



**HAL**  
open science

## Emergence of a cholecystokinin/sulfakinin signalling system in Lophotrochozoa

Julie Schwartz, Marie-Pierre Dubos, Jeremy Pasquier, Céline Zatylny-Gaudin,  
Pascal Favrel

► **To cite this version:**

Julie Schwartz, Marie-Pierre Dubos, Jeremy Pasquier, Céline Zatylny-Gaudin, Pascal Favrel. Emergence of a cholecystokinin/sulfakinin signalling system in Lophotrochozoa. *Scientific Reports*, 2018, 8 (1), pp.16424. 10.1038/s41598-018-34700-4. hal-01927201

**HAL Id: hal-01927201**

**<https://hal.sorbonne-universite.fr/hal-01927201>**

Submitted on 19 Nov 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SCIENTIFIC REPORTS



OPEN

## Emergence of a cholecystokinin/sulfakinin signalling system in Lophotrochozoa

Julie Schwartz, Marie-Pierre Dubos, Jérémy Pasquier, Céline Zatylny-Gaudin & Pascal Favrel

Chordate gastrin/cholecystokinin (G/CCK) and ecdysozoan sulfakinin (SK) signalling systems represent divergent evolutionary scenarios of a common ancestral signalling system. The present article investigates for the first time the evolution of the CCK/SK signalling system in a member of the Lophotrochozoa, the second clade of protostome animals. We identified two G protein-coupled receptors (GPCR) in the oyster *Crassostrea gigas* (Mollusca), phylogenetically related to chordate CCK receptors (CCKR) and to ecdysozoan sulfakinin receptors (SKR). These receptors, Cragi-CCKR1 and Cragi-CCKR2, were characterised functionally using a cell-based assay. We identified di- and mono-sulphated forms of oyster Cragi-CCK1 (pEGAWDY(SO<sub>3</sub>H)DY(SO<sub>3</sub>H)GLGGGRF-NH<sub>2</sub>) as the potent endogenous agonists for these receptors. The Cragi-CCK genes were expressed in the visceral ganglia of the nervous system. The Cragi-CCKR1 gene was expressed in a variety of tissues, while Cragi-CCKR2 gene expression was more restricted to nervous tissues. An *in vitro* bioassay revealed that different forms of Cragi-CCK1 decreased the frequency of the spontaneous contractions of oyster hindgut. Expression analyses in oysters with contrasted nutritional statuses or in the course of their reproductive cycle highlighted the plausible role of Cragi-CCK signalling in the regulation of feeding and its possible involvement in the coordination of nutrition and energy storage in the gonad. This study confirms the early origin of the CCK/SK signalling system from the common bilaterian ancestor and delivers new insights into its structural and functional evolution in the lophotrochozoan lineage.

Neuropeptides include a diverse group of small neuron-signalling peptides that govern most animal biological functions and play a crucial role in the elaboration of adapted physiological and behavioural responses to environmental constraints. They activate specific receptors, which predominantly belong to the G protein-coupled receptor (GPCR) family. Global analyses including neuropeptide precursors and GPCR sequences from a variety of animal phyla have highlighted a long-term co-evolution of metazoan receptor-ligand families<sup>1,2</sup>. Until recently, knowledge on neuropeptide signalling systems was chiefly limited to well-studied vertebrate species and to ecdysozoan model species such as *Drosophila melanogaster* and *Caenorhabditis elegans*. Despite initial studies<sup>3,4</sup>, neuropeptide signalling has so far remained largely elusive in Lophotrochozoa, the protostomian sister clade of Ecdysozoa. Lophotrochozoa represent one of the most diverse and evolutionarily highly successful bilaterian lineages<sup>5</sup>. The neuropeptide repertoire of several lophotrochozoan species has now been established<sup>6–9</sup>, and large-scale genomic and transcriptomic resources are available<sup>10,11</sup>. This new situation offers the opportunity to investigate neuropeptide receptor coupling in a lophotrochozoan species and thus gain insight into the evolution of neuropeptide signalling in Bilateria. Accordingly, new signalling systems have recently been discovered in the annelid *Platynereis dumerilii*<sup>12,13</sup> and in the mollusc *Crassostrea gigas*<sup>14–16</sup>. The present study investigates the evolution of the gastrin/cholecystokinin (G/CCK)/sulfakinin (SK) signalling system in Lophotrochozoa, using the mollusc *C. gigas* as a representative species.

In chordates, the G/CCK family includes structurally conserved, frequently sulphated regulatory peptides<sup>17</sup>. Similarly to gut hormones, gastrin stimulates gastric acid secretion from the parietal cells of the stomach and plays a central role in gastric mucosa growth<sup>18,19</sup>, whereas CCK stimulates pancreatic enzyme secretion, induces the contraction and thus the emptying of the gallbladder, controls gut motility<sup>20</sup>, and stimulates satiety<sup>21</sup>. Besides these canonical functions in the regulation of digestive processes, CCK peptides now emerge as ubiquitous messengers due to the widespread expression of their encoding gene in a number of organs, including the brain

Normandie Université, UNICAEN, Sorbonne Universités, MNHN, UPMC, UA, CNRS 7208, IRD 207, Biologie des Organismes et Ecosystèmes Aquatiques (BOREA), CS14032, 14032, Caen, Cedex 5, France. Correspondence and requests for materials should be addressed to P.F. (email: [pascal.favrel@unicaen.fr](mailto:pascal.favrel@unicaen.fr))

where they are major neurotransmitters<sup>22</sup> (for a review). The first protostomian member of the G/CCK family was biochemically characterised from head extracts of the cockroach (*Leucophaea maderae*)<sup>23</sup>. This peptide was named leucosulfakinin (SK), and it exhibits around 55% of sequence identity with chordate G/CCKs. Moreover, the presence of a conserved sulphated tyrosyl residue further emphasises the resemblance with chordate G/CCKs. SKs have been identified in a number of insect species<sup>24,25</sup> and crustaceans<sup>26</sup>. This family also includes the related nematode neuropeptide NLP12<sup>27</sup>. Like their vertebrate counterparts, they display potent myotropic activity in several tissues, especially the gut<sup>28</sup> and the heart<sup>29</sup>. Similarly, they can regulate the release of digestive enzymes<sup>30,31</sup> and play a role as satiety regulators<sup>28,31–33</sup>. In addition to their structural homology and their convergent biological activities with G/CCKs, SKs specifically activate receptors (SKRs) that share a common ancestral gene with CCK receptors (CCKRs)<sup>34</sup>.

Apart from initial studies reporting the occurrence of CCK-like immunoreactivity in the central nervous systems of *Aplysia californica*<sup>35</sup> and *Nereis diversicolor*<sup>36</sup>, or revealing that insect SKs elicit the release of  $\alpha$ -amylase from the gut of the scallop *Pecten maximus*<sup>37</sup>, the CCK/SK signalling system has remained unexplored in Lophotrochozoa.

The present study investigates a CCK/SK signalling system in Lophotrochozoa for the first time.

