University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

8-30-2022

AI protein structure prediction-based modeling and mutagenesis of a protostome receptor and peptide ligands reveal key residues for their interaction

Shi-Qi Guo

Ya-Dong Li

Ping Chen

Guo Zhang

Hui-Ying Wang

See next page for additional authors

Authors

Shi-Qi Guo, Ya-Dong Li, Ping Chen, Guo Zhang, Hui-Ying Wang, Hui-Min Jiang, Wei-Jia Liu, Ju-Ping Xu, Xue-Ying Ding, Ping Fu, Ke Yu, Hai-Bo Zhou, James W. Checco, and Jian Jing

AI protein structure prediction-based modeling and mutagenesis of a protostome receptor and peptide ligands reveal key residues for their interaction

Shi-Qi Guo[1,‡], Ya-Dong Li[1,‡], Ping Chen[1,‡], Guo Zhang[1,‡], Hui-Ying Wang[1], Hui-Min Jiang[1], Wei-Jia Liu[1], Ju-Ping Xu[1], Xue-Ying Ding[1], Ping Fu[1], Ke Yu[1], Hai-Bo Zhou[2,3,*], James W. Checco[4,*], and Jian Jing[1,3,5,*]

From the [1]State Key Laboratory of Pharmaceutical Biotechnology, Institute for Brain Sciences, Chinese Academy of Medical Sciences Research Unit of Extracellular RNA, Jiangsu Engineering Research Center for MicroRNA Biology and Biotechnology, Advanced Institute for Life Sciences, Chemistry and Biomedicine Innovation Center, School of Life Sciences, Nanjing University, Nanjing, Jiangsu, China; [2]School of Electronic Science and Engineering, Nanjing University, Nanjing, Jiangsu, China; [3]Peng Cheng Laboratory, Shenzhen, China; [4]Department of Chemistry and the Nebraska Center for Integrated Biomolecular Communication (NCIBC), University of Nebraska-Lincoln, Lincoln, Nebraska, USA; [5]Department of Neuroscience and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, New York, USA

Edited by Kirill Martemyanov

The protostome leucokinin (LK) signaling system, including LK peptides and their G protein-coupled receptors, has been characterized in several species. Despite the progress, molecular mechanisms governing LK peptide–receptor interactions remain to be elucidated. Previously, we identified a precursor protein for *Aplysia* leucokinin-like peptides (ALKs) that contains the greatest number of amidated peptides among LK precursors in all species identified so far. Here, we identified the first ALK receptor from *Aplysia*, ALKR. We used cell-based IP1 activation assays to demonstrate that two ALK peptides with the most copies, ALK1 and ALK2, activated ALKR with high potencies. Other endogenous ALK-derived peptides bearing the FXXWX-amide motif also activated ALKR to various degrees. Our examination of cross-species activity of ALKs with the *Anopheles* LK receptor was consistent with a critical role for the FXXWX-amide motif in receptor activity. Furthermore, we showed, through alanine substitution of ALK1, the highly conserved phenylalanine (F), tryptophan (W), and C-terminal amidation were each essential for receptor activation. Finally, we used an artificial intelligence–based protein structure prediction server (Robetta) and Autodock Vina to predict the ligand-bound conformation of ALKR. Our model predicted several interactions (*i.e.*, hydrophobic interactions, hydrogen bonds, and amide-pi stacking) between ALK peptides and ALKR, and several of our substitution and mutagenesis experiments were consistent with the predicted model. In conclusion, our results provide important information defining possible interactions between ALK peptides and their receptors. The workflow utilized here may be useful for studying other ligand–receptor interactions for a neuropeptide signaling system, particularly in protostomes.

Neuropeptides are the most diverse class of neuromodulators in both protostomes and deuterostomes (1–5). Neuropeptides primarily act on G protein-coupled receptors (GPCRs) to influence a variety of behaviors and physiological processes, including feeding, locomotion, and reproduction. Although a growing number of neuropeptides and their receptors have been characterized (6–11), molecular-level details of how peptide ligands engage their receptors remain poorly understood in many cases. Ideally, this information is obtained through high-resolution structures of a GPCR in a bound state with its peptide ligand (*e.g.*, by X-ray crystallography, NMR, or cryogenic-EM) (12–15). However, obtaining high-resolution structures remains a daunting task for most GPCRs, particularly for neuropeptide signaling systems that are only present in protostomes (most invertebrates) that lack well-studied homologs (12, 13, 15). Thus, previous work has often used amino acid substitution and other experiments to characterize the roles of specific residues in a ligand that may be critical for receptor activity (10). Efforts are also made to infer receptor activity of ligands based on ligands' structure in solution (16–24), although these analyses generally do not allow one to draw conclusions about specific interactions between ligands and the receptor. Moreover, previous work has analyzed the ligand–receptor interactions based on homology modeling of the structures of protostome receptors with their vertebrate homologs (*e.g.*, insect receptors for cholecystokinin (CCK) (25) and Neuromedin U (26)). However, few have explored the contributions of specific residues or other properties of ligands to receptor activity based on the structure of a protostome's receptor that has no known homologs in deuterostomes (vertebrates and some invertebrates), partly because a protein structure cannot be obtained using a homology modeling approach. In the present work, we utilize a molluscan model system, *Aplysia californica* (11, 27–59), to study this issue using *Aplysia* leucokinin peptides (ALKs) (60) and their receptor. In particular, recent successful efforts (14) have been made to

predict protein structure based on the amino acid sequence of a protein, particularly template-free modeling (61) using artificial intelligence (AI) deep machine learning algorithms such as Robetta (62) and AlphaFold (63). Presumably, the AI prediction approach could be applied to protostome proteins that have no homologs in deuterostomes. Thus, we sought to demonstrate this applicability using the ALK signaling system by taking advantage of the Robetta server, which is freely available, to obtain a predicted receptor structure. We then used Autodock Vina (64, 65) to predict the bound conformations of the ligands (including their analogs) with the receptor.

The leucokinin (LK) signaling system is known to be present only in protostomes (6, 7, 66, 67). LK peptides were first identified in cockroach *Leucophaea maderae* (now named *Rhyparobia maderae*) through bioassays on hindgut contractions (68–72). Subsequently, a number of LK peptides and some of their precursor proteins have been identified in arthropods, tardigrades, annelids, and molluscs. The C-termini of LK peptides share a FXXWX-amide motif (Fig. 1). LKs play diverse roles in the regulation of ion and water homeostasis, feeding, sleep–metabolism interactions, state-dependent memory formation, as well as modulation of gustatory sensitivity and nociception (66, 67). Interestingly, the first LK receptor (LKR) was found in the mollusc *Lymnaea* (73) and subsequently in several insect species (74–80). For all LKRs thus far identified, there is only a single receptor in each species. Prior studies have investigated activity of ligands on the LKR in each species, but no study has characterized the mechanisms of the ligand–receptor interactions at molecular resolution.

We have previously identified an *Aplysia* leucokinin (ALK) precursor protein that encodes up to 40 putative ALK peptides (60), representing one of the longest neuropeptide precursors known. The diverse ALK peptides generated from this precursor share the FXXWX-amide motif present in LKs in other invertebrates (Fig. 1). Here, we describe a newly identified receptor for ALK peptides, termed ALK receptor (ALKR). We took advantage of the diverse ALK ligands and showed that all the native ALKs could activate the receptor, albeit with different potencies. Mutagenesis studies of the ligands and receptor demonstrated that the conserved amino acids and amidation in different LKs are critical for receptor activity and gave insight into the roles of specific receptor residues critical for ligand-induced activation. Together with an AI-predicted model of peptide–receptor complex, these computational and experimental analyses elucidated the specific roles of several residues in both the peptide and receptor in this interaction. Overall, the results provide detailed information on the ALK–ALKR interaction and support the effectiveness of AI prediction of structures of protostome proteins.

Results

Identification of a putative ALKR

The ALK precursor protein has been previously characterized (60). Here, we sought to identify a receptor in *Aplysia* for ALK peptides. We used the receptor for lymnokinin, which is

A *Aplysia* **LK peptides**

Peptides	Sequences
ALK1 (20 copies)	PA**F**HS**W**S-**amide**
ALK2 (7 copies)	PA**F**HA**W**S-**amide**
ALK3 (2 copies)	PA**F**SA**W**S-**amide**
ALK4	AG**F**AP**W**A-**amide**
ALK5	PA**F**SA**W**N-**amide**
ALK6	SEKQPA**F**HA**W**S-**amide**
ALK7	PR**F**HA**W**S-**amide**
ALK8	PA**F**NA**W**G-**amide**
ALK9	PA**F**SP**W**S-**amide**
ALK10	PA**F**HA**W**G-**amide**
ALK11	PAA**F**HA**W**S-**amide**
ALK12	LAA**F**HA**W**S-**amide**
ALK13	PA**F**QA**W**S-**amide**

ALK4 - ALK13 are all single copies.
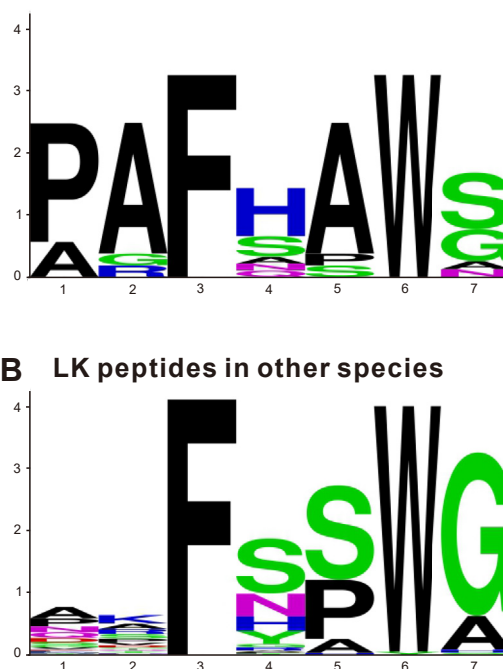

B **LK peptides in other species**

Figure 1. Comparison of ALK peptides and LK peptides in other species. *A*, primary sequences and sequence logo plot for the C-terminus of ALK peptides relevant to this study. All ALK peptides have a conserved FXXWX-amide motif. The shared residues and amidation are shown in *bold*. Note that the consensus sequence is identical to ALK2. *B*, sequence logo plot of LK peptide C-terminal conserved sequences in other species (excluding *Aplysia*) (see Table S4 for information on the selected sequences). The sequence logo plots in (*A*) and (*B*) show that shared residues in ALKs and LKs are similar. ALK, *Aplysia* leucokinin-like peptide; LK, leucokinin.

the mollusc *Lymnaea stagnalis* homolog of LK peptides (NCBI accession: U84499.1), as a query to perform a BLASTn search of NCBI GeneBank, and found a sequence (XM_013090833.1). The protein it encodes contains 205 amino acids (XP_012946287.1) but appears to be incomplete (see below). It is located on the *Aplysia* genomic sequence NW_004798839.1, which has 72,762 bp (Fig. 2A).

