Published in final edited form as:

Methods Enzymol. 2007 ; 423: 3–24. doi:10.1016/S0076-6879(07)23001-0.

The PICM Chemical Scanning Method for Identifying Domain–Domain and Protein–Protein Interfaces: Applications to the Core Signaling Complex of *E. coli* Chemotaxis

Randal B. Bass, Aaron S. Miller, Susan L. Gloor, and Joseph J. Falke

Abstract

The number of known protein structures is growing exponentially (Berman *et al.*, 2000), but the structural mapping of essential domain–domain and protein–protein interaction surfaces has advanced more slowly. It is particularly difficult to analyze the interaction surfaces of membrane proteins on a structural level, both because membrane proteins are less accessible to high-resolution structural analysis and because the membrane environment is often required for native complex formation. The Protein-Interactions-by-Cysteine-Modification (PICM) method is a generalizable, *in vitro* chemical scanning approach that can be applied to many protein complexes, in both membrane-bound and soluble systems. The method begins by engineering Cys residues on the surface of a protein of known structure, then a bulky probe is coupled to each Cys residue. Next, the effects of both Cys substitution and bulky probe attachment are measured on the assembly and the activity of the target complex. Bulky probe coupling at an essential docking site disrupts complex assembly and/or activity, while coupling outside the site typically has little or no effect. PICM has been successfully applied to the core signaling complex of the *E. coli* and *S. typhimurium* chemotaxis pathway, where it has mapped out essential docking surfaces on transmembrane chemoreceptor (Tar) and histidine kinase (CheA) components (Bass and Falke, 1998; Mehan *et al.*, 2003; Miller *et al.*, 2006). The approach shares similarities with other important scanning methods like alanine and tryptophan scanning (Cunningham and Wells, 1989; Sharp *et al.*, 1995a), but has two unique features: (1) functional effects are determined for both small volume (Cys) and large volume (bulky probe) side chain substitutions in the same experiment, and (2) nonperturbing positions are identified at which Cys residues and bulky probes can be introduced for subsequent biochemical and biophysical studies, without significant effects on complex assembly or activity.

Introduction

Domain–domain and protein–protein interactions are essential to the functions of many, if not most, proteins. Such molecular contacts are especially crucial in signaling pathways, including the two-component signaling pathways of prokaryotic organisms. Signal transduction through a cellular circuit typically requires both intramolecular and intermolecular contacts. Domain–domain interactions within a single pathway component are often essential for the transmission of internal conformational signals, while protein–protein interactions between pathway components are needed for transmission of information throughout the cellular circuit. Thus, a molecular understanding of signal transduction requires methods capable of mapping and analyzing domain–domain and protein–protein contacts. More generally, such mapping methods can provide useful information about a wide array of cellular processes involving interactions between domains or different proteins, ranging from interdomain allostery within enzymes, to cooperative interactions between the subunits of homo-oligomers, to the assembly and regulation of multiprotein complexes.

Comparison of the PICM Method with Other Scanning Approaches

Several scanning methods have proven useful in analyzing the location, function, and physical-chemical parameters of protein interaction surfaces, including alanine scanning, tryptophan scanning, and the Protein-Interactions-by-Cysteine-Modifications (PICM) method (Bass and Falke, 1998; Cunningham and Wells, 1989; Mehan *et al.*, 2003; Miller *et al.*, 2006; Sharp *et al.*, 1995a). These methods are most useful when they are applied to proteins of known structure, so that engineered mutations can be targeted exclusively to the protein surface where effects on the native fold of the modified protein are minimal, while effects on the docking interaction are maximal. The methods can all be applied effectively to soluble proteins, but, unlike many other structural methods, are also useful in the analysis of membrane proteins even in their native bilayer environments.

Alanine scanning substitutes Ala at selected surface positions, then measures the effects of each Ala substitution on the affinity of the docking interaction or on the activity of the docked complex (Cunningham and Wells, 1989; Wells, 1996). With the sole exception of Gly positions, substitution of Ala for a docking site residue truncates a larger side chain while having minimal effects on backbone flexibility. To a first approximation, then, the resulting effect of Ala substitution on docking affinity or activity reveals the contribution of an individual native side chain to the docking interaction. Alanine scanning is often used to analyze the physical chemistry of docking when the structure of the assembled complex is already known, but it can also map out the location of an unknown docking site on the surface of an isolated domain or protein as long as its docking partner is available for affinity and/or activity studies.

Tryptophan scanning substitutes Trp at selected surface positions, then determines the effect of each Trp substitution on complex assembly or activity (Sharp *et al.*, 1995a,b). Since Trp is the largest natural side chain, this substitution always increases side chain volume, thereby maximizing the probability of a dramatic effect on docking affinity. It follows that Trp substitutions within a docking site will generally yield measurable perturbations, making Trp scanning an efficient method of mapping out unknown docking sites on protein surfaces. For protein surfaces buried within a membrane bilayer, Trp scanning is particularly useful because the membrane environment often prevents the covalent coupling of extrinsic bulky probes, thereby largely eliminating the use of extrinsic probes in mapping the docking site.

The PICM method is complementary to the alanine and tryptophan scanning approaches, combining some of their strengths and offering additional advantages, particularly in systems where further biochemical and biophysical studies of the purified components are planned. PICM makes use of the unique chemical and physical properties of the Cys side chain (Bass and Falke, 1999; Falke *et al.*, 1986), which the method introduces at water-exposed surface positions scattered throughout the region where an unmapped docking site could reside. Subsequently, these engineered Cys residues are covalently modified with a bulky probe, and the effects of both the Cys substitution and bulky probe modification are determined on complex activity and assembly (Bass and Falke, 1998; Mehan *et al.*, 2003; Miller *et al.*, 2006). The engineered Cys side chain is smaller than all others except Gly, Ala, and Ser. Thus, Cys substitution typically replaces a native residue with a smaller side chain much as alanine scanning does. The bulky probe chosen for subsequent coupling to the Cys side chain is significantly larger than tryptophan, and thus replaces all native residues with a larger side chain that yields even better disruption of docking interactions than does tryptophan scanning. It follows that the PICM approach simultaneously determines the effects of smaller and larger side chain substitutions at most positions. Moreover, the PICM method yields an optimized labeling library of non-perturbing positions at which Cys substitution and bulky probe incorporation have little or no effect on docking interactions and activity. Such a library is of

great utility in further biochemical and biophysical studies requiring sulfhydryl chemistry or the attachment of large spectroscopic probes or crosslinkers.