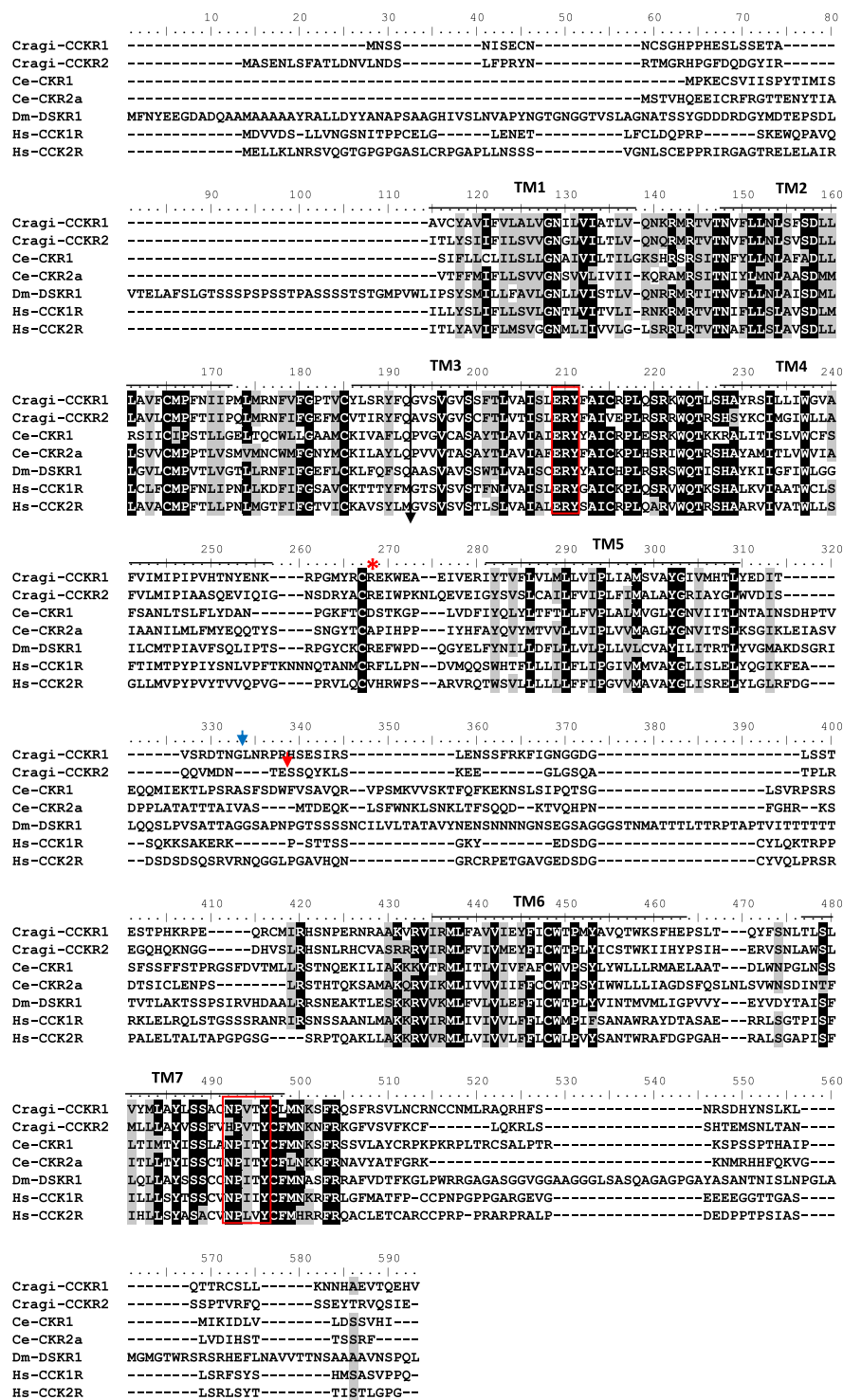
## Results

**Molecular characterization Cragi-CCKRs.** Two sequences displaying homologies with vertebrate and ecdysozoan CCK receptors were retrieved from GigaTON, an oyster comprehensive transcriptomic database<sup>11</sup>. These sequences named Cragi-CCKR1 and Cragi-CCKR2 share 43.5% of amino acid sequence identity. Cragi-CCKR1 and Cragi-CCKR2 also display around 31.5%–26.7% sequence identity with human CCKRs. Surprisingly, Cragi-CCKRs show slightly less identity with *C. elegans* CK-Rs (24.7–19.6%) and *Drosophila* DSKR1 (29.5–27.9%) (Fig. 1). In addition to the seven transmembrane domains characteristic of GPCRs, Cragi-CCKRs also hold the E/DRY and NPXXY (HPXXY for Cragi-CCKR2) motifs typical of family I GPCRs and critical for respectively receptor activation and G $\alpha$ q dependent signalling pathways as well as STAT activation<sup>38</sup>. A phylogenetic analysis clearly revealed that Cragi-CCKR1 and Cragi-CCKR2 cluster with receptors from annelids and molluscs but as a separate branch from the insect SKRs. All chordate CCKRs form a distinct clade but with two branch clusters separating type 1 and type 2 receptors. *C. elegans* CKRs appeared more distant and emerged as a detached clade (Fig. 2). Alignment of the Cragi-CCKR cDNAs with the *C. gigas* genome sequence (<http://www.oysterdb.com>)<sup>10</sup> identified Cragi-CCKR1 gene (scaffold1301, CGI\_10027768) and Cragi-CCKR2 gene (scaffold962, CGI\_10027668) and revealing the existence of an intron at a conserved position across vertebrate CCKR and insect SKR genes (Fig. 1)<sup>1</sup>.

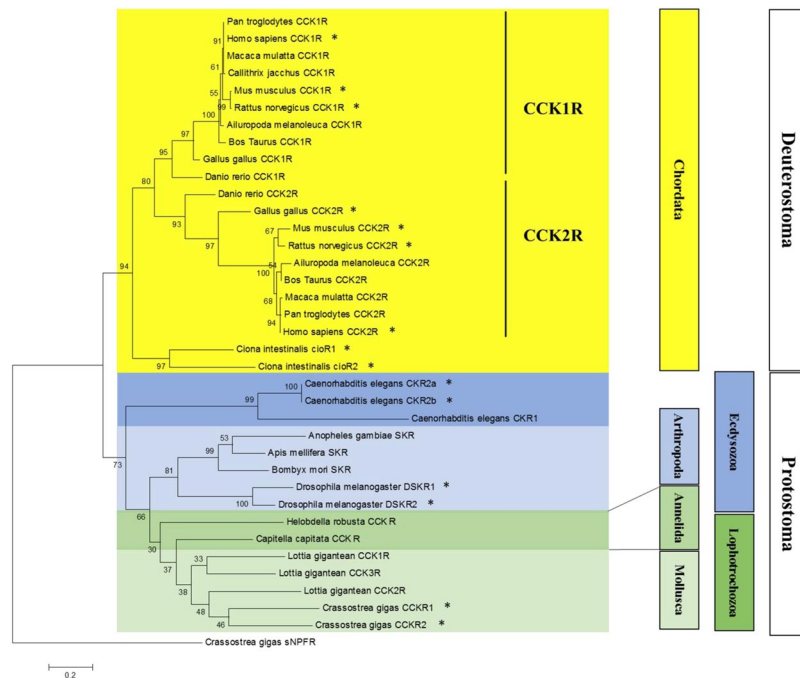
**Molecular characterization of Cragi-CCKs.** Within the *C. gigas* neuropeptide repertoire<sup>7</sup>, two potential CCK-like peptides, Cragi-CCK1 (pEGAWDYDYGLGGGRFamide) and Cragi-CCK2 (pEFDYGGGRWamide)<sup>39</sup>, encoded by a single precursor (Fig. 3A) were detected from visceral ganglia (VG) extracts by mass spectrometry<sup>7</sup>. Alignment of Cragi-CCKs with CCK and SK peptides from other species shows that Cragi-CCK1 displays the C-terminal RFamide sequence feature common to insect SKs and distinct from the DFamide motif found in all chordate CCKs (Fig. 3B). Instead, Cragi-CCK2 comprises a C-terminal RWamide motif. Membership of *C. gigas* peptides to the CCK/SK family is also suggested by the conserved position of a key tyrosine residue which sulphated forms are often critical to elicit biological activity. In addition to the conserved tyrosine (Y<sup>8</sup>, from the N-terminus), Cragi-CCK1 shows an extra tyrosine residue (Y<sup>6</sup>) at a position corresponding to a second sulphated tyrosine present in the white shrimp *Litopenaeus vannamei* SK1 and the protochordate *Ciona intestinalis* peptide Cionin<sup>26,40</sup>. Singularly, Cragi-CCK1 also displays a N-terminal sequence similarity to that of human gastrin-17. Using liquid chromatography (RP-HPLC) and mass spectrometry analysis, the occurrence of the non-sulphated form and of the mono-sulphated and disulphated forms of CCK1 was unambiguously confirmed in oyster visceral ganglia (Fig. 3C,D, Supplementary Fig. 1).

**Cragi-CCK1 forms specifically activate Cragi-CCKRs.** Transiently transfected HEK293T cells expressing Cragi-CCKR1 or Cragi-CCKR2 with or without the promiscuous G $\alpha$ 16 protein were challenged with rather high concentrations (10<sup>-5</sup>M) of synthetic non-sulphated Cragi-CCKs. No signal was obtained with cells transfected with an empty vector or with a G $\alpha$ 16 expressing vector. Activation of Cragi-CCKRs with Cragi-CCKs appeared slightly more efficient in absence of G $\alpha$ 16 (Supplementary Fig. 2). Thus a series of synthetic sulphated (s) or non-sulphated (ns) forms of Cragi-CCKs were tested as potential ligands for Cragi-CCKRs. Non-sulphated Cragi-CCK1 activates Cragi-CCKR1 and Cragi-CCKR2 with half maximal effective concentrations (EC<sub>50</sub>) of 0.19.10<sup>-6</sup>M and 0.46.10<sup>-6</sup>M respectively. The mono-sulphated form [Y<sup>8</sup>S] Cragi-CCK1 was, with respect to Cragi-CCKR1 and Cragi-CCKR2, 90 and 50 times more potent than its non-sulphated counterpart (Fig. 4, Table 1). Interestingly, the di-sulphated form [Y<sup>6</sup>S-Y<sup>8</sup>S] Cragi-CCK1 was even more potent (2 and 5 folds respectively). The synthetic mono-sulphated form [Y<sup>6</sup>S] Cragi-CCK1 was less potent on Cragi-CCKR1 and inactive on Cragi-CCKR2 at doses as high as 10<sup>-5</sup>M (Table 1). Furthermore, the sulphated form [Y<sup>3</sup>S] Cragi-CCK2 activated only Cragi-CCKR1 though at relatively high concentrations (EC<sub>50</sub>=5 10<sup>-5</sup>M). No response was observed with high concentrations of the oyster GALRF-amide unrelated neuropeptide used as negative control<sup>14</sup> (Fig. 4). A possible transduction via G $\alpha$ s was investigated using a cAMP luminescence assay but none of the synthetic peptides activated the cAMP signalling pathway even at concentrations as high as 10<sup>-5</sup>M.

**Gene expression of Cragi-CCK signalling components.** The expression of Cragi-CCKR1, Cragi-CCKR2 and Cragi-CCK genes was investigated by RT-qPCR in several adult tissues and in fed and starved animals. Cragi-CCK gene was almost exclusively expressed in the visceral ganglia and to a lesser extent in the mantle (Fig. 5A). Cragi-CCKR1 gene was ubiquitously expressed in adult tissues though with a higher expression



**Figure 1.** Alignment of the amino acid sequence of Cragi-CCKR1 (MF787221), Cragi-CCKR2 (MF787222) with CCKR family members from *Caenorhabditis elegans* (Ce-CCKR1: NP\_491918.3; Ce-CCKR2a: ACA81683.1), *Drosophila melanogaster* (Dm-DSKR1: NP\_001097021.1), *Homo sapiens* (Hs-CCKR1: NP\_000721.1; Hs-CCKR2: NP\_795344.1) using CLUSTALW<sup>65</sup>. Bars indicate the seven putative TM domains. Identical amino acid residues are highlighted in dark grey and similar residues in light grey. Black arrow indicates the position of a conserved intron, blue and red arrows indicate the respective position of Cragi-CCKR1 and Cragi-CCKR2 specific introns. Red boxes indicate the conserved “ERY” motif for receptor activation and “NPXXY” motif for Gαq and STAT –dependent signalling pathways. Red asterisk, the Arginine residue important for interaction with the sulphated tyrosine residue of CCK.



**Figure 2.** Phylogenetic representation of the relationship between Cragi-CCKRs and other CCKR family members. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 6<sup>61</sup> based on the maximum likelihood method. The accession members of the sequences used to construct the tree are listed in Supplementary Table 2. The Cg-sNPFR<sup>14</sup> was chosen as outgroup. \* indicates functionally characterized receptors.

