wall through the OM. Therefore we do not suggest that viruses alone were the driving force in the formation of OM—they were merely an important factor.

We can argue that if there were no viruses present, then the second membrane of Gram-negative bacteria might be less likely to have emerged. It has been suggested that the organism that initially gave rise to mitochondria in modern eukaryotes was a double membrane alpha-proteobacterium (Vesteg et al. 2006). The OM of this prokaryote possibly became the membrane of all the following eukaryote organisms and thus was a crucial point in the evolution of life on Earth. Whether this endosymbiosis of alpha-proteobacterium could have taken place in a virus-free world is, of course, highly hypothetical, but nevertheless an interesting scenario.

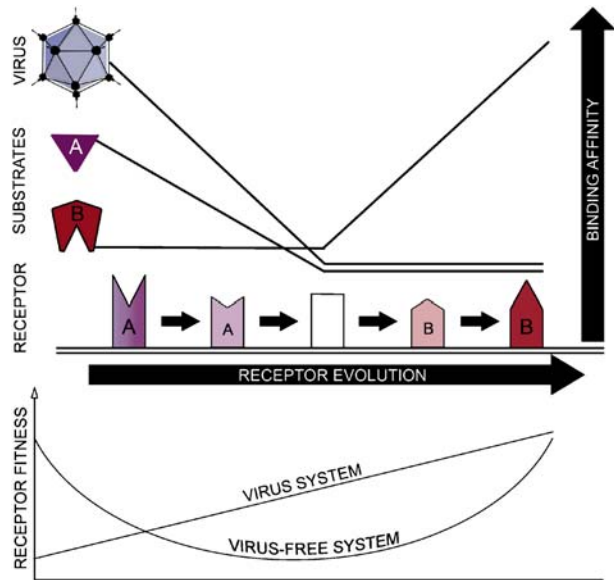
### The Evolution of Single Membrane Receptors

Viruses exploit different structures on the cell surface in recognizing suitable hosts. The evolution of these structures has been studied before since viruses quickly evolve to bind cell receptors which have mutated beyond recognition (Buckling and Rainey 2002). Also computational research has been performed in order to understand the faith bacteria and viruses in an ecosystem (Weitz et al. 2005). Here we suggest how virus driven evolution can alter the structure of a single membrane protein and how viruses may force cellular evolution into another direction.

Viruses bind to cell components which are crucial for the host organisms because binding to irrelevant structures would quickly result in a loss of the element in question from the cell (Tailor et al. 2003). Cells have a desperate need to modify these components in order to lose the chasing viruses. Therefore we often see the host recognition proteins of viruses evolving fastest as natural selection completely favors the viruses that can remain infectious (Buckling and Rainey 2002; Saren et al. 2005). This chase may have created various membrane proteins with different functions.

Let us imagine a membrane receptor-transporter which recognizes and transports a nutrient molecule inside the cell. There is also a virus in the ecosystem which exploits the same protein in attaching itself to the host. If the selective pressure to avoid the viral infection is much higher than the gain of nutrient transporting, this receptor should constantly alter its structure in order to remain unrecognizable. This altering of structure could lead into evolution of different kinds of membrane receptors recognizing totally new nutrients from the environment (or performing other tasks). Figure 2 represents how viruses make the intermediates in the evolution of a receptor favorable when the receptor mutates to recognize a new nutrient. In absence of viruses the receptor would retain its original form.

**Fig. 2** The evolution of a cell surface receptor in a virus containing system. The intermediates in the evolution of a receptor (to recognize alternative substrate) are not favorable in virus-free system as there are no viruses to favor the changes in receptor structure



As the membrane is the interface for prokaryotic cells in communicating with the external world, the emergence of different surface structures can provide means for more complex interactions with the environment. If the viral exploitation of a receptor is too harsh, the original nutrient transport function might be lost. This loss can, however, promote the evolution of internal synthesis of this nutrient or an alternative function to fulfill the task of the nutrient. In absence of viruses, the cells might remain totally dependent on external resources and the favorability needed to develop the machinery for intercellular metabolism might not be reached.

### The Proto-Viruses of the RNA-World, the Evolution of Cell Membrane Components and the Dynamics of Early Life

#### The First Self-Replicating Molecules Might Have Been Surrounded by Membranes

It is possible that life might have emerged inside a hydrothermal vent system (Baross and Hoffman 1985; Maher and Stevenson 1988; Nisbet and Sleep 2001), although this view has also been criticized (Miller and Bada 1988) and other possibilities have been suggested (Wächtershäuser 2003). In the early stages of the hypothesized RNA-world the proto-cells might have been surrounded by abiotically formed lipid membranes (Monnard and Deamer 2002). Simulation studies propose that nucleotides could accumulate inside the hydrothermal vent compartments (Baaske et al. 2007) and into lipid vesicles (Monnard and Deamer 2001), and therefore the very first steps of life may have already been bordered by membranes. However, it is still poorly known how the hypothesized first auto-replicating molecules finally evolved into the last common ancestor of cells. The structural approach used above seemed to provide some feasible answers to this matter and therefore we modeled a scenario to provide possible understanding of the early development of life. It should be noted that the scenario presented here is not dependent on the hydrothermal

vents, but rather on inorganic compartmentalization of biomolecules and on formation of abiotic vesicles or vesicle like structures.