In practice, the PICM approach is limited primarily to the analysis of water-exposed docking sites on purified proteins or on the extracellular domains of proteins in living cells. Because the PICM method requires coupling of a bulky probe to a Cys residue, typically via an alkylation or disulfide exchange reaction involving the Cys sulfanion, the approach is easily applied to aqueous docking sites accessible to the probe. By contrast, the PICM method is less useful for (a) lipid-exposed docking sites where the coupling reaction proceeds slowly because the low dielectric environment raises the sulfhydryl pKa, and (b) cytoplasmically exposed docking sites in living cells where the plasma membrane barrier and high cytoplasmic glutathione concentration typically interfere with probe coupling. In such cases, alanine and tryptophan scanning methods are generally preferred.

PICM Studies of the Core Signaling Complex of Bacterial Chemotaxis

The PICM method was originally developed in studies of the core signaling complex of bacterial chemotaxis (Bass and Falke, 1998; Mehan *et al.*, 2003; Miller *et al.*, 2006). This complex transduces attractant binding into a transmembrane signal that regulates the activity of a cytoplasmic histidine kinase (Baker *et al.*, 2006; Bourret and Stock, 2002; Falke *et al.*, 1997; Parkinson *et al.*, 2005). The oligomeric receptor serves as the framework for the formation of the core complex, by providing docking surfaces on its cytoplasmic domain where the soluble cytoplasmic components dock to form a stable assembly. The core complex components required for reconstitution of receptor-regulated kinase activity are the transmembrane receptor, the histidine kinase CheA, and the coupling protein CheW. Once the complex is formed, the apo state of the receptor stimulates CheA autophosphorylation, while the attractant-occupied receptor inhibits CheA. Thus, the standard measure of core complex activity, termed the “reconstituted core complex kinase assay,” begins by reconstituting the core complex under conditions where the CheA autophosphorylation reaction is the rate-limiting step in signal transduction.

The first PICM studies, using the initial version of the method now termed PICM- α , identified docking sites on the surface of the *Salmonella typhimurium* transmembrane aspartate receptor by introducing Cys residues at positions scattered uniformly over the protein surface (Bass and Falke, 1998; Bass *et al.*, 1999; Mehan *et al.*, 2003). These studies employed the standard reconstituted core complex kinase assay to determine the functional effects of each engineered surface Cys, both in its unmodified and bulky probe-labeled states. Essential docking surfaces were identified as regions where bulky probe attachment or, in fewer cases, the Cys substitution itself, blocked receptor-mediated kinase activation, indicating that the modification either prevented core complex assembly or prevented the flow of information from the receptor to the docked CheA kinase. The surface positions where modifications caused large losses of receptor-mediated kinase activation were found to be clustered in two specific regions. Several perturbing positions were located in the vicinity of the adaptation sites, which are known to modulate kinase activity. The great majority of perturbing positions, however, were found within a large docking surface that contains the contact sites essential for receptor oligomerization as well as the CheA and the identified docking surfaces remain consistent with all current evidence.

In 2006, a PICM study employed an enhanced version of the method, termed PICM- β , to directly determine the location of four docking sites on the surface of the *S. typhimurium* CheA histidine kinase (Miller *et al.*, 2006). This PICM- β analysis utilized both enzyme assays and direct binding measurements to determine the effects of engineered surface cysteines and bulky probe modifications on the assembly and activity of the multiprotein complex. The use of direct

protein–protein binding measurements greatly simplified the interpretation of the results, enabling direct identification of positions in or near essential docking sites. The resulting combination of activity assays and direct protein–protein binding measurements revealed the locations of four distinct docking sites on the surface of CheA: (1) the docking site on the substrate domain that associates with the catalytic domain during autophosphorylation, (2) the docking site on the catalytic domain that associates with the substrate domain during autophosphorylation, (3) the docking site for CheW on the regulatory domain, and (4) the putative docking site for the transmembrane chemoreceptor—a large site spanning portions of the CheA regulatory, catalytic, and dimerization domains. An independent X-ray crystal structure solved for the complex between a *Thermatoga thermophilus* CheA fragment and its CheW partner has confirmed that PICM- β correctly identified the location of the CheW docking site on the surface of the CheA (Park *et al.*, 2006), providing strong support for the accuracy of the PICM- β approach (Miller *et al.*, 2006). Thus, while the PICM- β approach cannot map out docking sites to atomic resolution, it can accurately map the general location of docking sites on the surfaces of proteins in their native environment, whether it be aqueous solution, a lipid bilayer, or a supermolecular complex.

Generalizing the PICM Method to Map Docking Sites in Other Systems

The present chapter focuses on the generalizable steps of the PICM procedure that can be used to map out domain–domain and protein–protein docking sites in a wide variety of systems. The initial steps in applying PICM to a specific protein are to generate a fully functional, Cysless version of the protein and to select sites for surface Cys introduction. Next, high-throughput methods are used to introduce the engineered Cys residues by site-directed mutagenesis and to purify the resulting single-Cys containing proteins. Each single-Cys protein is then labeled with a bulky probe, and the level of probe incorporation is quantitated. The effects of Cys substitution and bulky probe attachment on docking affinity and activity are measured in functional assays. Finally, the PICM data are used to map out the location of the essential docking site(s). The remaining sections discuss the generalization of these steps to PICM studies of new protein systems, using published studies of the aspartate receptor and CheA kinase of bacterial chemotaxis as examples.

Incorporation of an Affinity Tag and Creation of a Cysless Protein

Overview

Ideally, the protein selected for PICM analysis should contain an affinity tag to facilitate rapid purification of modified proteins. Moreover, the protein should lack intrinsic Cys residues, or at least lack Cys residues accessible to the bulky probe. Alternatively, intrinsic Cys residues can be replaced to create a Cysless background for subsequent incorporation of single Cys residues. Generally, a Cysless construct that retains full activity can be engineered. The advantage of such a Cysless construct is that each Cys introduced at a selected surface position becomes a unique site for covalent attachment of the sulfhydryl-specific bulky probe, with no complications due to probe incorporation at alternate Cys positions. In previous PICM studies, the transmembrane aspartate receptor needed no affinity tag since isolation of *E. coli* membranes containing the overexpressed receptor yielded sufficient purity for PICM- α analysis (Bass and Falke, 1998; Bass *et al.*, 1999; Mehan *et al.*, 2003). Moreover, this receptor lacks intrinsic Cys residues, so the native protein was itself a perfect Cysless background in which to incorporate surface Cys residues (see following text). By contrast, the native CheA protein was prepared for PICM- β analysis both by fusing it to an N-terminal 6-His affinity tag for ease of isolation and by replacing three intrinsic Cys residues to generate a functional Cysless construct (Miller *et al.*, 2006). These modifications of CheA were carried out using standard subcloning and site-directed mutagenesis procedures; thus, the present discussion focuses on the design aspects most useful in applications to other systems.

