An evolutionary perspective on signaling peptides: toxic peptides are selected to provide information regarding the processing of the propeptide, which represents the phenotypic state of the signaling cell [version 1; referees: 2 approved]

Keith Daniel Harris¹, Ari Barzilai²,³, Amotz Zahavi¹,³

¹Department of Zoology, Tel-Aviv University, Tel Aviv, 69978, Israel
²Department of Neurobiology, Tel Aviv University, Tel Aviv, 69978, Israel
³Sagol School of Neuroscience, Tel Aviv, 69978, Israel

Abstract

Structurally similar short peptides often serve as signals in diverse signaling systems. Similar peptides affect diverse physiological pathways in different species or even within the same organism. Assuming that signals provide information, and that this information is tested by the structure of the signal, it is curious that highly similar signaling peptides appear to provide information relevant to very different metabolic processes. Here we suggest a solution to this problem: the synthesis of the propeptide, and its post-translational modifications that are required for its cleavage and the production of the mature peptide, provide information on the phenotypic state of the signaling cell. The mature peptide, due to its chemical properties which render it harmful, serves as a stimulant that forces cells to respond to this information. To support this suggestion, we present cases of signaling peptides in which the sequence and structure of the mature peptide is similar yet provides diverse information. The sequence of the propeptide and its post-translational modifications, which represent the phenotypic state of the signaling cell, determine the quantity and specificity of the information. We also speculate on the evolution of signaling peptides. We hope that this perspective will encourage researchers to reevaluate pathological conditions in which the synthesis of the mature peptide is abnormal.

Corresponding author: Keith Daniel Harris (keithhar@post.tau.ac.il)

How to cite this article: Harris KD, Barzilai A and Zahavi A. An evolutionary perspective on signaling peptides: toxic peptides are selected to provide information regarding the processing of the propeptide, which represents the phenotypic state of the signaling cell [version 1; referees: 2 approved] F1000Research 2015, 4:512 (doi: 10.12688/f1000research.6874.1)

Copyright: © 2015 Harris KD et al. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Competing interests: No competing interests were disclosed.

Introduction
Signaling peptides are amino acid chains with diverse structures that serve as signaling molecules. The lengths of signaling peptides vary greatly from less than ten amino acids (such as oxytocin and vasopressin) to over 100 amino acids (such as the neurotrophic factors). The mature signaling peptide which is secreted is processed from a longer propeptide which contains other domains (prodomains) which are not part of the mature form (Figure 1).

Mature peptides of similar structure may function as a neurotransmitter, an endocrine or a paracrine signal within a multicellular organism, and also as a signal between unicellular organisms. For instance, the gonadotropin-releasing hormone (GnRH), which has a significant structural similarity to the yeast mating factor-alpha peptide, serves as both a hormone in mammals and as a mating pheromone in yeast. It also serves as a paracrine signal in the periphery of the multicellular organism.

Thus, though the structure of the mature signal of signaling peptides such as GnRH is conserved, its specific signaling role is not, and their prodomains differ markedly. Moreover, in the same organism, structurally similar signaling peptides may regulate a diverse range of signaling pathways, such as the structurally similar oxytocin and vasopressin, which also function as signals in unicellular organisms.

Assuming that signals elicit a response because they provide specific information that benefits the organism, how may similar peptides provide information regarding such diverse metabolic processes?

While the mature peptides of GnRH, oxytocin and vasopressin are short (9-10 amino acids), their propeptides are large proteins (100-160 amino acids). The cleavage of the propeptide to form the comparatively short mature peptide is often dependent on the completion of post-translational modifications, such as sequential enzymatic modification, glycosylation, glycosulfation or the pairing of S-S bonds.

As the cleavage of the mature peptide depends on the propeptide completing its various post-translational modifications, and as there is a fixed stochastic relationship between the mature peptide and the propeptide that is determined by the number of repeats of the mature domain within the propeptide, it is reasonable to assume that the ability of the signaling cell to complete the synthesis of the propeptide is the information provided by the mature peptide.

We suggest that while the synthesis and modifications of the propeptide are related to the phenotypic state of the signaling cell, the role of the mature peptide is to stimulate cells to be attentive to this information. In this opinion paper we briefly review a number of signaling peptides to support our suggestion and, in addition, speculate why and how a mature peptide is selected to serve as a stimulating molecule.