Viruses are selfish organisms that survive only by exploiting the extraneous resources. That is because selfish exploiting is a very good surviving strategy and natural selection favors those exploiting the resources most efficiently and smartly (at least while there are resources available). However, it must be noted that the cell killing capacity of a virus does not always correlate with the virus fitness (Herrera et al. 2007) and persistent virus life strategies are important for virus survival (Villarreal 2005). The modern viruses are thought to have originated at the earliest during the RNA-protein world as all the known viruses have protein capsids enclosing the viral genome (that being either RNA or DNA) inside. At least some RNA and DNA viruses might be related due to similar beta-barrels of certain major capsid proteins (Benson et al. 2004). Also the transition from RNA to DNA can be viral induced change (Forterre 2005). However, virus-like selfish elements should have existed even before the emergence of proteins since the exploitation of resources had to be favorable also then. As a matter of fact, as models for the early stages of life are being built, selfishness should be taken into consideration seriously. All the intermediates of emerging life on the early Earth have to be favorable in some sense (and to something) and selfish benefits are a good way of building logical support for favorability. Because even the simplest functional cell must be a complex system, the selfish elements in the primordial world that are being favored in the Darwinian manner should be more like virus-like exploiters rather than complete, reproducing cells. Similar reasoning has also been suggested before (Koonin and Martin 2005; Koonin et al. 2006). Credible models cannot be made out of long evolutionary jumps where many new genes or attributes have to be formed before the result finally provides benefits for the organisms (whatever the organisms might be). Our suggestion here provides a possible solution which makes the intermediates of primordial life logical.

It has been suggested that since viruses have several unique genes that have originated independently within the ancient gene pool, viruses are older than the LUCA of cells (Koonin et al. 2006). Some of the first RNA elements were virus-like exploiters while others were more cooperative agents. Altruistic elements capable of for example binding nucleotides spread around hydrothermal compartments and slowly more complex and better combinations of selfish actors arose. We explore here the possibility that even the very early auto-replicating molecules were (or became) enclosed inside a lipid vesicle as this supposition makes the understanding of the further evolution of RNA world quite plausible. Whether the LUCA was a membrane bound or membraneless organism is a topic of intense debate as there are credible points to support both views (Koonin and Martin 2005; Jekely 2006).

### The Emergence of Proto-Viruses in the Hydrothermal Vent System

Let us hypothesize that the first self-replicating molecule (SRM) was a RNA molecule (or a set of molecules) enclosed inside an abiotically formed lipid vesicle within the hydrothermal vent. The molecule probably replicates itself to the point where there are no resources left as the initial natural selection favors those who can make copies of themselves. Different theoretical models predict the faith of the SRMs (Scheuring et al. 2003; Fontanari et al. 2006). These models are very mathematical and therefore are not likely to explain why certain new biochemical attributes arise. Our approach here is closer to the one Darwin originally used as we look for the reasons how and why something could and should have evolved. However, this approach does not necessarily interfere with the mathematical models but rather suggests new factors to be introduced into future computations.



The basic idea of this hypothesis is that simple proto-viruses appeared soon after the emergence of the SRMs. The proto-viruses were capable of inducing the budding of a membrane vesicle out of the mother vesicle and thus they quickly spread around the hydrothermal world where a lot of non-living vesicles full of suitable biomolecules were available. The proto-viruses continued to exploit the resources of the hydrothermal world and more efficient proto-cells produced more proto-viruses. Thus they were favored over those proto-cells that produce less or less effective viruses. We suggest that in order to produce more selfish proto-viruses, the viruses actually invented many of the biochemical abilities of proto-cells as these abilities are required in effective metabolism. Any single new function that promoted virus production, be that a way to collect nucleotides from the environment or the invention of a harsh gene regulation system for establishing more efficient use of resources, was beneficial to selfishness.