Incorporation of a 6-His Affinity Tag

The simplest affinity tag to use is the 6-His tag, which can be attached at the N- or C-terminus of the target protein (Bornhorst and Falke, 2000), although an array of alternative, equally effective affinity tags are available (Corbin and Falke, 2004). Standard subcloning methods are used to insert the target protein gene into a vector containing the 6-His tag and an appropriate linker. To create the N-terminal 6-His derivative of CheA, for example, the CheA gene was subcloned into the pET28 vector (Novagen) providing both the 6-His sequence and a long, flexible linker inserted between the tag and the native CheA sequence, yielding the N-terminal sequence **MGSSHHHHHSSGLVPRGSHMASGGGGGGGVSM**D (where the unbolded, nonnative residues derived from the vector were inserted into the bolded native sequence between Met 1 and Ser 2). The 21 residue linker includes a 7-residue poly-glycine region, ensuring that the tag is flexible and able to reach out into solution, away from the protein surface. This design maximizes the interaction between the tag and the Ni-NTA affinity matrix during purification, while minimizing potentially perturbing interactions between the highly charged tag and the surface of the protein. In a new system, it is best to create and compare N-terminal and C-terminal affinity tags, since the location of the tag may affect function. For example, while CheA possessing the above N-terminal 6-His tag is fully functional (Miller *et al.*, 2006), an analogous construct placing the tag at the C-terminus has little or no measurable activity (Bornhorst and Falke, unpublished). To generate affinity tag fusions, an array of vectors containing 6-His and other affinity tags at the N- and C-terminal ends of polylinkers are commercially available, and standard subcloning methods are used to insert target protein genes into these vectors (pET vectors, Novagen).

Creation of a Functional Cysless Protein

Following introduction of the affinity tag, if the target protein contains intrinsic Cys residues, a systematic approach is used to generate a functional Cysless construct (Frillingos *et al.*, 1998; Miller *et al.*, 2006). At each of the intrinsic Cys positions, standard site-directed mutagenesis is used to convert each Cys to both Ala and Ser. The functional effects of these single-Ala and single-Ser substitutions are determined by isolating each point mutant and subjecting them to activity studies. For CheA, which contains three intrinsic Cys residues, the resulting three Ala and three Ser point mutants were tested in the reconstituted core complex kinase assay. The three point mutations yielding the smallest effects on receptor-mediated CheA kinase activation were combined to give a Cysless triple mutant, CheA C120S/C218A/C432A (Miller *et al.*, 2006). In other systems, the same approach is generally useful, using standard methods to mutate the intrinsic Cys residues. Commercially available site-directed mutagenesis kits are available and provide efficient mutagenesis protocols (QuikChange kits, Stratagene).

Testing the Function of the Affinity-Tagged, Cysless Protein

It is important to compare the function of the final construct to the native protein, to ensure that the addition of the affinity tag and the replacement of the intrinsic Cys residues do not significantly alter the docking interactions and activities being studied. To test the function of the 6-His-tagged, Cysless CheA construct, this modified protein was compared to native CheA in both *in vitro* and *in vivo* functional assays (Miller *et al.*, 2006) (and Bornhorst and Falke, unpublished). First, the purified proteins were compared in the reconstituted core complex kinase assay wherein the modified and wild type proteins were found to be indistinguishable, within experimental error, in their abilities to be activated by the receptor apo state and inhibited by attractant binding. Next, the abilities of the two proteins to restore normal pathway function in an *E. coli* strain lacking CheA were compared in the swarm plate assay for cellular chemotaxis. Both proteins were found to restore chemotaxis *in vivo*, although the swarm rate of the 6-His-tagged, Cysless CheA construct was 0.7-fold that of the native construct. It follows

that the 6-His-tagged, Cysless CheA protein is fully functional in the reconstituted core signaling complex *in vitro*, and its activity is only slightly perturbed *in vivo*. Since the 6-His-tagged, Cysless CheA protein is fully functional in the reconstituted core complex with CheA, CheW, and CheY, this modified construct is suitable for PICM studies of its interactions with these proteins (Miller *et al.*, 2006). In other systems, appropriate assays should be used to test the effects of the affinity tag and the removal of intrinsic Cys residues on target protein structure and function. If the replacement of a buried intrinsic Cys yields significant perturbations, or if two intrinsic Cys residues form an essential disulfide bond, such Cys residues can often be left in place since they typically react poorly with the bulky probe and thus do not interfere with the PICM method.

Choice of Positions for Cys Incorporation and Creation of a Mutant Library

Overview

The most rigorous type of PICM analysis scans the entire surface of the target protein for essential interaction sites. If additional information is available that limits the interaction site of interest to a specific domain or surface region, then the extent of the PICM analysis can be reduced. Either way, the approach is the same. Using the known structure of the target protein, sites are selected on the protein surface for Cys incorporation. High throughput methods are then employed to generate the single Cys mutants, express them, and purify them.

Selection of Engineered Cys Positions: General Considerations

When examining the candidate positions for Cys incorporation, it is important to select positions for which the side chains are fully exposed on the protein surface rather than partially buried in the protein interior. The goal is to avoid effects on protein folding and stability while focusing on surface residues that could be directly involved in a docking interaction. The positions selected should yield a distribution of engineered Cys residues scattered over the surface of the protein as uniformly as possible, with a spacing between Cys positions of approximately 10Å or less to ensure that a typical docking site would contain several Cys positions. Conserved surface positions hypothesized to be involved in docking should not be avoided but rather should be included as targets for Cys substitution, since exclusion of these conserved positions could cause the PICM approach to miss an important docking site, while the design of the PICM method will ensure that side chains essential for docking will be directly identified. To be safe, in a multidomain protein for which the structure was determined by crystallographic methods, engineered Cys residues should be targeted even to the surfaces of domains that appear to be buried at domain–domain contacts, since such contacts can be artifacts of crystal constraints or of nonspecific associations between hydrophobic docking sites on different domains. (Domains are often coupled by flexible linkers, such that their locations and contacts can be easily perturbed in a crystal structure (Zhang *et al.*, 1995)). The final PICM- α analysis of the aspartate receptor selected 52 positions for Cys incorporation scattered over two distinct domains (Mehan *et al.*, 2003), while the PICM- β analysis of CheA utilized 70 positions scattered over five domains (Miller *et al.*, 2006).

