

Research article

Open Access

Ancient conserved domains shared by animal soluble guanylyl cyclases and bacterial signaling proteins

Lakshminarayan M Iyer*, Vivek Anantharaman and L Aravind*

Address: National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA

Email: Lakshminarayan M Iyer* - lakshmin@ncbi.nlm.nih.gov; Vivek Anantharaman - ananthar@ncbi.nlm.nih.gov;

L Aravind* - aravind@ncbi.nlm.nih.gov

* Corresponding authors

Published: 3 February 2003

Received: 4 December 2002

BMC Genomics 2003, 4:5

Accepted: 3 February 2003

This article is available from: <http://www.biomedcentral.com/1471-2164/4/5>

© 2003 Iyer et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Soluble guanylyl cyclases (SGCs) are dimeric enzymes that transduce signals downstream of nitric oxide (NO) in animals. They sense NO by means of a heme moiety that is bound to their N-terminal extensions.

Results: Using sequence profile searches we show that the N-terminal extensions of the SGCs contain two globular domains. The first of these, the HNOB (Heme NO Binding) domain, is a predominantly α -helical domain and binds heme via a covalent linkage to histidine. Versions lacking this conserved histidine and are likely to interact with heme non-covalently. We detected HNOB domains in several bacterial lineages, where they occur fused to methyl accepting domains of chemotaxis receptors or as standalone proteins. The standalone forms are encoded by predicted operons that also contain genes for two component signaling systems and GGDEF-type nucleotide cyclases. The second domain, the HNOB associated (HNOBA) domain occurs between the HNOB and the cyclase domains in the animal SGCs. The HNOBA domain is also detected in bacteria and is always encoded by a gene, which occurs in the neighborhood of a gene for a HNOB domain.

Conclusion: The HNOB domain is predicted to function as a heme-dependent sensor for gaseous ligands, and transduce diverse downstream signals, in both bacteria and animals. The HNOBA domain functionally interacts with the HNOB domain, and possibly binds a ligand, either in cooperation, or independently of the latter domain. Phyletic profiles and phylogenetic analysis suggest that the HNOB and HNOBA domains were acquired by the animal lineage via lateral transfer from a bacterial source.

Background

Binding and recognizing diverse small molecules is a central aspect of signal transduction mechanisms in all cellular life forms. A variety of environmental small molecules, namely, nutrients, xenobiotics and first messengers are recognized by cells, with the help of specialized protein sensors on their surfaces [1,2]. Additionally cells also use protein-bound small molecules, such as FAD, cinnamic acid, tetrapyrroles and heme as sensors of photons and

the ambient redox states [3,4]. Within cells, small molecules, such as cyclic nucleotides, are used as intracellular messengers to transduce signals arising from a variety of stimuli [1,2]. Over the past few years a combination of protein sequence analysis and biochemical studies have revealed several unifying principles that govern the recognition of small molecules by cells [4–9]. A significant component of small molecule-protein interactions is mediated via a relatively small set of ancient conserved

protein domains that bind their ligands using specialized pockets [4,10]. Certain protein folds have given rise to several large superfamilies of ligand-binding domains. These include the PAS-like fold, which is the scaffold of the PAS [11,12], GAF [13], and probably the CACHE [14] superfamilies, and the ACT-like fold, which is seen in ACT, ferredoxin and related ligand-binding domains [4,15]. Identification of such conserved domains has, often, improved our understanding of the general structural and biochemical features that underlie protein-small molecule interactions.

Typically, these small molecule-binding domains (SMBDs) are combined, within the same polypeptide, to a number of conserved domains that are directly involved in signal transduction. These include enzymatic domains such as nucleotide cyclases, histidine and serine/threonine kinases, cNMP phosphodiesterases, receiver domains and several different kinds of ATPases [3,4]. Non-catalytic domains, like DNA- or RNA-binding modules (eg. the helix-turn-helix domains), methyl-accepting domains, and FHA domains that bind phosphorylated peptides, are also often fused to SMBDs [3,4].