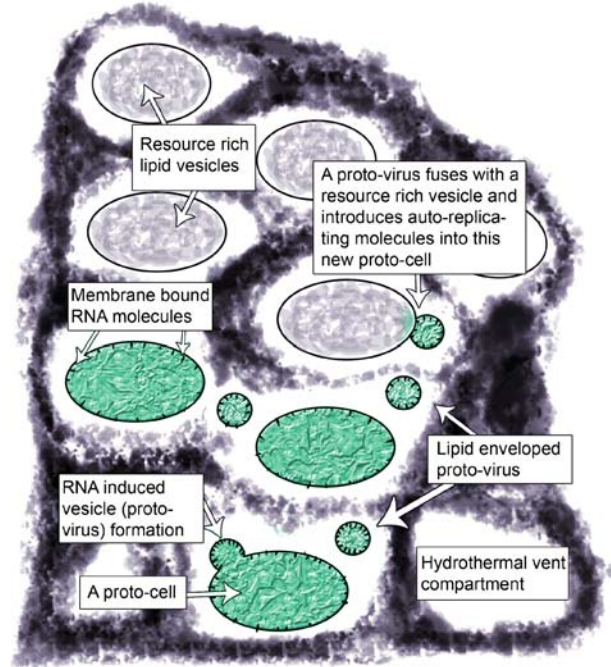
The pathway for the emergence of the first proto-virus is very simple. A hydrothermal vent vesicle full of SRMs might split and move to a new vesicle where unused resources are still available. Thus the split vesicle spreads life around. This membrane fission may occur through a random budding out of the first vesicle and be followed by the fusion with a new vesicle (for review, see Gesteland et al. 2006). However, if the SRM produces an alternative RNA segment which integrates into the membrane and catalyzes the vesicle budding, the vesicle could easily encapsulate some of the SRM inside and subsequently fuse with another vesicle. This way the new proto-cell would reproduce the pudding catalyst and SRMs and thus form new proto-viruses. The catalytically produced vesicles (i.e. proto-viruses) would spread around in the “population” of resourceful hydrothermal vent compartments and, in a sense, infect the “non-living” vesicles with life. Our models differ from the previous ones by assuming that RNA can catalyze the vesicle budding and that budding need not to be an entirely random event or result of physical stress. This assumption leads to the conclusion that randomly budded vesicles would be quickly outnumbered by induced vesicles and thus the genetic material producing induced vesicles would become general.

Interestingly, studies show that RNA is able to bind into membranes with a great variety of different sequences (Janas et al. 2006), which indicates that these RNA and lipid containing proto-viruses could have emerged very early after the formation of the first SRM. A set of genes that can catalyze the budding of the vesicles spread effectively in the system, and therefore these selfish virus-like structures would dominate the hydrothermal world. If you took a sample from one of the vesicles in the system, you would most probably see RNA molecules binding and bending the membrane rather than simply dull auto-replicating molecules. Proto-viruses might easily emerge in the primordial world (after the emergence of the first SRM) as their formation requires only two genes: one for the replication of the genetic material and one for catalyzing the vesicle formation. Thus no evolutionary jumps are needed to explain their origin. The basic idea of the model is represented in Fig. 3.

### Subsequent Evolution After the Emergence of Proto-Viruses

The presented scenario provides several benefits for understanding the consequent stages of evolution within the hydrothermal vent. Firstly, the number of proto-cells (i.e. resourceful vesicles) containing SRMs would be high and therefore the collective evolution in these proto-cells would be much faster than in a smaller population of proto-cells. When all the vesicles contain the first SRM and no initial spreading of life occurs anymore, the proto-viruses still continue their traffic between cells because genes that are able to travel from one cell to another (i.e. form vesicles) stay common in the population. Due to the continued

**Fig. 3** Proto-viruses could spread the life in the hydrothermal compartments by infecting resource rich vesicles. The proto-viruses make the membrane bound RNA favorable, which then again accelerate the evolution of an effective membrane of the first cells



traffic the functional genes spread around and eliminate poor and possibly collapsing systems. While the proto-cells themselves can not divide, it would be the cell's ability to produce more viruses which would favor one proto-cell over another in the manner explained below.

It is important to realize that the proto-viruses do not kill their hosts, and therefore the natural selection of proto-cells does not mean that the cells themselves are dying while others prevail, but that the properties of the cells are replaced by better ones. Proto-viruses (i.e. sets of genes) that are able to exploit the resources of cells more efficiently also multiply more rapidly. However, it should also be noted that in early life most viruses probably were not even harmful factors for the proto-cells but that efficient viruses favored today-like cellular abilities since these abilities made the production of viruses easier. The formation of, e.g., membrane transporters could allow energy molecules and nucleotides to be transferred inside the cell, which in turn would promote the multiplying of the virus. A beneficial function might spread with the virus membranes as occasionally a virus might take the membrane bound transporter along the formation of proto-virus envelope and then transfer the transporter to another cell due to the fusion of membranes. As the hypothetical RNA transporter acts as an enzyme and serves as a template for its own replication, it would therefore transfer the gene along with the function. There is also experimental evidence of RNA serving as a membrane transporter (Janas et al. 2004). Our scenario is compatible with the ideas presented by Carl Woese about the early stages of evolution where the genetic material moved sideways and not from parent to offspring as assumed in the traditional models of Darwinian evolution (Woese 2002).