High Throughput Cys Mutagenesis Protocol

1. Once a set of positions has been chosen for PICM analysis, it is useful to design a high-throughput strategy for site-directed mutagenesis. Typically, the most practical strategy is to run multiple mutagenesis reactions in parallel. The authors have found that 15 is a convenient number of reactions to run simultaneously using standard commercial mutagenesis protocols. Currently, we use the QuikChange II XL Site-Directed Mutagenesis Kit from Stratagene, which provides reagents for 30 mutagenesis reactions. After all the necessary reagents are obtained, an experienced researcher can generate at least two sets of 15 mutants per week.

2. Mutagenic primers are designed closely following the mutagenesis kit recommendations and ordered from an appropriate commercial DNA synthesis facility. We order primers in 25 nmol quantities, desalted but with no further purification, from Integrated DNA Technologies. Each primer arrives from the manufacturer as a lyophilized film, and the manufacturer specifies the total mass of lyophilized DNA. An appropriate volume of sterile, deionized distilled H₂O is added to bring each primer up to a concentration of 1.25 $\mu\text{g}/\mu\text{l}$, which serves as a 10 $\times$ stock for the subsequent mutagenesis reaction.
3. Template plasmid DNA is generated in *E. coli* strain DH5 α , then is isolated using a standard commercial miniprep protocol. Currently, we use the Qiagen QIAprep Spin Miniprep Kit, which for our plasmids typically yield 50 μl of 300 ng/ μl DNA. The miniprep DNA concentration is measured by its absorbance at 260 nm (1 A₂₆₀ = 50 $\mu\text{g}/\text{ml}$), then a portion of the miniprep is diluted in H₂O to give a template stock of 10 ng/ μl .
4. After obtaining both template and primers, parallel PCR mutagenesis reactions and transformations are carried out using the standard procedure recommended by the mutagenesis kit. Our PCR thermocycler is an Eppendorf MasterCycler, and we use the thermocycler settings specified by the mutagenesis kit for the length of the plasmid.
5. Following mutagenesis and transformation, mutations are confirmed by DNA sequencing. At least two colonies from each mutagenesis transformation are randomly selected for minipreps of plasmid DNA, carried out using the miniprep kit protocol, then quantitated as in step 3. We submit at least 250 ng DNA (typically, 1 to 2 μl of the miniprep) to a university facility for automated DNA sequencing. Sequencing primers are chosen to ensure that the region containing the intended mutation is adequately covered. Particularly when using PCR mutagenesis methods, it is important also to confirm that no unintended mutations have occurred elsewhere in the gene. Thus, one mutant isolate is chosen for full gene sequencing; alternatively, the sequenced region containing the desired mutation can be subcloned back into the starting plasmid.

Selection of a Cys-Specific Probe for Chemical Modification

Overview and General Considerations

A bulky probe suitable for PICM analysis is chosen based on probe size, charge, ease of incorporation, specificity for the Cys sulfhydryl, and reversibility. In general, the two most useful classes of Cys-specific coupling chemistries are (a) the maleimide alkylation reaction, which combines high sulfhydryl specificity and rapid reaction rate, and (b) the methanethiosulfonate disulfide exchange reaction, which combines high sulfhydryl specificity and reversibility.

In typical PICM applications where reversibility is not required, a suitable bulky probe is 5-fluorescein maleimide (5FM), which is significantly bulkier than any natural side chain (Fig. 1), reacts rapidly, specifically, and irreversibly with exposed Cys sulfhydryls, and yields bright, easily quantitated fluorescence on standard ultraviolet (UV) lightboxes. However, 5FM is highly sensitive to bleaching so care should be taken at all stages not to expose this probe, or proteins labeled with this probe, to UV light before the final quantitation step. Both the original PICM studies of the aspartate receptor and CheA kinase of bacterial chemotaxis used 5FM as the primary probe (Mehan *et al.*, 2003; Miller *et al.*, 2006).

For detailed PICM studies, it can be useful to investigate whether it is the size of the probe, or rather its charge, that dominates its effects on the docking interaction. For example, a set of

three maleimide-based probes that can be utilized in such a study would include (a) 5FM, which is large and possesses two negative charges at neutral pH, (b) 5-tetramethylrhodamine maleimide (5TRM), which is approximately the same dimensions as 5FM but is a zwitterion with one positive and one negative charge, and (c) N-ethylmaleimide (NEM), which is neutral and considerably smaller than both 5FM and 5TRM, being closer in size to a typical side chain (Fig. 1). Our most detailed PICM analysis of the aspartate receptor utilized these three probes to identify positions where probe size, charge, or both altered the interaction between the receptor and CheA kinase (Mehan *et al.*, 2003). We purchase 5FM and 5TRM from Invitrogen/Molecular Probes and NEM from Sigma.

In some cases, a reversible, disulfide-linked probe is preferable, for example, when the same sample will be used in activity studies of the labeled and unlabeled states. This approach can be utilized when all of the activity studies planned are insensitive to the presence of reducing agent, so that the labeled protein can be divided into two aliquots—one treated with reducing agent to reverse the labeling and the other untreated to retain the label. A wide selection of methanethiosulfonate (MTS) probes, varying in size, shape, and charge, is available from Toronto Research Chemicals. The MTS disulfide exchange reaction is significantly slower than the maleimide alkylation reaction, which must be considered when designing the labeling protocol (Frazier *et al.*, 2002; Malmberg *et al.*, 2003).