The soluble guanylyl cyclases are important signaling molecules in animals that transduce signals mediated by the gaseous first messenger, nitric oxide [16–19]. In animals, NO functions as a neurotransmitter both in the central and peripheral nervous system. NO is released by a cell through the action of the nitric oxide synthase (NOS) on its substrate, the amino acid arginine [20]. NO diffuses through the neighboring cells and elicits its functions by activating a soluble guanylyl cyclase (SGC) that synthesizes cGMP using GTP as a substrate [16,18]. The animal SGCs are dimeric enzymes comprised of two homologous subunits, α and β . The sequence similarity between these subunits extends throughout their entire length, and they contain a long extension N-terminal to the cNMP generating catalytic domain. The N-terminal region of the β subunit forms a covalent link with heme via a histidine (H105 in human SGC1 β) [21]. Though the porphyrin is identical to that found in hemoglobin, it only binds NO and carbon monoxide, and not oxygen, even though O₂ is present at much higher concentrations within the cell [22,23]. This ligand specificity of the NO binding domain of animal SGCs is very similar to bacterial cytochrome c' that likewise binds only NO and CO [24]. The N-terminal region of the α subunit has been shown to bind certain synthetic pyrazolopyridine ligands, which act as potent NO-independent agonists of cyclase activity [25,26].

Using sensitive sequence profile analysis we show that the heme-binding domain of the animal SGCs defines a novel family of ligand binding domains that are also widely distributed in bacteria. The SGCs also contain a second glob-

ular domain situated in between the heme-binding and cyclase domains. We show that this domain is also present in the bacteria, and present evidence that these two domains are likely to function together both in eukaryotes and bacteria, and activate a diverse set of downstream signals.

Results and Discussion

Identification of the HNOB and HNOBA domains

Though both the α and β subunits of SGCs are homologous throughout their entire length, only the β subunit contains the critical heme-binding residue. As this long homologous N-terminal extension of SGC subunits does not map to any previously characterized domains, we sought to investigate its provenance through sequence analysis methods. The length and conservation pattern of the N-terminal domain suggested that the N-terminal extension of the SGCs is likely to comprise of two globular domains in the least. This conjecture was also supported by an analysis of sequence complexity of the polypeptide using the SEG program [27].

A PSI-BLAST search was initiated with the conserved N-terminal extension of the SGC (human SGC1 β , gi: 4504215, region 1–360), using an inclusion threshold of .01, and compositional bias based statistics to eliminate false positives arising due to peculiarities of sequence composition. Both the N- and the C-terminal parts of this extension gave several distinct hits to different bacterial proteins, supporting the presence of two distinct globular domains in this extension. Based on these hits we divided the extension into N- and C-terminal parts and initiated separate PSI-BLAST searches with them. Searches with the N-terminal part of the extension gave significant hits to bacterial proteins of the length 180–195 residues within the first 3 iterations (eg. Mdge1313 from *Microbulbifer degradans* is detected with an expect-value (e) of 10⁻⁴ in the first iteration). This region of similarity encompasses the entire length of these bacterial proteins, and includes the heme-liganding histidine of the SGC β subunits, and is likely to define a distinct globular domain. In order to probe this domain further we isolated this region from SGC1 β (1–185) and seeded a PSI-BLAST search that was run to convergence. At convergence, this search detected, in addition to the soluble guanylyl cyclases from the animals, several bacterial proteins, including certain methyl-accepting chemotaxis receptors from bacteria such as *Desulfovibrio* and *Clostridium*. Transitive searches initiated with the small bacterial proteins that entirely correspond to the N-terminal-most domain of the SGCs also recovered the same set of proteins as observed in the earlier search, with $e < .01$. These observations suggest that the N-terminal-most domain of the SGCs is evolutionarily mobile: it occurs in several distinct contexts with other domains in the same polypeptide, or in a stand-alone form.