The whole hydrothermal system would evolve as a community due to proto-viruses. The viruses spread the beneficial functions to new cells after their emergence because the function helps the new cell to produce more viruses. While there are numerous proto-cells where

different functions can originate, the complexity slowly accumulates into individual proto-cells. When a function has spread to every proto-cell, it would be unnecessary for the proto-viruses to carry it along and thus proto-viruses themselves would remain rather simple organisms all the time. Large enough a community of proto-cells might be able to fight off destructive parasites studied in some computational models (Hogeweg and Takeuchi 2003; Fontanari et al. 2006), and since the early proto-cells would be very primitive, the malevolent functions are simply less favorable as they do not help the proto-virus propagation. However, if a very effective epidemic appears, only one proto-cell is needed to overcome the threat because consequently its proto-viruses would spread the cure around. The proto-viruses survive in the system because nothing favors any functions of cells before the cells can completely replicate themselves. Only viruses do not let cells to lose the gained functions as any lost but useful ability will soon be re-established by incoming viruses, which then of course also replaces the genes responsible for virus formation. Thus we suggest that proto-cells were nothing more than virus factories for a long time in the primordial world.

Another contribution of the presented scenario is that it would make the formation of membrane bounded RNA favorable since the very beginning (as RNA molecules catalyze the vesicle budding). The RNA could then rather easily generate mutants which could eventually develop to serve in the multiple functions needed in an effective membrane of the cell. Complex membrane enzymes would be less likely to evolve if there is no reason for the RNA molecules to bind into the membranes in large numbers. As many of the cell-like functions in the RNA world could be explained by their advantages to proto-viruses, we listed some of the hypothetical functions along with their suggested explanations in Table 1.

However, let us take an alternative view to this matter. What if life begun inside the hydrothermal vents, but outside the lipid vesicles? The original SRMs might have formed in the compartments exterior to vesicles as these vesicles might have been rare in the hydrothermal system. SRMs could have replicated vigorously and the more effective SRMs would have outstripped the poorer ones due to basic natural selection. SRMs could develop cooperative actors that might allow for example the synthesis of nucleotides in order to provide more material for selfish RNAs to replicate. Then random diffusion of elements from one hydrothermal vent compartment to another spread the useful abilities around as explained by Koonin and Martin in 2005. This kind of system could evolve into some complexity.

However, if there were lipid vesicles around, then their exploitation would quickly become very useful. A SRM could diffuse into a membrane vesicle and develop a proto-virus inducing gene. These genes would spread as described above. If a vesicle evolved a transporter, it could begin to pump useful nutrients into the vesicle from the external part of the compartment and thus these vesicles could begin to exploit the benefits developed by the membrane less compartments. Since the pump was associated with the membrane, it could travel from the compartment to compartment along with the induced vesicles. Every now and then a beneficial gene would diffuse inside a vesicle and thus improve its capability. It is difficult to see how the external world of non-vesicle bound life could fight back the evolving vesicles. Transporters that move the resources out of the lipid vesicles into the exterior environment could help a little, but then again these bumps would be attached to the vesicles and would not spread so easily (and it would not be favorable for proto-viruses to carry it around). We suggest that eventually membrane boundaries would surpass the purely external agents due to the benefits of membrane boundaries. The membrane less origin for life is possible and even likely, but further evolution of primordial world probably occurred inside vesicles if these vesicles just were available.

Using selfish abusers in the models may close the gap which currently exists between pre-cellular life and Darwin's model for natural selection, since Darwin's model assumes

**Table 1** Possible cell functions in the primordial RNA or RNA-protein world: hypothetical reasons for their emergence and distribution