Probe Labeling and Purification of the Single Cys Mutants

Overview

Assuming that an affinity tag has been incorporated into the protein, it is straightforward to design a parallel purification procedure that will simultaneously label multiple single Cys mutants with a selected probe, then will purify both the labeled and unlabeled proteins. Here, we focus on the maleimide and MTS coupling chemistries for chemical labeling, and the 6-His affinity tag for purification (Bornhorst and Falke, 2000). The protocol used to express, label, and purify a given cloned, 6-His tagged protein should be optimized for the specific protein of interest, but the following procedure provides a generally useful starting point. Except where noted otherwise in this and other protocols, all chemicals, biochemicals, and reagents were obtained from Sigma.

6-His Affinity Tag Purification Protocol

1. The expression plasmid for the 6-His tagged protein is transformed into an appropriate *E. coli* line, then induction conditions are optimized for the given plasmid and protein. Using the optimized conditions, expression cultures are grown and induced in 500 ml of the appropriate media. If necessary for proper aeration, this volume is divided among multiple growth flasks. We express and purify 6 to 12 mutant proteins in parallel, limited primarily by the number of bottle positions (typically, 6) in the centrifuge rotor, and the number of centrifuges available for simultaneous cell pelleting spins.
2. Cells are pelleted from 500 ml of media by spinning for 15 min at 6000×g in a 500 ml centrifuge bottle at 4°. The pellet is resuspended by shaking on ice in 10 ml of buffer A (300 mM NaCl, 10% (v/v) glycerol, 50 mM NaH₂PO₄, pH to 8.0 with NaOH) to which are added fresh reducing agent and protease inhibitors (10 mM β -mercaptoethanol or β ME, 1 mM phenylmethylsulfonylfluoride or PMSF, which hydrolyzes rapidly in aqueous solutions and thus is added from a 200 mM stock in absolute EtOH, 1 μ g/ml aprotinin, and 1 μ g/ml leupeptin). If storage is necessary before beginning purification, the resuspended cells are best frozen by pouring them into liquid nitrogen, straining out the frozen cells, and immediately placing in a -80° freezer.

3. When isolating, labeling, and purifying single Cys mutants from frozen cells, we typically prepare 3 to 6 mutants in parallel, limited again by the number of centrifuge rotor positions. Frozen cells are thawed by shaking on ice for approximately 30 min following addition of 10 ml ice cold buffer A, the latter containing β ME and fresh protease inhibitors. Lysozyme is then added to 3 mg/ml. The mixture is incubated on ice for 30 min to allow the enzyme to hydrolyze the bacterial cell wall, with a single hand mixing at 15 min. Brij-58 detergent (polyoxyethylene-20-cetyl-ether) is added as a 7% solution (w/v) diluted 70-fold into the mixture to yield 0.1% final, then the lysate is spun for 15 min at $28,300\times g$ in a 35 ml centrifuge tube at 4° . The resulting supernatant (approximately 20 ml) is carefully removed and placed on ice. Note that cell lysis exposes the target protein to high concentrations of active proteases, even in the presence of the protease inhibitors, so subsequent steps should be carried out as rapidly as possible.
4. The supernatant is split into two aliquots (approximately 10 ml each), one for labeling with bulky probe and the other for unlabeled protein. Each aliquot is loaded into a BioRad EconoPac column containing 2.5 ml of Qiagen Ni-NTA beads pre-equilibrated in binding/wash buffer (15 mM imidazole in buffer A). Following loading, the column is sealed and placed on a rotator at 22° for 15 min to maximize mixing and binding of the 6-His-tagged protein to the beads. The column is then moved to the cold room (4°) and washed thrice with 15 ml aliquots of cold binding/wash buffer to remove the contaminating native proteins lacking the His tag, and also to remove the reducing agent and protease inhibitors. Following the third wash, the beads (2.5 ml) are moved out of the cold room and resuspended in 2.5 ml of room temperature binding/wash buffer together with an appropriate volume of 20 mM probe in solvent (dimethylformamide is suitable for many probes), yielding a final probe concentration of 200 μ M. To the other aliquot is added 2.5 ml of binding/wash buffer and the same volume of pure solvent. The columns are again sealed and placed on a rotator at 22° for 15 min (maleimide probes) or 1 hr (methanethiosulfonate probes) to allow the labeling reaction to proceed to completion. Then, for maleimide reactions, a 25 mM reduced glutathione stock is added to 250 μ M final concentration and mixed to rapidly quench the remaining free probe (this step is omitted for methanethiosulfonate probes, to avoid reversing the labeling reaction). The columns containing the labeled and unlabeled proteins are then moved back to the cold room, the buffer is drained, and each column is washed twice more with 15 ml of cold binding/wash buffer.
5. Proteins are eluted from the beads by 3 to 5 washes with 2.5 ml of cold elution buffer (500 mM imidazole in buffer A) in the cold room. The washes from a given column are combined on ice, EDTA is added to 1 mM, and, for unlabeled and maleimide-labeled proteins, dithiothreitol (DTT) is also added to 10 mM (the last addition is omitted for methanethiosulfonate probes to avoid reduction of the disulfide linkage). Subsequently, the protein is concentrated by spinning $3220\times g$ at 4° in a 15 ml Amicon Ultra Centrifuge Filter Device (with an appropriate membrane MW cutoff). Spinning is carried out in 30 min increments until the desired final volume of 0.5 ml is achieved (typically, 30 to 120 min total).
6. The concentrated protein is dialyzed overnight at 4° in a Slide-A-Lyzer Extra Strength Dialysis Cassette (0.5 ml capacity, appropriate membrane MW cutoff) against one liter of final buffer (50 mM Tris, pH to 7.5 with HCl, 10% glycerol (v/v), 1 mM EDTA, and, for unlabeled proteins or proteins labeled with reduction-proof probes, 10 mM DTT. Subsequently, the protein is dialyzed twice for 4 hours against 1 liter of final buffer and, for unlabeled proteins or proteins labeled with reduction-proof probes, 10 μ M DTT. Finally, the dialyzed protein is spun in an ultracentrifuge at 4° for 10 min

at 500,000×g to remove any precipitates, and then is aliquoted and frozen in liquid nitrogen.

