ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/7716010

Identification and Mapping of Protein−Protein Interactions by a Combination of Cross-Linking, Cleavage, and Proteomics

ARTICLE *in* BIOCONJUGATE CHEMISTRY · JULY 2005

Impact Factor: 4.51 · DOI: 10.1021/bc050043a · Source: PubMed

CITATIONS
88

READS
104

3 AUTHORS, INCLUDING:

Michael A Trakselis
Baylor University
39 PUBLICATIONS **787** CITATIONS
SEE PROFILE

Faoud Ishmael
Penn State Hershey Medical Center and Pen…
54 PUBLICATIONS **715** CITATIONS
SEE PROFILE

Available from: Faoud Ishmael
Retrieved on: 22 January 2016

JULY/AUGUST 2005
Volume 16, Number 4

© Copyright 2005 by the American Chemical Society

Bioconjugate Chemistry

REVIEWS

Identification and Mapping of Protein−Protein Interactions by a Combination of Cross-Linking, Cleavage, and Proteomics

Michael A. Trakselis,[†] Stephen C. Alley,*[,‡] and Faoud T. Ishmael[§]

Medical Research Council, Cancer Cell Unit, Hutchison MRC Research Centre, Hills Road, Cambridge CB2 2XZ, United Kingdom, Seattle Genetics, Inc., 21823 30th Drive Southeast, Bothell, Washington 98021, and The Pennsylvania State University, College of Medicine, Department of Medicine, H039, 500 University Drive, P.O. Box 850, Hershey, Pennsylvania 17033-0850. Received February 15, 2005; Revised Manuscript Received May 2, 2005

Protein−protein interactions are vital for almost all cellular functions, and many require the formation of multiprotein complexes. Identification of the macroscopic and microscopic protein interactions within these complexes is essential in understanding their mechanisms, both under physiologic as well as pathologic conditions. This review describes the current technology available to investigate interactions between proteins utilizing chemical cross-linking and site-directed cleavage reagents, outlining the necessary steps involved in identifying interacting proteins both in vitro and in vivo. Once interacting proteins are identified, more information about the architecture of the assemblies is necessary. Unique separation techniques coupled with cross-linking and mass spectrometry can now identify specific interaction sites and lead to the development of quaternary structural protein models. Furthermore, combination of these methods with proteomic approaches enables the identification and analysis of complex interactions in vivo. Finally, future directions in cross-linking methodologies are discussed.

INTRODUCTION

Protein−protein interactions are of extraordinary importance in biology. From multiprotein complexes, such as those found in DNA replication and RNA transcription, to transient interactions, such as those found in signal transduction networks, biological systems contain interaction networks that are critical for their function. The completed sequence information of a significant number of genomes and the availability of an increasing number of X-ray and NMR structures of key individual proteins have precipitated many advances in describing protein−protein interactions.

Numerous techniques have been developed to probe protein−protein interactions, with some methods providing very high structural resolution with low throughput, such as X-ray crystallography and multidimensional NMR, while others offer very high throughput with low structural resolution, such as phage display, yeast two-hybrid screens, and protein microarrays. Quantitative descriptions of the strength of protein−protein interac-

* To whom correspondence should be addressed: Tel: 425-527-4752; fax: 425-527-4001; e-mail: SAlley@seagen.com.
[†] Hutchison MRC Research Centre.
[‡] Seattle Genetics, Inc.
[§] The Pennsylvania State University.

10.1021/bc050043a CCC: $30.25 © 2005 American Chemical Society
Published on Web 07/01/2005

tions with low structural resolution can be determined by methods such as gel filtration, Biacore, and fluorescence spectroscopy. Chemical cross-linking does not have the throughput of techniques such as phage display or yeast two-hybrid systems, where libraries of interacting partners can be rapidly screened, nor does it provide information about the strength of an interaction, but it does allow higher resolution structural data to be generated where protein−protein interactions can be mapped to specific domains or amino acids. Most importantly, cross-linking allows noncovalent protein−protein interactions, which may be transient or dependent on specific physiological conditions, to be captured into long-lived covalent complexes that maintain structural information during subsequent purification, enrichment, and analysis.