Different propeptides produce similar mature peptides
Similar signaling peptides are used in different species to affect diverse metabolic processes; however, in many cases these similar peptides are processed from entirely different propeptides. Such is the case of the 10-amino acid GnRH, a hormone produced by the hypothalamus and also by cells in the periphery in vertebrates, which is structurally similar to the mature peptide of the yeast mating-alpha factor. Also within the yeast genome, a similar mature peptide (mating-alpha factor) is produced by two different propeptides encoded by the genes MFAL-1 and MFAL-2. MFAL-1 has four tandem repeats of the mature domain, while MFAL-2 contains two repeats of the mature domain with a slight variation in sequence (Figure 2).
Even in the same species, the structurally similar forms of the GnRH peptide are connected to different prodomains. There are two genes in humans that encode the sequence of GnRH: *GNRH-I* and *GNRH-II*. Each gene has a highly similar mature peptide, but a different prodomain (Figure 3). The synthesis of the mature form (GnRH) requires an intricate series of enzymatic modifications.

Oxytocin and vasopressin are structurally similar peptides of nine amino acids. The mature peptides of oxytocin and vasopressin are highly similar, and they share an accessory protein, neurophysin, yet the other domains have little similarity in sequence (Figure 4).

The mature peptides of oxytocin and vasopressin are also active in unicellular organisms. When comparing oxytocin and vasopressin propeptides across the phylogenetic tree, it is evident that the mature domain is more conserved than the prodomain.

In addition, within the same species, within a conserved family of signal peptides, such as the neurotrophic factors BDNF, NT-3 and NGF, the variation of the sequence of the mature peptide between the different peptides is significantly lower than the variation among the prodomains of the respective propeptides (Figure 5). The variation in glycosylation sites is depicted schematically in Figure 6.
This variation is also evolutionarily conserved: the sequence of the prodomain is unique while the sequence of the mature peptide is common to the family.

The phenotypic state of the signaling cell affects the synthesis of the mature peptide
This variation is not only genetic, but may also be phenotypic. In the case of BDNF, there are several alternative transcripts that produce isoforms of BDNF. These isoforms differ in their prodomain but not in their mature domain, and this variation depends on metabolism and physiological parameters of the signaling cell. The production of specific isoforms of BDNF has been correlated to various pathological conditions in which the synthesis of BDNF is altered.

BDNF, NGF and NT-3 each have a conserved N-glycosylation site in the prodomain that is proximal to the processing site at which the propeptide is cleaved to form the mature peptide. When N-glycosylation is blocked, cleavage of the propeptide is affected, and less of the mature form is synthesized. What accumulate in the Golgi are truncated forms of proBDNF. The secreted mature form of BDNF is therefore a representation of the properly processed and cleaved proBDNF.

Propeptide non-signaling functions
We know of two cases in which a whole functional protein that has a non-signaling role serves as the propeptide of a mature peptide signal. One of them is the sex pheromone system of Enterococcus faecalis, the other the extracellular death factor (EDF) of Escherichia coli.

As far as we are aware, there is no known non-signaling function for the propeptides of the neurotrophic factors. Likewise is the case for oxytocin and vasopressin, and also GnRH. However, cases in which a non-signaling function is known may illustrate how peptide signaling systems evolved.

Enterococcus faecalis is a bacterium that has a sophisticated mechanism of plasmid transfer governed by signaling peptides of 7-8 amino acids in length. These peptides are produced from specific membranal proteins that perform non-signaling functions in the cell that are unrelated to the plasmid (Table 1). The mature peptide of the Enterococcus faecalis pheromones is part of the sequence of the propeptide which anchors it to the membrane. The cleavage of the propeptide from the membrane releases the mature peptide (the pheromone), which provides via a complex transduction mechanism reliable information that the signaling cell does not possess the plasmid.

Another example of a signaling peptide whose propeptide serves a non-signaling function in the cell is the extracellular death factor (EDF) pentapeptide that activates the mazEF pathway in Escherichia coli. The sequence of the pentapeptide is NNWNN, and it is synthesized by the proteolytic cleavage of the enzyme glucose-6-phosphate dehydrogenase, which includes the sequence NNWDN. The mature peptide requires the modification of the aspartic acid in the propeptide to asparagine.

As the propeptide of the pentapeptide is a functional protein, the relationship between its synthesis and degradation links the production of the pentapeptide directly to cell metabolism, and specifically the metabolism of glucose. Since the pentapeptide appears in the Escherichia coli genome only within this enzyme, and is an essential component of the enzyme, its secretion from Escherichia coli is reliable information that the signaling bacterium does not need the enzyme for its current metabolism.