7. The purified protein is quantitated via a standard protein assay.

Quantitation of Probe Coupling

Overview

Following labeling and purification of the Cys mutants, it is important to measure the purity of the final proteins and their extent of labeling (Mehan *et al.*, 2003; Miller *et al.*, 2006). In particular, for each mutant, accurate PICM analysis requires virtually complete modification (preferably over 90%) of the engineered Cys residue. The protocol for measuring the extent of labeling carries out another labeling step with a fluorescent probe under denaturing, high-temperature conditions to ensure that this second labeling reaction goes to completion. Subsequently, the protein is resolved on SDS-PAGE and the labeling efficiency is determined by quantitative fluorescence imaging. This procedure also checks the purity of the final protein.

Procedure to Quantitate the Extent of Probe Coupling

1. Three reactions are prepared for each mutant protein: one for the unlabeled protein and two for the labeled protein, respectively. Typically, each of these reactions is run in triplicate to generate statistics. For each reaction, the protein stock is diluted into final buffer to a concentration of 5 μ M. To the unlabeled and one of the labeled samples, 20 mM 5FM in DMF is added to a final concentration of 125 μ M, and to the other labeled sample, the same quantity of pure DMF is added. Each sample is diluted 1:1 with 2× Laemmli sample buffer, mixed quickly and heated at 95° for 1 min to denature the protein and drive the labeling reaction to completion. Typically, we analyze up to 8 mutants in parallel, which requires 8 SDS-PAGE gels.
2. The reactions for a given mutant are run side-by-side on a Laemmli SDS-PAGE gel of the appropriate acrylamide composition to yield a relative migration of 0.5, thereby maximizing resolution. Before staining with Coomassie or another protein stain, which typically quenches fluorescence, the fluorescent bands are visualized by placing the gel on a standard UV light box and imaged by digital photography, ensuring that none of the bands is saturated. Imaging software is used to quantitate the relative fluorescence of the bands. Following quantitation of fluorescence, the gel can be stained with Coomassie or other reagent and used to determine the purities and relative concentrations of the different protein samples by quantitation of the band absorption intensities. Care should be taken that the absorption spectrum of the bulky probe does not overlap that of the protein stain at the wavelength(s) used for imaging the stained bands.
3. To determine the extent of labeling, the relative fluorescence intensities of the three related bands are compared. These three bands are denoted LL (labeled protein that was subsequently subjected to a 5FM labeling reaction under denaturing conditions); UL (unlabeled protein that was subsequently subjected to a 5FM labeling reaction under denaturing conditions); and LU (labeled protein subjected to the control second labeling reaction without added label). The extent of labeling, which ranges between 0 for no labeling to 1 for complete labeling, can then be calculated from these fluorescence intensities. When the first and second labeling reactions both use the same fluorophore, the extent of labeling can be calculated as LU/LL or, alternatively, as LU/UL. Generally, the LU/LL calculation is preferred since both of its fluorescence values are determined for the same protein sample treated in two different ways. In the case where the first label is nonfluorescent and the second label is a fluorophore, the extent of labeling is calculated as 1 – (LL/UL). All these calculations make use

of the fact that the labeling reaction with a maleimide probe goes rapidly to completion under denaturing conditions at elevated temperatures.

Measuring Functional Effects of Cys Substitution and Bulky Probe Coupling

Overview and General Considerations

Once the library of single Cys mutants has been created and modified with the bulky probe, the effects of both the Cys substitution and bulky probe attachment on protein activity are determined using appropriate activity assays (Mehan *et al.*, 2003; Miller *et al.*, 2006). Often, the Cys substitution itself is relatively nonperturbing, since the Cys side chain is relatively small and sterically accommodating, while the sulfhydryl group can adapt to different electrostatic environments by varying between an apolar, protonated state and an anionic, deprotonated state (Falke *et al.*, 1988, 1986). Thus, even when located within a surface docking site, Cys substitutions are often tolerated unless they replace an essential docking residue. By contrast, the bulky probe, which is significantly larger than even the largest native side chain, is designed to sterically prevent docking when introduced at a docking site. Thus, the PICM approach typically yields two categories of surface Cys positions: those outside of functionally important contact sites where neither the cysteine substitution nor bulky probe is perturbing, and those within essential contact sites where the Cys may or may not be perturbing but the bulky probe is always perturbing. To resolve these two classes of positions, and thereby map out docking surfaces, it is essential to carry out suitable activity assays.

The activity assays selected to measure the functional effects of Cys substitution and bulky probe attachment are specialized for the system of interest. Ideally, the chosen assays would include a *stability or activity assay* for the modified proteins in their isolated state, to identify modifications that perturb the intrinsic folding or activity of the target protein itself. Modifications that perturb folding or stability are unsuitable for PICM analysis; thus, it is important to identify modified proteins exhibiting such perturbations so they can be excluded. Modifications that do not alter intrinsic folding or stability, but still block function, typically identify an essential domain–domain contact surface and thus provide useful information that can be gathered with the appropriate measurement. The chosen assays would also include a *complex formation assay* that quantitates the assembly and, if possible, the affinity of the complex formed by the target protein and its docking partner(s). This key measurement determines the effect of surface modifications on protein–protein interactions and is thus critical for mapping out docking sites. Finally, the chosen assays could include a *complex activity assay* to determine the effect of surface modifications on the activity of the multiprotein complex. This assay can double-check the results of the complex formation assay, since complexes that do not assemble should have no activity. More importantly, this assay can identify modifications that allow complex assembly but block complex function, usually by perturbing an essential regulatory surface or interaction.

The PICM- β analysis of CheA docking sites illustrates the selection of appropriate assays for a given system (Miller *et al.*, 2006). Since isolated CheA exhibits autokinase activity, a rapid and convenient ^{32}P autophosphorylation assay was chosen as an initial screen for native folding and activity of the isolated Cys mutants, both in their unmodified and bulky probe-modified states (Miller *et al.*, 2006). This assay revealed that most modified CheA proteins retained autophosphorylation activity, indicating that they were properly folded, functional kinases. The inactive proteins were eliminated from further PICM analysis, but the locations of their modifications provided useful information. All the perturbing modifications were found on two surfaces surrounding the ATP binding site on the CheA catalytic domain and the His 48 phosphorylation site on the CheA substrate domain, respectively, suggesting that contacts between these surfaces are essential for phosphotransfer from ATP to His 48 during the autophosphorylation reaction. To compare the intermolecular protein–protein interactions of

the modified CheA proteins, two types of binding assays were chosen: a CheW binding assay utilizing fluorescence anisotropy to quantitate the docking of CheW to free CheA (Boukhvalova *et al.*, 2002), and a core complex formation assay that used centrifugation to detect the incorporation of modified CheA proteins into membrane-bound complexes with the transmembrane receptor and CheW (Levit *et al.*, 2002). Finally, to measure the ability of the modified CheA kinases to be activated by receptor in the reconstituted, membrane-bound core complex, a receptor-regulated CheA kinase assay monitoring the formation of [³²P]-phospho-CheY was utilized (Borkovich *et al.*, 1989; Chervitz *et al.*, 1995; Ninfa *et al.*, 1991). Together, these assays enabled resolution of four docking sites involved in (1) the contacts between the CheA catalytic and substrate domains during autophosphorylation, (2) the contact between CheA and CheW formed during core complex assembly, and (3) the contact between CheA and the receptor formed during core complex assembly.

