

Rowan University

Rowan Digital Works

School of Osteopathic Medicine Faculty
Scholarship

School of Osteopathic Medicine

8-29-2013

Evolutionary Interrelationships and Insights into Molecular Mechanisms of Functional Divergence: An Analysis of Neuronal Calcium Sensor Proteins

Jeffrey Viviano

Rowan University School of Osteopathic Medicine

Hao Wu

Rowan University School of Osteopathic Medicine

Venkat Venkataraman

Rowan University School of Osteopathic Medicine

Follow this and additional works at: https://rdw.rowan.edu/som_facpub



Part of the [Biochemistry Commons](#), [Cell Biology Commons](#), [Evolution Commons](#), [Molecular Biology Commons](#), and the [Molecular Genetics Commons](#)

Recommended Citation

Viviano J, Wu H, Venkataraman V (2013) Evolutionary Interrelationships and Insights into Molecular Mechanisms of Functional Divergence: An Analysis of Neuronal Calcium Sensor Proteins. *J Phylogen Evolution Biol* 1:117. Doi: 10.4172/2329-9002.1000117

This Article is brought to you for free and open access by the School of Osteopathic Medicine at Rowan Digital Works. It has been accepted for inclusion in School of Osteopathic Medicine Faculty Scholarship by an authorized administrator of Rowan Digital Works.

Evolutionary Interrelationships and Insights into Molecular Mechanisms of Functional Divergence: An Analysis of Neuronal Calcium Sensor Proteins

Jeffrey Viviano[#], Hao Wu[#] and Venkat Venkataraman^{*}

Department of Cell Biology, Rowan University School of Osteopathic Medicine, Graduate School of Biomedical Sciences, Rowan University, Stratford, NJ 08084, USA
[#]these authors contributed equally to the work

Abstract

The normal function of any organism, its organizational complexity notwithstanding, depends on the interaction of its proteins with their targets. Thus, analysis of target site interaction is an essential part of all biology. At the protein level, such analyses are critical to both mechanistic knowledge and potential clinical applications such as drug discovery. Approaches to map amino acid residues involved in target site interaction typically are experimental or are based on three-dimensional structures obtained through crystallography. Here we test a novel approach that combines phylogenetic analyses with mining of experimental data using neuronal calcium sensor proteins. The proteins fall into three groups based on sequence comparison. One interaction was taken up for analysis from each group. Using the sequence divergence to evaluate the role of amino acids identified experimentally to form the interface with the target, we demonstrate that it is possible to predict residues that are likely to contribute to the specificity of the interaction and, therefore, the functional divergence. Thus, evolutionary analyses of proteins provide an important addition in approaches to generate refined maps of target site interactions in proteins. This approach is especially useful in delineating the functional divergence in a family of closely related proteins.

Keywords: Functional divergence; Evolution; Neuronal calcium sensor proteins; Guanylate cyclase; GCAP; KChIP; VILIP

Abbreviations: NCS: Neuronal Calcium Sensor; KChIP: Kv Potassium-channel Interacting Proteins; GCAP: Guanylate Cyclase Activating Proteins; VILIP: Visinin-like Proteins; mGC: Membrane Guanylate Cyclase; MS: Mass Spectrometry

Introduction

It is axiomatic that conservation during evolution implies irreplaceable function, as selection pressure constantly operates to nurture functions that improve adaptation to the environment. While variation is generated at the genetic level through multiple mechanisms such as mutation, gene duplication and other rearrangements of genetic information, regions that encode essential functions remain largely conserved. Such conservation is observable both at the level of entire gene/protein sequences and at the level of individual domains and motifs within a family of genes/proteins: for example, catalytic domains of enzymes such as adenylatecyclase [1-4]. However, the typical approach is one where the function of a protein is experimentally determined with some analysis of the evolutionarily conserved residues critical for its function. Is it possible, with the explosion of sequence information available for several entire genomes, to gain molecular insights into functional divergence of a related group of proteins through phylogenetic analyses? In this study, we have investigated this possibility through the analyses of a subfamily of proteins – the neuronal calcium sensor (NCS) proteins.

The family of NCS proteins is contained within the super family of proteins that are characterized by the presence of the calcium-binding EF-hand motif [5,6]. Members of the family are generally about 200 aa long and have four EF-hand domains one of which has lost the ability to bind to calcium (with the exception of recoverin which has only two calcium-binding EF hands). Another feature is an acyl modification at the N-terminus that was identified first in recoverin [7,8] and characterized as the calcium-myristoyl switch [9].

The members of the NCS family perform a variety of diverse functions. NCS proteins have been identified in all vertebrate animals, most invertebrate animals, in plants and in lower eukaryotes such as fungi [10,11]. It is currently estimated that 14 genes encode NCS proteins

in the human genome [12]; however, generation of additional diversity is already documented through splice-variants. Many of these proteins were initially characterized in neurons, where they carry out critical functions in many species. For example, loss of function of specific members has been demonstrated to cause loss of memory formation in nematodes [13] and mice [14], while others are linked to inherited visual degeneration [15-17]. In many instances, the defect has been traced to the loss of ability to sense and respond to changes in calcium in the neuron [15,18-20], which is the primary functional feature of this family of proteins. However, many of these proteins are also expressed outside the nervous system and interact with a variety of targets in multiple cell types [21-24], suggesting that the calcium-dependent modulation by these proteins is not restricted to neurons. Many of the individual members have also been characterized extensively at the biochemical and molecular level [25-28]. Given the experimental and sequence information available and the increasing divergence of these proteins due to the demand of an advancing neural network during evolution, we chose this family for our analyses. Our goal was to determine if analyses of evolutionary sequence divergence would provide insight into the molecular mechanism of functional divergence and if they would be of predictive value for experimental design.