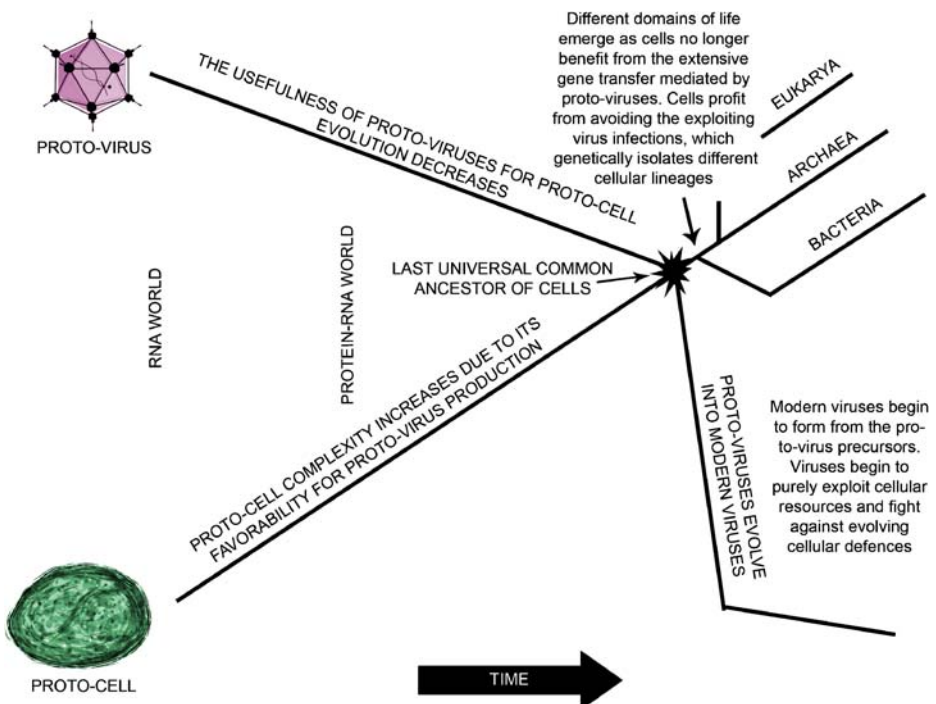
Cell function/attribute	Benefits for proto-viruses	Why did it spread in the population?	How it might have emerged
Membrane transporter	More resources for selfish exploiters	Allows more viruses to be built, thus it is favorable for a virus to spread the attribute in the proto-cell population	Membrane bound RNA molecules that caused the proto-virus formation may have evolved a transporter like structure
Ancient cell division apparatus: the genome division	Ensures genome inclusion within the budded vesicle	The proto-viruses which are able to ensure their genome inclusion inside the vesicle infect new proto-cells successfully and therefore allow new cycles of virus production to occur	Membrane integrated structures recognize and bind the genetic material
Ancient cell division apparatus: the membrane division	Membrane bending catalysts are critical for the formation of proto-viruses as proto-viruses are made of small vesicles	Proto-viruses spread around the population of membraned vesicles and therefore quickly became the dominates of the inanimate world	One of the very first replicated RNAs bound to membrane and catalyzed its bending
Gene regulating systems	Allows the more optimized use of resources	Optimizes the replication rate of catalysts in order to produce more viruses	Certain genes coded additional siRNA like segments that turned off unwanted products
Early energy metabolism apparatuses	Allows the systematic use of energy storage molecules to perform operations	The harness of chemical energy allows viruses to perform operations that would otherwise be impossible. This, in turn, can provide various advantages	Membrane bound structures modified to catalyze the formation of ATP from ADP
Protein building machinery	Even simplest proteins may be highly effective in bending membranes, which effectively boosts the formation of enveloped viruses	Proteins are ultimate enzymes compared to their RNA equivalents and therefore protein viruses are superior in abusing the resources and performing effective infections. The first protein might have been a virus capsid protein like those of today which boosted the efficiency of the virus and therefore spread rapidly within the RNA-world	Ribosome like RNA molecule begun to catalyze the peptide-bond formation based on a RNA template

the existence of units that can have direct offspring, which proto-cells were not capable of. Anything that was beneficent to selfish abusers was favored in the natural selection of early life and it might be profitable for the ancient selfish exploiters to use the proto-cells as their “platforms” of proto-virus production.

## The Origin of Modern Viruses

We assume the existence of a hydrothermal system in which a proto-cell community evolves due to the selfishness of proto-viruses as described above. At some point a proto-cell comes up with a way to synthesize a protein (Wolf and Koonin 2007). This protein is favored by its unbeatable competence to perform enzymatic tasks and therefore proto-viruses spread this information around the proto-cell community. Slowly, numerous RNA genes which encode beneficial proteins accumulate into proto-cells. Eventually the proteins are able to perform highly sophisticated metabolic tasks and it is no longer favorable for proto-viruses to spread metabolic genes around. Cells become highly independent: they can collect energy and biomolecules from the exterior world and synthesize some needed elements out of various precursors. At this point viruses generally turn from beneficial organisms into counterproductive parasites because now they can start to just exploit cell resources and the spreading of profitable genes no longer helps their replication. The transition of the virus nature and the origin of modern viruses are presented in Fig. 4.

Different virus capsid proteins emerged quickly after the emergence of proteins since proteins can make up effective viruses (as we can see today). The first protein viruses spread around the proto-cell community and delivered beneficial functions around. During this period it was not favorable for proto-cells to develop any countermeasures against virus infections as viruses mainly boosted the cellular complexity. Viruses were also free to develop their genome packaging machinery as those viruses that were able to carry the good genes around were favored over those that could not. This is probably the reason why



**Fig. 4** The transition in the nature of viruses: the origin of modern viruses and major domains of life

the conserved innate viral self (Bamford et al. 2002; Krupovic and Bamford 2007) of viruses contains only the capsid and genome packaging genes.