Many protein cross-linking tools are commercially available, enabling wide access to investigation of protein−protein interactions. New reagents and mass spectrometry (MS)[1] have significantly increased the power of these tools and the ability to investigate more complicated systems, including protein−protein interactions in vivo. Interestingly, the two areas of protein cross-linking and proteomics have taken inspiration from each other, and several hybrid tools have emerged.

The investigation of protein−protein interactions may have several goals. The first may simply be the identification of two proteins that interact and may start with the two proteins in either purified form or in more complex mixtures. Detecting a new, higher molecular-weight band on an SDS−PAGE gel after cross-linking and Coomassie or silver staining is sufficient to identify an interaction between two purified proteins. However, the absence of a new band is not sufficient to exclude the possibility that two proteins do interact. This limitation is identical for many other techniques for investigating protein−protein interactions, in that the absence of a signal only excludes protein−protein interactions under the specific set of conditions used, which may be very different from those in vivo or lack additional factors necessary for forming the interaction. Detection of cross-links in the presence of many other proteins requires affinity tagging the protein or using antibodies against the protein for visualization. In the absence of these reagents, more sophisticated separation and/or visualization techniques, such as using affinity handles on the cross-linker or tandem mass spectrometry, will be required for identification of cross-linked proteins. Once interacting partners have been identified, the next goal may be to map the location of the interaction, which may also employ these sample visualization and identification tools with MS. Importantly, the yield from cross-linking and any separation technique employed must be high enough so that the subsequent MS analysis will be successful in identifying points of interaction. Significant effort in the choice of cross-linking reagent, scale-up, and/or cross-linking reaction optimization may be necessary in advance of MS investigation. Ultimately, sophisticated

[1] Abbreviations: ANB-NOS, *N*-5-azido-2-nitrobenzoyloxysuccinimide; BMHT, betaine-homocysteine methyltransferase; DSP, dithiobis[succinimidylpropionate]; DSS, disuccinimidyl suberate; DTT, dithiothreitol; ECAT, element-encoded affinity tag; FeBABE, 1-(*p*-bromoacetmidobenzyl)-EDTA; IAC, *N*-(3-iodo-4-azidosalicyl)cysteine; ICAT, isotope-coded affinity tag; MBP, 4-maleimidobenzophenone; MS, mass spectrometry, SAED, sulfosuccinimidyl 2-[7-amino-4-methylcoumarin-3-acetamido]-ethyl-1,3′dithiopropionate; sulfo-SBED, sulfosuccinimidyl[2-6-(biotinamido)-2-(*p*-azidobenzamido)-hexanoamido]-ethyl-1,3′-dithiopropionate; VICAT, visible isotope-coded affinity tag.

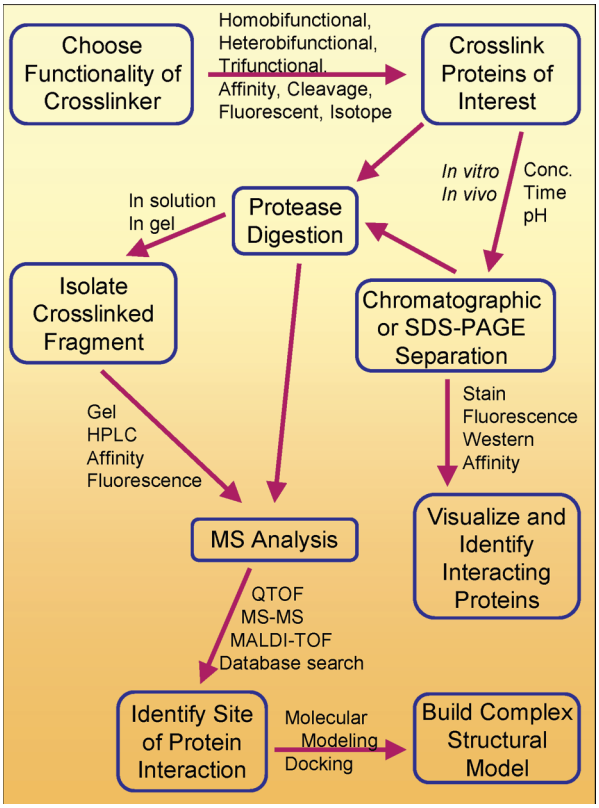

Figure 1. Cross-linking methodology flowchart demonstrating the procedures involved in identification of the specific site of cross-linking between two or more proteins and construction of quaternary complex models.