We then performed BLASTn search of *Aplysia* transcriptome databases (http://aplysiatools.org:4567/) and
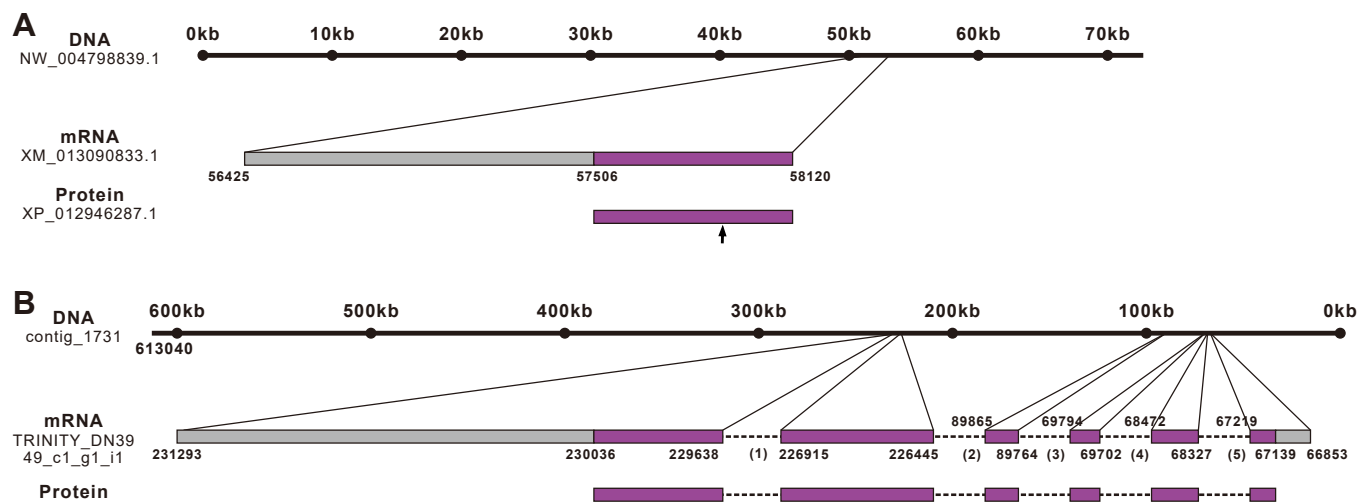
Figure 2. Gene expression mapping of ALKR. *A*, BLASTn search result of Genbank showing a DNA that encodes a partial sequence of the ALKR. There is no intron in the DNA, but the *arrow* indicates where a possible intron is located as shown in (*B*). *B*, BLASTn search result of *Aplysia* transcriptome databases and genome databases showing a DNA that encodes a complete sequence of ALKR. The DNA sequence was a complementary sequence of the genome, so the nucleotide number on *top* starts from the *right*. Exons were drawn to scale but introns were not. The length of intron 1: 2723 bp; intron 2: 136,580 bp; intron 3: 19,970 bp; intron 4: 1230 bp; intron 5: 1108 bp. Approximate locations of the exons and introns on the genome are marked. ALK, *Aplysia* leucokinin-like peptide; ALKR, *Aplysia* leucokinin-like peptide receptor.

genome databases (http://aplysiatools.org:8080/) using the *Lymnaea* lymnokinin receptor. This search returned an mRNA with an ORF of 1290 bp encoding 429 amino acids in the transcriptome database. In *Aplysia* genome database, this mRNA was located in contig_1731 and consisted of six exons and five introns (Fig. 2*B*). There is a long intron between the second and third exons, about 136 kb. To verify whether the sequence was predicted to be a complete GPCR, we analyzed the sequence using NCBI-conserved domain database (81) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) and TMHMM 2.0 (82, 83) (https://services.healthtech.dtu.dk/service.php?TMHMM-2.0). The analysis showed that the identified protein sequence was indeed predicted to be a GPCR with seven transmembrane domains (TMs) (Fig. S1). In addition, the putative receptor also contains the conserved Asp-Arg-Tyr (DRY) motif (84), located in the second intracellular loop, and the conserved Asn-Pro-Xaa-Xaa-Tyr (NPXXY) motif (85) located in the seventh TM helix, suggesting that it is a Class A GPCR (86) (Fig. S2). This sequence had 76.9% similarity with the *Lymnaea* lymnokinin receptor (Table S1), therefore, we tentatively named it an ALKR. Note that the 205-residue protein sequence from NCBI (protein: XP_012946287.1) is identical to the first 205 residues from this putative ALKR, indicating the receptor sequence deposited on NCBI is only a partial sequence. We also analyzed sites for possible posttranslational modifications in the ALKR (Supporting Results and Discussion and Table S2).

We generated a phylogenetic tree of LKRs from selected species in *Arthropoda*, *Mollusca*, *Annelida*, and *Tardigrade* (Fig. 3). Among the selected species, the LKRs of *Lymnea stagnalis*, *Drosophila melanogaster*, *Aedes aegypti*, *Anopheles stephensi*, and *Rhipicephalus microplus* have been functionally characterized (73–80) (Table S1). We performed BLASTp search of NCBI databases using the above verified sequences from the five species and found several additional sequences that are not annotated as putative LKRs, but we putatively named them LKR (Table S1). Indeed, when we used these putative LKRs to perform BLASTp search of NCBI databases, the most similar sequence is actually one of the above five sequences functionally characterized. Currently, there is only one type of LKR in the vast majority of known species, and no other subtypes exist (see (67)) with the possible exception of *C. secundus* (see Discussion). The phylogenetic tree showed that the ALKR is closely related to LKRs from other molluscs, including *Lymnaea* (*i.e.*, lymnokinin receptor: 76.9%) and *Plakobranchus ocellatus* (Similarity: 72.6%).

Finally, we sought to clone the ALKR from the *Aplysia* complementary DNA (cDNA). We designed primers (Table S3) using the putative ALKR sequence and obtained a PCR product (Fig. S3) that was identical with the putative ALKR in transcriptome database.

Activation of the putative ALKR by ALK peptides

As shown previously (60), the ALK precursor is predicted by NeuroPred (http://stagbeetle.animal.uiuc.edu/cgi-bin/neuropred.py) (87) to code 40 unique putative neuropeptides, with a total of 66 copies including duplicate sequences. These 40 peptides were also detected using Matrix-Assisted Laser Desorption Ionization mass spectrometry (Fig. S4, see also Table S2 of (60)). Among them, 13 peptide sequences have the conserved FXXWX-amide motif at the C-terminus, and are predicted to be fully processed ALKs. The two peptides with the largest number of copies were named ALK1 (20 copies) and ALK2 (seven copies). Except ALK1 and ALK2, we named the neuropeptide with two copies (P505-S511) ALK3, and the other ten peptides were named ALK4-ALK13 according to their appearance order on the precursor (Fig. 1*A*). Among all ALKs, ALK11 and ALK12 are octapeptides, ALK6 is an undecapeptide, and the others are heptapeptides. A sequence logo plot of the peptides (Fig. 1*A*,
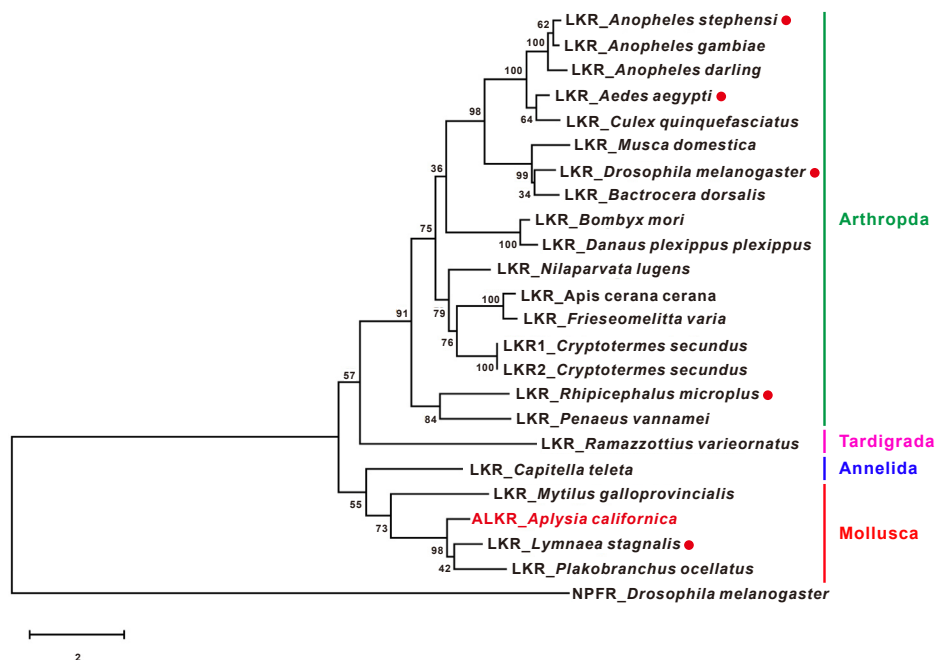
Figure 3. A phylogenetic tree of leucokinin receptors in protostomes. *Drosophila* neuropeptide F receptor (NPFR) was used as an out-group to root the tree. Amino acid sequences of full-length receptors were used for the analysis (see Table S1 for information on the selected sequences). Sequences were aligned using the Clustal W. Maximum-likelihood trees were constructed by MEGA X software using JTT+G method. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The numbers at the nodes of the branches represent the percentage bootstrap support (1000 replications) for each branch. LKRs that have been functionally characterized are indicated by a *red dot* after the species name. LKR, leucokinin receptor.

bottom panel) showed that aromatic phenylalanine and tryptophan residues are completely conserved in ALK sequences. C-terminal amidation is also conserved in all ALK peptides. These conserved elements are also found in LKs from other species (Fig. 1*B* and Table S4). Notably, the amino acid sequence with the highest frequency at each site of all ALKs was the same as ALK2 sequence. We also generated a sequence logo plot for most of the known LKs from different species other than ALKs (Fig. 1*B* and Table S4). Interestingly, the two frequency plots indicate that the amino acids that are most frequent at most positions are similar for the ALKs and LKs from other species. Only at the second position was there some minor difference (AGR for ALKs *versus* KAR for other species). This suggests that ALKs may be a good representation of LKs in different species.

To examine whether the native ALKs could activate the putative ALKR, we expressed the ALKR in Chinese hamster ovary K1 (CHO-K1) cells and examined ALK-mediated changes in the concentration of inositol monophosphate (IP1), a degradation product of the second messenger (inositol trisphosphate) in the G_q signaling pathway upon ligand-induced activation (88) (Figs. 4 and S5). In these experiments, we did not need to cotransfect with a promiscuous $G\alpha_q$ protein (see (9–11, 89)) to elicit IP1 accumulation upon ALK peptide stimulation, suggesting that ALKR can associate with native $G\alpha_q$ proteins in CHO-K1 cells. All endogenous ALK peptides could activate ALKR in a dose-dependent manner (Fig. 4, *A−C*). Nine of the 13 endogenous ALKs exhibited high potency, with EC_{50} values ranging from 10 nM to 22 nM. The lowest EC_{50} values were for ALK8 (EC_{50} = 10 nM) and ALK9 (EC_{50} = 10 nM). The EC_{50} values for ALK1, ALK2, ALK3, ALK4, ALK5, ALK10, ALK13 were slightly higher than that of ALK9, but there was no

statistically significant difference between them (Table S5). ALK6, ALK11, and ALK7 had a somewhat lower potency, with EC_{50} value of 32 nM, 62 nM, and 92 nM, respectively. ALK12 (EC_{50} = 330 nM) had the lowest activity and was significantly different than all other ALKs (Table S5).

To determine the selectivity of ALKs on the ALKR, we tested the effects of ALK1 on a different *Aplysia* receptor (9, 59), that is, the receptor for *Aplysia* allatotropin-like peptide (45). ALK1 did not show any activation of *Aplysia* allatotropin-like peptide receptor (Fig. S6*A*, *n* = 3). Conversely, *Aplysia* allatotropin-like peptide also did not have any effect on ALKR (Fig. S6*B*, *n* = 3).