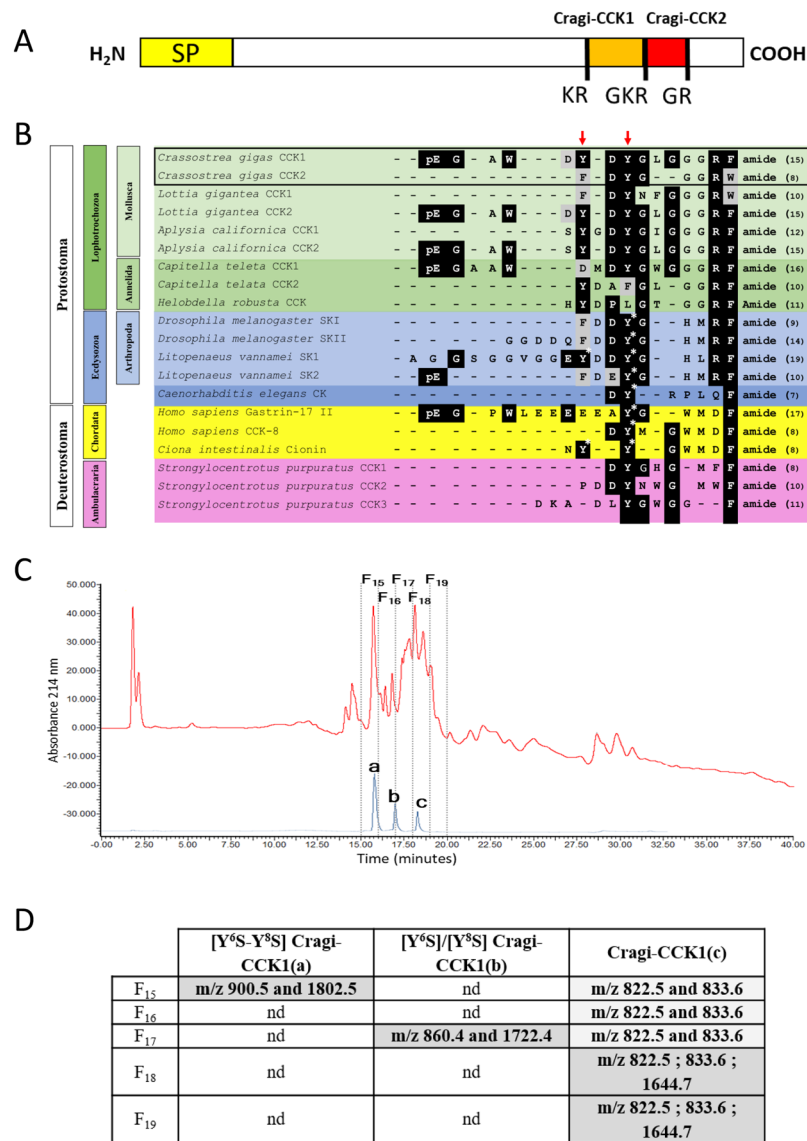
in the heart, the mantle and the visceral ganglia (Fig. 5B). In comparison, Cragi-CCKR2 was overall expressed at a lower level in adult tissues though at higher levels in the visceral ganglia, the mantle and the adductor muscle (Fig. 5C). To investigate the possible involvement of the oyster CCK signalling system in feeding regulation, we assessed the expression of the genes encoding the CCK signalling components including the major enzyme implicated in protein sulfation (TPST: Tyrosyl Protein sulfotransferase) in fed and four weeks starved animals. Both Cragi-CCK and Cragi-CCKR2 genes were significantly more expressed in VG of fed animals than in starved animals. Expression of all the other tested genes also slightly declined in starved animals yet not significantly (Fig. 6). Outside the VG, Cragi-CCKR1 decreased significantly in the gonad of starved animals (Supplementary Table 4). Given the expression of Cragi-CCKR1 in the gonad, we investigated the expression of receptor and ligand transcripts along a reproductive cycle. Expression of Cragi-CCK and Cragi-CCKRs did not fluctuate significantly along a reproductive cycle in VG. In contrast Cragi-CCKR1 gene showed a higher expression in the gonad during the first stages of reproduction and a significant decline of expression at the maturity stage (stage 3) in both males and females (Fig. 7).

**Myotropic activity assays.** Bioassays were carried out on the oyster hindgut. The basal activity is cyclic and composed of a succession of contraction and relaxation phases of an average period of 30 minutes (Fig. 8). Different concentrations of [Y<sup>6</sup>S] Cragi-CCK1, [Y<sup>8</sup>S] Cragi-CCK1 and [Y<sup>6</sup>S-Y<sup>8</sup>S] Cragi-CCK1 were administered. All peptides except [Y<sup>6</sup>S] Cragi-CCK1 induce organ contraction though with distinct threshold concentrations. A minimal threshold concentration of 10<sup>-10</sup>M was obtained for [Y<sup>8</sup>S] Cragi-CCK1, 10<sup>-9</sup>M for [Y<sup>6</sup>S-Y<sup>8</sup>S] Cragi-CCK1 and to 10<sup>-7</sup>M for Cragi-CCK1 (Table 2). No response was obtained with [Y<sup>6</sup>S] Cragi-CCK1 at a concentration up to 10<sup>-5</sup>M. Active peptides induce a significant increase of the contraction and relaxation duration (except for [Y<sup>8</sup>S] Cragi-CCK1) and a significant increase of amplitude of contraction when a threshold concentration of [Y<sup>6</sup>S-Y<sup>8</sup>S] Cragi-CCK1 is administered. No effect was observed with FMRF-amide at a concentration of 10<sup>-7</sup>M.

## Discussion

Structural and functional homology between peptide members of both the chordate G/CCK and arthropod SK groups, together with the phylogenetic proximity of their cognate receptors, support a common origin of these two signalling systems in the bilaterian ancestor. The present report reveals for the first time the existence of a functional CCK/SK signalling system in Lophotrochozoa, one of the two major clades of protostome animals.

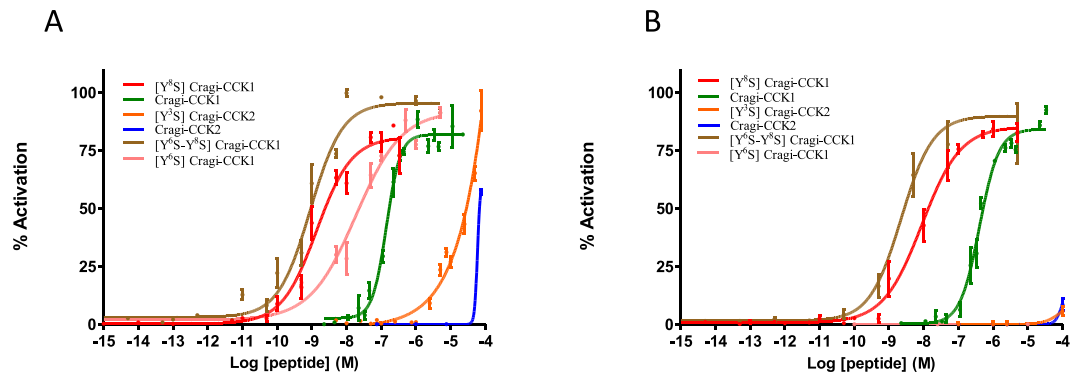
The occurrence of a lophotrochozoan CCK/SK signalling system was first investigated by mining the transcriptomic and genomic resources of the oyster *C. gigas*. Then a wide-ranging phylogenetic analysis including a large number of functionally characterised chordate CCKRs and ecdysozoan SKRs was performed, and clearly identified Cragi-CCKR1 and Cragi-CCKR2 as orthologues of ecdysozoan SKRs and chordate CCKRs. Both oyster receptor genes harboured an intron at a position conserved among all chordate and arthropod receptor



**Figure 3.** (A) Schematic representation of Cragi-CCK precursor<sup>7</sup> (SP: Signal peptide). (B) Sequence alignment of Cragi-CCK peptides with Deuterostome G/CCK family members<sup>44</sup>, ecdysozoan sulfakinins (SK)/CK<sup>27</sup> and lophotrochozoan CCK/SK<sup>39,42</sup>. The sulphated tyrosine residues (Y\*) are labelled with an asterisk. Red-colored arrows indicate putative sulphated tyrosine residues in Cragi-CCKs or in putative CCK/SK peptides from other species. Black-colored and grey-colored shadings indicate respectively identical or similar amino acids. (C) HPLC separation: In red, chromatogram of the second purification step of visceral ganglia extract. In blue, chromatogram of synthetic CCKs. a, [Y<sup>6</sup>S-Y<sup>8</sup>S] Cragi-CCK1; b, [Y<sup>6</sup>S]/[Y<sup>8</sup>S] Cragi-CCK1; c, Cragi-CCK1. (D) Detection with SIM mode of singly charged and doubled charged *ions* of different forms of Cragi-CCK1s in the different chromatographic fractions. nd: not detected.

genes, further strengthening the argument that they evolved from a common bilaterian ancestor<sup>1,34</sup>. At least two CCKRs or SKRs are present in most animal lineages. The type-1 and type-2 CCKR groups are clearly separated within vertebrates. Our phylogenetic analysis revealed that the two receptor types present in some arthropods and Lophotrochozoa not only clustered separately from the vertebrate receptors but also grouped in a phylum-specific way, consistent with independent duplication events that occurred at least after Ecdysozoa and Lophotrochozoa diverged. The situation was similar for the two related CioRs of the urochordate *Ciona intestinalis*<sup>41</sup>. In *C. gigas*, the two Cragi-CCKRs clustered together and separately from putative CCK/SK receptors from other molluscs or from annelids. As Cragi-CCKR genes were positioned in separate genome contigs, the duplication event was probably not recent and likely occurred in an ancestor of the bivalve lineage.