PSI-BLAST searches initiated with the C-terminal part of the SGC-specific extension (the region in between the above-detected N-terminal-most domain, and the C-terminal cyclase domain; human SGC1 β , region: 200–370) recovered homologous regions from all other animal soluble guanylyl cyclases, and, additionally, N-terminal regions of histidine kinases from *Nostoc* and *Anabaena* species (eg. gi: 17229771, $e = 10^{-4}$ in iteration 1) and a diguanylate cyclase from *Rhodobacter sphaeroides* (gi: 22958462 $e = 10^{-3}$ in iteration 3). The region of similarity shared by these bacterial proteins and animal SGCs encompassed more-or-less the entire middle region, which is present between the two other globular domains. Reciprocal searches with the corresponding region from the cyanobacterial histidine kinases recovered the animal SGCs, thereby confirming the evolutionary relationship between these regions. These observations suggest that the middle region of the animal SGCs defines a second evolutionarily mobile globular domain that is shared with several distinct bacterial proteins.

A multiple alignment of the N-terminal-most domain of the SGCs and its bacterial homologs was prepared using the T_coffee program and adjusted manually based on the PSI-BLAST HSPs (Fig. 1). The multiple alignment shows that the heme-binding histidine of the SGC- β subunits is contained within this domain. We accordingly termed this domain the HNOB (Heme-Nitric Oxide Binding) domain. The histidine is predicted to be positioned next to the N-terminus of an α -helix, and is conserved throughout this superfamily, with the exception of α -subunits of the SGCs (Fig. 1) and two bacterial versions from *Rhodobacter sphaeroides* and *Magnetococcus* species. Though SGC α -subunits lack the histidine, the region encompassing the HNOB domain in human SGC1 α has been shown to be important for functional heme binding by the dimeric enzyme [28]. Hence, in addition to covalent heme binding, HNOB domains may also interact with it non-covalently. The N-terminal part of the domain contains a universally conserved acidic residue that may play a critical role creating the local environment for NO selectivity. The rest of the sequence conservation maps mainly to the cores of secondary structure elements and some well conserved turns (Fig. 1). Secondary structure prediction shows that the HNOB domain is likely to adopt a predominantly α -helical fold with a few interspersed extended elements. Further, the predicted trans-membrane topologies of HNOB domain proteins suggest an intracellular localization for all occurrences of this domain. However, neither direct comparisons, nor sequence-structure threading revealed any relationship with the two other well-characterized α -helical heme-binding domains, namely globin and cytochrome C. Hence, it is likely that the HNOB domain adopts a structure distinct from these domains.

Because the middle domain of the SGCs is always associated with the HNOB domains (see below), we named it the HNOBA (HNOB-associated) domain. A multiple alignment of this domain revealed, that unlike the HNOB domain, it is likely to adopt an $\alpha + \beta$ fold with segregated α and β domains (Fig. 2). Its core contains 5 conserved strands, followed, at the extreme C-terminus, by a long helix that is likely to form a coiled-coil linker region involved in dimerization (Fig. 2). Within the core the first two strands are predicted to form a β -hairpin followed by a helical region, which in turn is followed by 3 consecutive strands and single α -helix. A similar secondary structure pattern is also observed in several other ligand-binding domains such as the PAS, GAF, profilin and probably even the CACHE domains [10]. Sequence-structure threading using the 3DPSSM method detected the PAS domain (Photoactive yellow protein PDB: 1drm) with marginally significant scores (.5, 50–70% probability of correct fold prediction). This suggests that the HNOBA domain could potentially contain a PAS-like fold.

All versions of the HNOBA domain, except one from *Rhodobacter sphaeroides*, contain a conserved histidine between the core region and the long C-terminal α -helix that links it to a kinase or cyclase domain (Fig. 2). However, studies on the SGCs suggest no evidence for a second covalent heme-binding site, implying that this histidine could have some other function. Interestingly, binding studies on a synthetic pyrazolopyridine SGC agonist, BAY 41–2272 have suggested a role for the region 236–290 of the human SGC1 α -subunit in interacting with this ligand [25,26]. Though actual cross-linking occurred with cysteine residues situated just N-terminal to the HNOBA domain, it is likely that the rest of domain is involved in contact with the ligand. This may suggest the possibility of the HNOBA domain binding some, as yet unknown, endogenous ligand. In such a scenario it could participate in allosteric regulation of the SGCs, similar to the GAF domains of the cyclic nucleotide phosphodiesterases [6,13].