Once a suitable set of assays has been selected, it is generally possible to design a streamlined experimental flow diagram that minimizes the number of measurements required. A streamlining procedure was implemented during the PICM- β analysis of CheA, yielding a significant reduction in the number of assays carried out on the 140 different modified proteins (Miller *et al.*, 2006). To develop such a streamlined procedure, first, all of the modified proteins are tested in the assay for intrinsic folding or function, enabling elimination of perturbed proteins from further study. Second, the remaining proteins are tested for activity in the reconstituted multiprotein complex. Those exhibiting full activity need not be examined further, since they can assemble normally with the other components to form the active complex. Third, the modified proteins that exhibit perturbed activities in the reconstituted complex are subjected to binding assays, in order to ascertain which perturbing modifications directly inhibit a protein docking reaction. Overall, this approach provides the maximal unique information from the minimal number of measurements. Alternatively, when it is feasible to carry out all of the assays on each modified protein, this approach is preferred since it generates the maximum level of redundant, overdetermined data that can be used to check for self-consistency.

Interpretation of Results—Mapping Out Docking Sites

Overview and General Considerations

Once the PICM data has been obtained, the resulting information is used to map out one or more docking sites on the surface of the target protein. The analysis assumes that each docking site is located on a defined face or region of the protein surface, so that residues located within the same docking site will be clustered together. The key to developing an accurate PICM map of each docking site is the accurate identification of those positions at which Cys substitution and/or bulky probe modification generates a significant perturbation in a given assay. In general, a perturbation threshold must be defined for each assay to enable classification of one subset of modifications as nonperturbing (their activities lie above the threshold) and a second subset of modifications as perturbing (their activities lie below the threshold). The distribution of perturbing modifications on the protein surface will then map out the extent of the docking surface of interest. Since setting the perturbation threshold is a subjective procedure, a logical, systematic approach should be used to define this threshold. It is useful to discuss separately the procedures used to identify essential intramolecular domain–domain contacts and those used to map out essential protein–protein docking sites.

Essential intramolecular, domain–domain contacts can exist within an isolated protein, and successful formation of these contacts can be required for native folding, stability, or activity. In such a case, there will be two distinct contact surfaces located within the protein, one on each of the contacting domains. These surfaces can be identified by focusing on the PICM data obtained for the isolated protein, which define the effects of Cys substitution and bulky probe

modification on the stability or activity of the isolated protein in the absence of other components. To use PICM data to map out the two surfaces, the threshold that defines the perturbing modifications is first set quite high, so that the positions operationally defined as perturbing are widely distributed on the target protein surface. Since the domain–domain contact is assumed to be localized to two limited surfaces, the threshold is lowered until two distinct clusters of perturbing positions are apparent. These two clusters identify the contact sites. Typically, the observed effects of perturbing modifications within these sites are similar, since they inhibit the same domain–domain contact. Positions within the docking sites at which the Cys substitution itself is perturbing are strong candidates for essential docking residues. To double check the mapping process, additional PICM sites can be engineered and analyzed within the newly identified contact regions, if desired. The use of PICM to map out the contact surfaces involved in a functionally essential, intramolecular domain–domain interaction was illustrated by the 2006 PICM analysis of CheA kinase, which successfully identified the contact surfaces on the catalytic and substrate domains that must associate during the autophosphorylation reaction (Miller *et al.*, 2006).

The most important application of PICM is to identify docking sites involved in intermolecular, protein–protein contacts. Essential protein–protein contact surfaces are identified using the subset of PICM perturbations that have little or no effect on the stability or folding of the isolated protein. Again, the key to accurate identification of the docking site is the setting of the perturbation threshold for each binding and activity assay to the appropriate level. Since a protein–protein docking site is assumed to be localized to a single surface region or face, the threshold is lowered until a single cluster of perturbing positions appears, thereby mapping the docking site. Typically, perturbing modifications at different positions within the docking site have similar effects on the folding and activity assays, since they all inhibit the same protein–protein interaction. Positions within the docking site where Cys substitution itself perturbs the docking interaction are strong candidates for essential docking residues. Overall, this conservative method of analyzing PICM data tends to underestimate the total area of the docking site, since the actual docking site could possess a larger area than that detected. Once the general location of a docking site has been determined, additional PICM sites can be engineered and analyzed to more accurately determine the boundaries of the site if desired.

The use of PICM to map out protein–protein interaction surfaces has been illustrated by applications to the aspartate receptor, which have identified a large docking surface involved in the docking of CheA and CheW to the receptor, and in the assembly of receptor oligomers from isolated dimers (Mehan *et al.*, 2003). However, the fact that only activity assays were used in these applications, rather than both pairwise binding and activity assays, prevented the assignment of specific regions within the large docking surface to specific contacts with CheA, CheW, and other receptor dimers. The combined use of both binding and activity assays has been illustrated by an analysis of docking sites on CheA, which identified the specific docking sites for CheW and the receptor (Miller *et al.*, 2006). The CheW docking site defined by PICM analysis on the surface of CheA has been confirmed by an independent X-ray crystal structure of a complex formed between a thermophilic CheA fragment and its CheW partner (Park *et al.*, 2006), illustrating the ability of PICM to correctly map out a docking surface.

Finally, the PICM analysis typically identifies a large number of surface positions at which Cys substitution and bulky probe coupling have little or no effect on target protein stability, activity, and docking interactions. These nonperturbing Cys positions are useful in further biochemical and biophysical studies employing sulfhydryl chemistry, spectroscopic probes, or crosslinking chemistries to analyze the structure and mechanism of the assembled protein complex.