Methods

Computation of sequence interrelationships

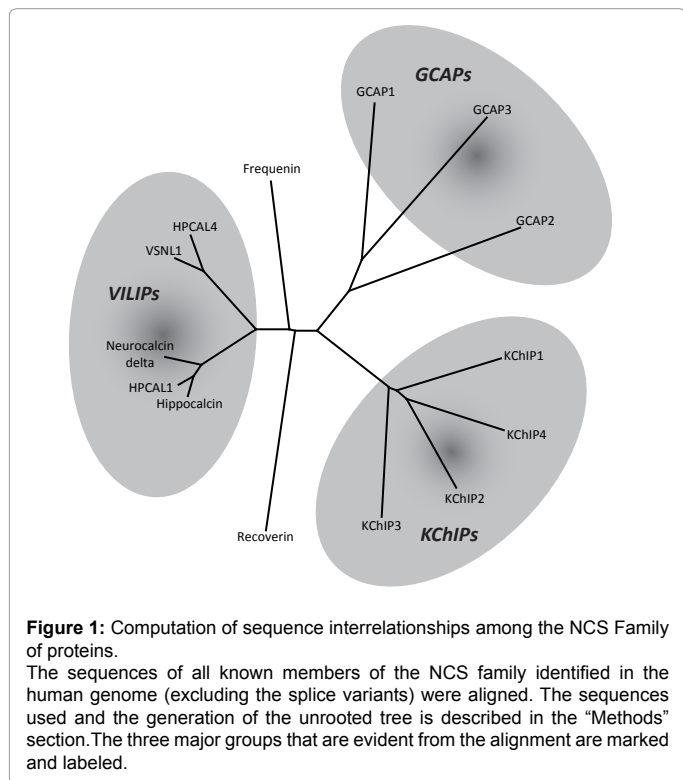
The sequences of NCS proteins used for analyses are present-

***Corresponding author:** Venkat Venkataraman, Department of Cell Biology, SC 220, Rowan University School of Osteopathic Medicine, USA, Tel: 856-566-6418; E-mail: vvenkat2007@gmail.com

Received July 15, 2013; Accepted August 27, 2013; Published August 29, 2013

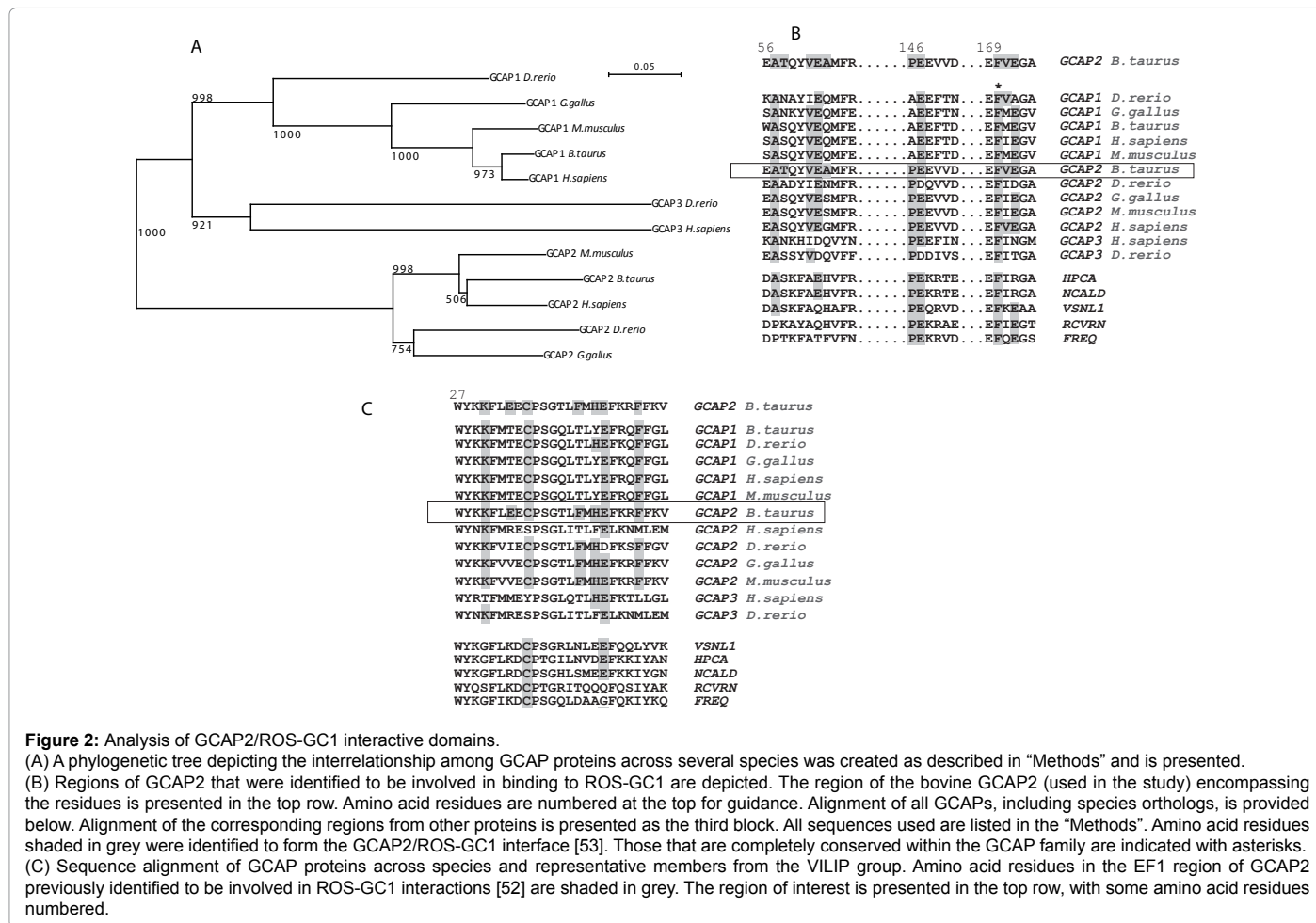
Citation: Viviano J, Wu H, Venkataraman V (2013) Evolutionary Interrelationships and Insights into Molecular Mechanisms of Functional Divergence: An Analysis of Neuronal Calcium Sensor Proteins. J Phylogen Evolution Biol 1: 117. doi:10.4172/2329-9002.1000117