The numerous originally formed capsids and genome packaging machineries became the foundation for the modern virus structures when the cells no longer benefited from virus infections. This makes the origin of modern viruses very understandable because the most crucial components of modern viruses were initially favorable for the evolution of both cells (even thought not in direct selection but through the benefits for viruses) and viruses, and therefore complete viruses did not need to be formed in a single evolutionary event. For the same reason we can also see a wide variety of different viruses infecting cells today as these beneficial organisms originated multiple times in history. It is hard to understand why rather complex parasites would originate several times if some intermediates in their formation were not useful to anyone. Similarly, all higher parasites (made out of cells) seem to have been initially free living organisms and only afterwards evolved to the parasitic life style.

At first most viruses had a membrane envelope surrounding their capsid, but when cells developed cell walls (as described above), viruses required new ways of penetrating into cytoplasm. Viruses came up with numerous ways of battling against the defenses of cells and thus the host recognition and genome delivery apparatuses of homologous viruses (which had a common ancestor from the beneficial times) are totally unlike. The virus core structures (the capsids and genome packaging apparatuses) are noted to be older than the LUCA of cells, but other parts of viruses show less deep connections (Hendrix et al. 1999; Benson et al. 2004; Khayat et al. 2005; Krupovic and Bamford 2007; Akita et al. 2007). This fits perfectly into our scenario as the virus cores originated at the time when viruses were still improving the cellular complexity. The other parts of viruses developed only after cells became more or less independent organisms because it then became favorable for them to start avoiding viral infections. We suggest that this transition, where viruses generally turned from beneficial agents into counterproductive exploiters, was the point when the LUCA of cells existed. This is because the amount of horizontal gene transfer decreased and cells begun to isolate as an indirect result of evolved countermeasures against virus infections, thus cellular lineages (the precursors of bacteria and archaea) were able to separate.

## Conclusions

Here we have explored the role of viruses in the evolution and origin of life. First we plunged into the emergence of cell walls, second membranes and the evolution of membrane proteins. Then we treated the proto-viruses of RNA-world as a way of explaining the emergence of actual cells and modern viruses.

The constant siege of viruses might have been the selective force in causing the formation of the first cell walls because the walls kept viruses outside the cells. This event might have eventually freed cells from their hydrothermal vent hatchery to move inhabit different environments of Earth. Afterwards viruses evolved to breach the wall and bacteria formed a second membrane to mask the virus exploited sites. Also the membrane proteins might have evolved a higher diversity due to the selective pressure induced by viruses. Sometimes it might have been safer for cells to evolve synthetic pathways to form some nutrients than acquire them from the environment as the receptors for recognition and transportation of nutrients are often exploited by viruses.

On the ancient Earth even the very first replicating molecules might have been bordered by abiotically formed membranes. If a possible primordial hydrothermal vent was full of resource-rich vesicles, it is probable that the auto-replicating molecules evolved ways of inducing vesicle

budding since budding would have allowed them to spread faster in the vent compartments (and thus quickly dominate the system over the randomly budding vesicles). Due to the abusive nature of these proto-viruses, many beneficial attributes arose in the proto-cells (i.e. in the major vesicles) as these attributes made the harsher exploitation of resources possible. The proto-viruses delivered these features to new compartments and slowly the overall complexity of proto-cells increased. We therefore suggested that proto-viruses actually favored the invention of the majority of cellular abilities in the primitive world as it allowed proto-cells to produce more viruses. In the end proto-cells became highly independent organisms and viruses just began to exploit cellular resources. At this point cells began to fight off the viral infections and viruses transformed into counterproductive agents. This was possibly the phase when the LUCA of cells existed and modern viruses began to emerge.

Testing the ideas and hypotheses presented here might be possible in the near future due to the increased capability of computers and virtual simulations. When the first auto-replicating molecules are artificially produced, the direct examination of the evolution of the hydrothermal vent-like system could be studied. Our model could be tested by recreating a system with resourceful vesicles and introducing an auto-replicating molecule inside one. Before such progress is possible, computational models could explore the effects of induced vesicle formation and lifeless, resourceful vesicles on information threshold and theoretical evolution. Studying whether RNA can bend lipid membranes is also crucial for the model presented here. Better general understanding of the conditions and dynamics of early Earth, naturally, provides us boundaries in limiting the scenarios and thus basic research on the different elements in the origin of life is critically important.

Viruses are sometimes neglected in evolutionary models due to their non-living nature. Yet, the question whether viruses are either living or inanimate organisms is only about schematics and relying on such arguments is not fruitful in producing true comprehension to biology. Viruses might have had a tremendous effect on the life on Earth, thus situating the viewpoint to structural co-evolution between cells and viruses in order to provide insights to the scientific understanding of nature seemed lucrative. Philosophically it is important to realize the driving forces of evolution as it unravels the needless mysteries of the series of instances that led into the present world. In the absence of reason, many events in the history of life might seem impossible and miraculous to an ignorant spectator while in reality they are only logical consequences if being supported by the right factors.