In line with the receptor/ligand co-evolution theory, we inferred that putative oyster neuropeptides previously identified based on their slight sequence homology with CCK/SK peptides<sup>7</sup> might represent the cognate ligands for Cragi-CCKRs. These peptides, i.e. Cragi-CCK1 and Cragi-CCK2, are generated from a single precursor<sup>7</sup> displaying an organization similar to their mollusc<sup>8</sup>, annelid<sup>42</sup>, nematode<sup>43</sup> and arthropod counterparts<sup>34</sup>. But

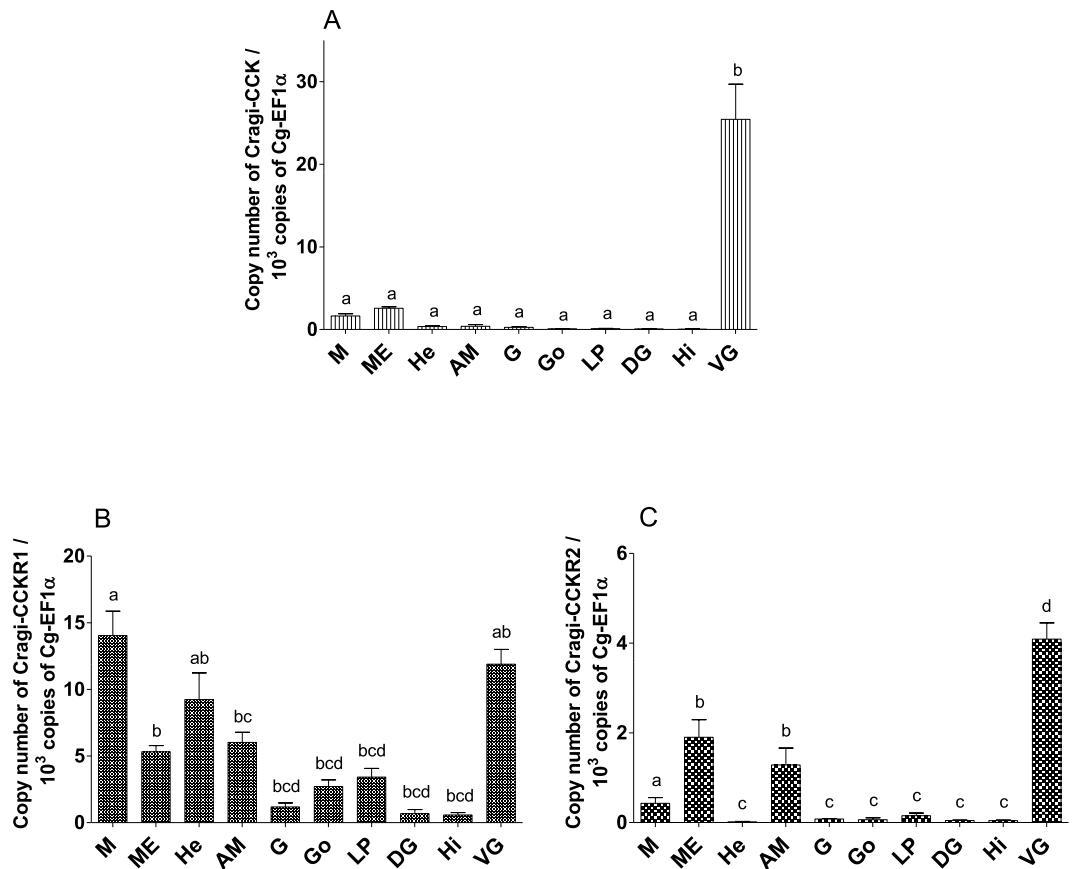


**Figure 4.** Dose-dependent activity of Cragi-CCK peptides on Cragi-CCKR1 (A) and on Cragi-CCKR2 (B) expressed in HEK293T cells. Data are shown as relative (%) to the highest value (100% activation) for a given peptide and represent the mean of an experiment ( $n = 4$ ) performed in triplicate. Vertical bars represent the standard error of the mean (SEM).

Peptide Name	Cragi-CCKR1	Cragi-CCKR2
Cragi-CCK1	$19.5 \pm 2.0 \cdot 10^{-8} \text{M}$	$46.2 \pm 6.7 \cdot 10^{-8} \text{M}$
[Y <sup>6</sup> S] Cragi-CCK1	$1.64 \pm 0.01 \cdot 10^{-8} \text{M}$	No activation
[Y <sup>8</sup> S] Cragi-CCK1	$0.229 \pm 0.155 \cdot 10^{-8} \text{M}$	$1 \pm 0.19 \cdot 10^{-8} \text{M}$
[Y <sup>6</sup> S-Y <sup>8</sup> S] Cragi-CCK1	$0.11 \pm 0.025 \cdot 10^{-8} \text{M}$	$0.198 \pm 0.025 \cdot 10^{-8} \text{M}$
Cragi-CCK2	$>10^{-4} \text{M}$	$\gg 10^{-4} \text{M}$
[Y <sup>3</sup> S] Cragi-CCK2	$>5 \cdot 10^{-5} \text{M}$	$\gg 10^{-4} \text{M}$

**Table 1.** Cragi-CCK respective EC<sub>50</sub> for receptor activation.

their precursors are distinct from the chordate G/CCK precursors which only contain one peptide that can be processed into a variety of N-terminally extended mature forms<sup>34,38</sup>. All chordate G/CCKs share the common C-terminal GWMDF-amide motif, which is not conserved in the other animal groups. For example, non-chordate deuterostomian putative CCK/SK-like peptides display a GXXF-amide motif<sup>44</sup>, the *C. elegans* homologue a QF-amide motif<sup>27</sup>, although all arthropod SKs display the C-terminal core sequence YGHMRF-amide motif<sup>44</sup>. Thus, in addition to the basically conserved (D/E)Y(G) sequence, the F-amide terminal motif appears as the common feature shared by bilaterian CCK/SK family members<sup>44</sup>. Hence Cragi-CCK1 resembles SKs because it exhibits a C-terminal RF-amide. In contrast, Cragi-CCK2 represents a singular exception because it displays an atypical C-terminal RW-amide motif. A comparison of Cragi-CCKs with other family members revealed the presence of one or even two (for Cragi-CCK1) conserved tyrosine residues that are frequently post-translationally sulphated<sup>22</sup>. Since both tyrosine residues are located in a consensus sequence compatible for recognition as TPST substrates<sup>45</sup>, Cragi-CCKs were synthesized in their sulphated and non-sulphated forms. This report clearly demonstrates that the different sulphated forms of Cragi-CCK1 specifically activated both Cragi-CCKR1 and Cragi-CCKR2 and were far more potent than their non-sulphated counterparts. Interestingly, similarly to the cionin signalling system<sup>41</sup>, the disulphated form of Cragi-CCK1 was more potent than its monosulphated derivatives. The sulphate moiety also appeared critical for receptor activation in other CCK/SK systems<sup>38,46,47</sup>, although it was not found essential for the activation of *C. elegans* CKRs<sup>27</sup> or mammalian gastrin/CCK2R<sup>22</sup>. Both Cragi-CCKRs possess a conserved arginine residue in the extra-cellular loop 2. In mammalian CCK1R, it participates in the interaction with the negatively charged sulphate moiety of CCK<sup>48</sup>. Interestingly, CCK2R does not share this arginine residue, and neither do *C. elegans* CKRs (Fig. 1). The biochemical characterisation of Cragi-CCKs presented in this study clearly proves the existence of both mono and di-sulphated forms in the VG, consistent with the expression of the TPST gene in this tissue. Mass spectrometry signals associated with the sulphated forms were rather weak as compared to those of the non-sulphated form, suggesting that the former were expressed in lower quantities. However, mass spectrometry was not quantitative, and moreover this situation may have resulted from partial desulphation due to the acidic conditions used in the initial extraction process and in source dissociation<sup>45</sup>. Our experimental conditions did not allow us to discriminate between [Y<sup>6</sup>S] Cragi-CCK1 and [Y<sup>8</sup>S] Cragi-CCK1. Both forms were likely to be present, resulting either from partial desulphation of the disulphated form or representing intermediary molecular species of the synthesis of the di-sulphated form. Of the two monosulphated forms, [Y<sup>8</sup>S] Cragi-CCK1 was more potent than [Y<sup>6</sup>S] Cragi-CCK1, suggesting that the sulphated moiety located on the tyrosine residue with the most conserved position could be an ancestral feature, as suggested for cionin<sup>41</sup>. Thus, the sulphated forms of Cragi-CCK likely represent physiologically active peptides in oyster, even though the non-sulphated Cragi-CCK1 form may also be biologically relevant given its potency to activate Cragi-CCKRs. Indeed, in the context of local release from nerve terminals, a rather high concentration of peptide can be expected *in vivo* at the vicinity of the receptors. This assumption is fully compatible with a role of Cragi-CCK1 as a neurotransmitter, as formerly demonstrated for CCK<sup>22</sup> and SK<sup>49</sup> peptides. Regulatory



**Figure 5.** Distribution of mRNAs encoding Cragi-CCK (A), Cragi-CCKR1 (B) and Cragi-CCKR2 (C) in adult tissues. Each value is the mean + SEM of 5 pools of 6 animals. Expression levels were calculated as the number of copies of each specific transcript per  $10^3$  copies of Elongation Factor 1  $\alpha$  (EF1 $\alpha$ ) mRNA. Results were statistically tested with a one-way ANOVA  $p < 0,05$ . Significantly different means are indicated by different letters. M: Mantle; ME: Mantle Edge; He: heart; AM: Adductor Muscle; G: Gills; Go: Gonad; LP: Labial Palps; DG: Digestive Gland; Hi: Hindgut; VG: Visceral Ganglia.