Cross-activity with an insect (*A. stephensi*) receptor

Considering the similarity of leucokinin-like peptides in different species (Fig. 1), we sought to evaluate cross-activity between leucokinin peptides and their receptors in different species. We selected *A. stephensi* because it is a pest, and this would also provide an opportunity to determine if *Aplysia* LKs could be potentially used as insecticides (67). Specifically, we examined leucokinin 1 (LK1, the most active LK in *Anopheles*) and the LKR of *A. stephensi* for experiments with ALKR and the three most potent peptides (ALK1, ALK8, ALK9). In preliminary experiments, we found that the *Anopheles* receptor was not responsive to any of the LKs when there was no cotransfection with a promiscuous $G\alpha_q$ protein but was responsive when there was cotransfection with promiscuous $G\alpha_q$. Thus, we cotransfected promiscuous $G\alpha_q$ with the *Anopheles* receptor, whereas there was no cotransfection of promiscuous $G\alpha_q$ for ALKR. The results (Fig. 5) showed that both *Aplysia* and *A. stephensi* LKRs
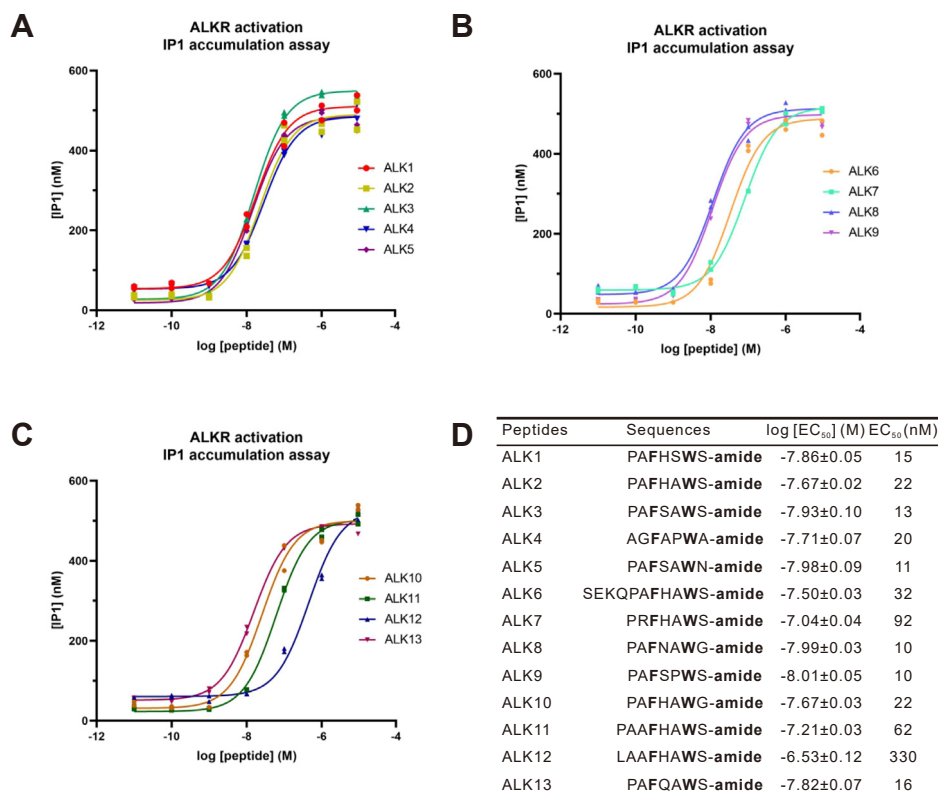
Figure 4. Activation of the ALKR by native ALK peptides. *A–C*, dose-response curves for ALK peptides on CHO-K1 cells transfected with ALKR, as determined by IP1 accumulation assay. Each point represents the value from an individual well on the plate, with each condition run in duplicate. *D*, sequences of all peptides tested and a summary of the average log [EC_{50}] and EC_{50} on ALKR. The log [EC_{50}] values are reported as the mean ± SEM. For ALK1, $n = 13$; for ALK2, $n = 4$; for other native peptides, $n = 3$. ALK, *Aplysia* leucokinin-like peptide; ALKR, *Aplysia* leucokinin-like peptide receptor; IP1, inositol monophosphate.

displayed low EC_{50} values of 3.8 to 15 nM when activated by their own endogenous peptides. The EC_{50} values of the cross-species neuropeptides and receptor pairs were significantly higher (ranging from 46 to 740 nM).

The roles of specific residues and amidation of the ALKs to receptor activation based on exploration of ligand–receptor interactions

To determine the influence of C-terminal amidation and each residue on activity of ALKs, we synthesized ALK analogs with each residue of ALK1 substituted by Ala (Fig. 6). The dose-response curves showed that the ALKR was not activated by ALK analogs when F3 or W6 (Fig. 1) were replaced by A or the C-terminal amidation was replaced with a carboxylic acid. In contrast, ALKR could be effectively activated by the other analogs. However, potency was significantly reduced when P1 or H4 of ALK1 was replaced with A (Fig. 6C). P1 substitution had the largest influence, and the EC_{50} value was 140 nM after it was replaced (ALK1_A1). H4 substitution had a smaller effect on ALK activity, and EC_{50} value was 51 nM after this residue was replaced (Fig. 6, *C* and *D*).

To examine the ligands' interaction with the receptor and possibly explain the activity of these ALKs and their analogs, we used Robetta protein prediction server (https://robetta. bakerlab.org/) (62) to generate a receptor structure (Figs. 7A and S7, A and B) and evaluated that the predicted model was appropriate and reliable based on the Ramachandran plot (90)

and QMEAN score (91) (Supporting Results and Discussion and Figs. S8–S11). This model had conserved bound conformations in extracellular loops 1 and 2 among peptide GPCRs, with extracellular loop 2 forming a β hairpin (15) (Fig. S7C). This structure was then imported into Autodock Vina (65) to generate docking results with ALK1 and its analogs (Fig. 7, *B–E*, Table S6, and Fig. S7). The molecular model predicted that ALK1, ALK2, and all of the active ALK analogs from the alanine scan bound to the ALKR putative-binding pocket in a similar conformation (Fig. 7B). In this conformation, L157 and I329 of the receptor formed two hydrophobic interactions with the peptide analogs (Table S6). Y213, Q317, and Q133 in ALKR formed three hydrogen bonds (H-bonds) with each of these effective analogs (Table S6). Q311 and Q250 in the ALKR formed amide-pi stacking interactions (for bound distances and angles, Table S6) (92, 93) with the effective analogs. Interestingly, the docking model also predicted that inactive analogs ALK1_A6 (W6 replaced by A) (Fig. 7C) and ALK1-OH (no C-terminal amidation) (Fig. 7D) bound to the receptor in the similar conformation as the active analogs, although some interactions differed (Supporting Results and Discussion). In contrast, ALK1_A3 (F3 replaced by A) did not bind to ALKR in the common conformation (Fig. 7E). Comparison of the ALKR and 12 LKRs in other species (Fig. S12) suggest that Q133, Q250, Q311, and I329 are completely conserved, while L157 and Q317 are moderately conserved (Supporting Results and Discussion).
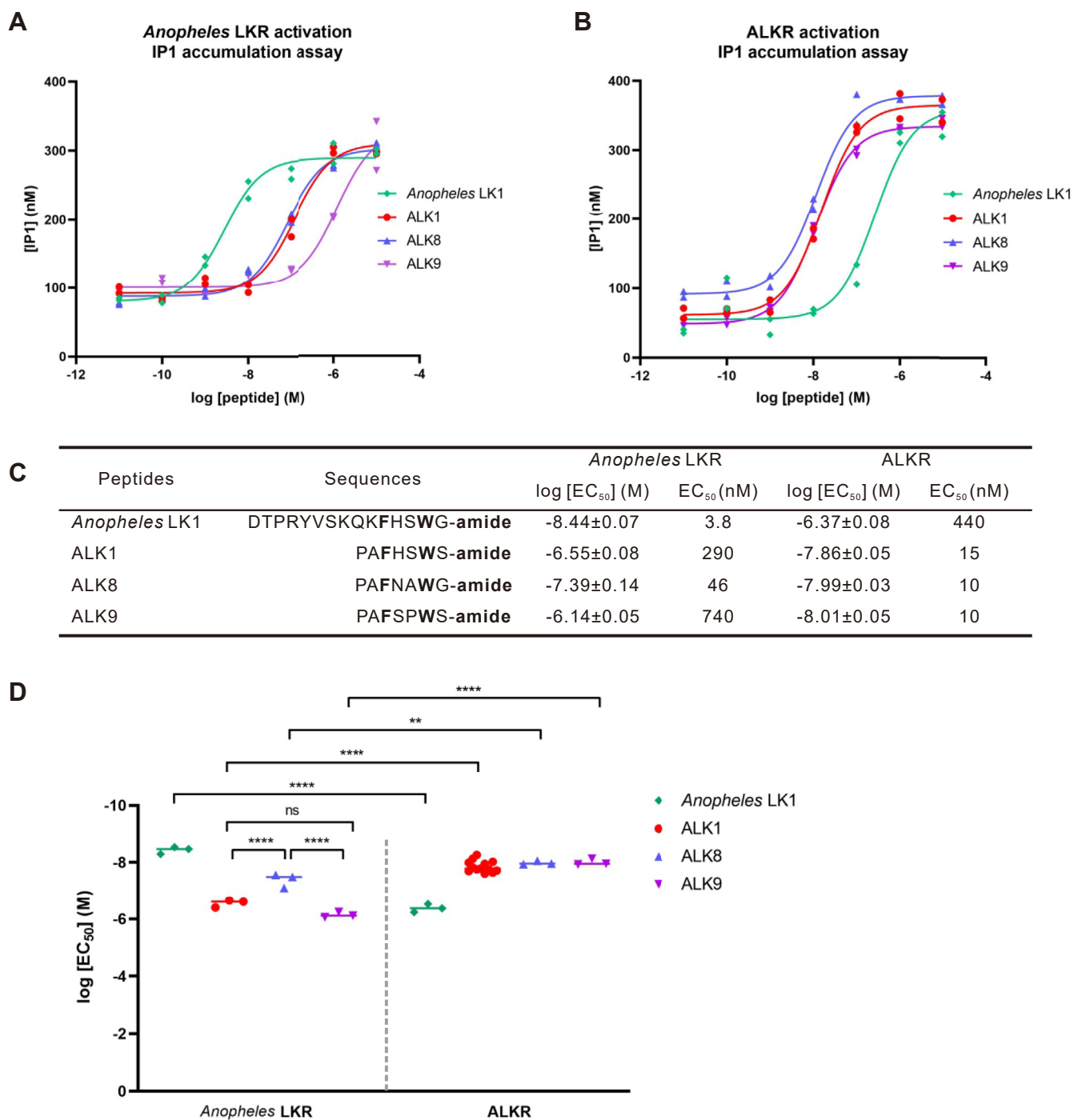
Figure 5. Activation of *Anopheles* LKR and ALKR by *Anopheles* LK1 and three ALKs. *A* and *B*, dose-response curves showing the ability of *Anopheles* LK1, ALK1, ALK8, and ALK9 to activate *Anopheles* LKR (*A*) or ALKR (*B*) transfected in CHO-K1 cells, as determined by IP1 accumulation assay. Each point represents the value from an individual well on the plate, with each condition run in duplicate. *C*, sequences of all peptides tested and summary of the average log [EC_{50}] and EC_{50} on ALKR. For active peptides, log [EC_{50}] values are reported as the mean ± SEM from at least three independent experiments. For ALK1 acting on ALKR, $n = 13$; in other cases, $n = 3$. *D*, for peptides shown in (*C*), log [EC_{50}] values showed overall significant difference (*F* (7, 26) = 88.06, $p < 0.0001$). Bonferroni post-hoc test: **$p < 0.01$; ****$p < 0.0001$. ALK, *Aplysia* leucokinin-like peptide; ALKR, *Aplysia* leucokinin-like peptide receptor; IP1, inositol monophosphate; LK, leucokinin; LKR, leucokinin receptor.

Mutagenesis of the ALKR

We next sought to determine whether the ALKR residues that are predicted to make contact with the peptide ligands in our model (L157, I329, Q133, Y213, Q317, Q250, Q311) are critical for receptor activation. To gain insight into the effect of each of these residues on receptor function, we generated