**Acknowledgements** This study was supported by the Finnish Centre of Excellence Program of Academy of Finland (1213467, 2006-2011).

## References

- Akita F, Chong KT, Tanaka H et al (2007) The crystal structure of a virus-like particle from the hyperthermophilic archaeon *pyrococcus furiosus* provides insight into the evolution of viruses. *J Mol Biol* 368(5):1469–1483
- Anderson JS, Matsuhashi M, Haskin MA et al (1965) Lipid-phosphoacetyl-muramyl-pentapeptide and lipid-phosphodisaccharide-pentapeptide: presumed membrane transport intermediates in cell wall synthesis. *Proc Natl Acad Sci USA* 53:881–889
- Baaske P, Weinert FM, Duhr S et al (2007) Extreme accumulation of nucleotides in simulated hydrothermal pore systems. *Proc Natl Acad Sci USA* 104:9346–9351
- Bamford DH (2003) Do viruses form lineages across different domains of life? *Res Microbiol* 154:231–236
- Bamford DH, Burnett RM, Stuart DI (2002) Evolution of viral structure. *Theor Popul Biol* 61:461–470

- Baross JA, Hoffman SE (1985) Submarine hydrothermal vents and associated gradient environments as sites for the origin and evolution of life. *Orig Life* 15(4):327–345
- Battistuzzi FU, Feijao A, Hedges SB (2004) A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol Biol* 4:44
- Benson SD, Bamford JK, Bamford DH et al (1999) Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. *Cell* 98:825–833
- Benson SD, Bamford JK, Bamford DH et al (2004) Does common architecture reveal a viral lineage spanning all three domains of life? *Mol Cell* 16:673–685
- Bergh O, Borsheim KY, Bratbak G et al (1989) High abundance of viruses found in aquatic environments. *Nature* 340:467–468
- Boucher Y, Kamekura M, Doolittle WF (2004) Origins and evolution of isoprenoid lipid biosynthesis in archaea. *Mol Microbiol* 52:515–527
- Brown JR, Doolittle WF (1997) Archaea and the prokaryote-to-eukaryote transition. *Microbiol Mol Biol Rev* 61:456–502
- Brussow H, Canchaya C, Hardt WD (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* 68:560–602
- Buckling A, Rainey PB (2002) Antagonistic coevolution between a bacterium and a bacteriophage. *Proc Biol Sci* 269:931–936
- Burggraf S, Olsen GJ, Stetter KO et al (1992) A phylogenetic analysis of *Aquifex pyrophilus*. *Syst Appl Microbiol* 15:352–356
- Canchaya C, Fournous G, Chibani-Chennoufi S et al (2003) Phage as agents of lateral gene transfer. *Curr Opin Microbiol* 6:417–424
- Comeau AM, Krusch HM (2005) War is peace—dispatches from the bacterial and phage killing fields. *Curr Opin Microbiol* 8:488–494
- Doerfler W (1975) Integration of viral DNA into the host genome. *Curr Top Microbiol Immunol* 71:1–78
- Doerrler WT (2006) Lipid trafficking to the outer membrane of gram-negative bacteria. *Mol Microbiol* 60:542–552
- Dworkin JP, Lazcano A, Miller SL (2003) The roads to and from the RNA world. *J Theor Biol* 222:127–134
- Fontanari JF, Santos M, Szathmari E (2006) Coexistence and error propagation in pre-biotic vesicle models: a group selection approach. *J Theor Biol* 239:247–256
- Forterre P (2005) The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie* 87:793–803
- Gesteland RF, Cech TR, Atkins FJ (2006) The RNA world, 3rd edn. Cold Spring Harbor Laboratory Press, New York, USA
- Gupta RS (2001) The branching order and phylogenetic placement of species from completed bacterial genomes, based on conserved indels found in various proteins. *Int Microbiol* 4:187–202
- Hendrix RW, Smith MC, Burns RN et al (1999) Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc Natl Acad Sci USA* 96:2192–2197
- Herrera M, Garcia-Arriaza J, Pariente N et al (2007) Molecular basis for a lack of correlation between viral fitness and cell killing capacity. *PLoS Pathog* 3:e53
- Higashi Y, Strominger JL, Sweeley CC (1967) Structure of a lipid intermediate in cell wall peptidoglycan synthesis: a derivative of a C55 isoprenoid alcohol. *Proc Natl Acad Sci USA* 57:1878–1884
- Hogeweg P, Takeuchi N (2003) Multilevel selection in models of prebiotic evolution: Compartments and spatial self-organization. *Orig Life Evol Biosph* 33:375–403
- Holzel F, Sokol F (1974) Integration of progeny simian virus 40 DNA into the host cell genome. *J Mol Biol* 84:423–444
- Iyer LM, Balaji S, Koonin EV et al (2006) Evolutionary genomics of nucleocytoplasmic large DNA viruses. *Virus Res* 117:156–184
- Janas T, Janas T, Yarus M (2004) A membrane transporter for tryptophan composed of RNA. *RNA* 10:1541–1549
- Janas T, Janas T, Yarus M (2006) Specific RNA binding to ordered phospholipid bilayers. *Nucleic Acids Res* 34:2128–2136
- Jekely G (2006) Did the last common ancestor have a biological membrane? *Biol Direct* 1:35
- Khayat R, Tang L, Larson ET et al (2005) Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *Proc Natl Acad Sci USA* 102:18944–18949
- Kidambi SP, Ripp S, Miller RV (1994) Evidence for phage-mediated gene transfer among *Pseudomonas aeruginosa* strains on the phyloplane. *Appl Environ Microbiol* 60:496–500
- Koch AL (1994) Development and diversification of the last universal ancestor. *J Theor Biol* 168:269–280
- Koch AL (1998) How did bacteria come to be? *Adv Microb Physiol* 40:353–399
- Koch AL (2003) Were gram-positive rods the first bacteria? *Trends Microbiol* 11:166–170