Copyright: © 2013 Viviano J, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



ed below and grouped according to the figures. They were: Freqenin [AAF01804.1], Recoverin [NP_002894.1], Neurocalcin delta [NP_001035720.1], Hippocalcin [BAA74456.1], VSNL1 [P62760.2], HPCAL1 [P37235.3], HPCAL4 [Q9UM19.3], GCAP1 [P43080.3], GCAP 2 [Q9UMX6.4], GCAP 3 [NP_005450.3], KChIP1 [Q9NZI2.2], KChIP2 [Q9NS61.3], KChIP3 [Q9Y2W7.1] and KChIP4 [Q6PIL6.1] in Figure 1; GCAP1 *Homo sapiens* [P43080.3], GCAP 2 *Homo sapiens* [Q9UMX6.4], GCAP 3 *Homo sapiens* [NP_005450.3]), GCAP1 *Bos taurus* [P46065.2], GCAP1 *Mus musculus* [NP_032215.2], GCAP1 *Gallus gallus* [AAB47111.1], GCAP1 *Danio rerio* [AAL05602.1], GCAP2 *Gallus gallus* [AAB47112.1], GCAP2 *Musmusculus* [NP_666191.1], GCAP2 *Bos taurus* [NP_777211.1], GCAP2 *Danio rerio* [NP_571946.1], and GCAP3 *Danio rerio* [AAK95949.1] in Figure 2; Freqenin [AAF01804.1], Recoverin [NP_002894.1], Neurocalcin delta [NP_001035720.1], Hippocalcin [BAA74456.1], VSNL1 [P62760.2], HPCAL1 [P37235.3], HPCAL4 [Q9UM19.3], NCS1 *Caenorhabditis elegans* [AAA85631.1], NCS1 *Danio rerio* [AAO34710.1], NCS1 *Xenopus laevis* [NP_001084088.1], NCS1 *Saccharomyces cerevisiae* [Q06389.2], and NCS1 *Drosophila melanogaster* [AAH04856.1] in Figure 3, and; KChIP1 [Q9NZI2.2], KChIP2 [Q9NS61.3], KChIP3 [Q9Y2W7.1], KChIP4 [Q6PIL6.1], KChIP1 isoform 3 [NP_001030010.1] denoted as KChIP1.3, KChIP2 isoform 2 [NP_775283.1] denoted as KChIP2.2, KChIP3 isoform 2 precursor [NP_001030086.1] denoted as KChIP3.2, and KChIP 4 isoform 5 [NP_001030175.1] denoted as KChIP4.5 in Figure 4.

All alignments were performed using ClustalX 2.1 with the default



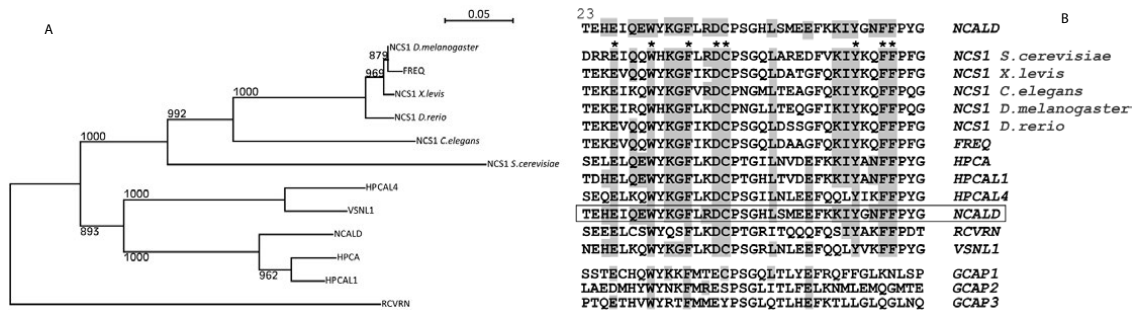


Figure 3: Analysis of NCALD/ROS-GC1 interactive domains. (A) A phylogenetic tree of the proteins of the VILIP group along with recoverin and frequenin (NCS-1) across several species was created after Clustal alignment as described in "Methods". (B) Residues that are involved in NCALD/ROS-GC1 interactions were identified [46]. The region of interest in bovine NCALD is presented in the top row, with some amino acid residues numbered. Alignment of the corresponding region from all VILIPs (including species orthologs), recoverin and frequenin is provided below. Alignment of the region from GCAPs is presented as the third block. All sequences used are listed in the "Methods". Those that are completely conserved within the VILIP family are indicated with asterisks.

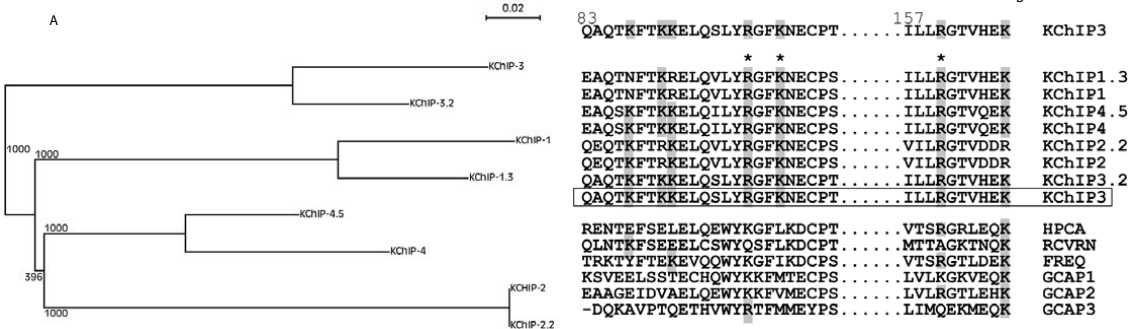


Figure 4: Analysis of KChIP/DRE sequence interactions. (A) A phylogenetic tree created from a Clustal sequence alignment of the KChIP family of proteins along with isoforms of each variant is presented. The sequences used and the generation of the phylogenetic tree is described in the "Methods" section. (B) Sequence alignment of the NCS proteins corresponding to the region identified in KChIPs to interact with DRE sequence [48]. The region of interest in KChIP3 is presented in the top row, with some amino acid residues numbered. Those that are completely conserved within the KChIP family are indicated with asterisks. Alignment of the corresponding regions from other proteins is also presented.

setting: Gap opening 10, Gap extension 0.1, Protein weighted matrix –Gonnet 250 [29]. The unrooted tree was compiled using NJPlot [30] and enhanced in Adobe Illustrator. For the construction of phylogenetic trees, the NJ Bootstrapping method with a random number generator seed set at 111 and 1000 reiterations was used. The tree was displayed with NJPlot [30] and enhanced with Adobe Illustrator.

