The cytoplasmic helical linker domain of receptor histidine kinase and methyl-accepting proteins is common to many prokaryotic signalling proteins

L. Aravind a, Chris P. Ponting b,*

a Department of Biology, Texas A&M University, College Station, TX 70843, USA
b National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA

Received 6 April 1999; received in revised form 13 April 1999; accepted 13 April 1999

Abstract

Mutations in the cytoplasmic linker regions of receptor histidine kinase and chemoreceptor proteins have been shown previously to significantly impair receptor functions. Here we demonstrate significant sequence similarities between these regions in numerous histidine kinases, methyl-accepting proteins, adenylyl cyclases and other prokaryotic signalling proteins. It is suggested that these ‘HAMP domains’ possess roles of regulating the phosphorylation or methylation of homodimeric receptors by transmitting the conformational changes in periplasmic ligand-binding domains to cytoplasmic signalling kinase and methyl-acceptor domains. © 1999 Published by Elsevier Science B.V. All rights reserved.

Keywords: Histidine kinase; Chemotaxis; PP2C-like phosphatase; Adenylyl cyclase; PAS domain; Regulator of receptor function; Sequence analysis

1. Introduction

The intracellular response of prokaryotic cells to environmental change is mediated by phosphorylation- and methylation-dependent signalling circuits. Transmembrane receptors of extracellular ligands include molecules with histidine kinase or methyl-accepting functions. Homodimeric histidine kinases, such as Escherichia coli EnvZ, catalyze transphosphorylation of specific histidine residues and the phosphoryl groups are subsequently transferred to specific aspartyl residues, usually on response regulator domains homologous to E. coli CheY [1–3]. Methylation and demethylation of homodimeric chemoreceptors such as E. coli Tar, Tsr, Trg, Tap and Aer are catalyzed by methyltransferases and methyl-esterases homologous to E. coli CheR and CheB, respectively (reviewed in [4,5]).

Although considerable advances have been made in understanding the functions of these enzymes, much remains to be determined of the mechanisms that regulate the enzyme and pathway specificities. Assistance in these investigations is offered by the finding that histidine kinases and chemoreceptors...
are often accompanied by regulatory domains that occur within the same polypeptide. Sequence analysis has demonstrated the co-occurrence with histidine kinase and methyl-accepting domains of PAS predicted ligand binding and dimerization domains [6,7] and GAF putative cyclic nucleotide-binding/small ligand-binding domains [8]. These regulatory domains are in addition to the two known domains of histidine kinases, the $\alpha$-helical dimerization ‘A’ domain and the $\alpha$/C catalytic domain similar to the amino-terminal domains of a type II topoisomerase and a chaperone, Hsp90 [9–12].

Sequence similarities have been noted between two receptor histidine kinases (EnvZ and NarX) and chemoreceptor methyl-accepting proteins in cytoplasmic coiled coil-like regions that occur between the C-terminal transmembrane helix and the kinase or methyl-accepting domains [13–16]. However, a recent study was unable to detect similarities between large numbers of histidine kinase sequences and the chemoreceptor coiled coil-like domain [17]. The predicted $\alpha$-helical regions of chemoreceptors and histidine kinase have been demonstrated to be crucial for proper receptor function [13,18–20]. Consequently, it was our aim to quantify the significance of sequence similarities between these regions of chemoreceptors and histidine kinases, and to investigate whether this region occurs in other multidomain contexts.

2. Materials and methods

The position-specific, iterative version of BLAST (PSI-BLAST) [21] was used to search a non-redundant version (ftp://ncbi.nlm.nih.gov/blast/db) of current sequence databases. Subsequent analyses included comparison of a hidden Markov model (HMM) with sequence databases using HMMER2 (S. Eddy, unpublished; http://hmmer.wustl.edu/). The domain architectures of signalling proteins were established using SMART (http://smart.embl-heidelberg.de/SMART/) [22].