Discussion
The main purpose of this article is to propose a solution to the problem we faced when trying to understand how very similar short peptides may provide information that is relevant to receiver cells designed to serve very different roles. Mature peptides are often conserved across the phylogenetic tree, from unicellular organisms to mammals. Hence, it is tempting to attempt to identify what properties of these mature peptides cause them to be adapted for their role. Our perspective is derived from the assumption that signals provide reliable information regarding the behavior of the signaling cell.

We suggest that mature peptides were selected as optimal carriers for transferring information due to their ability to stimulate

<table>
<thead>
<tr>
<th>Plasmid function</th>
<th>Pheromone primary structure</th>
<th>Propeptide function</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAD1 – Haemolysin/bacteriocin and UV resistance</td>
<td>LFSVLVLAG</td>
<td>Membrane immunogen with FMN-binding domain*</td>
</tr>
<tr>
<td>pCF10 – Tetracycline resistance</td>
<td>LVTLVFV</td>
<td>YidC membrane insertase</td>
</tr>
<tr>
<td>pPD1 – Bacteriocin</td>
<td>FLVMFLSG</td>
<td>YidC membrane insertase</td>
</tr>
<tr>
<td>pAM373</td>
<td>AIFILAS</td>
<td>Peptidase</td>
</tr>
<tr>
<td>pOB1 – haemolysin/bacteriocin</td>
<td>VAVLVLGA</td>
<td>ABC Methionine transporter substrate-binding protein</td>
</tr>
</tbody>
</table>

* Based on similarity
the receiving cell to attend to the information they represent. We speculate that their advantage as stimulating agents is due to their harmful effects which force the receiver to attend to the information (Figure 7).

Predictions

Our suggestion that the mature peptide does not directly represent information regarding the signaling cell’s metabolism, but reflects information regarding the synthesis of the prodomain, can be tested. Here we suggest experiments that may test this suggestion:

1. Replacing the mature peptide of one propeptide with a similar mature peptide from the same family of peptides (such as oxytocin with vasopressin), may reveal that the mature peptide has the same effect due to its specificity for its receptor; however, we expect that the timing and quantity of the effect will be determined by the protein to which it is attached.

2. The harmful effects of the mature peptides may be assessed by increasing their quantities beyond the ability of the receiving cells to counter their noxious effects to which they are usually exposed or by removing any other mechanisms that under normal circumstances counter these harmful effects, such as enzymes that degrade the mature peptide or bind it.

Some speculations on the stages of the evolution of signal peptides

It is tempting to speculate how the sequence of mature peptides evolved, even though at present we have succeeded in collecting limited data to support this speculation.

It is reasonable to assume that the first generation of mature peptides were part of large proteins that had non-signaling functions in the cell. During the proteolysis of these large functional proteins, short peptides were secreted. Among the short peptides cleaved and secreted from cells, the peptides that harmed neighboring cells selected, in neighboring cells, for mechanisms that counter the harmful effects of the peptides. The level of the response to the harmful effects is correlated to the level of the secreted peptide and, hence, could be used by the neighboring cells as a source of information regarding the metabolism of the proteins from which they were cleaved in the signaling cells. These large proteins still serve their initial non-signaling function, yet they also serve as propeptides. These are the cases mentioned previously of the EDF pentapeptide of *E. coli* which is a functional element of glucose-6-phosphate dehydrogenase[16] and the enterococcus pheromones, which anchor the propeptide to the membrane[17].

We suggest that a signal which harms the receiver will force the receiver to respond to a smaller change in its concentration than a signal that provides a positive or neutral effect. Zahavi[19] and Zahavi & Zahavi[5] found this to be the case for many signals used by birds and humans. In addition, Harris *et al.*[20] pointed out that several non-peptide signals such as glutamate and dopamine may cause harm to cells that do not counter their harmful effects in relation to their level of release. They suggested that the toxicity ensures that the response of the receiver cell is correlated to the level of the release from the signaling cell.

We also suggest that in the case of signaling peptides, the toxicity of the mature peptide ensures that the response of the receiving cell is correlated to the concentration of the secreted peptide. At present

![Figure 7](image-url)
we are aware of only one case in which a peptide has a known direct toxicity that is crucial to its function, the EDF pentapeptide. The pentapeptide kills *E. coli* by binding to the mazF toxin and interfering with the ability of the mazE antitoxin protein to inhibit the activity of the mazF toxin\(^1\).