Evolutionary history, domain architectures, and gene neighborhoods of the HNOB and HNOBA domain proteins

The HNOB domain is found in various bacterial lineages such as cyanobacteria, proteobacteria and low GC Gram-positive bacteria. It is entirely absent in all archaeal proteomes available to date, and amidst the eukaryotes is only detectable in the animal lineage. In bacteria, the HNOB domain proteins display two principal architectures. They may occur as stand-alone proteins, entirely composed of the HNOB domain, in several bacterial lineages (Fig. 3). In the low GC Gram-positive bacteria and *Desulfovibrio desulfuricans* they occur in combination with the methyl accepting chemotaxis receptor proteins, with or without the HAMP domain [29,30]. Additionally, in

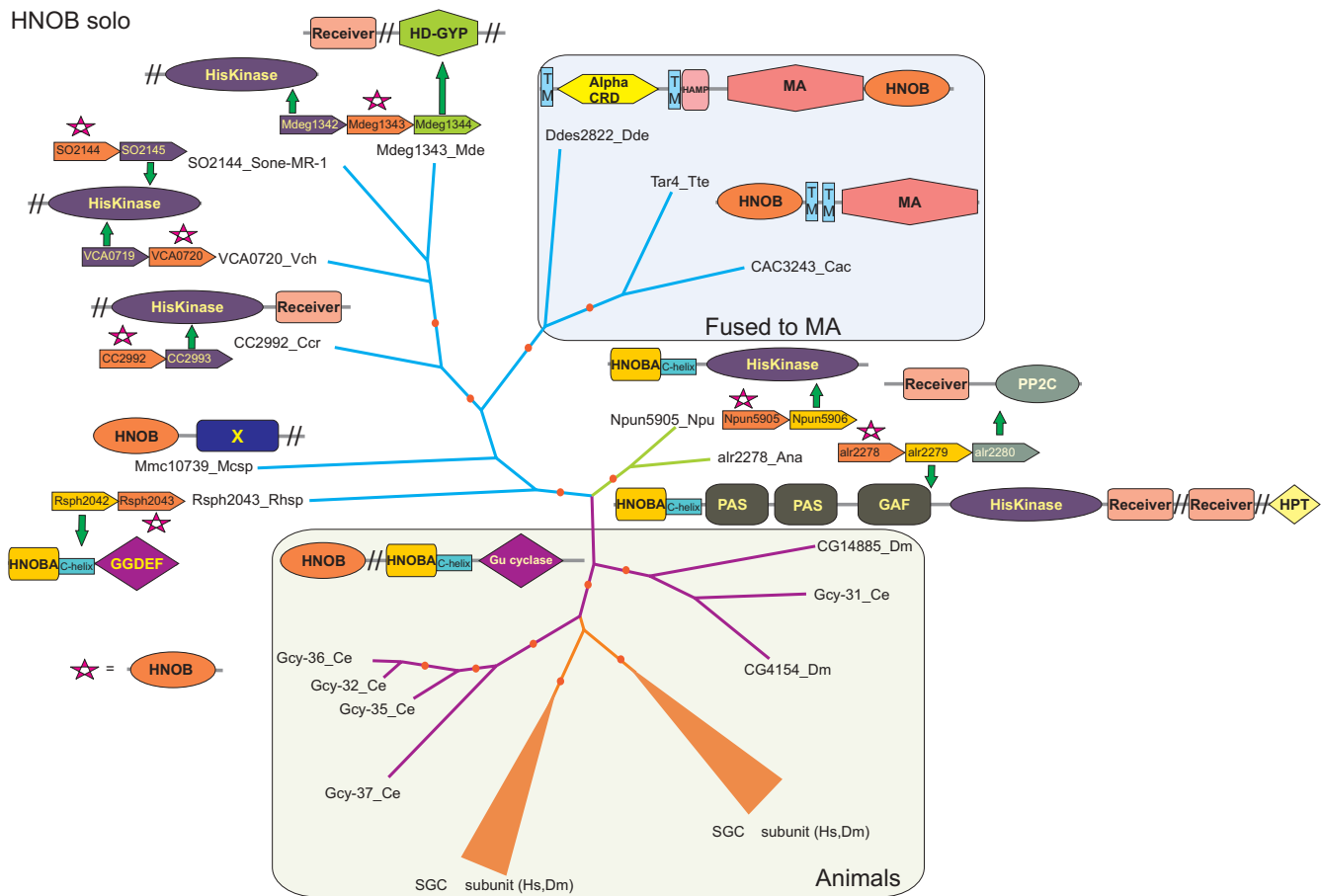


Figure 3
Phylogenetic tree, domain architectures and gene neighborhoods of proteins containing HNOB domains. Internal branches with RELL Bootstrap support >80% are indicated with red circles. Selected gene neighborhoods of HNOB domain proteins that provide contextual functional information are shown as box-arrows (The gene neighborhoods were determined as explained in the methods). The HNOB domain encoding genes are all colored dark orange, while other genes in the predicted operons are colored differently to match with the product they encode. All genes encoding proteins with a standalone HNOB domain are marked with red asterisks. Additionally, domain architectures are shown for all multidomain proteins with a HNOB domain and the products of their gene neighbors (latter are the indicated by green arrows). The globular domains are drawn approximately to scale, but low complexity linker regions are not shown (indicated by the "//"). Protein domain names generally follow accepted abbreviations. Additional domain abbreviations: TM: Transmembrane, MA: Methyl accepting chemotaxis receptor domain, HD-GYP: predicted cyclic diguanylate phosphodiesterase of the HD hydrolase superfamily, PP2C: PP2C-like phosphatase domain, HPT – Histidine Phosphotransfer domain, Receiver: Receiver domain of the two component system, alpha-CRD is a previously, unidentified α -helical receptor domain, with conserved aspartates, that is found in several prokaryotic chemotaxis receptors. C-helix: Conserved linker helix present at the C-terminus of the HNOBA domain. Gene names and species abbreviations are as in figure 1.