3. Results and discussion

A PSI-BLAST [21] search of current sequence databases was used to investigate the putative $\alpha$-helical region of *E. coli* EnvZ (amino acid residues 180–234). The search used a threshold for inclusion of sequences in subsequent iterations of $E < 0.01$, where $E$ is the number of sequences expected to be found in the search purely by chance. The first three iterations of PSI-BLAST revealed significant ($E < 10^{-3}$) similarities to similar regions of seven distinct receptor histidine kinases: *Bordetella bronchiseptica* RisS, *Pseudomonas aeruginosa* NarX and KinB, *Streptomyces coelicolor* AfsQ2, *Bacillus subtilis* YycG, *Mycobacterium tuberculosis* Rv0982, and *Myxococcus xanthus* SasS. The fourth iteration revealed significant similarities to numerous histidine kinases and methyl-accepting chemotaxis proteins including four from *B. subtilis*, TlpA, TlpB, McpA and McpB [23]. After nine PSI-BLAST iterations and convergence, 200 proteins were identified with significant similarities to the putative coiled coil region of *E. coli* EnvZ. From the domain architectures of these proteins determined using SMART [22], 119 are predicted to be histidine kinases, 66 to be methyl-accepting chemotaxis proteins, eight to be adenylyl cyclases and one to be a PP2C-type serine/threonine phosphatase. Thus, these results indicate that the $\alpha$-helical region common to chemoreceptors and histidine kinases represents an approximately 50-amino acid conserved domain that is present in multiple signalling contexts. Multiple alignments of this and other domains in prokaryotic signalling proteins are available from the SMART Web site (http://smart.embl-heidelberg.de/SMART/).

We term this homologous domain HAMP (domain present in histidine kinases, adenylyl cyclases, methyl-accepting proteins and phosphatases). Construction of a multiple alignment of HAMP domain sequences allowed additional domain homologues to be detected by comparison of a HMM with sequence databases using HMMER2 and an $E$ value threshold of 0.1; further homologues were detected using reciprocal PSI-BLAST searches and an $E$ value threshold of 0.01 (Fig. 1). Fifteen of the 30 known *E. coli* histidine kinases and all five methyl-accepting chemotaxis proteins contain a single HAMP domain. There is a strong correlation between the presence of a HAMP domain and transmembrane regions in a protein: in addition to the aforementioned 20 transmembrane *E. coli* proteins, all of the eight ini-
tially detected adenyl cyclases are predicted to be membrane-bound.

Single HAMP domains are not limited to methyl-accepting, histidine kinase and adenyl cyclase molecular contexts. Single domains were observed in a putative receptor, *Aquifex aeolicus* \(\text{aq}_1825\) and its homologues, and were found to accompany other signalling domains such as PAS, PP2C-type phosphatases, diguanylate cyclases, phosphodiesterases A, an ammonium transporter and a winged helix-turn-helix-type DNA binding domain (Fig. 2).

HAMP domains do not always occur as single repeats, nor are they restricted to prokaryotes and nor do they always occur in membrane-associated proteins. Soluble histidine kinases such as *S. coelicolor* SC7C7.03, *Neurospora crassa* Nik-1 [24] and *Candida albicans* COS1 [25] contain 12, six and five HAMP tandem repeats, respectively. Prokaryotic methyl-accepting receptors were found to contain a single HAMP. The exceptions to this are halobacterial methyl-accepting receptors [26,27], *Rhizobium* sp. NGR234 Y4FA and *Agrobacterium tumefaciens* McpA, that contain a pair of tandem HAMPs.