point mutants using site-directed mutagenesis (Fig. 8). When L157 and I329, which were each predicted to form hydrophobic interactions with the peptide ligands, were individually mutated to the polar Q, the EC_{50} values for ALK1 activation were significantly increased (77 nM and 6800 nM, respectively). When both of these two residues were simultaneously
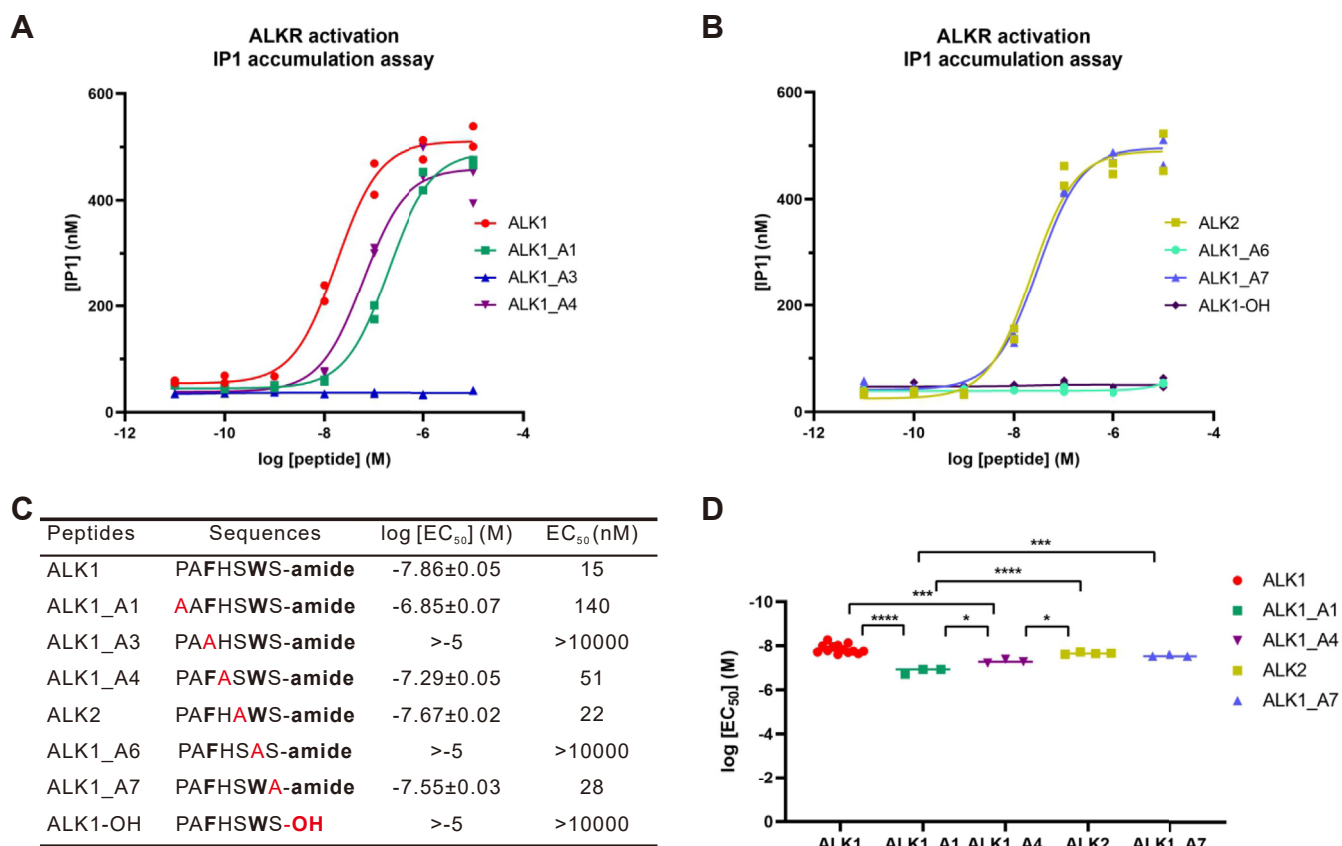
Figure 6. Activation of ALKR by ALK1 and its analogs generated by alanine substitution or by removing amidated C-terminus. *A* and *B*, dose-response curves for ALK1 analogs on CHO-K1 cells transfected with ALKR, as determined by IP1 accumulation assay. Each point represents the value from an individual well on the plate, with each condition run in duplicate. *C*, sequences of all peptides tested and a summary of the average log [EC_{50}] and EC_{50} on ALKR, as determined by IP1 accumulation assay. Alanine substitution or no C-terminal amidation are shown in *red*. For active peptides, log [EC_{50}] values are reported as the mean ± SEM. For inactive peptides, the log [EC_{50}] is listed as being greater than the highest concentration tested. For ALK1, $n = 13$; for ALK2, $n = 4$; for other ALK1 analogs, $n = 3$. *D*, for peptides shown in (*C*), log [EC_{50}] showed overall significant difference ($F (4, 21) = 27.49$, $p < 0.0001$). Bonferroni post-hoc test: *$p < 0.05$; ***$p < 0.001$; ****$p < 0.0001$. ALK, *Aplysia* leucokinin-like peptide; ALKR, *Aplysia* leucokinin-like peptide receptor; IP1, inositol monophosphate.

mutated to Q, the experimental EC_{50} value (35,000 nM) became much larger and significantly different from those of the single-site mutant receptors (Fig. 8*A*). The summation effect from the individual mutations was much lower than the combined effect of the simultaneous mutations, indicating a synergistic action between two hydrophobic interactions (Fig. 8*F*). For Q133, Y213, and Q317, which were predicted to form H-bonds with the peptide ligands, we found that EC_{50} from ALK1 was significantly increased to 78 nM, 370 nM, and 56 nM, respectively, when each of these residues were mutated to A. When all three residues were simultaneously mutated to A, the experimental EC_{50} (13,000 nM) became much larger and significantly different from those of the single-site mutated receptors (Fig. 8*B*). In the same fashion, the summation effect from the individual mutations were much lower than the combined effect of the simultaneous mutations indicating a synergistic action between three H-bonds (Fig. 8*G*). Note that when Q317 was mutated to N, which has similar properties as Q, the experimental EC_{50} of the mutated receptor was not significantly different from the ALKR (Fig. 8*C*). For residues Q250 and Q311, which were predicted to be involved in amide-pi stacking interactions, simultaneous mutation of both

of these residues to A significantly increased the EC_{50} value (19,000 nM) (Fig. 8*C*). These results indicate that the above residues play some roles in ligand activation of ALKR, and that the molecular docking results may be useful in predicting details of ligand–receptor interactions.

In addition to the residues predicted to directly interact with ALKs in the docking results, we also mutated the residue N314 to A (Fig. 8*C*). N314 corresponds to a relatively conserved residue (55% of N or Q) that might interact with ligands in other receptors for C-terminal amidated peptides (13) (see Table S7). In our model, N314 is within 3.3 to 3.4 Å of H4 in the effective peptide ligands, suggesting that these residues have the potential to interact (*e.g.*, *via* H-bonds). Consistent with this prediction, mutation of N314 to A significantly increased EC_{50} (120 nM), suggesting that N314 plays a role in receptor activity for ALKR (Supporting Results and Discussion).

To determine if mutagenesis could affect the expression of the receptors on the surface of CHO-K1 cells, we added FLAG tags to the ALKR and the mutant receptors (Figs. S13 and S14) to facilitate cell surface expression analysis. EC_{50} values for receptors with the FLAG tags (Fig. S13) showed a similar
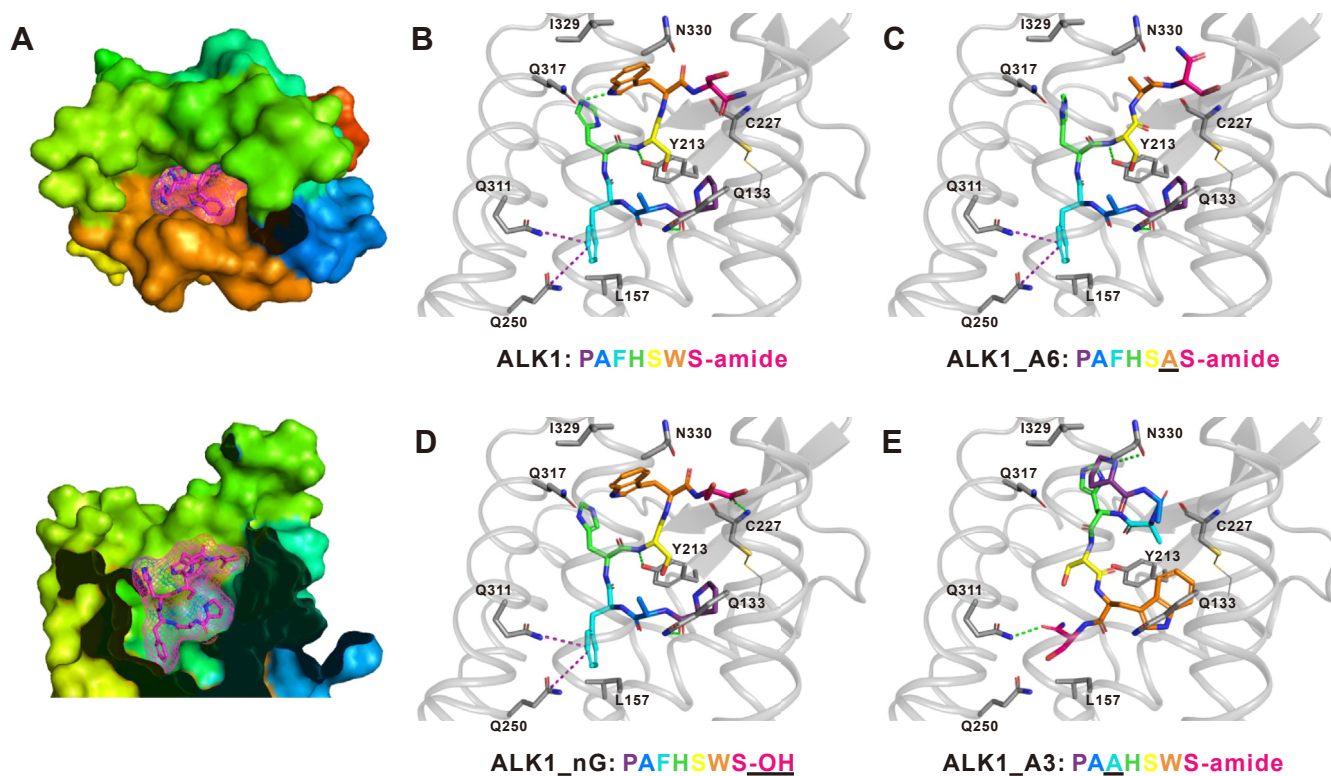
Figure 7. Docking results of ALK1 and ALK1 analogs with ALKR. *A*, global diagrams of docking of ALK1 with ALKR. The ALKR is colored coded with N-terminus shown in *dark blue*, the C-terminus shown in *red*. The ligand (ALK1) is shown in *purple*. *Upper panel*, top view after removing residues on ALKR N-terminus; *lower panel*, side view after removing residues in ALKR that are in front of the ALK1. *B–E*, binding regions of ALK1 (*B*), ALK1_A6 (*C*), ALK1-OH (*D*), ALK1_A3 (*E*) with ALKR. ALK1 is active, whereas the other three analogs are not active on the ALKR. For the three inactive analogs, the residue or the C-terminus that is different from ALK1 is *underlined*. ALK1, ALK1_A6, and ALK1-OH have a similar bound conformation, while ALK1_A3 had a different bound conformation. H-bonds are shown in *dotted green lines* and amide-pi interactions are shown in *dotted purple lines*. For the residue C227, the entire amino acid is shown because the N atom on its backbone formed an H-bond with ALK1-OH (*D*). For the other residues highlighted in the ALKR, only side chains are displayed. ALK, *Aplysia* leucokinin-like peptide; ALKR, *Aplysia* leucokinin-like peptide receptor.

pattern of changes for mutant receptors without the FLAG tags (Fig. S13, *D–F*). In addition, cell surface receptor expression experiments showed that expression of the mutant receptors did not change significantly from that of the ALKR (Fig. S14), suggesting that changes in EC$_{50}$ of mutant receptors (Figs. 8 and S13) compared to the ALKR were not due to changes in receptor expression.

Discussion

We have previously characterized an LK precursor protein in *Aplysia* and showed that LK signaling plays a significant role in modulating the feeding circuit (60). In this work, we have identified the ALKR in *Aplysia* for the first time and through bioinformatics, cell-based assays, computer modeling, and mutagenesis, demonstrated how diverse ALK ligands or their analogs may interact with the receptor.

Actions of ALKs on ALKR and cross-activity with Anopheles LKR

Although LK peptides were first identified in insects, previous work has shown that the LK signaling system is also important in molluscs, such as playing a major role in modulating feeding circuit in *Aplysia* (60). Interestingly, the first LKR was found in pond snail *L. stagnalis* (73). The LK

precursor in *Aplysia*, perhaps a longest neuropeptide precursor that is known, generates a number of ligands (60) with shared motifs that are representative of LKs in other species, including arthropods (Fig. 1). Our present identification of ALKR demonstrates multiple notable features in the receptor. First, there is a long intron of about 136 kb between the second and third exons, which is uncommon in other species. In fact, the incomplete sequence currently deposited in NCBI only contains the first exon and some partial sequence in the second exon (Fig. 2). Perhaps, the long intron after the second exon might be partly the reason why the ALKR sequence in NCBI is incomplete. Second, there are a number of sites for various potential posttranslational modifications (94) (See Supporting Results and Discussion). Finally, it should be noted that only a single LKR sequence (no subtypes) appears to exist in the vast majority of known species, including all six verified LKRs (Fig. 3) (67). (Note that *C. secundus* might have two sequences, LKR1 and LKR2, which are nearly identical except that LKR2 has 19 fewer amino acids at the N-terminus than LKR1).