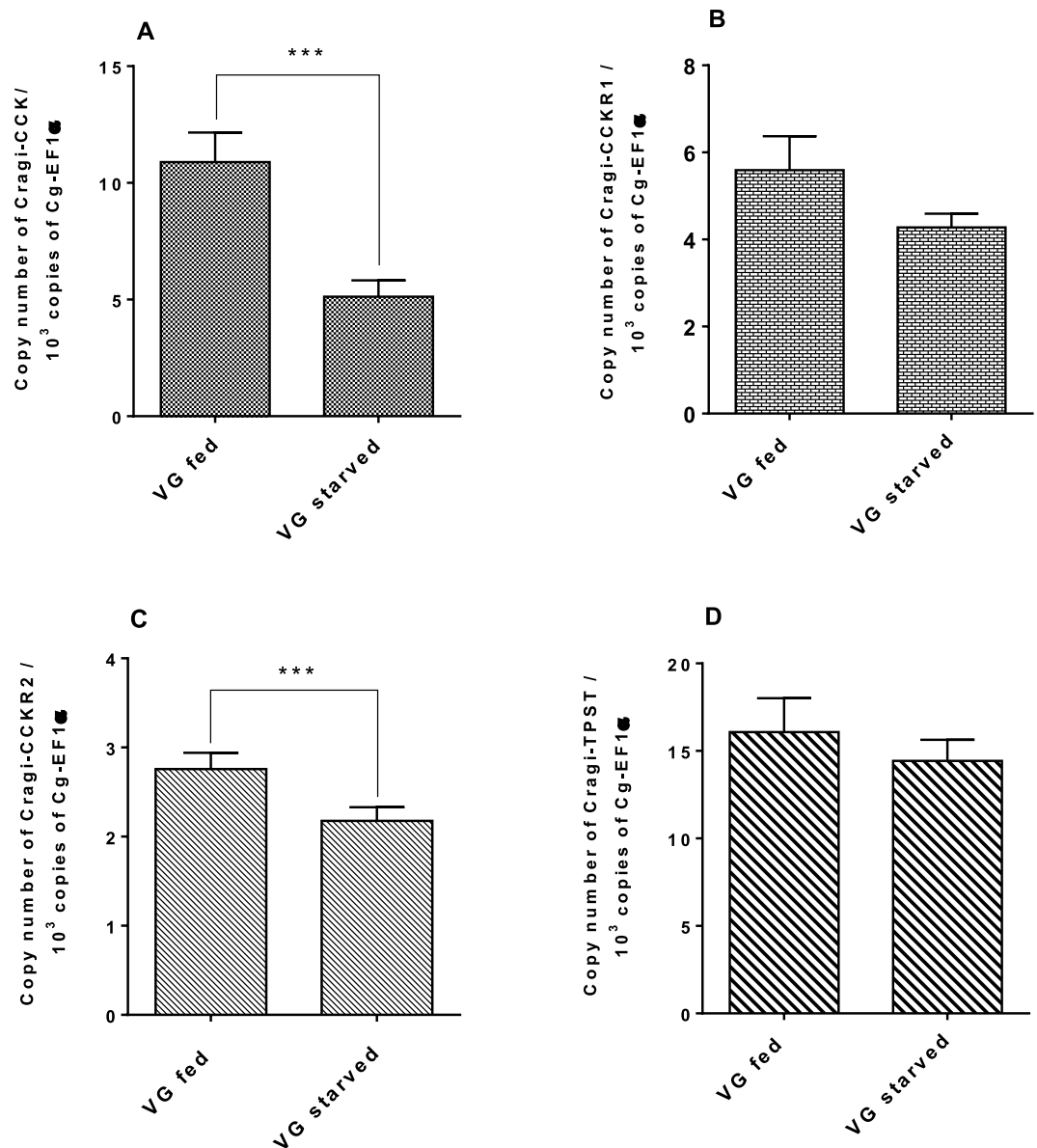
activity from non-sulphated SKs was also shown in insects<sup>50,51</sup>. All oyster peptides activated calcium accumulation in Cragi-CCKR-expressing cells, suggesting that these receptors couple to G proteins of the  $G_{\alpha_q/11}$  family. Similar findings were observed for chordate CCKRs<sup>38,41</sup> and SK family receptors<sup>27,46,47</sup>, even though CCK1R<sup>52</sup>, *C. elegans* CKR<sup>27</sup>, and the red flour beetle SKR<sup>47</sup> also activated the cAMP transduction pathway. Activation of other signaling pathways cannot be ruled out, e.g. the JAK/STAT-dependent pathway: the activation of this pathway for CCK2R requires  $G_{\alpha_q}$  proteins and the NPXXY motif, which is also present in the Cragi-CCKR1 sequence<sup>38</sup>. Taken together, these results clearly show that the neuropeptide/receptor pairs Cragi-CCK1/Cragi-CCKR1 and Cragi-CCKR2 represent a Lophotrochozoan orthologue of the CCK/SK signalling systems of chordates and ecdysozoa.

In other respects, Cragi-CCK2 forms did not act as specific ligands for Cragi-CCKR1 and Cragi-CCKR2 since the concentrations required to elicit a partial activation of the receptors appear not physiologically relevant except conceivably in the specific context of synaptic transmission. This raises the issue of the nature of the cognate receptor(s) for this neuropeptide form which logically should display a relative phylogenetic vicinity to Cragi-CCKR1/2. The type of interaction between the two Cragi-CCK signalling systems is particularly intriguing since both peptides represent end products of the same precursor.

CCK/SK signalling systems regulate gut functioning, food ingestion, satiety, as well as a variety of other biological processes, so that these neuropeptides are now recognized as multifunctional regulators<sup>22,53</sup>. In oyster, the Cragi-CCK gene was predominantly expressed in nervous tissues, as expected for a neuropeptide-encoding gene. In contrast, Cragi-CCKRs were widely expressed in many tissues but with distinct patterns. Cragi-CCKR1 was expressed in almost all tissues though at variable levels, Cragi-CCKR2 was less expressed and mostly in the VG, in richly innervated tissues such as the adductor muscle and the mantle edge or in tissues containing neurons such as the mantle. These observations support a multifunctional activity for Cragi-CCKR1 signalling, and a more restricted role possibly associated with neuroregulation for Cragi-CCKR2 signalling.

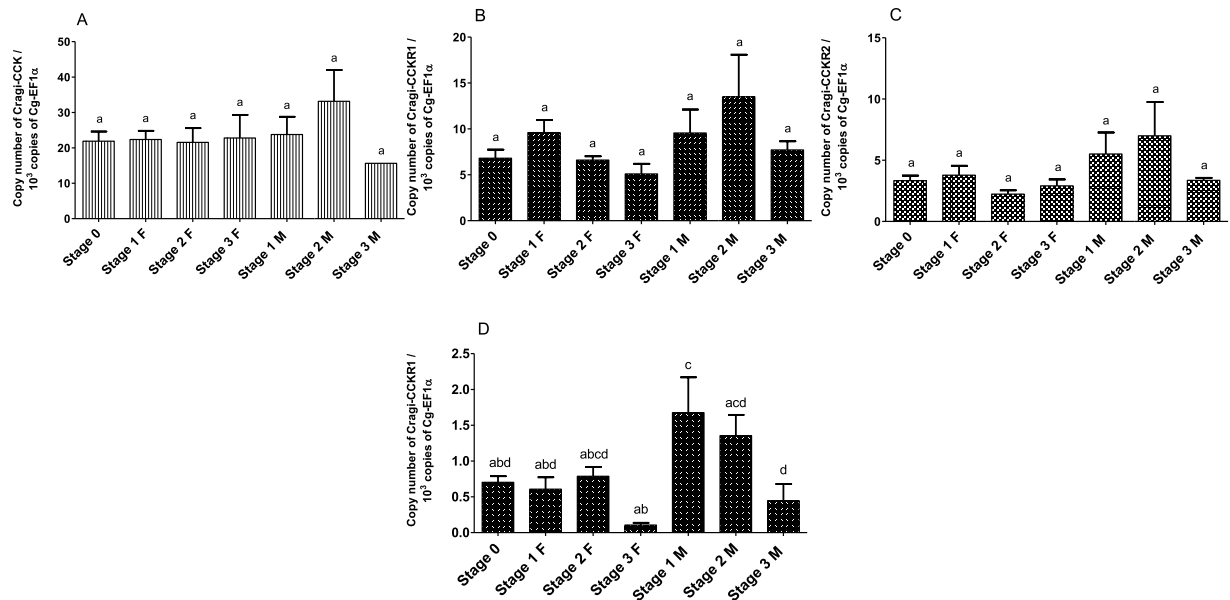
Predominant expression of the Cragi-CCKR1 gene in oyster heart is reminiscent of the role of SKs in heart activity in insects<sup>29</sup>. In the same way, the expression of the Cragi-CCKR1 gene in oyster digestive tissues, albeit at moderate levels, is in agreement with the fact that SKs induce the release of  $\alpha$ -amylase from the scallop digestive gland<sup>37</sup> or with the recorded activity of both sulphated and non-sulphated Cragi-CCK1 peptides on oyster



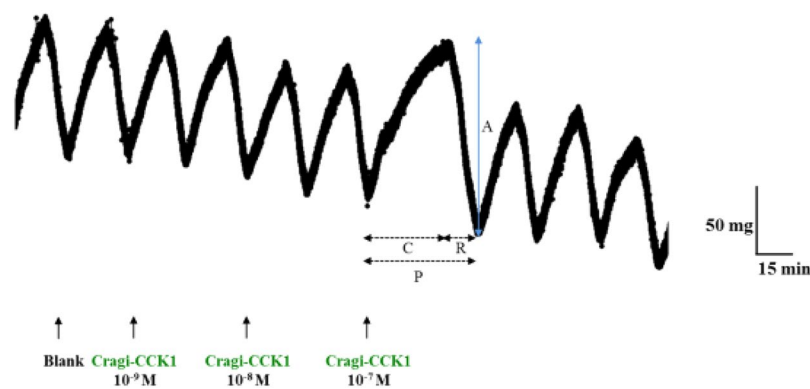


**Figure 6.** Expression levels of Cragi-CCK (A), Cragi-CCKR1 (B), Cragi-CCKR2 (C), Cragi-TPST (D) mRNAs in VG of four weeks *Isochrysis galbana* fed or starved oysters. Each value is the mean + SEM of 15 independent animals (VG after conditioning with or without food). Results were statistically tested with a student's t test. Significantly different means of samples from fed and starved animals are indicated by \* ( $p < 0,05$ ) (C) and \*\*\* ( $p < 0,001$ ) (A). No significant statistical difference was observed for (B,D).