models of multiprotein complexes can be assembled from these results. Figure 1 characterizes the uses of cross-linking discussed above and outlines general strategies for the identification of specific protein interaction sites that can be used to develop quaternary protein models.

This review will illustrate the techniques that are currently being used in the identification of protein−protein interactions. Many commercially available reagents have been used in these studies, and novel probes and mass spectrometry techniques have come into play as the complexity of the multiprotein complexes under investigation have increased. Figure 2 provides an overview of the types of protein modification groups that are available. Modern mass spectrometry and proteomics techniques have significantly increased the scope of these studies and have begun to allow investigation of protein−protein interactions under more physiologically relevant conditions than for two purified and isolated proteins, including cell lysates and in vivo.

IDENTIFYING PROTEIN−PROTEIN INTERACTIONS

Homobifunctional and Heterobifunctional Cross-Linkers. The most common type of protein−protein interaction is the interaction of a protein with itself to form multimers. Often this can regulate the activity of the protein and provide a powerful switch between active and inactive forms. There are numerous techniques for studying oligomerization, and cross-linking provides the ability both to determine the degree of oligomerization and to map the network of interactions at the protein−protein interface. Many commercially available reagents can facilitate these studies by forming cross-links in proteins between pairs of lysines, lysines and aspartic or glutamic acids, lysines and cysteines, or pairs of

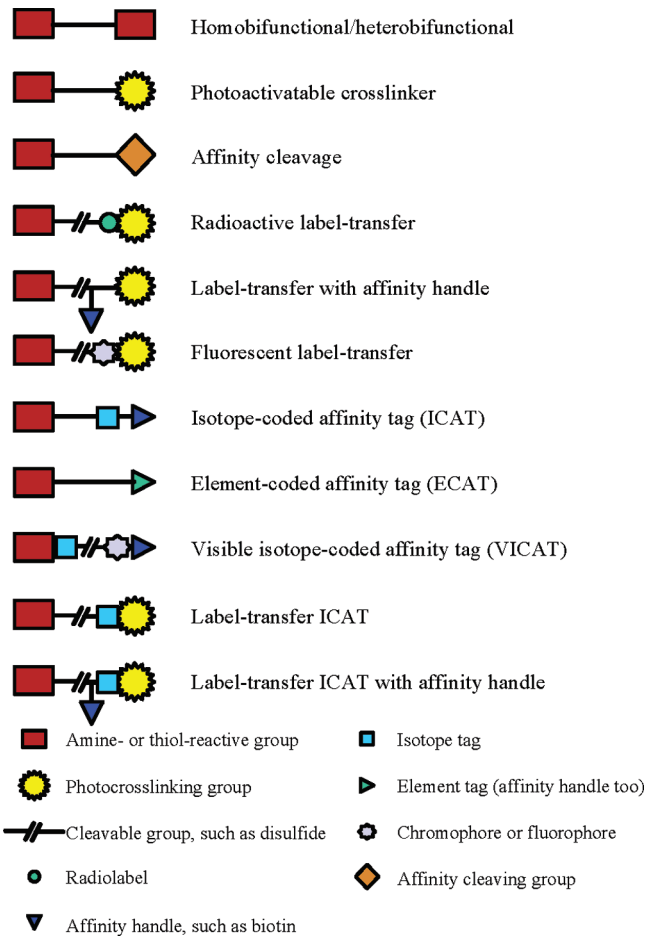

Figure 2. Schematic structures of the protein modification groups discussed in this review.

cysteines. The following examples show that these types of cross-links can be formed under a variety of biologically relevant conditions, including inside of cells and membrane preparations.