Receptor activity studies using cell-based assays with recombinant receptor demonstrated that ALKR is indeed the receptor for ALK peptides. Notably, all 13 native ALK peptide ligands are able to activate the receptor with various potencies, with EC$_{50}$ values ranging from 10 to 330 nM (Figs. 4 and S4).
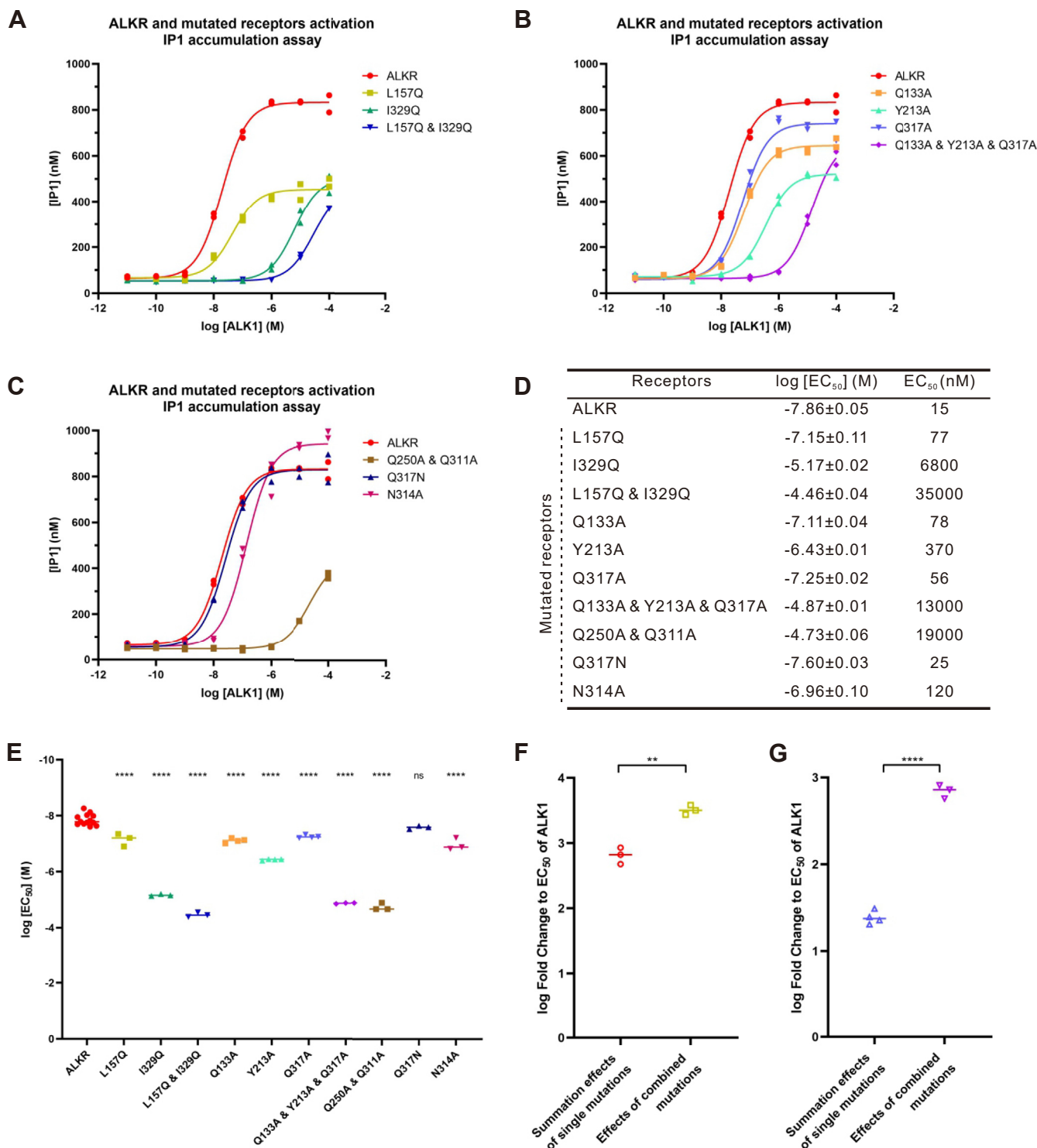
Figure 8. Mutagenesis of specific residues in the ALKR. These residues are predicted to be involved in hydrophobic interactions (L157 and I329), H-bonds (Q133, Y213, and Q317) or amide-pi stacking (Q250 and Q311), or are suspected to interact with the ligand (N314). *A–C*, dose-response curves showing the ability of ALK1 to activate the ALKR and the ALKR mutants expressed in CHO-K1 cells, as determined by IP1 accumulation assay. *D*, summary of the results shown in (*A–C*): ALKR ($n = 13$), L157Q ($n = 3$), I329Q ($n = 3$), L157Q & I329Q ($n = 3$), Q133A ($n = 4$), Y213A ($n = 4$), Q317A ($n = 4$), Q317N ($n = 4$), Q133A & Y213A & Q317A ($n = 3$), Q250A & Q311A ($n = 3$), and N314A ($n = 3$). *E*, group data comparing the effects of mutations ($F (10, 35) = 307.5$, $p < 0.0001$). Bonferroni post-hoc test: ns, not significant; ****$p < 0.0001$ (All comparisons were made against ALKR). *F*, fold change of L157Q & I329Q combined effects was compared with that of summation effects of L157/I329 single-site mutations using *t* test: **$p < 0.01$. *G*, fold change of Q133A & Y213A & Q317A combined effects was compared with that of summation effect of Q133/Y213/Q317 single-site mutations using *t* test: ****$p < 0.0001$. ALK, *Aplysia* leucokinin-like peptide; ALKR, *Aplysia* leucokinin-like peptide receptor; IP1, inositol monophosphate.

Because all of the native ALK peptides share the F3, W6 (using ALK1 numbering), and C-terminal amidation, these results suggested that variations in other amino acids can moderately affect ALK activity.

The cross-species activity between the LK peptides and receptors of *Aplysia* and *A. stephensi* (77) also provides information on the roles of various residues. LKs do activate each other's receptor, with EC_{50} ranging from 46 to 740 nM (Fig. 5), consistent with the idea that the shared residues and C-terminal amidation among LKs are important for receptor activity. However, the EC_{50} values for cross-species activity are significantly higher than the EC_{50} values for LKs activating native receptors, indicating that other residues in LK peptides also play roles in receptor potency. Interestingly, although ALK1, ALK8, ALK9 had similar EC_{50} values with the ALKR, their EC_{50} values with *Anopheles* LKR are significantly different, with ALK8 having the lowest EC_{50} value. This result further supports the observation that the residues other than F and W play roles in receptor activity. Given that LKs have been considered for pest control (67) and ALK8 had the lowest EC_{50} on *Anopheles* LKR and is shorter than *Anopheles* leucokinin 1 (7 *versus* 15 amino acid), ALK8 could potentially act as an economical insecticide. Finally, a notable difference between these receptors is that the *Aplysia* receptor does not need cotransfection of a promiscuous $G\alpha_q$ protein to couple to phospholipase C in CHO-K1 cells, while the *Anopheles* receptor does. This suggests that although they are orthologous, the molluscan and insect receptors might differ in their association with G proteins and intracellular signaling pathways.

Key residues involved in ligand–receptor interactions

The above analysis and previous work on the five receptors in insects and *Lymnaea* (73–80) provided some basis for understanding the roles of each residue and the amidation on the receptor activity, but further work is needed to clarify their roles. Here, we first evaluated what elements of the ALK1 peptide sequence is critical for receptor activity by removing amidation or substituting each residue with alanine. The findings provided direct evidence that amidation and highly conserved residues (F and W) are indeed essential for receptor activity, whereas the other residues play a lesser role (Fig. 6).

Importantly, we have taken a further step in studying ligand–receptor interactions using computer modeling and mutagenesis of the receptor. We used the protein structure prediction server Robetta (62) to generate an ALKR structure. Subsequently, we used Autodock Vina (65) to predict the bound conformations, H-bonds, hydrophobic interactions, and amide-pi stacking interactions between ALK1 or ALK1 analogs with ALKR (Fig. 7, and Supporting Results and Discussion). Specifically, two amino acid residue side chains in ALKR are predicted to directly engage in hydrophobic interactions with the ligand: L157 and I329. Three amino acid residues in ALKR are predicted to form hydrogen bonds with ALK peptides: Q133, Y213, and Q317. Finally, two amino acid residues from ALKR form amide-pi stacking interactions

(92, 93) with the ligands: Q250 and Q311. Moreover, comparison with other LKRs showed that several of the above important residues, including Q317, are largely conserved (Fig. S12), suggesting that similar ligand–receptor interacting mechanisms might operate in other LK signaling systems.

Notably, our ALKR mutagenesis experiments are consistent with the idea that hydrophobic interactions, H-bonds, and amide-pi stacking interactions may be involved in receptor activity and support the effectiveness of our molecular modeling. Our cell surface receptor expression experiments (Fig. S14) showed no significant changes for the expression of the mutant receptors compared with that of the ALKR, indicating that the changes in the activation of mutant receptors by ALK1 (Figs. 8 and S13) were not a result of changes in receptor expression. Overall, our computational and experimental evidence provided an account for specific roles of F, W, and amidation (affecting the bound confirmation) (Fig. 7, C–E). Specifically, F3 (Fig. 1) appears to be essential for the maintenance of the active conformation of ligands, as ALK1_A3 is not predicted to adopt the same conformation as active analogs when in the binding pocket. Similarly, the critical W6 residue of the ALK peptide is predicted to be involved in hydrophobic interactions (with I329) and H-bonds (with Q317) with ALKR. In cell-based assays, when either the nonpolar I329 was mutated to a polar Q or the polar Q317 was mutated to a nonpolar A, the experimental EC_{50} was increased. The I329Q mutation (EC_{50} = 6800 nM) had a much larger effect on EC_{50} than the Q317A mutation (EC_{50} = 56 nM), suggesting hydrophobic interactions involving W6 may be more critical for activity than possible H-bonding interactions.

Amidation is the most common posttranslational modification of neuropeptides and peptide hormones (*e.g.*, see (13)). For many peptides, this modification not only improves the stability and delivery of the peptide but is also necessary for its activity. At present, there are few crystal structures of a GPCR bound to an amidated peptide demonstrating direct interactions between the peptide C-terminal amide and the receptor. A validated homology model of amidated CCK4 with CCKR2 (*i.e.*, cholecystokinin B receptor) suggests that the amidated C-terminus forms an H-bond with the receptor (13, 95). The only currently available crystal structure is the bound conformation of amidated CCK8 with cholecystokinin A receptor, which directly demonstrated the presence of this H-bond (96). By contrast, our docking results did not find a specific interaction of the amidated C-terminus with ALKR. However, ALK1-OH, an ALK1 analog lacking C-terminal amidation, was not able to activate ALKR and does have difference in some interactions with ALKR, for example, missing the H-bond with Q317. Thus, the C-terminal amidation of different peptide ligands might have distinct roles during interactions with their receptors.

To examine the relevance of our work to other amidated peptide receptors, we also compared our findings with previous receptor comparison studies (13, 15) (See Supporting Results and Discussion and Table S7). Our results are consistent with a bound orientation of ALK peptides in which

both the N- and C-termini of the peptide ligand are within the putative-binding pocket of the receptor, with the N-terminus pointing deep in the pocket. These results contrast with several other known examples of neuropeptide–receptor interactions where one terminus is critical for receptor activation and this terminus is buried deep in the binding pocket (13). In these cases, the critical terminus is oriented toward the receptor while the other terminus points toward solution. In our case, we showed that amidation at the C-terminus is essential for ALK's activity in cell-based assays. Modeling then predicted that although the C-terminus is found in the binding pocket, it is oriented toward the extracellular region of the receptor whereas the N-terminus extends deeper into the pocket. There are a few other neuropeptides whose amino acids are all within the receptor-binding pockets, for example, thyrotropin-releasing hormone, which might be similar to ALKs. In the future, it will be interesting to determine how broadly the predicted bound conformations of ALKs are applicable to other peptide–receptor interactions.

We do note that although our work is relatively complete in its own right, there are outstanding issues that warrant future investigation. For example, can ALKR couple to G proteins other than those in the Gα_q family? Would ALKR behave similarly in cell lines other than CHO-K1 cells we used, for example, in *Drosophila* Schneider 2 cells (see also "IP1 accumulation assay" section in Experimental procedures)? What is the distribution of ALKR-positive neurons in the CNS of *Aplysia*? The answers to these and other questions could add important information to the mechanisms and functional roles of the LK signaling system in protostomes in general and in *Aplysia* in particular.

In summary, we have identified a LKR in *Aplysia* and studied ligand–receptor interactions using AI-based protein structure prediction, docking software, and mutagenesis of the receptor. Our results demonstrate how specific residues in both the peptide ligands and the receptor influence receptor activity, and molecular modeling predicts hydrophobic interactions, H-bonds, and amide-pi stacking that may be mediating this peptide–receptor interaction. We expect that our approach could be readily applied to other neuropeptide signaling systems, particularly to protostome peptide signaling systems that do not have homologs in deuterostomes. As more signaling systems in both protostomes and deuterostomes are being studied with an improved paradigm, we will gain a better insight on how broadly applicable the specific interactions we identified here are and on how diverse the operations of neuromodulatory systems in both protostomes and deuterostomes are.