- Koch AL (2006) The exocytoskeleton. *J Mol Microbiol Biotechnol* 11:115–125
- Koonin EV, Martin W (2005) On the origin of genomes and cells within inorganic compartments. *Trends Genet* 21:647–654
- Koonin EV, Senkevich TG, Dolja VV (2006) The ancient virus world and evolution of cells. *Biol Direct* 1:29
- Krupovic M, Bamford DH (2007) Putative prophages related to lytic tailless marine dsDNA phage PM2 are widespread in the genomes of aquatic bacteria. *BMC Genomics* 8:236
- Maher KA, Stevenson DJ (1988) Impact frustration of the origin of life. *Nature* 331:612–614
- Miller SL, Bada JL (1988) Submarine hot springs and the origin of life. *Nature* 334:609–611
- Monnard PA, Deamer DW (2001) Nutrient uptake by protocells: a liposome model system. *Orig Life Evol Biosph* 31:147–155
- Monnard PA, Deamer DW (2002) Membrane self-assembly processes: steps toward the first cellular life. *Anat Rec* 268:196–207
- Nisbet EG, Sleep NH (2001) The habitat and nature of early life. *Nature* 409:1083–1091
- Orgel LE (2004) Prebiotic chemistry and the origin of the RNA world. *Crit Rev Biochem Mol Biol* 39:99–123
- Poole AM, Penny D (2007) Evaluating hypotheses for the origin of eukaryotes. *Bioessays* 29:74–84
- Saren AM, Ravanti JJ, Benson SD et al (2005) A snapshot of viral evolution from genome analysis of the tectiviridae family. *J Mol Biol* 350:427–440
- Scheuring I, Czaran T, Szabo P et al (2003) Spatial models of prebiotic evolution: soup before pizza? *Orig Life Evol Biosph* 33:319–355
- Tailor CS, Lavillette D, Marin M et al (2003) Cell surface receptors for gammaretroviruses. *Curr Top Microbiol Immunol* 281:29–106
- Vesteg M, Krajcovic J, Ebringer L (2006) On the origin of eukaryotic cells and their endomembranes. *Riv Biol* 99:499–519
- Weinbauer MG (2004) Ecology of prokaryotic viruses. *FEMS Microbiol Rev* 28:127–181
- Weinbauer MG, Rassoulzadegan F (2004) Are viruses driving microbial diversification and diversity? *Environ Microbiol* 6:1–11
- Weitz JS, Hartman H, Levin SA (2005) Coevolutionary arms races between bacteria and bacteriophage. *Proc Natl Acad Sci USA* 102:9535–9540
- Villarreal LP (2005) Viruses and the evolution of life. ASM Press, Washington, DC
- Witzany G (2006) Natural genome-editing competences of viruses. *Acta Biotheor* 54:235–253
- Woese CR (2002) On the evolution of cells. *Proc Natl Acad Sci USA* 99:8742–8747
- Wolf YI, Koonin EV (2007) On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. *Biol Direct* 2:14
- Wommack KE, Hill RT, Kessel M et al (1992) Distribution of viruses in the Chesapeake Bay. *Appl Environ Microbiol* 58:2965–2970
- Wächtershäuser G (2003) From pre-cells to Eukarya—a tale of two lipids. *Mol Microbiol* 47:13–22
- Young R (1992) Bacteriophage lysis: mechanism and regulation. *Microbiol Rev* 56:430–481
- Zorzopulos J (2003) Birth of the domains bacteria, archaea and eucarya and of major taxa within them: a hypothesis. *Rev Argent Microbiol* 35:175–182