Tissue expression analysis

The analyses were carried out *in silico*. Expression at mRNA as well as protein level was analyzed. Two different databases – Genevestigator [31] and BioGPS [32]–were used to obtain mRNA expression profiles across different human tissues (brain, heart, liver, lung and skin) for indicated proteins.

Protein expression level in whole human brain and liver was obtained for indicated proteins from two independent databases: (a) Human Protein Atlas database [33-35], which is based on immunohistochemical analyses, and scores expression levels as negative (0), weak (1), moderate (2) or strong (3), and (b) PaxDB [36], which compiles standardized and integrated data on protein expression levels based on MS/MS. The data from the Human Protein Atlas was also utilized to investigate the relative expression levels of the indicated proteins within different neurons of the human brain.

Results and Discussion

As a first step towards analyzing the link between evolutionary and functional diversity, we investigated the relationship among the NCS family members in the human genome. Several additional sequences (truncated or splice variants or without additional documentation) were left out of the analysis. There were 14 members of the NCS family: Frequenin (also known as NCS1), Recoverin, Kv Potassium-Channel Interacting Proteins types 1 through 4 (KChIP 1-4), Guanylate Cyclase Activating Proteins types 1 through 3 (GCAP 1-3), Visinin-like Protein type 1 (VILIP1; gene ID VSNL1), Neurocalcin Delta (NCALD), Hippocalcin (HPCA) and hippocalcin-like proteins type 1 and 4 (HPCAL1 & HPCAL4). The result presented in Figure 1 shows that these members generate three major groups, which we have identified as GCAPs, VILIPs and KChIPs (indicated by shaded areas in Figure1). Frequenin and recoverin were out grouped consistent with earlier analysis from other laboratories [21,37].

The grouping is also consistent with known functions of the members in each: All GCAPs are intimately linked with the maintenance of the autoregulatory feedback loop in phototransduction [38-40] through their interaction with membrane guanylate cyclase (mGC) [22,39]; GCAP1, in addition, may also contribute to olfaction [41]. Some members of the VILIP group also share this ability to

modulate mGC activity [26,27,42-46]. However, in addition to mGC, each member of the VILIP group is known to interact with multiple targets, generating perhaps the most functionally diverse group among the three [21,22,25]. The KChIPs were initially identified as modulators of potassium currents [47]. It has now been demonstrated that all KChIPs share DNA-binding ability also [48]. Interestingly, between the two outgrouped members, frequenin activates mGC [44] but recoverin does not; instead, it regulates rhodopsin kinase [49] and contributes to the feedback loop in phototransduction [50,51]. Thus, members of the GCAP and VILIP groups as well as the outgrouped frequenin share the ability to modulate mGC activity. Further details of this shared function are presented below.

Intricacy of NCS/mGC interactions

A closer investigation reveals an extraordinarily intricate relationship between the NCS protein modulators and mGCs. There are seven genes in the human genome that encode mGCs. Of the seven, the one encoded by GUCY2D (also known as ROS-GC1/RetGc1) is modulated by GCAP1, GCAP2, GCAP3, and NCALD [27,46]; (reviewed in: [22,39,49,51]). As an additional level of intrigue, the GCAPs inhibit mGC activity at calcium concentrations greater than 500 nM, while NCALD stimulates it ([27,46]; reviewed in:[39,51]). All the above proteins also modulate the GUCY2F product (ROS-GC2/RetGc2) in a similar fashion, albeit at differing efficiencies as assessed by fold-change in activity and EC₅₀ values [39,51]. Frequenin [44] and HPCA [45] activate the olfactory guanylate cyclase (ONE-GC/GC-D) at elevated calcium concentrations; interestingly, so does GCAP1 [41]. VSNL1, on the other hand, activates the product of NPR2B (atrial natriuretic peptide receptor type B/GC-B) [26,42,43]. The overlapping specificities of NCS proteins to one cyclase, the ability of one NCS protein to modulate different cyclases and opposing responses to elevated calcium are all observed in this interaction, which is spread between two groups and the two outliers. Information on the binding site on some NCS proteins necessary for their modulation of mGC is also available. Therefore, we chose to analyze this function in terms of evolutionary interrelationships among these proteins. Two laboratories have independently mapped amino acid residues on GCAP2 that mediate its interaction with ROS-GC1 [52,53]. Similar information is available for NCALD, which regulates ROS-GC1 in an opposite fashion compared to GCAP2 in their calcium-loaded states [46]. A correlation between the molecular interactions leading up to functional divergence and phylogenetic relationships among these proteins are presented below.

GCAP2 modulation of ROS-GC1

For these analyses, results from two different laboratories were used. Pettelkau et al. [53] have recently identified residues on GCAP2 that are in close proximity to and, presumably, form the contact-points with ROS-GC1. This was accomplished through a combination of chemical cross-linking, photo-affinity labeling and MS. Based on these studies, a model for the calcium-loaded GCAP2 interaction with ROS-GC1 is also proposed involving specific amino acid residues in the interface. Ermilov et al. [52] identified residues on GCAP2 necessary for calcium-dependent regulation of ROS-GC1 using a completely different approach. Between the two studies, residues on GCAP2 that are critical for the interaction as well as modulation of ROS-GC1 have been mapped. Given that all GCAPs interact with ROS-GC1, along with at least one member of the VILIPs, it was of interest to investigate the presence of these identified residues within the GCAP group and, then, compare with the VILIP group members. The results are presented in Figure 2. As a first step, several orthologs for GCAP1 and GCAP2

from different vertebrate species as well as the two known orthologs of GCAP3 were aligned as described in Methods and the results are presented (Figure 2A). It is noted that for GCAP3 these are the only reported sequences available although it has been suggested that it is expressed in multiple species. None of the GCAPs are completely conserved among the species, variation does exist between species and the grouping is as expected: sequences for each of the GCAPs from different species cluster together (Figure 2A). Are they divergent on the amino acid residues that interface with ROS-GC1? An alignment of all GCAP sequences (species indicated in *italicized font*) with the amino acid residues spanning the interface is presented in Figure 2B. In addition, they are aligned with members from the VILIP group. It is noted that all members of the VILIP groups are conserved across the vertebrate species (data not shown).