Once the sequences of mature peptides evolved as carriers of information due to their ability to stimulate cells to attend to information, natural selection could transpose a mature peptide to be a part of other proteins whose synthesis reflected information regarding the phenotypic state of the signaling cell, granted that this information benefitted the organism through its effect on the receiving cells. Small modifications in the mature peptide of the first generation were required to prevent the binding of the new mature peptide to the receptors of the first generation peptides, if the first generation and the second generation mature peptide were to function in the same organism (Figure 8).

The third generation of mature peptides is secreted from central glands and from the brain to activate peripheral cells. Wessler *et al.*\(^{21}\) suggested that acetylcholine, which serves as a paracrine signal to coordinate activities between epithelial cells in the airpipe, evolved first as a paracrine signal and was only later adopted by the brain to activate peripheral cells that already respond to acetylcholine. Following this suggestion, we recently proposed a general model for the evolution of non-peptide signals\(^{20}\). It is reasonable to assume that signals which function between cells in the periphery, and are also used by the brain and central glands for their similar effect on peripheral cells, evolved first as paracrine signals in the periphery\(^{18,21}\) (i.e., neurons and endocrine cells adopted the mature peptides that served as signals in the periphery to activate cells that were already adapted to respond to them).

In several cases it is known that the mature peptide by which the brain stimulates peripheral cells is also synthesized in small quantities by the cells that respond to it. Oxytocin and GnRH for example are synthesized in peripheral cells that respond to the same peptides that are secreted in the brain. The present function of the synthesis of the small amounts of mature peptides in peripheral cells is unclear.

It is reasonable to assume that the function of the propeptides in the brain is not to provide information relating to the metabolism of the secreting neuron, but to allow the synthesis of the mature peptide in large quantities to regulate and synchronize the activities of various organs to respond to decisions made in the brain. Hence, we expect that the structure of the propeptides may differ between the brain and peripheral cells, as each serves a different need: neurons need to synthesize large quantities of the mature peptide, while in the periphery the propeptide reflects the phenotypic state of the signaling cell.

**How our perspective changes the focus of treatments for pathological conditions in which the mature peptide is lacking**

Many neuropathologies, including amyotrophic lateral sclerosis, Alzheimer’s disease and Parkinson’s disease, have been associated with a reduction in the synthesis of neurotrophic factors\(^{22-24}\). Current treatments of these diseases often involve administering synthesized mature peptides (the neurotrophic factors). However, if the inability to process the propeptide correctly (such as the inability

---

**Figure 8. Mutations may transfer mature toxic peptides to other functional proteins and evolve new signals.**

A A function protein (Protein A) contains a toxic signaling peptide (M) which has a complementary receptor (Receptor A) B A mutation causes the toxic signaling peptide M to be transferred into a different functional protein (Protein B), so that both proteins produce the same toxic peptide that interacts with Receptor A. The toxic peptide M no longer provides accurate information regarding the processing of Protein A. C A mutation in the toxic peptide M in Protein B (yielding the toxic peptide M*) prevents its binding to the same receptor as M, and a receptor specific to M* (Receptor A*) evolves to counter its toxicity, which may provide information on the processing of Protein B.
to perform the glycosylations) reflects reliably the phenotypic state of the signaling cell, then these treatments provide false information to the neuron. The false information may mask the underlying problem, which may be the deterioration of the signaling cell’s ability to serve its function. If the mature peptide provides retrograde information to the neuron on the ability of the peripheral cell to be activated by the neuron, then supplementing it may mislead the organism that the system is functioning correctly even though it is not. In addition, if, as we suggest, the mature peptide functions due to its harmful effects, then providing an abnormal concentration of it to the receiver cells may cause unnecessary damage to the signaling system. Hence, the attempts to counter the disease should focus on attempting to improve the phenotypic state of the signaling cells, for instance, by providing anti-oxidants.

Methods
All sequences were taken from the UniProt database. Alignments were calculated using Clustal Omega implemented in UniProt. Jalview v2.8.2 was used for the graphical representation of the alignments.

References

Author contributions
Keith D. Harris collected information from the literature, conducted the bioinformatic work and co-wrote the manuscript. Ari Barzilai advised and provided guidance regarding the content and composition of the manuscript. Amotz Zahavi supervised the research and co-wrote the manuscript. All authors have seen and agreed to the final content of the manuscript.

Competing interests
No competing interests were disclosed.

Grant information
The author(s) declared that no grants were involved in supporting this work.