hindgut motility. Interestingly, Cragi-CCK1 peptides increase the period of spontaneous hindgut contractions *in vitro* but chiefly did not affect their amplitude, thus contrasting with the stimulatory effect of SKs on insect hindgut motility<sup>54,55</sup>, but in line with the *in vivo* inhibition of foregut or larval anterior midgut contraction by SKs in *Drosophila*<sup>51</sup>. Cragi-CCK1 peptides seem to decrease the emptying process of the intestine in oyster. Thus, it would be interesting to test whether the anterior parts of the gut also respond to these peptides since different responses were noticed in different insect species<sup>51,56</sup>. By negatively controlling hindgut activity, Cragi-CCK1s probably affect food intake in oyster. The increased expression of Cragi-CCK1 and Cragi-CCKR2 in fed oysters as compared to starved oysters also pleads for a possible role of the Cragi-CCK system in satiety signalling, as also revealed in *Drosophila*<sup>57</sup> and other insects<sup>31,33,58</sup>. Besides the pleiotropic activities of CCK/SK peptides<sup>22,53</sup>, it was intriguing to find a differential expression of Cragi-CCKR1 in the gonad area along the reproductive cycle and according to the nutritional status. Cragi-CCKR1 expression declined as gametogenesis progressed. Therefore, this receptor probably plays a role in glycogen storage cells since this tissue almost completely disappears in the mature gonad. SK signalling in Ecdysozoa is involved in the regulation of fat<sup>27</sup> and energy storage<sup>50</sup>, and expression of Cragi-CCKR1 is higher in the gonad of fed oysters. Therefore it is tempting to propose a role for the Cragi-CCK signalling system as an additional regulator of the storage metabolism that might connect feeding and storage processes, as recently proposed for the sNPF-like signalling system in *C. gigas*<sup>14</sup>.



**Figure 7.** Level of expression of Cragi-CCK/Cragi-CCKR1/Cragi-CCKR2 mRNAs in VG (A–C) and of Cragi-CCKR1 mRNA in the gonads (D) along an annual reproductive cycle. Each value is the mean + SEM of 5 pools of 6 animals. Results were statistically tested with a one-way ANOVA,  $p < 0,05$ . Samples with significant statistical difference are marked with distinct letters. F: Female; M: Male; 0: stage 0 (sexual resting stage); 1: stage 1 (gonial multiplication stage); 2: stage 2 (tubule development and maturation stage); 3: stage 3 (sexual maturity stage).



**Figure 8.** (A) Representative record of the hindgut-stimulating activity of Cragi-CCK1. Arrows indicate the application of saline solution (Blank) or Cragi-CCK1 peptide fractions at different concentrations. The different parameters taken into account to measure the biological response are indicated (C: Contraction time, R: Relaxation time, P: Period (min.) A: Amplitude of contraction (mg).

Peptide Name	Treatment	Duration (min)				Amplitude (A) (g)	
		Period (P)	Contraction (C)	Relaxation (R)			
Cragi-CCK1	No peptide	30,67 ± 0,67	20,18 ± 0,33	10,48 ± 0,78	0,161 ± 0,038	ns	
	Peptide 10 <sup>-7</sup> M	51,27 ± 0,15	35,23 ± 1,32	16,03 ± 1,13	0,203 ± 0,038		
[Y <sup>6</sup> S] Cragi-CCK1	No peptide	31,53 ± 1,45	20,18 ± 0,98	11,35 ± 0,85	0,129 ± 0,019	ns	
	Peptide 10 <sup>-5</sup> M	32,23 ± 1,53	21 ± 1,42	11,23 ± 0,60	0,104 ± 0,021		
[Y <sup>8</sup> S] Cragi-CCK1	No peptide	29,66 ± 2,22	18,58 ± 1,03	11,08 ± 2,27	0,121 ± 0,025	ns	
	Peptide 10 <sup>-10</sup> M	54,75 ± 1,97	37,13 ± 3,28	17,62 ± 5,27	0,173 ± 0,023		
[Y <sup>6</sup> S-Y <sup>8</sup> S] Cragi-CCK1	No peptide	30,57 ± 2,22	18,76 ± 1,07	11,80 ± 1,63	0,103 ± 0,006	*	
	Peptide 10 <sup>-9</sup> M	62,39 ± 2,50	38,57 ± 3,53	23,82 ± 1,35	0,153 ± 0,016		

**Table 2.** Effect of Cragi-CCK peptide variants on hindgut contraction parameters. Each value is the mean ± SEM of four experiments. Results were statistically tested with a student's t test, \* $p < 0,05$ , \*\*\* $p < 0,001$ , ns: not significant.

This study outlines the lophotrochozoan evolutionary version of the CCK/SK signalling system and underlines its remarkable conservation at the structural and functional levels, implying a central role of CCK/SK in physiological regulation in bilateria.

## Material and Methods

**Animals and tissue sampling.** Two-year old adult oysters *C. gigas*, purchased from a local farm (Normandie, France), were used for peptide characterization and transcription analyses. Stages of reproduction (Stage 0: resting undifferentiated stage, Stage (1): gonial multiplication stage, Stage (2): maturation stage, Stage (3): sexual maturity) were determined by histological analysis of gonad sections as described previously<sup>59</sup>. To study the influence of trophic conditions, one-year-old adult oysters were reared in water tanks either in absence of food or in presence of *Isochrysis galbana* (clone T-Iso) maintained at a concentration of 6 million of cells/mL during (4) weeks. Adult tissues (mantle, gills, labial palps, digestive gland, hindgut, gonad (mix of all stages), heart, adductor muscle) were sampled, the visceral ganglia (VG) were carefully dissected out, thus limiting any contamination from the adjacent adductor muscles. All the samples were either placed in TriReagent (Sigma) or stored at  $-80^{\circ}\text{C}$  until use. For expression studies, adult tissues or VG and gonads during gametogenesis from (6) animals were mixed to generate (5) pools of each tissue. Individual VG from 15 fed or starved animals were used to study gene expression.

**Peptide synthesis.** All peptides were custom-synthesized by GeneCust (Luxemburg) using a standard Fmoc solid-phase protocol (Supplementary Table 1). Sulphated peptide variants were synthesized by coupling Fmoc-Tyr(SO<sub>3</sub>Na)-OH amino acids instead of Fmoc-tyr(tBu)-OH amino acids. The sequences of *C. gigas* peptides were obtained from an in-house peptide database yielded by MS analyses of tissue extracts and data mining<sup>7</sup>.

**In silico analyses.** Multiple sequence alignment was performed with CCKRs from various species (Supplementary Table 2) using MUSCLE<sup>60</sup>. The alignment was manually trimmed that spans from the first to the seventh transmembrane domains. To determine the relationship between Cragi-CCKRs and CCKRs from other species, phylogenetic and molecular evolutionary analyses were conducted using MEGA version 6<sup>61</sup> based on the maximum likelihood method. The reliability of the inferred trees was estimated by applying the bootstrap procedure with 1000 replications.

**Reverse endocrinology.** *Molecular cloning of the Cragi-CCKRs and transfection of mammalian cells.* BLAST analysis of *C. gigas* transcriptomic database “GigaTon”<sup>11</sup> using *Drosophila* SKR as query resulted in the identification of two full length cDNAs encoding Cragi-CCKR1 and Cragi-CCKR2 (CHOYP\_CCKAR.3.6 and CHOYP\_CCKAR.4.6) respectively. The CDS of the Cragi-CCKR genes were amplified by PCR (Pfu DNA polymerase, Promega) using gene-specific sense primers harbouring a Kozak consensus sequence and antisense primers (Supplementary Table 3). Ten nanogram of plasmid DNA (Pal 17.3 vector, Evrogen) from a *C. gigas* “all developmental stages and adult central nervous system” directional and normalized cDNA library<sup>62</sup> was used as template. The resulting PCR products were directionally cloned into the pcDNA3.1 expression vector (Invitrogen). The correct insertion of the PCR products was confirmed by sequencing. Human Embryonic Kidney (HEK293T) cells were transiently transfected with the Cragi-CCKR/pcDNA3.1 constructs using Eugene HD (Promega) according to the manufacturer’s instructions. As a first step, co-transfection was done with a pcDNA3.1 expression construct for the human G $\alpha_{16}$  subunit, a promiscuous G protein that can direct intracellular signalling of GPCRs to the release of calcium via the phospholipase C $\beta$  pathway, regardless of the endogenous G protein coupling of the receptor<sup>63</sup>. To assess receptor activity independent of G $\alpha_{16}$ , calcium responses were measured in cells expressing only the Cragi-CCKRs. Cells for negative control experiments were transfected with empty pcDNA3.1 and G $\alpha_{16}$ /pcDNA3.1 constructs.

*Calcium fluorescence assay.* Activation of the Cragi-CCKRs by candidate peptide ligands was monitored using a fluorescence-based calcium mobilization assay. Briefly, transfected HEK293T cells were loaded with Fluo-4 Direct (Invitrogen) plus probenecid (qsp 2.5 mM final in the cell) (Molecular Probes) for 1 hour (45 min at  $37^{\circ}\text{C}$  and 15 min at room temperature). Excitation of the fluorophore was done at 488 nm. The calcium response was measured for 2 min at 525 nm using a FLEXstation 3 (Molecular Devices) at  $37^{\circ}\text{C}$ . Data were analysed using SoftMax Pro (Molecular Devices). Candidate ligands were first tested at a final concentration of  $10^{-5}$  M. Concentration-response measurements of activating ligands were conducted in quadruplicate and for at least three independent experiments. Half maximal effective concentrations (EC<sub>50</sub> values) were calculated from concentration-response curves that were constructed using nonlinear regression analysis with a sigmoidal dose-response equation using Prism 5.0 (GraphPad software, USA).