proteins contain HNOBA domains at their extreme N-terminus (Fig. 3). Thus, even in the bacteria, the HNOBA domain always appears to function in association with the HNOB domain. Given that these cyanobacterial HNOB proteins are also the closest bacterial relatives of the animal HNOB domains, it appears likely that the animal lineage acquired a related assemblage of HNOB and HNOBA domain as a single piece. Thus, the HNOB and

HNOBA domains resemble functionally similar CACHE and CHASE domains, which have also been acquired by certain eukaryotic lineages via lateral transfer from bacteria [14,36,37]. A phylogenetic analysis of the guanylyl cyclase domains of the animal SGCs shows that their closest relatives are cyclases from various bacteria such as cyanobacteria, and *Leptospira* (data not shown). Taken together, these observations support a potential bacterial origin for

these components of the nitric oxide signaling pathway. This is also consistent with the presence of orthologs of the animal NO synthases in several bacteria [38,39].

The co-occurrence of the HNOB and the HNOBA domains in either the same protein or proteins encoded by the same operon suggests a strong functional interaction between them. Studies on the human SGC1 suggest that the synthetic pyrazolopyridine ligand binds at site distinct from NO and heme, but requires the heme-binding site for its action [25,26]. This could imply a synergy or cooperation between these domains in heme interaction and cyclase activation. Majority of the bacterial HNOB domain proteins are predicted to covalently bind heme. Those versions that lack the heme-binding histidine might either non-covalently interact with heme, as in the case of the SGC α -subunits, or bind some other unknown ligand. They are likely to function as sensors of diffusible gaseous ligands that activate at least three distinct, downstream signaling pathways: namely phosphotransfer through the two-component systems, methyl ester-dependent chemotactic response and cyclic nucleotide signaling through the diguanylate cyclases/phosphodiesterases. However, further experimental studies would be required to determine if they function as NO or CO sensors, like the animal proteins, or as oxygen sensors.