Acknowledgements
We would like to thank Adva Yeheskel from the Bioinformatics Unit, George S. Wise Faculty of Life Sciences, Tel Aviv University, for her assistance with the bioinformatics.
Open Peer Review

Current Referee Status: ✔️ ✔️

Version 1

Referee Report 12 October 2015

doi:10.5256/f1000research.7401.r10773

David Granot
Department of Vegetable Research, Institute of Plant Sciences, Agricultural Research Organization, Volcani Centre, Bet Dagan, Israel

This opinion article addresses basic questions about the nature of signals in biological systems. The underlying assumption is the brilliant handicap theory of Amotz Zahavi that signals must have some burden on the signaling cell to be perceived as a reliable signal.

The central question that this opinion article is dealing with is how very similar (mature) signaling peptides control different metabolic pathways in different species or even within the same organism.

The authors suggest that the message (signal) resides in the propeptides that yield the mature signaling peptides. The amount, location and timing of the very similar (yet not identical) signals provide a reliable and specific information on the capability of the producing cells to synthesize, modify, cleave and secrete the mature signaling peptides.

They also suggest that the mature signaling peptides might have toxic effects on the receiving cells that enforce them to respond.

The authors also offer an evolutionary prediction of how such signaling peptides might have evolved from propeptides that might have functions which are not necessarily related to the signal. Finally they offer a perspective on the potential impact of their theory on evaluation of pathological treatments, in which the current practice is increasing the amount of the signaling peptides rather than dealing with malfunction of the source of the peptides.

As usual, the theories stemming from Zahavi’s and his colleagues (Harris and Barzilai this time) are stimulating, but require validation, as emphasized by the authors themselves. Without validation, the appeal of theories is dependent on how reasonable they are, and if they make (new) sense of disturbing questions. The suggested theories should also be evaluated in comparison with alternative theories. It is not clear if alternative theories exist, and if they do, I would suggest discussing them.

The major question raised by the authors is how very similar signaling peptides have different functions?

Although very similar, even small signaling peptides are not identical, and it is advised not to exclude the option that the specificity resides within the different amino acids. The examples brought by the authors of cross species effects of similar peptides support the authors' theory, but can also be explained by the modular nature of biological structures of signaling peptides and receptors. As suggested by the authors,
shuffling mature peptides between propeptides may provide an answer.

An important part of the author’s hypothesis is that signals might (should?) be toxic to the receiving cell to enforce a response. The only example is the *E. coli* EDF pentapeptide that has a direct toxicity that is crucial to its function.

The biological logic of toxicity to the receiving cells as a way to enforce response is not clear. Toxicity means that the outcome could be harmful, such as killing or deactivating the receiving cells. Do the authors suggest that in such cases the signaling peptide has also a selection function?

The authors also discuss a case in which neurons within the brain produce a signaling peptide that affects peripheral cells and suggest that although processed from propeptides in the brain it does not provide information relating to the phenotype of the secreting neuron.

It is not clear why these neurons are excluded from the “handicap principle”? Furthermore, dealing with pathological conditions related to the administration of overdose of the signaling peptides (in Alzheimer disease), the authors suggest that it might mask the “inability to process the propeptide correctly by the signaling cell” which could be the cause of the disease. Wouldn’t that mean that production of signaling peptide from the propeptide in neurons should provide information about the phenotype of the secreting neuron?

As hinted above, brain stimulation by various scientific hypotheses is usually positive, and may indicate the excellent “phenotypic state” of the current “signaling” group of authors. Yet, I hope there is some “toxicity” that may enforce few readers to respond and test the hypotheses

**I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

**Competing Interests:** No competing interests were disclosed.

---

**Referee Report 01 September 2015**

doi:10.5256/f1000research.7401.r10173

**Uzi Motro**
Department of Ecology, Evolution and Behavior, The Hebrew University of Jerusalem, Jerusalem, Israel

A stimulating paper, suggesting a new perspective on the nature of signaling peptides.

The main idea can be summarized as follows:

1. The mere structure of mature signaling peptides does not necessarily convey the information they carry, but rather their sheer existence, which is due to successful processes undergone by their propeptides, reliably indicates the phenotypic state of the signaling cell.

2. The information given by signaling peptides is heard by the receiving cell because of the toxic nature of the signaling peptides – their toxicity cannot be ignored by the receiving cell.

This is a truly novel idea, which seems to be a solid consequence of natural selection. The authors present a few examples, yet – as they point out – sound experimental evidence is necessary. As a theoretician of evolutionary biology (but by no means an expert in cell biology!) I am not in a position to
comment or to suggest such an experimentation. If this idea is valid, the consequences to some medical
treatments can be substantial, as indicated in the paper's last section.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that
it is of an acceptable scientific standard.

**Competing Interests:** No competing interests were disclosed.