Cross-linking through the ϵ-amino group of lysines can demonstrate oligomerization due to the large number of lysines present on the surface of most proteins. It is highly likely that two lysines will be close to each other at the protein−protein interface and form cross-links. Szegedi and Garrow (*1*) used a 30-min incubation with glutaraldehyde to cross-link betaine-homocysteine methyltransferase (BHMT). Coupled with gel filtration, the oligomeric state of BHMT and several mutants could be assigned. Wild-type BHMT and some mutants were tetramers, while other mutants were dimers. Dimeric mutants were enzymatically inactive, demonstrating that the tetrameric form is required for activity.

Cysteines are less common on the surface of proteins, but can provide a higher resolution investigation of the protein−protein interface. Bunn and Ordal (*2*) made 30 single cysteine mutants (2 runs of 15 contiguous amino acids) of the homodimeric *Bacillus subtilis* chemoreceptor McpB, a member of a two-component signal transduction cascade. Using a 5-min incubation with copper-phenanthroline to oxidize properly orientated cysteines from each monomer to disulfides (a zero-length and reversible cross-link), contacts within the two α-helical transmembrane domains were mapped. Cross-linking was carried out using cells transformed with plasmids containing mutant *mcpB* and cross-links visualized by Western blots probed with anti-McpB antibodies. As would be expected for an α-helical bundle, cross-linking was maximized (up to 100% yield) at a spacing of n to $n+3/n+4$ in each domain, establishing the dimerization interface. However, cross-linking was also observed at amino acids on the opposite side of the α-helical bundle from the dimerization interface, leading the authors to suggest that in vivo, McpB can also exist as trimers of dimers.

The tether length of the cross-linker can be varied to map the distance between two cross-linked cysteines. Kaback and co-workers used 5 and 21 Å tether length bis-methane thiosulfonates to cross-link single cysteine mutants of *Escherichia coli* lactose permease (LacY). LacY is a membrane protein with 12 transmembrane helices and loops on the periplasmic and cytoplasmic sides of the membrane. Two hundred fifty transmembrane mutants (*3*) and 143 extramembrane mutants (*4*) were investigated for their importance in homodimer formation. Cross-linking was carried out for 3 min to 1 h using sonicated *E. coli* membrane suspensions made from cells transformed with plasmids containing mutant *lacY*, with anti-C-terminal antibodies used for visualization of cross-links on Western blots. Cross-linking yields of greater than 40% were observed for some mutants. Only 4% of the transmembrane mutants formed cross-links with the 5 Å cross-linker, while 53% of the extramembrane mutants formed cross-links with the 5 and 21 Å cross-linkers. Just over half of the extramembrane cross-links formed with only the 21 Å cross-linker. An interaction map of the homodimer interface was constructed from these results, with domains from both the periplasmic and cytoplasmic faces contributing significant contacts.

Photoactivatable Cross-Linking. Cross-links can also be formed between proteins using reagents that are activated by light. Heterobifunctional benzophenones and aryl azides can first be attached to proteins through lysines or cysteines, and then multiprotein complexes can be assembled under safe lighting conditions. Cross-links are only formed upon exposure to highly intense, short wavelength (250−350 nm) light. This offers the researcher a significant advantage, in that there is very good control for the start and end of cross-linking since light is the exclusive trigger for the cross-linking reaction. Photo-cross-linking times are generally shorter (5−15 min with aryl azides) than chemical cross-linking times, sometimes significantly shorter (down to 1 s with metal catalysis). However, photo-cross-linking can often be a process with poor yields and a diversity of possible products, especially with aryl azides, potentially making the identification of the location of photo-cross-linking very difficult. Numerous commercially available reagents, such as 4-maleimidobenzophenone (MBP) and *N*-5-azido-2-nitrobenzoyloxysuccinimide (ANB-NOS) (Figure 3), have been used to form photo-cross-links.