Mutagenesis studies [13,19,28] provide evidence that receptor histidine kinase HAMP domains negatively regulate receptor activation rather than mediate...
ating dimerization, as was proposed for some of the predicted coiled coil regions [17]. Amino acid substitutions at eight positions in the HAMP alignment (Fig. 1) produce constitutively active receptor histidine kinases [13,19,28]. How this is achieved remains unknown. However, we propose from the predicted secondary structure of HAMP of two α-helices (α1, α2) (Fig. 1) that the bihelical HAMP domain of monomeric histidine kinases binds intramolecularly to the bihelical dimerization domain [12] thereby forming a four-helix bundle similar to that seen in dimeric CheA [12]. This internal interdomain association would prevent spontaneous signalling in the absence of cognate extracellular ligands by obstructing molecular dimerization involving pairs of bihelical dimerization domains. This proposal implies that ligand-bound monomeric receptor histidine kinases alter their conformations allowing dissociation of the HAMP domain, and subsequent formation of the four-helix bundle within histidine kinase dimers.

We propose a similar protein interaction function for HAMP domains in chemoreceptors, that is modulated by conformational changes initiated by periplasmic ligand-binding events. Mutagenesis studies of the chemoreceptor Tar [18,20,29,30] demonstrate that the Tar HAMP domain is required to form functional heterodimers, although it is clear that, as with the receptor histidine kinases, it is not necessary for dimerization of the Tar cytoplasmic domains. Eight Tar variants with single amino acid substitutions in the HAMP domain have been shown to be unable to generate CheA histidine kinase acti-

Fig. 2. Schematic representation of the multidomain architectures of several HAMP (domain present in histidine kinases, adenyl cyclases, methyl-accepting proteins and phosphatases) domain-containing proteins. Domains possessing diguanylate cyclase (CYC) and phosphodiesterase A (cGDE) activities have been previously named GGDEF and EAL [34,35]. The regulatory domains shown are: Per-Arnt-Sim (PAS) domains [6,7], a winged helix-turn-helix-type DNA-binding domain (wHTH) [36] and CheY-like receiver phosphoaccepting domains [37].
vation in vitro [20] (Fig. 1). Thus it would appear that the two HAMP α-helices α1 and α2 adopt two contrasting conformational states dependent on the ligand-bound state of the chemoreceptor. Switching between these states is likely to result from the known piston-like displacement of the transmembrane ‘signalling helix’ that follows aspartate binding [31]. In this model, aspartate binding results in the bihelical HAMP domain adopting a conformation within Tar homodimers or oligomers that inhibits autophosphorylation of CheA within the Tar-Chew-CheA multimolecular complex.

The presence of HAMP domains in prokaryotic signalling molecules other than histidine kinases and chemoreceptors suggests that similar regulatory mechanisms may operate for HAMP-containing catalytic adenyl cyclase, diguanylate cyclase/phosphodiesterase and sigma-regulating PP2C phosphatase proteins. However, it is apparent that this does not account for the functions of single HAMP domains that are unaccompanied by catalytic domains, such as the HAMP in the putative receptor, *A. aeolicus* aq_1825. An intriguing possibility is that these domains act in a manner similar to the truncated Tar receptor molecules [29,30] and provide new functional specificities by associating with other receptors.

One further possible function of the HAMP domain is suggested by the presence of multiple, tandem domain homologues in fungal and actinomycete histidine kinases such as *C. albicans* Cos1 and *N. crassa* Nik-1 (Fig. 2). Here multiple HAMP domains alternate with an intervening predicted α-helical conserved sequence. Consequently, a serial intramolecular association of these HAMP domains might form a regular superstructure that is necessary for mediating interactions of these proteins with target molecules.

In conclusion, a ~50-amino acid α-helical domain has been identified with statistical significance in multidomain proteins that participate in a variety of signal transduction processes. The HAMP domain differs from ‘input’ sensing domains such as PAS, and ‘output’ receiver domains such as CheY homologues, and in receptor contexts is likely to act not as an adapter module, mediating association of signalling complexes, but rather as a sensor of conformational change and inhibitor of constitutive histidine kinase signalling. Recognition that HAMP domains are widespread among prokaryotic signalling proteins should assist in the understanding of the molecular mechanisms that mediate the intricate regulation of prokaryotic membrane-associated signalling arrays [5].

References