Conclusions

Using sequence profile searches we identify two conserved domains in the N-terminal extensions of the soluble guanylyl cyclases from animals. One of these, the HNOB domain, contains the heme-binding site of the SGCs and defines a family of small-molecule binding domains exclusively occurring in animals and bacteria. In bacteria the HNOB domain occurs either as a standalone version or fused to the HAMP and methyl accepting domains of chemotaxis receptors. The bacterial solo versions co-occur in predicted operons encoding genes for two component systems and diguanylate cyclases or phosphodiesterases, and are likely to transmit signals to these downstream molecules. The second conserved domain of SGCs, the HNOBA domain is always associated with the HNOB domain either in the same polypeptide or in polypeptide encoded by neighboring genes in an operon. It is predicted to adopt an $\alpha + \beta$ fold that is possibly related to the PAS-like fold. The HNOB and HNOBA domains are likely to function synergistically or cooperatively. The latter may either cooperate in binding heme or might bind distinct ligands.

The phyletic pattern of these domains, and phylogenetic analysis of the HNOB domain suggest that the animal versions were probably acquired through lateral transfer from a bacterial source. Identification of the HNOB and HNOBA domains could help in further investigations of

the SGCs that are critical components of the nitric oxide signaling pathway in animals. Furthermore, investigating their role in bacteria, such as *Vibrio cholerae*, could be important in understanding the sensory mechanisms of human pathogens.

Methods

The non-redundant (NR) database of protein sequences (National Center for Biotechnology Information, NIH, Bethesda) was searched using the BLASTP and PSI-BLAST programs [40]. Profile searches using the PSI-BLAST program were conducted either with a single sequence or an alignment used as the query, with a profile inclusion expectation (E) value threshold of 0.01, and were iterated until convergence. Multiple alignments were constructed using the T_Coffee program [41], followed by manual correction based on the PSI-BLAST results. Protein secondary structure was predicted using a multiple alignment as the input for the JPRED and PHD programs [42–44]. Preliminary clustering of proteins was done using the BLASTCLUST program with empirically determined length and score threshold cut off values (For documentation see <ftp://ftp.ncbi.nih.gov/blast/documents/README.bcl>). Previously known, conserved domains were identified using PSI-BLAST derived profiles for them, with the RPS-BLAST program [45]. Sequence-structure threading was performed using the 3DPSSM server <http://www.sbg.bio.ic.ac.uk/~3dpssm/>. Gene neighborhoods were obtained by isolating all conserved genes, in the neighborhood of the gene under consideration, which showed a separation of less than 70 nucleotides between their termini. Genes fulfilling this criterion and occurring in the same direction were considered likely to form operons.

Phylogenetic analysis was performed using the neighbor joining or least square method followed by local rearrangements using the maximum likelihood algorithm to predict the most likely tree. The robustness of tree topology was assessed with 10,000 Resampling of Estimated Log Likelihoods (RELL) bootstrap replicates. The MOLPHY and Phylip software packages were used for phylogenetic analyses [46,47].

Authors' contributions

LMI contributed to the discovery process, and preparation of figure 3. VA contributed to the preparation of the multiple alignments and figure 1 and 2. LA contributed to the discovery process and prepared the manuscript. All authors read and approved the final manuscript.

References

1. Lodish H, Berk A, Zipursky SL and Matsudaira P **Molecular Cell Biology** New York, NY: W H Freeman & Co 1999,
2. Alberts B, Johnson A, Lewis J, Raff M, Roberts K and Walter P **Molecular Biology of the Cell** New York, NY: Garland Pub 2002,