Experimental procedures
Subjects

Experiments were performed on mollusc *A. californica* (100–350 g) obtained from Marinus. *Aplysia* are hermaphroditic (*i.e.*, each animal has reproductive organs normally associated with both male and female sexes). Animals were maintained in circulating artificial seawater at 14 to 16 °C and the animal room was equipped with a 12:12 h light-dark cycle with light period from 6:00 AM to 6:00 PM. All chemicals were purchased from Sigma-Aldrich unless otherwise stated.

Bioinformatic analysis of peptide precursors and receptors

Initially, we used NCBI to search specific sequences of interests. In addition, we also searched AplysiaTools databases (Dr Thomas Abrams, University of Maryland (58)) to obtain additional sequences for comparison. These latter databases (http://aplysiatools.org) include databases for *Aplysia* transcriptome and *Aplysia* genome.

The ORFs of the putative receptor full-length cDNA sequence was obtained using ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/). We compared the LK peptides and receptor sequences with those of other species using BioEdit software (https://bioedit.software.informer.com/7.2/). Sequence logo plots (aligned from C-terminus) for LK peptides were generated using a Weblogo software (http://weblogo.berkeley.edu/logo.cgi). For the putative ALKR, TMs were predicted using TMHMM Server v. 2.0 (82, 83) (http://www.cbs.dtu.dk/services/TMHMM/). The phylogenetic trees of sequences from different species were constructed by MEGA X software (https://www.megasoftware.net/) using alignment by Clustal W and the maximum likelihood method with 1000 replicates, and JTT+G method was performed (Fig. 3). The selection of the models was based on the results of MEGA analysis.

Cloning of receptor mRNA in Aplysia
RNA extraction

After anesthesia with 30 to 50% of the body weight with 333 mM MgCl$_2$, *Aplysia* cerebral, pleural-pedal, buccal, and abdominal ganglia were dissected out and maintained in artificial seawater containing the following: 460 mM NaCl, 10 mM KCl, 55 mM MgCl$_2$, 11 mM CaCl$_2$, and 10 mM Hepes, pH 7.6, in a dish lined with Sylgard (Dow Corning). RNA was prepared from the *Aplysia* ganglia using the Trizol reagent method. Specifically, the dissected ganglia were placed into 200 µl Trizol (Sigma, T9424) and stored at −80 °C until use. The frozen ganglia in Trizol were thawed and homogenized with a plastic pestle, then Trizol was added to a total volume of 1 ml, which were incubated at room temperature for 10 min. Then, 200 µl chloroform was added, and the solution was mixed thoroughly by shaking and incubated on ice for 15 min. The solution was centrifuged (12,000g, 4 °C, 15 min), and the supernatant was added to an equal volume of isopropanol. The tube was shaken gently by hand and let stand at −20 °C for 2 h. After 2 h, the solution was centrifuged (12,000g, 4 °C, 15 min) again, the supernatant was discarded, and 1 ml of 75% ethanol/water was added, and the centrifuge tube was shaken gently by hand to suspend the pellet. The tube was centrifuged (12,000g, 4 °C, 10 min), the supernatant discarded, and the precipitant was dried at room temperature for 5 to 10 min. Finally, 30 µl of nuclease-free water was added to dissolve the RNA pellet, and the RNA concentration was determined with a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific).

Reverse transcription

Using the above extracted RNA as a template, cDNA was synthesized by reverse transcription using PrimeScript RT Master Mix Kit (Takara, RR036A) according to the instructions and then stored at –20 °C until use. The synthesized first-strand cDNA serves as a template for PCR.

PCR

The synthesized cDNA above was used as a template for PCR. Each pair of specific primers was designed (Table S3) in Primer Premier 6 and Oligo7, based on protein coding sequences for the putative receptor. The PCR reaction was performed with 98 °C/2 min predenaturing, 98 °C/10 s denaturing, ~64 °C/15 s annealing, 72 °C/30 s extension, and 72 °C/5 min re-extension for 35 cycles. The PCR products were subcloned into vector pcDNA3.1(+) and sequenced to ensure the sequences were correct.

IP1 accumulation assay

IP1 accumulation assay measures concentration of IP1 (88), that is hydrolyzed from the second messenger, inositol trisphosphate, generated by $G\alpha_q$ pathway when GPCR expressed in CHO-K1 cells is activated by an appropriate ligand. We note two reasons that we selected mammalian CHO-K1 cells for assays for ligand-receptor (GPCR) activity: (1) there are no established molluscan cell lines available; (2) CHO-K1 cells have been extensively utilized not only for vertebrate GPCRs (*e.g.*, (97)) but also for invertebrate/protostome GPCRs, including other molluscs (9, 11, 73), annelids (8), jellyfish (98), *Caenorrhabditis elegans* (99), and arthropods (76, 78–80). Given that LKRs are only present in protostomes (6, 7, 66, 67), an additional advantage of using mammalian cells to assay ligand-receptor activity is that the effects of a possible native LKR in protostome cell lines could be excluded.

In order to express the *Aplysia* putative receptor transiently in CHO-K1, the cDNA was cloned into the mammalian expression vector pcDNA3.1(+). CHO-K1 cells (Procell, CL-0062) were cultured in F-12K medium (Gibco, 21127-022) with 10% fetal bovine serum (FBS, Genial, G11-70500) at 37 °C in 5% CO_2. Transfection experiments were performed when the cells were grown to 70 to 90% confluence. For the ALKR, in each dish (60-mm diameter), 4 μg of the putative receptor plasmids [in pcDNA3.1(+)] were mixed with 400 μl of Opti-MEM (Gibco, 11058021), followed by the addition of 15 μl of Turbofect (Thermo Fisher Scientific, R0531). For the *Anopheles* LKR, we could not obtain IP1 responses with the above procedure, suggesting that this receptor may not associate with the native $G\alpha_q$ in CHO-K1 cells. Thus, 3 μg of the putative receptor plasmids [in pcDNA3.1(+)] and 3 μg of promiscuous $G\alpha_q$ protein (89) [in pcDNA3.1(+)] were cotransfected in the above procedure for the *Anopheles* LKR (8, 10). The CHO-K1 cells with the reagents added above were mixed gently and incubated at room temperature for 15 min. The DNA/Turbofect mixture dropwise was then added to the dish, and the cells were incubated at 37 °C in 5% CO_2 overnight. The next day, the cells were trypsinized and reseeded in opaque white 96-well half-area (Corning, 3688) or

384-well tissue culture–treated plates (Corning, 3570) at a density of 20,000 cells/well in F-12K and 10% FBS and incubated at 37 °C in 5% CO_2 overnight. On the third day, the activation of the receptors was detected by monitoring IP1 accumulation using IP1 detection kit (Cisbio, 62IPAPEB) with a Tecan Spark plate reader. Except for using 0.5 × reagents for the anti-IP1-cryptate and IP1-d2 reagents, all other procedures were performed in accordance with the IP1 detection kit manufacturer's instructions.

Molecular modeling of the interactions of peptide ligands and the receptor

The topology files of peptide ligands, that is, ALK1, ALK2, and ALK1 analogs, were synthesized with SYBYL X-2.0 and optimized with Amber FF99SB force field. Max Iteration = 100,000, gradient = 0.005. The topology file of the receptor was predicted by Robetta Server (http://robetta.bakerlab.org/). Five receptor structures which differ primarily at the N- and C-termini were obtained from Robetta, and we selected the best model. The quality of the generated model was evaluated using the online servers of QMEAN (https://swissmodel.expasy.org/qmean) and PROCHECK v.3.5.4 (https://saves.mbi.ucla.edu/) (Supporting Results and Discussion).

Autodock Vina is used for semi-flexible docking, and all ligands adopt the same parameters for docking. We set a grid box (37.5 Å × 37.5 Å × 37.5 Å) centered at (36.787, –2.312, –15.122) Å. Due to flexibility of the peptide ligands, we modified the following two parameters: modes = 1000, exh. = 10. Other parameters were set to default. We performed molecular docking five times for each peptide and made sure the conformations that had the lowest affinity energy are similar at least four times. We then used the similar one for further analysis.

Mutagenesis of the ALKR

In a first set of experiments, mutagenesis of the ALKR was performed without FLAG tag (Fig. 8). Specifically, construction of the ALKR mutants was performed employing the full-length ALKR cDNA cloned into the pcDNA3.1(+) plasmid. Site-directed mutagenesis was performed using the site-directed mutagenesis kit (Sangon Biotech), following the manufacturer's instructions. Briefly, forward and reverse primers containing the expected mutation were mixed with kit components, and 10 ng of pcDNA3.1(+)-ALKR was used as the mutation template. After 14 to 18 rounds of PCR amplification, 1 μl of DpnI was added and incubated at 37 °C for 1 h in order to digest the template. The primers used to obtain the mutants were designed based on the ALKR cDNA sequence and listed in Table S3. Mutants were confirmed *via* DNA sequencing. IP1 accumulation assay with ALKR and mutants without FLAG tags (Fig. 8) was performed using procedures described in "IP1 accumulation assay" section.

In order to determine receptor expression, we added a FLAG tag to the ALKR and mutant receptors by taking advantage of the recombinant plasmids of pcDNA3.1(+)-ALKR (ALKR or mutants) we constructed above. For each mutant, the pcDNA3.1(+)-ALKR recombinant plasmid was first cleaved

with the restriction enzymes HindIII and ApaI to obtain the DNA sequence of the receptor. The cleavage product was subcloned into vector pcDNA3.1(+)-FLAG to obtain pcDNA3.1-FLAG-ALKRs (ALKR or mutants) recombinant plasmids (FLAG tag at the N-terminus of each respective receptor). The final products were confirmed by DNA sequencing.

To perform IP1 accumulation assay with the receptors with FLAG tags (Fig. S13), transfection experiments with recombinant plasmids were performed in accordance with the jetPRIME (Polyplus Transfection, PT-114-15) manufacturer's instructions. In each well (6-well plates), 2 μg of the ALKR (ALKR or mutants) recombinant plasmids [in pcDNA3.1(+)-FLAG] were mixed with 200 μl of jetPRIME buffer, followed by the addition of 4 μl of jetPRIME and incubated at room temperature for 10 min. The DNA/jetPRIME mixture dropwise was then added to the dish. The cells were incubated at 37 °C in 5% CO_2 overnight. Other procedures followed "IP1 accumulation assay" section described earlier.

Cell surface expression analysis by ELISA

Cell surface expression determination of ALKR and mutant receptors, all with FLAG tags (Fig. S14), was performed using a procedure modified from previous work (100–102). Specifically, CHO-K1 cells (Procell, CL-0062) were maintained in F-12K medium (Procell, PM150910) with 10% FBS (Genial, G11-70500) at 37 °C in 5% CO_2. Cells were cultured in 6-well plates (BIOFIL, TCP011006). When the cells were grown to 70 to 90% confluence, transfection experiments were performed in accordance with the jetPRIME (Polyplus Transfection, PT-114-15) manufacturer's instructions. Briefly, in each well, 2 μg of the receptor plasmids [in pcDNA3.1(+)-FLAG] were mixed with 200 μl of jetPRIME buffer, followed by the addition of 4 μl of jetPRIME and incubated at room temperature for 10 min. The DNA/jetPRIME mixture dropwise was then added to the dish. The cells were incubated at 37 °C in 5% CO_2 overnight. After 24 h, transfected cells were plated in 96-well white clear-bottom cell culture plates (Corning, 3610) at a density of 20,000 cells in 100 μl per well and incubated overnight. The following day, culture media was aspirated and cells were washed twice with 200 μl of 1× PBS (Procell, PB180327). Then, 100 μl of 1× PBS containing 5% (w/v) bovine serum albumin was added to each well and incubated at room temperature. After 30 min, 100 μl of 1:10,000 anti-FLAG M2-HRP conjugate (Sigma-Aldrich, Cat A8592) was added to each well and incubated for 30 min at 37 °C. Cells were washed twice with 200 μl of 1× PBS and then incubated with 200 μl of TMB Chromogen Solution (Sangon Biotech, E661007) for 30 min at 37 °C in the dark. Finally, 50 μl of ELISA Stopping Solution (Sangon Biotech, E661006) was added to each well to stop the reaction. The absorbance at 450 nm was measured using a Tecan Spark microplate reader. In each experiment, expression of the mutant ALKRs was assessed and compared with that measured for the ALKR in the same experiment.

Peptide and DNA synthesis

Peptides were synthesized by Synpeptide Co, Ltd, Guoping Pharmaceutical Co, Ltd, or ChinaPeptides Co, Ltd (Fig. S15)

and were aliquoted in 50 nmol per microcentrifuge tubes, stored at –20 °C until use. The DNA sequence of *Anopheles* LKR (Fig. S15) was synthesized by Tsingke Biotechnology Co, Ltd and was stored at –20 °C until use.