Giron-Monzon and co-workers (*5*) investigated the ATP-dependent activation of the *E. coli* mismatch repair exonuclase MutH by MutL by photo-cross-linking. MutL has 7 native cysteines, and these were first mutated to serine, leucine, or phenylalanine. Thirteen single cysteine mutants were then constructed and labeled with MBP. Complexes with MutH in the presence or absence of ATP, the nonhydrolysable ATP analogue ADPnP, and DNA were formed and then photolyzed for 10−30 min. An interaction map was then constructed, and a mechanism for the ATP dependence of the MutH-MutL complex formation could be determined.

Protein−protein interactions can also be studied by metal-catalyzed chemical- and photo-cross-linking. Chelated Ni(II) can be used to introduce zero-length cross-links by treating with an oxidant such as magnesium

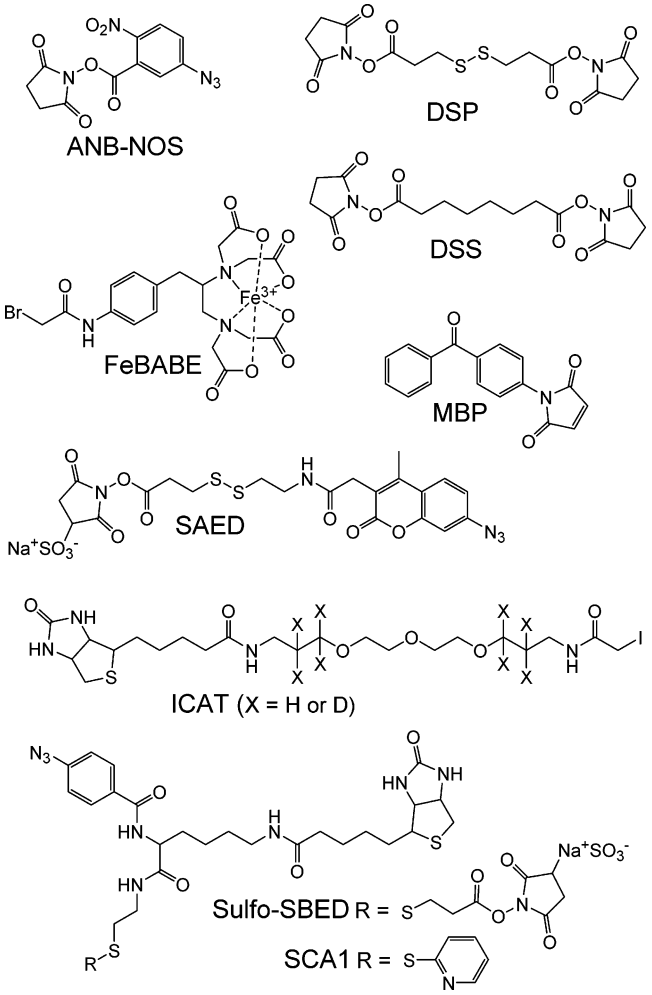

Figure 3. A selection of the structures of commonly used cross-linkers discussed in this review: *N*-5-azido-2-nitrobenzoyl-oxysuccinimide (ANB-NOS), dithiobis(succinimidyl propionate) (DSP), disuccinimidyl suberate (DSS), 1-(*p*-bromoacetmidobenzyl)-EDTA (FeBABE), 4-maleimidobenzophenone (MBP), sulfosuccinimidyl 2-(7-azido-4-methylcoumarin-3-acetamido)-ethyl-1,3′-dithiopropionate (SAED), isotope-coded affinity tag (ICAT), sulfosuccinimidyl 2-[6-(biotinamido)-2-(*p*-azidobenzamido)-hexanoamido]-ethyl-1,3′-dithiopropionate (Sulfo-SBED), and 1-[6-(biotinamido)-2-(*p*-azidobenzamido)-hexanoamido]-2-(2′-pyridyldithio)ethane (SCA1).

monoperoxyphthalic acid (*6*), while chelated Ru(II) (*7*) or Pd(II) (*8*) can introduce zero-length cross-links by photolysis in the presence of persulfate. Genetically encoded His_6-tags have been used for protein-bound Ni(II) chelators, while bodipy and porphyrins have been used for nonprotein bound chelation of Ru(II) and Pd(II), respectively. Depending on the system, photo-cross-linking times can be less than 1 s, and yields tend to be very high. Cross-linked amino acids are commonly tyrosines. These cross-linking reagents can also be used for affinity labeling with biotin-tyramine (*9*, *10*), which is chemically similar to tyrosine. The mechanisms of cross-linking and the protein complexes investigated with these reagents have been described in two excellent reviews (*11*, *12*).