*cAMP luminescence assay.* Cragi-CCKR transfected HEK 293 T cells were incubated with Glosensor cAMP reagent (qsp 4% final in the medium) (Promega) for 2 hours at room temperature prior to the injection of the candidate ligands. cAMP luminescence response was measured for 30 min after injection using a FLEX station 3 (Molecular Devices) at room temperature. Data were analysed using SoftMax Pro (Molecular Devices). Candidate peptide ligands were first tested at a final concentration of  $10^{-5}$  M.

**Reverse transcription and quantitative PCR.** RT-qPCR analysis was performed using the iCycler iQ<sup>©</sup> apparatus (Bio-Rad). Total RNA was isolated from adult tissues using Tri-Reagent (Sigma-Aldrich) according to the manufacturer’s instructions. Recovered RNA was then further purified on Nucleospin RNAII columns (Macherey-Nagel). After treatment during 20 min at  $37^{\circ}\text{C}$  with 1 U of DNase I (Sigma) to prevent genomic DNA contamination, 1  $\mu\text{g}$  of total RNA was reverse transcribed using 1  $\mu\text{g}$  of random hexanucleotidic primers (Promega), 0.5 mM dNTPs and 200 U MMuLV Reverse Transcriptase (Promega) at  $37^{\circ}\text{C}$  for 1 h in the appropriate buffer. The

reaction was stopped by incubation at 70 °C for 10 min. The GoTaq<sup>®</sup> qPCR Master Mix (Promega) was used for real time monitoring of amplification (5 ng of cDNA template, 40 cycles: 95 °C/15 s, 60 °C/15 s) with gene-specific primers (Supplementary Table 2). A parallel amplification of *C. gigas* Elongation Factor 1 $\alpha$  (EF1  $\alpha$ ) transcript (BAD15289) was carried out to normalize the expression data of the studied transcripts. EF1  $\alpha$  was found as a reliable normalization gene as no significant difference ( $p < 0.05$ ) of Ct values was observed between the different samples compared. Coefficient of variation of EF1  $\alpha$  was less than 5%. Thus, the relative level of each gene expression was calculated for one copy of the EF1  $\alpha$  reference gene by using the following formula:  $N = 2^{(Ct_{EF1\alpha} - Ct_{Cg-cDNA})}$ . The PCR amplification efficiency (E;  $E = 10^{(-1/\text{slope})}$ ) for each primer pair was determined by linear regression analysis of a dilution series to ensure that E ranged from 1.98 to 2.02. The specificity of the primer pairs was confirmed by melting curve analysis at the end of each qPCR run.

**Myotropic bioassay.** The myotropic bioassay was performed using oyster's hindgut according to the protocol described earlier for the cuttlefish oviduct<sup>64</sup>. The muscle chamber was perfused at a flow rate of 0.67 mL.min<sup>-1</sup>. Increasing concentrations of synthetic peptides (Cragi-CCK1, [Y<sup>6</sup>S] Cragi-CCK1, [Y<sup>6</sup>S] Cragi-CCK1, [Y<sup>6</sup>S-Y<sup>8</sup>S] Cragi-CCK1) were injected in the perfusing flow using a three-way valve to avoid mechanical stress. The flow of the samples into the muscle chamber was traced by adding phenol red.

**Statistical analysis.** Gene expression levels between different tissues and between samples at different reproduction stages were compared using one-way ANOVA followed by a Tukey post hoc test. Expression levels between fed and starved animals were compared using an unpaired Student's t test. Significance was set at  $p < 0.05$ .

**Purification and mass spectrometry analysis of endogenous Cragi-CCK1 forms.** One hundred visceral ganglia, frozen and crushed in liquid nitrogen, were extracted in methanol/water/acetic acid (90/10/1 v/v/v) 30 minutes at 4 °C. After centrifugation 20 minutes at 15 000 g at 4 °C, supernatant was concentrated on C18 Sep-Pak cartridges. HPLC analysis was performed with a VARIAN-9012 solvent delivery system coupled to a VARIAN-9050 wave-length UV-VIS detector set at 214 nm. The extract was eluted with a 55-min linear gradient of 0.75% ACN per min from 22 to 58% ACN with TFA 0.1% on a C18 nucleodur column (4 mm  $\times$  250 mm  $\times$  5  $\mu$ m). Fractions 45–48 corresponding to the retention times of the synthetic forms of CCK1 were pooled and separated on the same column with a 10-min linear gradient of 0.2% ACN per min from 45 to 47% ACN containing 10 mM ammonium acetate pH 5.2. Synthetic Cragi-CCK1, [Y<sup>6</sup>S] Cragi-CCK1 and [Y<sup>6</sup>S-Y<sup>8</sup>S] Cragi-CCK1 were also separated in the same conditions. Fractions 15 to 19 were analyzed by mass spectrometry. Analysis was realized in a triple quadrupole mass spectrometer with an electrospray interface (LCMS 8030Plus; Shimadzu), with selected ion monitoring (SIM) mode. Ions selected for non-sulphated Cragi-CCK1 in positive mode were [M + H]<sup>+</sup> + 1644.7, [M + 2H]<sup>2+</sup> + 822.5 and 833.6, the [M + 2H]<sup>2+</sup> + with sodium adduct. For sulphated peptides, negative mode was selected with respectively singly charged and doubled charged ions corresponding to 1722.4 and 860.4 for [Y<sup>6</sup>S] Cragi-CCK1 and 1802.5 and 900.5 for [Y<sup>6</sup>S-Y<sup>8</sup>S] Cragi-CCK1.

## References

- Mirabeau, O. & Joly, J. Molecular evolution of peptidergic signaling systems in bilaterians. *Proc Natl Acad Sci USA* **110**, 2028–2037 (2013).
- Jékely, G. Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc Natl Acad Sci USA* **110**, 8702–8707 (2013).
- Cox, K. J. *et al.* Cloning, characterization, and expression of a G-protein-coupled receptor from *Lymnaea stagnalis* and identification of a leucokinin-like peptide, PSFHSWSamide, as its endogenous ligand. *J Neurosci* **17**, 1197–1205 (1997).
- Tensen, C. P. *et al.* The *Lymnaea* cardioexcitatory peptide (LyCEP) receptor: a G-protein-coupled receptor for a novel member of the RFamide neuropeptide family. *J Neurosci* **18**, 9812–9821 (1998).
- Moroz, L. L. *et al.* Neuronal transcriptome of *Aplysia*: neuronal compartments and circuitry. *Cell* **127**, 1453–67 (2006).
- Conzelmann, M., Williams, E. A., Krug, K., Franz-wachtel, M. & Macek, B. The neuropeptide complement of the marine annelid *Platynereis dumerilii*. *BMC Genomics* **14**, 906 (2013).
- Stewart, M. J. *et al.* Neuropeptides encoded by the genomes of the Akoya pearl oyster *Pinctada fucata* and Pacific oyster *Crassostrea gigas*: a bioinformatic and peptidomic survey. *BMC Genomics* **15**, 840 (2014).
- Veenstra, J. A. Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects. *Gen. Comp. Endocrinol.* **167**, 86–103 (2010).
- Zatylny-Gaudin, C. *et al.* Neuropeptidome of the cephalopod *Sepia officinalis*: identification, tissue mapping, and expression pattern of neuropeptides and neurohormones during egg laying. *J. Proteome Res.* **15**, 48–67 (2016).
- Zhang, G. G. *et al.* The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* **490**, 49–54 (2012).
- Riviere, G. *et al.* GigaTON: An extensive publicly searchable database providing a new reference transcriptome in the pacific oyster *Crassostrea gigas*. *BMC Bioinformatics* **16**, 401 (2015).
- Bauknecht, P. & Jékely, G. Large-scale combinatorial deorphanization of *Platynereis* neuropeptide GPCRs. *Cell Rep.* **12**, 684–693 (2015).
- Conzelmann, M. *et al.* Conserved MIP receptor-ligand pair regulates *Platynereis* larval settlement. *Proc Natl Acad Sci USA* **110**, 8224–8229 (2013).
- Bigot, L. *et al.* Functional characterization of a short neuropeptide F-related receptor in a lophotrochozoan, the mollusk *Crassostrea gigas*. *J. Exp. Biol.* **217**, 2974–2982 (2014).
- Li, S. *et al.* Adipokinetic hormones and their G protein-coupled receptors emerged in Lophotrochozoa. *Scientific Reports* 1–13, <https://doi.org/10.1038/srep32789> (2016).
- Dubos, M.-P. *et al.* Characterization of a tachykinin signalling system in the bivalve mollusk *Crassostrea gigas*. *Gen. Comp. Endocrinol.* **266**, 110–118, <https://doi.org/10.1016/j.ygcen.2018.05.003> (2018).
- Johnsen, A. H. Phylogeny of the Cholecystokinin/Gastrin Family. *Front. Neuroendocrinol.* **19**, 73–99 (1998).
- Makhlouf, G. M., McManus, J. P. & Card, W. I. The action of gastrin II on gastric-acid secretion in man. Comparison of the maximal secretory response to gastrin II and histamine. *Lancet (London, England)* **2**, 485–9 (1964).
- Koh, T. J. & Chen, D. Gastrin as a growth factor in the gastrointestinal tract. *Regul. Pept.* **93**, 37–44 (2000).
- Schmidt, W. E. *et al.* Role of CCK in regulation of pancreaticobiliary functions and GI motility in humans: effects of loxiglumide. *Am. J. Physiol.* **260**, G197–206 (1991).