On GCAP2, two areas of interaction were identified [53]: one that spans the EF1 hand and the other spans the EF4 hand. The bovine form of GCAP2 was used in these experiments. The amino acid residues in the region are presented in the top row of Figure 2B and numbers for some amino acid residues are provided on top for guidance. The residues identified to be at the interface are shaded in grey. They are: A57, T58, V61, E62, A63 (all in EF1), P146, E147, F170, V171 and E172 (all in EF4). How are these residues conserved during the divergence of GCAPs and VILIPs? A57 is highly conserved in GCAPs as well as most VILIPs but is replaced by a Pro residue in recoverin and frequenin. Neither of these proteins interacts with ROS-GC1 ([44]; reviewed in: [49]). Therefore, could this residue be critical for ROS-GC1 interaction? Interestingly, HPCA and VSNL1, which are shown to interact with mGCs other than ROS-GC1, have preserved A57. Is it possible that these NCS proteins also interact with ROS-GC1? In order to answer that question, the conservation of residues at other positions was investigated. T58 is not conserved even in species orthologs of GCAP2. Given that GCAPs have been shown to operate across species (GCAP) from one species can modulate ROS-GC1 from another species [54], it is unlikely that this residue is critical for interaction. V61 and E62 are conserved in most GCAPs or replaced by Ile or Asp residues across the GCAPs. This suggests that both of the sites are likely to be the amino acids utilized for mGC activation in GCAPs. In VILIPs, there is an Ala at position corresponding to V61 and, therefore, supports a GCAP-specific role for this residue. On the other hand, E62 is preserved in HPCA and NCALD. Considering that VILIPs and GCAPs have opposite effects on mGC stimulation ([27,46]; reviewed in: [39,51]), could V61 be the reason for this functional divergence? Like T58, A63 is not conserved even within GCAP2 orthologs and, therefore, is likely not an interaction site [55].

The amino acid residues in EF4 that are located in the interface are: P146, E147, F170, V171 and E172 [53]. P146 is interesting because it is conserved throughout the NCS family but is substituted with an Ala in all species orthologs of GCAP1, suggesting that this Ala is critical for the unique properties of GCAP1 such as binding at a different site [56]. E147 is almost completely conserved, with only the GCAP2 and GCAP3 proteins in *D. rerio* bearing a conserved substitution of Asp at the corresponding position. F170 is completely conserved (indicated with an asterisk; Figure 2B). Position corresponding V171 is highly variable even between the same proteins in different species with substitutions ranging from presumably conserved ones such as with an Ile to Met and positively charged Lys and Gln. There are two completely opposite inferences that may be drawn: (a) it is possible that this position may govern parameters of interaction of individual members with ROS-GC1, in which case mutation of these residues would affect the interaction of the protein with ROS-GC1 or (2) this residue may play no role, in

which case mutation would have no effect. The last identified residue on GCAP2 is E172. Conservation at this position shows an interesting pattern; a negatively charged residue is present in this position in all GCAPs except GCAP3 and *D. rerio* GCAP1; this also holds true for the VILIPs, except for HPCA and NCALD, where a positively charged Arg is found. Thus, the phylogenetic analyses have enabled to distinguish the role of amino acid residues identified at the GCAP2 interface with ROS-GC1: some are unlikely to play an important role either because they are not conserved at all (T58, A63) or completely conserved among all examined NCS sequences (F170). Further predictions could be made about residues that may determine NCS-specific interaction (Ala at P146 for GCAP1, Arg at position corresponding to E172 for HPCA and NCALD, V61 and E62 for GCAPs) and would enable more incisive experimental design to test these predictions.

A completely different approach was taken by Ermilov et al. [52] using site-directed mutagenesis and analyzing the effects on biological activity to map amino acid residues within the EF1 hand of GCAP2 that mediate its calcium-dependent regulation of ROS-GC1. Unlike the previous approach [53], the emphasis is on residues that contribute to the regulation of ROS-GC1. Several residues were investigated for a possible contribution to mGC regulation - K30, E33, C35, F41, H43, E44 and F48. Interestingly, all residues identified by this approach, reside within EF1 region, unlike the previous study which identified residues in both EF1 and EF4 (Figure 2B). The residues on bovine GCAP2 identified by Ermilov et al. [52] are shaded grey in Figure 2C. An alignment of all GCAP sequences as well as representative VILIPs are also depicted. Conserved residues are shaded in grey. Analyses of sequence divergence as before, lead to the following conclusions regarding the mapped region: F41 is GCAP2-specific and may be critical for the specificity of GCAP2 interaction with ROS-GC1; a similar argument could be made for H43, except that the residue is not preserved in the human orthologs. However, experimentally, mutations of these residues affect the ability of GCAP2 to activate mGC [52]. K30 is present in all GCAPs except GCAP3, suggesting that its role is common at least between GCAP1 and GCAP2. The observation that the K30G mutation showed decreased activation, therefore, is less likely due an effect on specific GCAP2/ROS-GC1 interaction and is more likely due to an effect on overall protein function common between GCAP1 and GCAP2. F48 is also conserved in all GCAPs except GCAP3 and, notably, the human GCAP2 ortholog. The scenario with C35 is identical to that of F48. However, C35 is also conserved in the VILIPs, suggesting that the residue may not be involved in unique GCAP2/ROS-GC1 interaction. E44 is conserved in all members analyzed except the two outgrouped members: recoverin and frequenin. The final site thought to be involved in mGC activation by GCAP2 is E33. Interestingly the sequence alignment shows that this site is variable even between species. We predict that E33 contributes little to ROS-GC1 activation, since it is not conserved even among species orthologs of GCAP2. Not surprisingly, the prediction is supported by experimental data [52].

In summary, evolutionary interrelationships were used to analyze amino acid residues determined to be at the GCAP2/ROS-GC1 interface or critical for activation of ROS-GC1 by GCAP2. As a result, residues could be predicted to fall into one three classes: (i) those that have little role in the interaction, (ii) those that are likely to contribute to the specificity of the NCS/ROS-GC1 interaction, and (iii) those that could contribute the total protein structure rather than specific interactions. Notably, some of these predictions have been substantiated by experimental data.