Data and statistical analyses

Dose-response curves and bar graphs for experimental data were plotted using Prism software (GraphPad). Data are expressed as the mean ± SEM. Final EC_{50} values are rounded to two significant figures. All experimental data were taken from individual animals or preparations, and n refers to the number of preparations unless otherwise stated. Statistical tests were performed using Prism software. They included Student's *t* test, one-way ANOVA, as appropriate. Data that showed significant effects in ANOVA were further analyzed in individual comparisons with Bonferroni's correction.

Data availability

Structural model and docking result files used to generate Figure 7 are available on github (https://github.com/li-yadong/ALKR-paper). All other data are included in this article and the supporting information. The nucleotide sequence(s) reported in this article has been submitted to the GenBank/EBI Data Bank with accession number(s): OP292655.

Abbreviations—The abbreviations used are: AI, artificial intelligence; ALK, *Aplysia* leucokinin-like peptide; ALKR, *Aplysia* leucokinin-like peptide receptor; CCK, cholecystokinin; cDNA, complementary DNA; CHO-K1 cells, Chinese hamster ovary K1 cells; FBS, fetal bovine serum; GPCRs, G protein-coupled receptors; IP1, inositol monophosphate; LK, leucokinin; LKR, leucokinin receptor; TMs, transmembrane domains.

References

1. Taghert, P. H., and Nitabach, M. N. (2012) Peptide neuromodulation in invertebrate model systems. *Neuron* **76**, 82–97
2. Nusbaum, M. P., Blitz, D. M., and Marder, E. (2017) Functional consequences of neuropeptide and small-molecule co-transmission. *Nat. Rev. Neurosci.* **18**, 389–403
3. Cropper, E. C., Jing, J., Vilim, F. S., Barry, M. A., and Weiss, K. R. (2018) Multifaceted expression of peptidergic modulation in the feeding system of *Aplysia*. *ACS Chem. Neurosci.* **9**, 1917–1927
4. Abid, M. S. R., Mousavi, S., and Checco, J. W. (2021) Identifying receptors for neuropeptides and peptide hormones: challenges and recent progress. *ACS Chem. Biol.* **16**, 251–263
5. Zhang, G., Guo, S. Q., Wang, H. Y., Li, Y. D., Jiang, H. M., Yu, K., *et al.* (2022) Functional studies on neuropeptides and receptors in model animal *Aplysia* (in Chinese). *Sci. Sin Vitae* **52**, 387–402
6. Jekely, G. (2013) Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 8702–8707
7. Mirabeau, O., and Joly, J. S. (2013) Molecular evolution of peptidergic signaling systems in bilaterians. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E2028–E2037
8. Bauknecht, P., and Jekely, G. (2015) Large-scale combinatorial deorphanization of *Platynereis* neuropeptide GPCRs. *Cell Rep.* **12**, 684–693
9. Checco, J. W., Zhang, G., Yuan, W. D., Le, Z. W., Jing, J., and Sweedler, J. V. (2018) *Aplysia* allatotropin-related peptide and its newly identified d-amino acid-containing epimer both activate a receptor and a neuronal target. *J. Biol. Chem.* **293**, 16862–16873
10. Checco, J. W., Zhang, G., Yuan, W. D., Yu, K., Yin, S. Y., Roberts-Galbraith, R. H., *et al.* (2018) Molecular and physiological characterization of a receptor for d-amino acid-containing neuropeptides. *ACS Chem. Biol.* **13**, 1343–1352
11. Jiang, H. M., Yang, Z., Xue, Y. Y., Wang, H. Y., Guo, S. Q., Xu, J. P., *et al.* (2022) Identification of an allatostatin C signaling system in mollusc *Aplysia*. *Sci. Rep.* **12**, 1213
12. Wu, F., Song, G., de Graaf, C., and Stevens, R. C. (2017) Structure and function of peptide-binding G protein-coupled receptors. *J. Mol. Biol.* **429**, 2726–2745
13. Tikhonova, I. G., Gigoux, V., and Fourmy, D. (2019) Understanding peptide binding in class A G protein-coupled receptors. *Mol. Pharmacol.* **96**, 550–561
14. Dhakal, A., McKay, C., Tanner, J. J., and Cheng, J. (2022) Artificial intelligence in the prediction of protein-ligand interactions: recent advances and future directions. *Brief Bioinform.* **23**, bbab476
15. Vu, O., Bender, B. J., Pankewitz, L., Huster, D., Beck-Sickinger, A. G., and Meiler, J. (2021) The structural basis of peptide binding at class A G protein-coupled receptors. *Molecules* **27**, 210
16. Edison, A. S., Espinoza, E., and Zachariah, C. (1999) Conformational ensembles: the role of neuropeptide structures in receptor binding. *J. Neurosci.* **19**, 6318–6326
17. Miskolzie, M., Lucyk, S., and Kotovych, G. (2003) NMR conformational studies of micelle-bound orexin-B: a neuropeptide involved in the sleep/awake cycle and feeding regulation. *J. Biomol. Struct. Dyn.* **21**, 341–351
18. Chakrabarti, R., Klibanov, A. M., and Friesner, R. A. (2005) Computational prediction of native protein ligand-binding and enzyme active site sequences. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10153–10158
19. Brylinski, M., and Skolnick, J. (2008) A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 129–134
20. McGann, M. (2011) FRED pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **51**, 578–596
21. Swann, S. L., Brown, S. P., Muchmore, S. W., Patel, H., Merta, P., Locklear, J., *et al.* (2011) A unified, probabilistic framework for structure- and ligand-based virtual screening. *J. Med. Chem.* **54**, 1223–1232
22. Lawrenz, M., Shukla, D., and Pande, V. S. (2015) Cloud computing approaches for prediction of ligand binding poses and pathways. *Sci. Rep.* **5**, 7918
23. Moretti, R., Bender, B. J., Allison, B., and Meiler, J. (2016) Rosetta and the design of ligand binding sites. *Methods Mol. Biol.* **1414**, 47–62
24. Do, T. D., Checco, J. W., Tro, M., Shea, J. E., Bowers, M. T., and Sweedler, J. V. (2018) Conformational investigation of the structure-activity relationship of GdFFD and its analogues on an achatin-like neuropeptide receptor of *Aplysia californica* involved in the feeding circuit. *Phys. Chem. Chem. Phys.* **20**, 22047–22057
25. Yu, N., Zotti, M. J., Scheys, F., Braz, A. S., Penna, P. H., Nachman, R. J., *et al.* (2015) Flexibility and extracellular opening determine the interaction between ligands and insect sulfakinin receptors. *Sci. Rep.* **5**, 12627
26. Kawai, T., Katayama, Y., Guo, L., Liu, D., Suzuki, T., Hayakawa, K., *et al.* (2014) Identification of functionally important residues of the silkmoth pheromone biosynthesis-activating neuropeptide receptor, an insect ortholog of the vertebrate neuromedin U receptor. *J. Biol. Chem.* **289**, 19150–19163
27. Kaang, B. K., Kandel, E. R., and Grant, S. G. N. (1993) Activation of cAMP-responsive genes by stimuli that produce long- term facilitation in *Aplysia* sensory neurons. *Neuron* **10**, 427–435
28. Borovikov, D., Evans, C. G., Jing, J., Rosen, S. C., and Cropper, E. C. (2000) A proprioceptive role for an exteroceptive mechanoafferent neuron in *Aplysia*. *J. Neurosci.* **20**, 1990–2002
29. Jing, J., and Weiss, K. R. (2001) Neural mechanisms of motor program switching in *Aplysia*. *J. Neurosci.* **21**, 7349–7362
30. Kandel, E. R. (2001) The molecular biology of memory storage: a dialogue between genes and synapses. *Science* **294**, 1030–1038
31. Jing, J., and Weiss, K. R. (2002) Interneuronal basis of the generation of related but distinct motor programs in *Aplysia*: implications for current neuronal models of vertebrate intralimb coordination. *J. Neurosci.* **22**, 6228–6238
32. Dembrow, N. C., Jing, J., Proekt, A., Romero, A., Vilim, F. S., Cropper, E. C., *et al.* (2003) A newly identified buccal interneuron initiates and modulates feeding motor programs in *Aplysia*. *J. Neurophysiol.* **90**, 2190–2204
33. Furukawa, Y., Nakamaru, K., Sasaki, K., Fujisawa, Y., Minakata, H., Ohta, S., *et al.* (2003) PRQFVamide, a novel pentapeptide identified from the CNS and gut of *Aplysia*. *J. Neurophysiol.* **89**, 3114–3127
34. Jing, J., Vilim, F. S., Wu, J. S., Park, J. H., and Weiss, K. R. (2003) Concerted GABAergic actions of *Aplysia* feeding interneurons in motor program specification. *J. Neurosci.* **23**, 5283–5294
35. Cropper, E. C., Evans, C. G., Hurwitz, I., Jing, J., Proekt, A., Romero, A., *et al.* (2004) Feeding neural networks in the mollusc *Aplysia*. *Neurosignals* **13**, 70–86
36. Dembrow, N. C., Jing, J., Brezina, V., and Weiss, K. R. (2004) A specific synaptic pathway activates a conditional plateau potential underlying protraction phase in the *Aplysia* feeding central pattern generator. *J. Neurosci.* **24**, 5230–5238
37. Due, M. R., Jing, J., and Weiss, K. R. (2004) Dopaminergic contributions to modulatory functions of a dual-transmitter interneuron in *Aplysia*. *Neurosci. Lett.* **358**, 53–57
38. Jing, J., and Weiss, K. R. (2005) Generation of variants of a motor act in a modular and hierarchical motor network. *Curr. Biol.* **15**, 1712–1721
39. Proekt, A., Vilim, F. S., Alexeeva, V., Brezina, V., Friedman, A., Jing, J., *et al.* (2005) Identification of a new neuropeptide precursor reveals a novel source of extrinsic modulation in the feeding system of *Aplysia*. *J. Neurosci.* **25**, 9637–9648
40. Jing, J., Vilim, F. S., Cropper, E. C., and Weiss, K. R. (2008) Neural analog of arousal: persistent conditional activation of a feeding modulator by serotonergic initiators of locomotion. *J. Neurosci.* **28**, 12349–12361
41. Mozzachiodi, R., Lorenzetti, F. D., Baxter, D. A., and Byrne, J. H. (2008) Changes in neuronal excitability serve as a mechanism of long-term memory for operant conditioning. *Nat. Neurosci.* **11**, 1146–1148
42. Sasaki, K., Jing, J., Due, M. R., and Weiss, K. R. (2008) An input-representing interneuron regulates spike timing and thereby phase switching in a motor network. *J. Neurosci.* **28**, 1916–1928
43. Jing, J., Gillette, R., and Weiss, K. R. (2009) Evolving concepts of arousal: insights from simple model systems. *Rev. Neurosci.* **20**, 405–427
44. Sasaki, K., Brezina, V., Weiss, K. R., and Jing, J. (2009) Distinct inhibitory neurons exert temporally specific control over activity of a motoneuron receiving concurrent excitation and inhibition. *J. Neurosci.* **29**, 11732–11744

45. Jing, J., Sweedler, J. V., Cropper, E. C., Alexeeva, V., Park, J. H., Romanova, E. V., *et al.* (2010) Feedforward compensation mediated by the central and peripheral actions of a single neuropeptide discovered using representational difference analysis. *J. Neurosci.* **30**, 16545–16558

46. Wu, J. S., Vilim, F. S., Hatcher, N. G., Due, M. R., Sweedler, J. V., Weiss, K. R., *et al.* (2010) Composite modulatory feedforward loop contributes to the establishment of a network state. *J. Neurophysiol.* **103**, 2174–2184

47. Jing, J., Sasaki, K., Perkins, M. H., Siniscalchi, M. J., Ludwar, B. C., Cropper, E. C., *et al.* (2011) Coordination of distinct motor structures through remote axonal coupling of projection interneurons. *J. Neurosci.* **31**, 15438–15449

48. Wan, Q., Jiang, X. Y., Negroiu, A. M., Lu, S. G., McKay, K. S., and Abrams, T. W. (2012) Protein kinase C acts as a molecular detector of firing patterns to mediate sensory gating in *Aplysia*. *Nat. Neurosci.* **15**, 1144–1152

49. Bai, L., Livnat, I., Romanova, E. V., Alexeeva, V., Yau, P. M., Vilim, F. S., *et al.* (2013) Characterization of GdFFD, a D-amino acid-containing neuropeptide that functions as an extrinsic modulator of the *Aplysia* feeding circuit. *J. Biol. Chem.* **288**, 32837–32851