21. Gibbs, J., Young, R. C. & Smith, G. P. Cholecystokinin elicits satiety in rats with open gastric fistulas. *Nature* **245**, 323–5 (1973).
22. Rehfeld, J. F. Cholecystokinin-From local gut hormone to ubiquitous messenger. *Front. Endocrinol. (Lausanne)*. **8**, 1–8 (2017).
23. Nachman, R. J., Holman, G., Haddon, W. F. & Ling, N. Leucosulfakinin, a sulfated insect neuropeptide with homology to Gastrin and Cholecystokinin. *Science (80-)*. **234**, 71–73 (1986).
24. Clynen, E. & Schoofs, L. Peptidomic survey of the locust neuroendocrine system. *Insect Biochem. Mol. Biol.* **39**, 491–507 (2009).
25. Zoepfel, J., Reiher, W., Rexer, K.-H., Kahnt, J. & Wegener, C. Peptidomics of the agriculturally damaging larval stage of the cabbage root fly *Delia radicum* (Diptera: Anthomyiidae). *PLoS One* **7**, e41543 (2012).
26. Torfs, P. *et al.* Isolation, identification, and synthesis of a disulfated sulfakinin from the central nervous system of an arthropod the white shrimp *Litopenaeus vannamei*. *Biochem. Biophys. Res. Commun.* **299**, 312–20 (2002).
27. Janssen, T. *et al.* Discovery of a cholecystokinin-gastrin-like signaling system in nematodes. *Endocrinology* **149**, 2826–2839 (2008).
28. Maestro, J. L. *et al.* Screening of antifeedant activity in brain extracts led to the identification of sulfakinin as a satiety promoter in the German cockroach. Are arthropod sulfakinins homologous to vertebrate gastrins-cholecystokinins? *Eur. J. Biochem.* **268**, 5824–30 (2001).
29. Nichols, R., Manoogian, B., Walling, E. & Mispelon, M. Plasticity in the effects of sulfated and nonsulfated sulfakinin on heart contractions. *Front. Biosci.* **14**, 4035–43 (2009).
30. Harshini, S., Nachman, R. J. & Sreekumar, S. *In vitro* release of digestive enzymes by FMRF amide related neuropeptides and analogues in the lepidopteran insect *Opisina arenosella* (Walk.). *Peptides* **23**, 1759–63 (2002).
31. Zels, S. *et al.* Sulfakinin is an important regulator of digestive processes in the migratory locust, *Locusta migratoria*. *Insect Biochem. Mol. Biol.* **61**, 8–16 (2015).
32. Wei, Z. *et al.* Sulfakinins reduce food intake in the desert locust, *Schistocerca gregaria*. *J. Insect Physiol.* **46**, 1259–1265 (2000).
33. Meyerling-Vos, M. & Müller, A. RNA interference suggests sulfakinins as satiety effectors in the cricket *Gryllus bimaculatus*. *J. Insect Physiol.* **53**, 840–848 (2007).
34. Yu, N. & Smagghe, G. CCK(-like) and receptors: Structure and phylogeny in a comparative perspective. *Gen. Comp. Endocrinol.* **209**, 74–81 (2014).
35. Vigna, S. R., Morgan, J. L. & Thomas, T. M. Localization and characterization of gastrin/cholecystokinin-like immunoreactivity in the central nervous system of *Aplysia californica*. *J. Neurosci.* **4**, 1370–7 (1984).
36. Dhainaut-Courtois, N., Dubois, M. P., Tramu, G. & Masson, M. Occurrence and coexistence in *Nereis diversicolor* O.F. Müller (Annelida Polychaeta) of substances immunologically related to vertebrate neuropeptides. *Cell Tissue Res.* **242**, 97–108 (1985).
37. Nachman, R. J. *et al.* Insect myosuppressins and sulfakinins stimulate release of the digestive enzyme  $\alpha$ -amylase in two invertebrates: The scallop *Pecten maximus* and insect *Rhynchophorus ferrugineus*. *Annals of the New York Academy of Sciences* **814**, (1997).
38. Dufresne, M., Seva, C. & Fourmy, D. Cholecystokinin and Gastrin Receptors. *Physiol. Rev.* **86**, 805–847 (2006).
39. Zatylny-Gaudin, C. & Favrel, P. Diversity of the RFamide peptide family in mollusks. *Front. Endocrinol. (Lausanne)*. **5**, 1–14 (2014).
40. Johnsen, A. H. & Rehfeld, J. F. Cionin: A disulfotyrosyl hybrid of cholecystokinin and gastrin from the neural ganglion of the protochordate *Ciona intestinalis*. *J. Biol. Chem.* **265**, 3054–3058 (1990).
41. Sekiguchi, T., Ogasawara, M. & Satake, H. Molecular and functional characterization of cionin receptors in the ascidian, *Ciona intestinalis*: the evolutionary origin of the vertebrate cholecystokinin/gastrin family. *J. Endocrinol.* **213**, 99–106 (2012).
42. Veenstra, J. A. Neuropeptide evolution: neurohormones and neuropeptides predicted from the genomes of *Capitella teleta* and *Helobdella robusta*. *Gen. Comp. Endocrinol.* **171**, 160–75 (2011).
43. McVeigh, P., Leech, S., Marks, N. J., Geary, T. G. & Maule, A. G. Gene expression and pharmacology of nematode NLP-12 neuropeptides. *Int. J. Parasitol.* **36**, 633–640 (2006).
44. Elphick, M. R. & Mirabeau, O. The Evolution and Variety of RFamide-Type Neuropeptides: Insights from Deuterostomian Invertebrates. *Front. Endocrinol. (Lausanne)*. **5**, 93 (2014).
45. Griffiths, J. R. & Unwin, R. D. Analysis of protein post-translational modifications by mass spectrometry. (Wiley and Sons, 2016).
46. Kubiak, T. M. *et al.* Cloning and functional expression of the first *Drosophila melanogaster* sulfakinin receptor DSK-R1. *Biochem. Biophys. Res. Commun.* **291**, 313–20 (2002).
47. Zels, S. *et al.* Signaling properties and pharmacological analysis of two sulfakinin receptors from the red flour beetle, *Tribolium castaneum*. *PLoS One* **9**, e94502. (2014).
48. Gigoux, V. *et al.* Arginine 197 of the cholecystokinin-A receptor binding site interacts with the sulfate of the peptide agonist cholecystokinin. *Protein Sci.* **8**, 2347–54 (1999).
49. Duve, H. *et al.* The sulfakinins of the blowfly *Calliphora vomitoria*. Peptide isolation, gene cloning and expression studies. *Eur. J. Biochem.* **640**, 633–640 (1995).
50. Slocinska, M., Marciniak, P., Jarmuszkiewicz, W. & Rosinski, G. New metabolic activity of the nonsulfated sulfakinin Zopat-SK-1 in the insect fat body. *Peptides* **68**, 157–163 (2015).
51. Nichols, R. The first nonsulfated sulfakinin activity reported suggests nsDSK acts in gut biology. *Peptides* **28**, 767–773 (2007).
52. Yule, D. I., Tseng, M. J., Williams, J. A. & Logsdon, C. D. A cloned CCK-A receptor transduces multiple signals in response to full and partial agonists. *Am. J. Physiol.* **265**, G999–1004 (1993).
53. Nässel, D. R. & Williams, M. J. Cholecystokinin-like peptide (DSK) in *Drosophila*, not only for satiety signaling. *Front. Endocrinol. (Lausanne)*. **5**, 1–5 (2014).
54. Nachman, R. J., Holman, G. M., Cook, B. J., Haddon, W. F. & Ling, N. Leucosulfakinin-II, a blocked sulfated insect neuropeptide with homology to cholecystokinin and gastrin. *Biochem. Biophys. Res. Commun.* **140**, 357–364 (1986).
55. Al-Alkawi, H., Lange, A. B. & Orchard, I. Cloning, localization, and physiological effects of sulfakinin in the kissing bug, *Rhodnius prolixus*. *Peptides* **98**, 15–22 (2017).
56. Haselton, A. T., Yin, C.-M. & Stoffolano, J. G. The effects of *Calliphora vomitoria* Tachykinin-I and the FMRFamide-related peptide Perisulfakinin on female *Phormia regina* crop contractions, *in vitro*. *J. Insect Physiol.* **52**, 436–441 (2006).
57. Söderberg, J. A. E., Carlsson, M. A. & Nässel, D. R. Insulin-producing cells in the *Drosophila* brain also express satiety-inducing cholecystokinin-like peptide, drosulfakinin. *Front. Endocrinol. (Lausanne)*. **3**, 1–13 (2012).
58. Yu, N., Nachman, R. J. & Smagghe, G. Characterization of sulfakinin and sulfakinin receptor and their roles in food intake in the red flour beetle *Tribolium castaneum*. *Gen. Comp. Endocrinol.* **188**, 196–203 (2013).
59. Rodet, F., Lelong, C., Dubos, M.-P., Costil, K. & Favrel, P. Molecular cloning of a molluscan gonadotropin-releasing hormone receptor orthologue specifically expressed in the gonad. *Biochim Biophys Acta* **1730**, 187–195 (2005).
60. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–7 (2004).
61. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729 (2013).
62. Fleury, E. *et al.* Generation and analysis of a 29,745 unique Expressed Sequence Tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase. *BMC Genomics* **15**, 1–15 (2009).
63. Offermanns, S. & Simon, M. I. G alpha 15 and G alpha 16 couple a wide variety of receptors to phospholipase C. *J. Biol. Chem.* **270**, 15175–80 (1995).
64. Endress, M. *et al.* Crustacean cardioactive peptides: Expression, localization, structure, and a possible involvement in regulation of egg-laying in the cuttlefish *Sepia officinalis*. *Gen. Comp. Endocrinol.* **260**, 67–79 (2018).
65. Thompson, J., Higgins, D. & Gibson, T. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–80 (1994).

## Acknowledgements

This work was funded by the ANR project “NEMO” (ANR 14CE02 0020). J. Schwartz PhD fellowship was co-financed by the NEMO project and by the European Union within the framework of the operational program FEDER/FSE 2014–2020. We are grateful to Professor Raphaël Delépée and the ICORE facility “PRISMM” for the mass spectrometry analyses and to Dr Emilie Réalis-Doyelle for her help in statistical analyses

## Author Contributions

J.S. and P.F. designed the experiments. J.S., M.P.D. and C.Z.G. performed the experiments, J.S. and J.P. performed the phylogenetic analyses. C.Z.G. analysed the mass spectrometry data. J.S., C.Z.G., M.P.D., J.P. and P.F. analysed the data. J.S. and P.F. wrote the paper.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-34700-4>.

**Competing Interests:** The authors declare no competing interests.

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018