3. Taylor BL, Zhulin IB and Johnson MS **Aerotaxis and other energy-sensing behavior in bacteria.** *Annu Rev Microbiol* 1999, **53**:103-128
4. Anantharaman V, Koonin EV and Aravind L **Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains.** *J Mol Biol* 2001, **307**:1271-1292
5. Morais Cabral JH, Lee A, Cohen SL, Chait BT, Li M and Mackinnon R **Crystal structure and functional analysis of the HERG potassium channel N terminus: a eukaryotic PAS domain.** *Cell* 1998, **95**:649-655
6. Martinez SE, Wu AY, Glavas NA, Tang XB, Turley S, Hol WG and Beavo JA **The two GAF domains in phosphodiesterase 2A have distinct roles in dimerization and in cGMP binding.** *Proc Natl Acad Sci U S A* 2002, **99**:13260-13265
7. Ho YS, Burden LM and Hurley JH **Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor.** *Embo J* 2000, **19**:5288-5299
8. Perutz MF, Paoli M and Lesk AM Fix L, **a haemoglobin that acts as an oxygen sensor: signalling mechanism and structural basis of its homology with PAS domains.** *Chem Biol* 1999, **6**:R291-297
9. Gong W, Hao B, Mansy SS, Gonzalez G, Gilles-Gonzalez MA and Chan MK **Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction.** *Proc Natl Acad Sci U S A* 1998, **95**:15177-15182
10. Aravind L, Mazumder R, Vasudevan S and Koonin EV **Trends in protein evolution inferred from sequence and structure analysis.** *Curr Opin Struct Biol* 2002, **12**:392-399
11. Ponting CP and Aravind L **PAS: a multifunctional domain family comes to light.** *Curr Biol* 1997, **7**:R674-677
12. Zhulin IB, Taylor BL and Dixon R **PAS domain S-boxes in Archaea, Bacteria and sensors for oxygen and redox.** *Trends Biochem Sci* 1997, **22**:331-333
13. Aravind L and Ponting CP **The GAF domain: an evolutionary link between diverse phototransducing proteins.** *Trends Biochem Sci* 1997, **22**:458-459
14. Anantharaman V and Aravind L **Cache – a signaling domain common to animal Ca(2+)-channel subunits and a class of prokaryotic chemotaxis receptors.** *Trends Biochem Sci* 2000, **25**:535-537
15. Aravind L and Koonin EV **Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches.** *J Mol Biol* 1999, **287**:1023-1040
16. Snyder SH and Brecht DS **Nitric oxide as a neuronal messenger.** *Trends Pharmacol Sci* 1991, **12**:125-128
17. Denninger JW and Marletta MA **Guanylate cyclase and the NO/cGMP signaling pathway.** *Biochim Biophys Acta* 1999, **1411**:334-350
18. Murad F **Regulation of cytosolic guanylyl cyclase by nitric oxide: the NO-cyclic GMP signal transduction system.** *Adv Pharmacol* 1994, **26**:19-33
19. Moncada S, Palmer RM and Higgs EA **Nitric oxide: physiology, pathophysiology, and pharmacology.** *Pharmacol Rev* 1991, **43**:109-142
20. Marletta MA **Nitric oxide synthase: aspects concerning structure and catalysis.** *Cell* 1994, **78**:927-930
21. Wedel B, Humbert P, Harteneck C, Foerster J, Malkewitz J, Bohme E, Schultz G and Koesling D **Mutation of His-105 in the beta 1 subunit yields a nitric oxide-insensitive form of soluble guanylyl cyclase.** *Proc Natl Acad Sci U S A* 1994, **91**:2592-2596
22. Gerzer R, Bohme E, Hofmann F and Schultz G **Soluble guanylate cyclase purified from bovine lung contains heme and copper.** *FEBS Lett* 1981, **132**:71-74
23. Stone JR and Marletta MA **Soluble guanylate cyclase from bovine lung: activation with nitric oxide and carbon monoxide and spectral characterization of the ferrous and ferric states.** *Biochemistry* 1994, **33**:5636-5640
24. Lawson DM, Stevenson CE, Andrew CR and Eady RR **Unprecedented proximal binding of nitric oxide to heme: implications for guanylate cyclase.** *Embo J* 2000, **19**:5661-5671
25. Becker EM, Alonso-Alija C, Apeler H, Gerzer R, Minuth T, Pleibeta U, Schmidt P, Schramm M, Schroder H, Schroeder W, Steinke W, Straub A and Stasch JP **NO-independent regulatory site of direct sGC stimulators like YC-1 and BAY 41-2272.** *BMC Pharmacol* 2001, **1**:13