50. Sasaki, K., Cropper, E. C., Weiss, K. R., and Jing, J. (2013) Functional differentiation of a population of electrically coupled heterogeneous elements in a microcircuit. *J. Neurosci.* **33**, 93–105

51. Wu, J. S., Wang, N., Siniscalchi, M. J., Perkins, M. H., Zheng, Y. T., Yu, W., *et al.* (2014) Complementary interactions between command-like interneurons that function to activate and specify motor programs. *J. Neurosci.* **34**, 6510–6521

52. Jing, J., Alexeeva, V., Chen, S. A., Yu, K., Due, M. R., Tan, L. N., *et al.* (2015) Functional characterization of a vesicular glutamate transporter in an interneuron that makes excitatory and inhibitory synaptic connections in a molluscan neural circuit. *J. Neurosci.* **35**, 9137–9149

53. Zhang, G., Yuan, W. D., Vilim, F. S., Romanova, E. V., Yu, K., Yin, S. Y., *et al.* (2018) Newly identified *Aplysia* SPTR-gene family-derived peptides: localization and function. *ACS Chem. Neurosci.* **9**, 2041–2053

54. Tam, S., Hurwit, I., Chiel, H. J., and Susswein, A. J. (2020) Multiple local synaptic modifications at specific sensorimotor connections after learning are associated with behavioral adaptations that are components of a global response change. *J. Neurosci.* **40**, 4363–4371

55. Zhang, G., Yu, K., Wang, T., Chen, T. T., Yuan, W. D., Yang, F., *et al.* (2020) Synaptic mechanisms for motor variability in a feedforward network. *Sci. Adv.* **6**, eaba4856

56. Evans, C. G., Barry, M. A., Jing, J., Perkins, M. H., Weiss, K. R., and Cropper, E. C. (2021) The complement of projection neurons activated determines the type of feeding motor program in *Aplysia*. *Front. Neural Circuits* **15**, 685222

57. Due, M. R., Wang, Y., Barry, M. A., Jing, J., Reaver, C. N., Weiss, K. R., *et al.* (2022) Convergent effects of neuropeptides on the feeding central pattern generator of *Aplysia californica*. *J. Neurophysiol.* **127**, 1445–1459

58. Orvis, J., Albertin, C. B., Shrestha, P., Chen, S., Zheng, M., Rodriguez, C. J., *et al.* (2022) The evolution of synaptic and cognitive capacity: insights from the nervous system transcriptome of *Aplysia*. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2122301119

59. Zhang, G., Guo, S. Q., Yin, S. Y., Yuan, W. D., Chen, P., Kim, J. I., *et al.* (2022) Exogenous expression of an allatotropin-related peptide receptor increased the membrane excitability in *Aplysia* neurons. *Mol. Brain* **15**, 42

60. Zhang, G., Vilim, F. S., Liu, D. D., Romanova, E. V., Yu, K., Yuan, W. D., *et al.* (2017) Discovery of leucokinin-like neuropeptides that modulate a specific parameter of feeding motor programs in the molluscan model, *Aplysia*. *J. Biol. Chem.* **292**, 18775–18789

61. Kuhlman, B., and Bradley, P. (2019) Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.* **20**, 681–697

62. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., *et al.* (2021) Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876

63. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589

64. Ferreira, L. G., Dos Santos, R. N., Oliva, G., and Andricopulo, A. D. (2015) Molecular docking and structure-based drug design strategies. *Molecules* **20**, 13384–13421

65. Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., and Olson, A. J. (2016) Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **11**, 905–919

66. Nassel, D. R. (2021) Leucokinin and associated neuropeptides regulate multiple aspects of physiology and behavior in *Drosophila*. *Int. J. Mol. Sci.* **22**, 1940

67. Nassel, D. R., and Wu, S. F. (2021) Leucokinins: multifunctional neuropeptides and hormones in insects and other invertebrates. *Int. J. Mol. Sci.* **22**, 1531

68. Holman, G. M., and Cook, B. J. (1983) Isolation and partial characterization of a second myotropic peptide from the hindgut of the cockroach, *Leucophaea maderae*. *Comp. Biochem. Physiol. C Comp. Pharmacol. Toxicol.* **76**, 39–43

69. Holman, G. M., Cook, B. J., and Nachman, R. J. (1986) Primary structure and synthesis of a blocked myotropic neuropeptide isolated from the cockroach, *Leucophaea maderae*. *Comp. Biochem. Physiol. C Comp. Pharmacol. Toxicol.* **85**, 219–224

70. Holman, G. M., Cook, B. J., and Nachman, R. J. (1986) Primary structure and synthesis of 2 additional neuropeptides from *Leucophaea-Maderae* - members of a new family of Cephalomyotropins. *Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol.* **84**, 271–276

71. Holman, G. M., Cook, B. J., and Nachman, R. J. (1987) Isolation, primary structure, and synthesis of Leukokinin-V and Leukokinin-VI - myotropic peptides of *Leucophaea-Maderae*. *Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol.* **88**, 27–30

72. Holman, G. M., Cook, B. J., and Nachman, R. J. (1987) Isolation, primary structure and synthesis of Leukokinin-VII and Leukokinin-VIII - the final members of this new family of cephalomyotropic peptides isolated from head extracts of *Leucophaea-Maderae*. *Comp. Biochem. Physiol. C Pharmacol. Toxicol. Endocrinol.* **88**, 31–34

73. Cox, K. J., Tensen, C. P., Van der Schors, R. C., Li, K. W., van Heerikhuizen, H., Vreugdenhil, E., *et al.* (1997) 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

74. Holmes, S. P., He, H., Chen, A. C., lvie, G. W., and Pietrantonio, P. V. (2000) Cloning and transcriptional expression of a leucokinin-like peptide receptor from the southern cattle tick, *Boophilus microplus* (Acari: Ixodidae). *Insect Mol. Biol.* **9**, 457–465

75. Radford, J. C., Davies, S. A., and Dow, J. A. (2002) Systematic G-protein-coupled receptor analysis in *Drosophila melanogaster* identifies a leucokinin receptor with novel roles. *J. Biol. Chem.* **277**, 38810–38817

76. Holmes, S. P., Barhoumi, R., Nachman, R. J., and Pietrantonio, P. V. (2003) Functional analysis of a G protein-coupled receptor from the southern cattle tick *Boophilus microplus* (Acari: Ixodidae) identifies it as the first arthropod myokinin receptor. *Insect Mol. Biol.* **12**, 27–38

77. Radford, J. C., Terhzaz, S., Cabrero, P., Davies, S. A., and Dow, J. A. (2004) Functional characterisation of the *Anopheles* leucokinins and their cognate G-protein coupled receptor. *J. Exp. Biol.* **207**, 4573–4586

78. Pietrantonio, P. V., Jagge, C., Taneja-Bageshwar, S., Nachman, R. J., and Barhoumi, R. (2005) The mosquito *Aedes aegypti* (L.) leucokinin receptor is a multiligand receptor for the three Aedes kinins. *Insect Mol. Biol.* **14**, 55–67

79. Taneja-Bageshwar, S., Strey, A., Zubrzak, P., Pietrantonio, P. V., and Nachman, R. J. (2006) Comparative structure-activity analysis of insect kinin core analogs on recombinant kinin receptors from southern cattle tick *Boophilus microplus* (Acari: Ixodidae) and mosquito *Aedes aegypti* (Diptera: Culicidae). *Arch. Insect Biochem. Physiol.* **62**, 128–140

80. Xiong, C., Kaczmarek, K., Zabrocki, J., Nachman, R. J., and Pietrantonio, P. V. (2020) Activity of native tick kinins and peptidomimetics on the cognate target G protein-coupled receptor from the cattle fever tick, *Rhipicephalus microplus* (Acari: Ixodidae). *Pest Manag. Sci.* **76**, 3423–3431

81. Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., *et al.* (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* **48**, D265–D268

82. Sonnhammer, E. L., von Heijne, G., and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **6**, 175–182

83. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580

84. Rovati, G. E., Capra, V., and Neubig, R. R. (2007) The highly conserved DRY motif of class A G protein-coupled receptors: beyond the ground state. *Mol. Pharmacol.* **71**, 959–964

85. Barak, L. S., Menard, L., Ferguson, S. S., Colapietro, A. M., and Caron, M. G. (1995) The conserved seven-transmembrane sequence NP(X)2,3Y of the G-protein-coupled receptor superfamily regulates multiple properties of the beta 2-adrenergic receptor. *Biochemistry* **34**, 15407–15414

86. Fredriksson, R., Lagerstrom, M. C., Lundin, L. G., and Schioth, H. B. (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **63**, 1256–1272

87. Southey, B. R., Amare, A., Zimmerman, T. A., Rodriguez-Zas, S. L., and Sweedler, J. V. (2006) NeuroPred: a tool to predict cleavage sites in neuropeptide precursors and provide the masses of the resulting peptides. *Nucleic Acids Res.* **34**, W267–W272

88. Trinquet, E., Fink, M., Bazin, H., Grillet, F., Maurin, F., Bourrier, E., *et al.* (2006) D-myo-inositol 1-phosphate as a surrogate of D-myo-inositol 1,4,5-tris phosphate to monitor G protein-coupled receptor activation. *Anal. Biochem.* **358**, 126–135

89. Sharma, S., and Checco, J. W. (2021) Evaluating functional ligand-GPCR interactions in cell-based assays. *Methods Cell Biol.* **166**, 15–42

90. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486

91. Benkert, P., Biasini, M., and Schwede, T. (2011) Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**, 343–350

92. Ferreira de Freitas, R., and Schapira, M. (2017) A systematic analysis of atomic protein-ligand interactions in the PDB. *Medchemcomm* **8**, 1970–1981

93. Krone, M. W., Travis, C. R., Lee, G. Y., Eckvahl, H. J., Houk, K. N., and Waters, M. L. (2020) More than pi-pi-pi stacking: contribution of amide-pi and CH-pi interactions to crotonyllysine binding by the AF9 YEATS domain. *J. Am. Chem. Soc.* **142**, 17048–17056

94. Patwardhan, A., Cheng, N., and Trejo, J. (2021) Post-translational modifications of G protein-coupled receptors control cellular signaling dynamics in space and time. *Pharmacol. Rev.* **73**, 120–151

95. Langer, I., Tikhonova, I. G., Travers, M. A., Archer-Lahlou, E., Escrieut, C., Maigret, B., *et al.* (2005) Evidence that interspecies polymorphism in the human and rat cholecystokinin receptor-2 affects structure of the binding site for the endogenous agonist cholecystokinin. *J. Biol. Chem.* **280**, 22198–22204

96. Liu, Q., Yang, D., Zhuang, Y., Croll, T. I., Cai, X., Dai, A., *et al.* (2021) Ligand recognition and G-protein coupling selectivity of cholecystokinin A receptor. *Nat. Chem. Biol.* **17**, 1238–1244

97. Tunaru, S., Kero, J., Schaub, A., Wufka, C., Blaukat, A., Pfeffer, K., *et al.* (2003) PUMA-G and HM74 are receptors for nicotinic acid and mediate its anti-lipolytic effect. *Nat. Med.* **9**, 352–355

98. Quiroga Artigas, G., Lapebie, P., Leclere, L., Bauknecht, P., Uveira, J., Chevalier, S., *et al.* (2020) A G protein-coupled receptor mediates neuropeptide-induced oocyte maturation in the jellyfish *Clytia*. *PLoS Biol.* **18**, e3000614

99. Reilly, D. K., McGlame, E. J., Vandewyer, E., Robidoux, A. N., Muirhead, C. S., Northcott, H. T., *et al.* (2021) Distinct neuropeptide-receptor modules regulate a sex-specific behavioral response to a pheromone. *Commun. Biol.* **4**, 1018

100. Lourenco, E. V., and Roque-Barreira, M. C. (2010) Immunoenzymatic quantitative analysis of antigens expressed on the cell surface (cell-ELISA). *Methods Mol. Biol.* **588**, 301–309

101. Newton, C. L., Anderson, R. C., Katz, A. A., and Millar, R. P. (2016) Loss-of-function mutations in the human luteinizing hormone receptor predominantly cause intracellular retention. *Endocrinology* **157**, 4364–4377

102. Zhuang, Y. W., Xu, P. Y., Mao, C. Y., Wang, L., Krumm, B., Zhou, X. E., *et al.* (2021) Structural insights into the human D1 and D2 dopamine receptor signaling complexes. *Cell* **184**, 931–942

≣ASBMB