NCALD activation of ROS-GC1

NCALD is a member of the VILIP group (Figure 1) and brings about

activation when it binds to ROS-GC1 in its calcium-loaded state [27,46]. Analysis of the VILIP group members across species is of interest in itself: several members such as NCALD, HPCA, FREQ and VSNL1 are conserved across vertebrates. The VILIP member in lower organisms is identified as NCS1 and has been included in the alignment (Figure 3A). All VILIPs that stimulate mGC activity do so in their calcium-loaded state [27,42,45,46]. Information on the residues on the NCS protein that may mediate the function is available only for NCALD: Venkataraman et al. [46] have identified residues on NCALD that may be critical for the interaction through a combination of biological assays, peptide competition and direct binding assays. The residues thus identified and localized to the EF1 hand region of NCALD are: H25, E26, Q28, E29, W30, K32, G33, F34, R36, D37, C38, L43, E47, K50, I51, Y52, N54, F55 and F56. They are identified by shading in the top row in Figure 3B. Comparison of residues that are conserved among VILIPs against those conserved in GCAPs (Figure 3B) reveals interesting patterns. Several residues are completely conserved across VILIP members even in lower organisms. These are E26, W30, F34, D37, C38, Y52, F55 and F56 (indicated by asterisks; Figure 3C). At corresponding positions in GCAPs, only E26 (D26 in GCAP2), W30 and F34 are conserved. This would lead to the prediction that these residues are important for overall protein integrity, while the rest of the conserved residues contribute to a VILIP-specific function. Furthermore, this function is most likely conserved all the way from lower eukaryotes to vertebrates. The vertebrate VILIPs all activate mGCs in their calcium-loaded state [27,42,45,46]. If this is the preserved function, the results would indicate that the function would be conserved in lower eukaryotes too. On the other hand, it may be a function waiting to be discovered. Further experimental analyses would enable validating one of the alternative scenarios.

The extraordinarily high conservation among VILIPs prompted the question whether these could be redundant proteins. Experimental evidence from loss-of-function scenarios [13,14,17,57] and the fact that they are completely conserved across vertebrates argue that it is not the case. This leads to two possibilities: (a) their expression is separated by space and/or time, or (b) despite the conservation, they are sufficiently divergent and perform critical functions. In order to distinguish between these two possibilities, we used a data-mining approach to scan the expression of three members of the VILIP family - two closely related members of the VILIP group (HPCA and NCALD) and an outgrouped member (FREQ) (Figure 1). mRNA expression levels across tissues based on microarrays, were integrated by two different websites (Genevestigator and BioGPS). The data for NCALD, HPCA and FREQ are compiled and presented in Figure S1A and Figure S1B. The results indicate that all three are expressed in the same tissue, although the expression levels might be different. Similar conclusions are reached at the protein level based on data compiled from Human Protein Atlas (Figure S1C) and PaxDB (Figure S1D). However, in liver, FREQ is expressed at a lower level while HPCA and NCALD are absent. Even among neurons from different regions of brain, Purkinje neuron shows a comparable expression level across all three proteins, while the presence of FREQ in cerebral cortex, HPCA in Hippocampus and HPCA/NCALD in lateral ventricle are more abundant compared to the rest. These results (Figure S1) argue that there is no spatial or temporal separation of the expression of the analyzed NCS proteins, either at the mRNA or at the protein level. Furthermore, in the brain, they seem to be co-expressed in different neuronal subtypes, albeit at different levels. This would support a non-overlapping, perhaps critical, function for each tested NCS protein. Thus, the approach presented in this study, we feel, will be particularly beneficial in delineating the functional divergence among members of this group, which exhibit a high degree

of evolutionary conservation. We would predict that the functional divergence is mediated by residues that lie outside of the region that interacts with ROS-GC1, which appears to be highly conserved (Figure 3B).

Interaction of KChIPs with the downstream regulatory element

Compared to GCAPs and VILIPs, members of the KChIP group are less well characterized and are not known to modulate mGC; their primary shared function appears to be interaction with Kv channels [47]. Only 4 genes have been documented in the human genome, although several splice variants have been reported. An alignment of the known KChIP genes and several of their splice variants is depicted in Figure 4A. For the purpose of our analyses of this closely-preserved unique group, we chose a function that every KChIP is believed to carry out: namely, binding to the Downstream Regulatory Element (DRE) [21,48]. Residues that mediate this interaction were identified through experimental analyses to be K87, K90, K91, R98, K101, R160 and K166 (shaded grey in Figure 4B). Sequence divergence analyses across members of the three groups were carried out along with some variants chosen at random. Surprisingly, KChIP2 and KChIP 2.2 are completely conserved, suggesting the possibility that one sequence was perhaps reported more than once. With regard to molecular insight into the KChIP interaction with DRE, however, the analyses were beneficial once again. K87 is replaced by Asn in KChIP 1 and its variant; R98 and R101 are completely conserved in all KChIPs (indicated by asterisks; Figure 4B), while the others exhibit a conserved substitution (Arg for Lys) (Figure 4B). When compared to members from other groups, R160 and K166 are conserved throughout and, therefore, are less likely to contribute to the specific interaction [58]. K87 is preserved in recoverin, which does not bind to DRE [48] and is also unlikely to contribute to the interaction. A conserved substitution (Lys) is observed in most members at position corresponding to R98 in KChIPs decreasing the likelihood of the residue being critical for the specific interaction. Thus, the prediction would be that the remaining residues – K90, K91 and K101 – are the ones that determine the DRE-binding function in KChIPs. Of the three, only K101 remains to be experimentally validated [48].

Conclusion

This study was designed to examine if analyses of phylogenetic relationships among closely related proteins would yield insights into molecular mechanism of functional divergence. We used the NCS protein family for this purpose. Sequence alignment led to the formation of three distinct groups. While GCAPs and VILIPs share a common function – modulation of mGC activity – they cause opposing effects in their calcium-loaded states. For KChIPs, their ability to bind DRE sequences was used as the function. Experimental data on residues important for each of these specific interactions provided the platform to evaluate relative roles of these residues in each interaction. Based on sequence divergence during evolution, predictions could be made whether a given residue could contribute to the specificity of the interaction, is more likely to play a role in general molecular integrity or least likely to contribute at all. Where experimental data was available, they did indeed validate the predictions. Thus, this approach is likely to enhance experimental design in analyzing functional divergence as well as molecular interactions that mediate target interactions of proteins, especially given the abundance of sequence information across genomes.