Acknowledgments

This work was supported by NIH R01 grant GM040731 to JJF.

References

- Baker MD, Wolanin PM, Stock JB. Signal transduction in bacterial chemotaxis. *Bioessays* 2006;28:9–22. [PubMed: 16369945]
- Bass RB, Coleman MD, Falke JJ. Signaling domain of the aspartate receptor is a helical hairpin with a localized kinase docking surface: Cysteine and disulfide scanning studies. *Biochemistry* 1999;38:9317–9327. [PubMed: 10413506]
- Bass RB, Falke JJ. Detection of a conserved alpha-helix in the kinase-docking region of the aspartate receptor by cysteine and disulfide scanning. *J. Biol. Chem* 1998;273:25006–25014. [PubMed: 9737956]
- Bass RB, Falke JJ. The aspartate receptor cytoplasmic domain: *In situ* chemical analysis of structure, mechanism, and dynamics. *Struct. Fold Des* 1999;7:829–840.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res* 2000;28:235–242. [PubMed: 10592235]
- Borkovich KA, Kaplan N, Hess JF, Simon MI. Transmembrane signal transduction in bacterial chemotaxis involves ligand-dependent activation of phosphate group transfer. *Proc. Natl. Acad. Sci. USA* 1989;86:1208–1212. [PubMed: 2645576]
- Bornhorst JA, Falke JJ. Purification of proteins using polyhistidine affinity tags. *Methods Enzymol* 2000;326:245–254. [PubMed: 11036646]
- Boukhvalova MS, Dahlquist FW, Stewart RC. CheW binding interactions with CheA and Tar. Importance for chemotaxis signaling in *Escherichia coli*. *J. Biol. Chem* 2002;277:22251–22259. [PubMed: 11923283]
- Bourret RB, Stock AM. Molecular information processing: Lessons from bacterial chemotaxis. *J. Biol. Chem* 2002;277:9625–9628. [PubMed: 11779877]
- Chervitz SA, Lin CM, Falke JJ. Transmembrane signaling by the aspartate receptor: Engineered disulfides reveal static regions of the subunit interface. *Biochemistry* 1995;34:9722–9733. [PubMed: 7626643]
- Corbin JA, Falke JJ. Affinity tags for protein immobilization and purification. *Encyclo. Biol. Chem* 2004;1:57–63.
- Cunningham BC, Wells JA. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* 1989;244:1081–1085. [PubMed: 2471267]
- Falke JJ, Bass RB, Butler SL, Chervitz SA, Danielson MA. The two- component signaling pathway of bacterial chemotaxis: A molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu. Rev. Cell Dev. Biol* 1997;13:457–512. [PubMed: 9442881]
- Falke JJ, Dernburg AF, Sternberg DA, Zalkin N, Milligan DL, Koshland DE Jr. Structure of a bacterial sensory receptor. A site-directed sulfhydryl study. *J. Biol. Chem* 1988;263:14850–14858. [PubMed: 3049592]
- Falke JJ, Sternberg DE, Koshland DE Jr. Site-directed sulfhydryl chemistry and spectroscopy: Applications in the aspartate receptor system. *Biophys. J* 1986;49:20a.
- Frazier AA, Wisner MA, Malmberg NJ, Victor KG, Fanucci GE, Nalefski EA, Falke JJ, Cafiso DS. Membrane orientation and position of the C2 domain from cPLA2 by site-directed spin labeling. *Biochemistry* 2002;41:6282–6292. [PubMed: 12009889]
- Frillingos S, Sahin-Toth M, Wu J, Kaback HR. Cys-scanning mutagenesis: A novel approach to structure function relationships in polytopic membrane proteins. *FASEB J* 1998;12:1281–1299. [PubMed: 9761772]
- Levit MN, Grebe TW, Stock JB. Organization of the receptor-kinase signaling array that regulates *Escherichia coli* chemotaxis. *J. Biol. Chem* 2002;277:36748–36754. [PubMed: 12119289]
- Malmberg NJ, Van Buskirk DR, Falke JJ. Membrane-docking loops of the cPLA2 C2 domain: Detailed structural analysis of the protein–membrane interface via site- directed spin-labeling. *Biochemistry* 2003;42:13227–13240. [PubMed: 14609334]

- Mehan RS, White NC, Falke JJ. Mapping out regions on the surface of the aspartate receptor that are essential for kinase activation. *Biochemistry* 2003;42:2952–2959. [PubMed: 12627961]
- Miller AS, Kohout SC, Gilman KA, Falke JJ. CheA kinase of bacterial chemotaxis: Chemical mapping of four essential docking sites. *Biochemistry* 2006;45:8699–8711. [PubMed: 16846213]
- Ninfa EG, Stock A, Mowbray S, Stock J. Reconstitution of the bacterial chemotaxis signal transduction system from purified components. *J. Biol. Chem* 1991;266:9764–9770. [PubMed: 1851755]
- Park SY, Borbat PP, Gonzalez-Bonet G, Bhatnagar J, Pollard AM, Freed JH, Bilwes AM, Crane BR. Reconstruction of the chemotaxis receptor-kinase assembly. *Nat. Struct. Mol. Biol* 2006;13:400–407. [PubMed: 16622408]
- Parkinson JS, Ames P, Studdert CA. Collaborative signaling by bacterial chemoreceptors. *Curr. Opin. Microbiol* 2005;8:116–121. [PubMed: 15802240]
- Sharp LL, Zhou J, Blair DF. Features of MotA proton channel structure revealed by tryptophan-scanning mutagenesis. *Proc. Natl. Acad. Sci. USA* 1995a;92:7946–7950. [PubMed: 7644518]
- Sharp LL, Zhou J, Blair DF. Tryptophan-scanning mutagenesis of MotB, an integral membrane protein essential for flagellar rotation in *Escherichia coli*. *Biochemistry* 1995b;34:9166–9171. [PubMed: 7619816]
- Wells JA. Binding in the growth hormone receptor complex. *Proc. Natl. Acad. Sci. USA* 1996;93:1–6. [PubMed: 8552582]
- Zhang XJ, Wozniak JA, Matthews BW. Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozyme. *J. Mol. Biol* 1995;250:527–552. [PubMed: 7616572]

Fig. 1. Space-filling structures of representative cysteine-specific probes and the cysteine side chain. All images to the same scale. For probes, the text indicates their full name, abbreviation, and net charge.