26. Stasch JP, Becker EM, Alonso-Alija C, Apeler H, Dembowski K, Feurer A, Gerzer R, Minuth T, Perzborn E, Pleiss U, Schroder H, Schroeder W, Stahl E, Steinke W, Straub A and Schramm M **NO-independent regulatory site on soluble guanylate cyclase.** *Nature* 2001, **410**:212-215
27. Wootton JC **Non-globular domains in protein sequences: automated segmentation using complexity measures.** *Comput Chem* 1994, **18**:269-285
28. Foerster J, Harteneck C, Malkewitz J, Schultz G and Koesling D **A functional heme-binding site of soluble guanylyl cyclase requires intact N-termini of alpha 1 and beta 1 subunits.** *Eur J Biochem* 1996, **240**:380-386
29. Le Moual H and Koshland DE Jr **Molecular evolution of the C-terminal cytoplasmic domain of a superfamily of bacterial receptors involved in taxis.** *J Mol Biol* 1996, **261**:568-585
30. Aravind L and Ponting CP **The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins.** *FEMS Microbiol Lett* 1999, **176**:111-116
31. Dandekar T, Snel B, Huynen M and Bork P **Conservation of gene order: a fingerprint of proteins that physically interact.** *Trends Biochem Sci* 1998, **23**:324-328
32. Wolf YI, Rogozin IB, Kondrashov AS and Koonin EV **Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context.** *Genome Res* 2001, **11**:356-372
33. Galperin MY, Natale DA, Aravind L and Koonin EV **A specialized version of the HD hydrolase domain implicated in signal transduction.** *J Mol Microbiol Biotechnol* 1999, **1**:303-305
34. Ausmees N, Mayer R, Weinhouse H, Volman G, Amikam D, Ben-Ziman M and Lindberg M **Genetic data indicate that proteins containing the GGDEF domain possess diguanylate cyclase activity.** *FEMS Microbiol Lett* 2001, **204**:163-167
35. Pei J and Grishin NV **GGDEF domain is homologous to adenylyl cyclase.** *Proteins* 2001, **42**:210-216
36. Anantharaman V and Aravind L **The CHASE domain: a predicted ligand-binding module in plant cytokinin receptors and other eukaryotic and bacterial receptors.** *Trends Biochem Sci* 2001, **26**:579-582
37. Mouguel C and Zhulin IB **CHASE: an extracellular sensing domain common to transmembrane receptors from prokaryotes, lower eukaryotes and plants.** *Trends Biochem Sci* 2001, **26**:582-584
38. Adak S, Bilwes AM, Panda K, Hosfield D, Aulak KS, McDonald JF, Tainer JA, Getzoff ED, Crane BR and Stuehr DJ **Cloning, expression, and characterization of a nitric oxide synthase protein from Deinococcus radiodurans.** *Proc Natl Acad Sci U S A* 2002, **99**:107-112
39. Makarova KS, Aravind L, Wolf YI, Tatusov RL, Minton KW, Koonin EV and Daly MJ **Genome of the extremely radiation-resistant bacterium Deinococcus radiodurans viewed from the perspective of comparative genomics.** *Microbiol Mol Biol Rev* 2001, **65**:44-79
40. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**:3389-3402
41. Notredame C, Higgins DG and Heringa J **T-Coffee: A novel method for fast and accurate multiple sequence alignment.** *J Mol Biol* 2000, **302**:205-217
42. Cuff JA, Clamp ME, Siddiqui AS, Finlay M and Barton GJ **JPred: a consensus secondary structure prediction server.** *Bioinformatics* 1998, **14**:892-893
43. Rost B and Sander C **Prediction of protein secondary structure at better than 70% accuracy.** *J Mol Biol* 1993, **232**:584-599
44. Rost B, Schneider R and Sander C **Protein fold recognition by prediction-based threading.** *J Mol Biol* 1997, **270**:471-480
45. Schaffer AA, Wolf YI, Ponting CP, Koonin EV, Aravind L and Altschul SF **IMPALA: matching a protein sequence against a collection of PSI-BLAST-constructed position-specific score matrices.** *Bioinformatics* 1999, **15**:1000-1011
46. Felsenstein J **Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods.** *Methods Enzymol* 1996, **266**:418-427
47. Adachi J and Hasegawa M **MOLPHY: Programs for Molecular Phylogenetics.** Tokyo: Institute of Statistical Mathematics 1992,