Acknowledgements

The authors would like to thank Prof. Ronald Ellis (RUSOM) for discussions

on phylogenetic relationships and Dr. Anuradha Krishnan for critical discussions. Financial support from the University of Medicine and Dentistry of New Jersey Foundation and the Osteopathic Heritage Foundation are gratefully acknowledged.

References

1. Dessauer CW, Tesmer JJ, Sprang SR, Gilman AG (1999) The interactions of adenylate cyclases with P-site inhibitors. *Trends Pharmacol Sci* 20: 205-210.
2. Seifert R, Lushington GH, Mou TC, Gille A, Sprang SR (2012) Inhibitors of membranous adenylyl cyclases. *Trends PharmacolSci* 33: 64-78.
3. Sinha SC, Sprang SR (2006) Structures, mechanism, regulation and evolution of class III nucleotidylcyclases. *Rev Physiol Biochem Pharmacol* 157: 105-140.
4. Tesmer JJ, Sprang SR (1998) The structure, catalytic mechanism and regulation of adenylyl cyclase. *Curr Opin Struct Biol* 8: 713-719.
5. Ikura M (1996) Calcium binding and conformational response in EF-hand proteins. *Trends Biochem Sci* 21: 14-17.
6. Moncrief ND, Kretsinger RH, Goodman M (1990) Evolution of EF-hand calcium-modulated proteins. I. Relationships based on amino acid sequences. *J Mol Evol* 30: 522-562.
7. Dizhoor AM, Ericsson LH, Johnson RS, Kumar S, Olshevskaya E, et al. (1992) The NH2 terminus of retinal recoverin is acylated by a small family of fatty acids. *J BiolChem* 267: 16033-16036.
8. Tanaka T, Ames JB, Harvey TS, Stryer L, Ikura M (1995) Sequestration of the membrane-targeting myristoyl group of recoverin in the calcium-free state. *Nature* 376: 444-447.
9. Zozulya S, Stryer L (1992) Calcium-myristoyl protein switch. *ProcNatAcadSci U S A* 89: 11569-11573.
10. Hilfiker S (2003) Neuronal calcium sensor-1: a multifunctional regulator of secretion. *Biochem Soc Trans* 31: 828-832.
11. Weiss JL, Hui H, Burgoyne RD (2010) Neuronal calcium sensor-1 regulation of calcium channels, secretion, and neuronal outgrowth. *Cell Mol Neurobiol* 30: 1283-1292.
12. Weiss JL, Burgoyne RD (2002) Sense and sensibility in the regulation of voltage-gated Ca(2+) channels. *Trends Neurosci* 25: 489-491.
13. Gomez M, De Castro E, Guarín E, Sasakura H, Kuhara A, et al. (2001) Ca²⁺ signaling via the neuronal calcium sensor-1 regulates associative learning and memory in *C. elegans*. *Neuron* 30: 241-248.
14. Noguchi H, Kobayashi M, Miwa N, Takamatsu K (2007) Lack of hippocalcin causes impairment in Ras/extracellular signal-regulated kinase cascade via a Raf-mediated activation process. *J Neurosci Res* 85: 837-844.
15. Dizhoor AM, Boikov SG, Olshevskaya EV (1998) Constitutive activation of photoreceptor guanylatecyclase by Y99C mutant of GCAP-1. Possible role in causing human autosomal dominant cone degeneration. *J Biol Chem* 273: 17311-17314.
16. Downes SM, Payne AM, Kelsell RE, Fitzke FW, Holder GE, et al. (2001) Autosomal dominant cone-rod dystrophy with mutations in the guanylate cyclase 2D gene encoding retinal guanylate cyclase-1. *Arch Ophthalmol* 119: 1667-1673.
17. Payne AM, Downes SM, Bessant DA, Taylor R, Holder GE, et al. (1998) A mutation in guanylatecyclase activator 1A (GUCA1A) in an autosomal dominant cone dystrophy pedigree mapping to a new locus on chromosome 6p21.1. *Hum Mol Genet* 7: 273-277.
18. Duda T, Krishnan A, Venkataraman V, Lange C, Koch KW, et al. (1999) Mutations in the rod outer segment membrane guanylatecyclase in a cone-rod dystrophy cause defects in calcium signaling. *Biochem* 38: 13912-13919.
19. Duda T, Venkataraman V, Goraczniak R, Lange C, Koch KW, et al. (1999) Functional consequences of a rod outer segment membrane guanylatecyclase (ROS-GC1) gene mutation linked with Leber's congenital amaurosis. *Biochem* 38: 509-515.
20. Duda T, Venkataraman V, Jankowska A, Lange C, Koch KW, et al. (2000) Impairment of the rod outer segment membrane guanylatecyclase dimerization in a cone-rod dystrophy results in defective calcium signaling. *Biochem* 39: 12522-12533.
21. Burgoyne RD (2007) Neuronal calcium sensor proteins: generating diversity in neuronal Ca²⁺ signalling. *Nat Rev Neurosci* 8: 182-193.

22. Burgoyne RD, Haynes LP (2012) Understanding the physiological roles of the neuronal calcium sensor proteins. *Mol Brain* 5: 2.
23. Burgoyne RD, Weiss JL (2001) The neuronal calcium sensor family of Ca²⁺-binding proteins. *Biochem J* 353: 1-12.
24. Palczewski K, Polans AS, Baehr W, Ames JB (2000) Ca²⁺-binding proteins in the retina: structure, function, and the etiology of human visual diseases. *BioEssays: news and reviews in molecular, cellular and developmental biology* 22: 337-350.
25. Ames JB, Hendricks KB, Strahl T, Huttner IG, Hamasaki N, et al. (2000) Structure and calcium-binding properties of Frq1, a novel calcium sensor in the yeast *Saccharomyces cerevisiae*. *Biochem* 39: 12149-12161.
26. Braunewell KH, Brackmann M, Schaupp M, Spilker C, Anand R, et al. (2001) Intracellular neuronal calcium sensor (NCS) protein VILIP-1 modulates cGMP signalling pathways in transfected neural cells and cerebellar granule neurones. *J Neurochem* 78: 1277-1286.
27. Krishnan A, Venkataraman V, Fik-Rymarkiewicz E, Duda T, Sharma RK (2004) Structural, biochemical, and functional characterization of the calcium sensor neurocalcin delta in the inner retinal neurons and its linkage with the rod outer segment membrane guanylatecyclase transduction system. *Biochem* 43: 2708-2723.
28. Spilker C, Gundelfinger ED, Braunewell KH (2002) Evidence for different functional properties of the neuronal calcium sensor proteins VILIP-1 and VILIP-3: from subcellular localization to cellular function. *Biochim Biophys Acta* 1600: 118-127.
29. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinform* 23: 2947-2948.
30. Perrière G, Gouy M (1996) WWW-query: an on-line retrieval system for biological sequence banks. *Biochimie* 78: 364-369.
31. Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, et al. (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinform* 2008: 420747.
32. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, et al. (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A* 101: 6062-6067.
33. Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, et al. (2010) Towards a knowledge-based Human Protein Atlas. *Nat Biotechnol* 28: 1248-1250.
34. Uhlén M, Björling E, Agaton C, Szigarty CA, Amini B, et al. (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics* 4: 1920-1932.
35. Pontén F, Gry M, Fagerberg L, Lundberg E, Asplund A, et al. (2009) A global view of protein expression in human cells, tissues, and organs. *Mol Syst Biol* 5: 337.
36. Wang M, Weiss M, Simonovic M, Haertinger G, Schrimpf SP, et al. (2012) PaxDb, a database of protein abundance averages across all three domains of life. *Mol Cell Proteomics* 11: 492-500.
37. Seaton G, Hogg EL, Jo J, Whitcomb DJ, Cho K (2011) Sensing change: the emerging role of calcium sensors in neuronal disease. *Semin Cell Dev Biol* 22: 530-535.
38. Detwiler P (2002) Open the loop: dissecting feedback regulation of a second messenger transduction cascade. *Neuron* 36: 3-4.
39. Koch KW, Duda T, Sharma RK (2010) Ca²⁺-modulated vision-linked ROS-GC guanylatecyclase transduction machinery. *Mol Cell Biochem* 334: 105-115.
40. Pugh EN Jr, Lamb TD (1990) Cyclic GMP and calcium: the internal messengers of excitation and adaptation in vertebrate photoreceptors. *Vision Res* 30: 1923-1948.
41. Pertzov A, Duda T, Sharma RK (2010) Ca²⁺ sensor GCAP1: A constitutive element of the ONE-GC-modulated odorant signal transduction pathway. *Biochem* 49: 7303-7313.
42. Brackmann M, Schuchmann S, Anand R, Braunewell KH (2005) Neuronal Ca²⁺ sensor protein VILIP-1 affects cGMP signalling of guanylylcyclase B by regulating clathrin-dependent receptor recycling in hippocampal neurons. *J Cell Sci* 118: 2495-2505.
43. Buttgerit J, Qadri F, Monti J, Langenickel TH, Dietz R, et al. (2010) Visinin-like protein 1 regulates natriuretic peptide receptor B in the heart. *Regul Pept* 161: 51-57.
44. Fik-Rymarkiewicz E, Duda T, Sharma RK (2006) Novel frequenin-modulated Ca²⁺-signaling membrane guanylatecyclase (ROS-GC) transduction pathway in bovine hippocampus. *Mol Cell Biochem* 291: 187-204.
45. Krishnan A, Duda T, Pertzov A, Kobayashi M, Takamatsu K, et al. (2009) Hippocalcin, new Ca²⁺ sensor of a ROS-GC subfamily member, ONE-GC, membrane guanylatecyclase transduction system. *Mol Cell Biochem* 325: 1-14.
46. Venkataraman V, Duda T, Ravichandran S, Sharma RK (2008) Neurocalcin delta modulation of ROS-GC1, a new model of Ca²⁺ signaling. *Biochem* 47: 6590-6601.
47. Wang K (2008) Modulation by clamping: Kv4 and KChIP interactions. *Neurochem Res* 33: 1964-1969.
48. Link WA, Ledo F, Torres B, Palczewska M, Madsen TM, et al. (2004) Day-night changes in downstream regulatory element antagonist modulator/potassium channel interacting protein activity contribute to circadian gene expression in pineal gland. *J Neurosci* 24: 5346-5355.
49. Ames JB, Lim S, Ikura M (2012) Molecular structure and target recognition of neuronal calcium sensor proteins. *Front Mol Neurosci* 5: 10.
50. Chen CK (2002) Recoverin and rhodopsin kinase. *Adv Exp Med Biol* 514: 101-107.
51. Koch KW (2012) Biophysical investigation of retinal calcium sensor function. *Biochim Biophys Acta* 1820: 1228-1233.
52. Ermilov AN, Olshevskaya EV, Dizhoor AM (2001) Instead of binding calcium, one of the EF-hand structures in guanylylcyclase activating protein-2 is required for targeting photoreceptor guanylylcyclase. *J Biol Chem* 276: 48143-48148.
53. Pettelkau J, Schröder T, Ihling CH, Olausson BE, Kölbl K, et al. (2012) Structural insights into retinal guanylylcyclase-GCAP-2 interaction determined by cross-linking and mass spectrometry. *Biochem* 51: 4932-4949.
54. Imanishi Y, Yang L, Sokal I, Filipek S, Palczewski K, et al. (2004) Diversity of guanylatecyclase-activating proteins (GCAPs) in teleost fish: characterization of three novel GCAPs (GCAP4, GCAP5, GCAP7) from zebrafish (*Danio rerio*) and prediction of eight GCAPs (GCAP1-8) in pufferfish (*Fugu rubripes*). *J Mol Evol* 59: 204-217.
55. Stephen R, Palczewski K, Sousa MC (2006) The crystal structure of GCAP3 suggests molecular mechanism of GCAP-linked cone dystrophies. *J Mol Biol* 359: 266-275.
56. Lange C, Duda T, Beyermann M, Sharma RK, Koch KW (1999) Regions in vertebrate photoreceptor guanylylcyclase ROS-GC1 involved in Ca²⁺-dependent regulation by guanylylcyclase-activating protein GCAP-1. *FEBS Lett* 460: 27-31.
57. Baehr W, Palczewski K (2007) Guanylate cyclase-activating proteins and retina disease. *Subcell Biochem* 45: 71-91.
58. Lusin JD, Vanarotti M, Li C, Valiveti A, Ames JB (2008) NMR structure of DREAM: Implications for Ca²⁺-dependent DNA binding and protein dimerization. *Biochem* 47: 2252-2264.