Site-Directed Cleavage Agents. Instead of producing covalent cross-links, site-directed cleavage agents can be used to cleave a target protein at the point of protein−protein interactions. Iron-EDTA complexes have been used to generate hydroxyl radicals via a Fenton reaction and cleave both nucleic acids (*13*) and proteins. 1-(*p*-Bromoacetmidobenzyl)-EDTA (FeBABE) (Figure 3) can be attached to protein cysteines or 2-iminothiolane-modified amino groups and activated with iron, sodium ascorbate, and hydrogen peroxide. First demonstrated by internally cleaving bovine serum albumin (*14*), this technique generates peptide fragments that are not chemically damaged at their termini and can be Edman-sequenced to identify the location of cleavage. This site-directed cleavage agent can cleave peptide bonds within a protein or on interacting proteins as well as phosphodiester bonds of polynucleotides (Figure 4).

FeBABE has been used to study protein−protein interactions in the *E. coli* RNA polymerase holoenzyme. The stationary phase σ factor (σ^s) was mutated such that unique cysteines were independently placed at nine locations. The FeBABE−σ^s conjugates were assembled into the $\alpha_2\beta\beta'$ RNA polymerase holoenzyme, and cleavage was observed on the β and β' subunits, but not on the α subunit (*15*). In this study, visualization of the individual RNA polymerase subunits was accomplished by probing a Western blot with antibodies to the N-termini of each of the subunits.

In an extension of this method, FeBABE was used to map interactions between *S. cerevisiae* TFIIB and RNA ploymerase II. Five mutants of TFIIB containing unique, alkylatable cysteines were labeled with FeBABE and incubated with yeast nuclear extracts containing RNA polymerase II subunits individually genetically flag-tagged. The location of cleavage was determined using flag-tagged size standards from in vitro transcription/translation reactions of truncated RNA polymerase II subunit genes as templates (*16*). The use of genetically encoded flag tags removed the requirement for production of antibodies against all of the potentially cleaved proteins. In principle, any type of genetically encoded affinity tag, such as His-tags, GST-tags, or peptide sequences that are biotinylated during expression (*17*), could provide both purification and visualization handles for any recombinant proteins without the need to produce antibodies to these proteins.

Visualization Techniques: Label Transfer Agents. Tracking changes by mobility shift on by SDS−PAGE with Coomassie or silver staining to identify interacting proteins becomes quickly intractable as the number of proteins in the reaction mixture increases. Protein−protein interactions found within large complexes, such as those found in DNA replication and RNA transcription complexes, often occur in the presence of up to 50 polypeptides. To study these interactions, researchers have developed novel reagents to monitor protein−protein interactions. These reagents incorporate isotope or element tags to allow identification by mass spectrometry, radiolabels or fluorophores to enable identification of interaction partners, or affinity tags to allow both purification and identification. One powerful application of these techniques is label transfer (Figure 5).

Label transfer requires that a single protein of interest (A in Figure 5) is first modified with the label transfer agent. Second, the multiprotein complex is assembled in the presence of the labeled protein and cross-links are formed, typically by photolysis of reactive groups such as benzophenones or aryl azides. The identification of the interacting protein or proteins (B in Figure 5) is accomplished by separation by SDS−PAGE followed by autoradiography or visualization of the affinity tag (for instance, probing a gel with streptavidin−horseradish peroxidase to identify biotinylated proteins). The resolution of these techniques can be significantly enhanced by coupling with molecular biology: engineering of unique cysteines allows the label to be incorporated site-

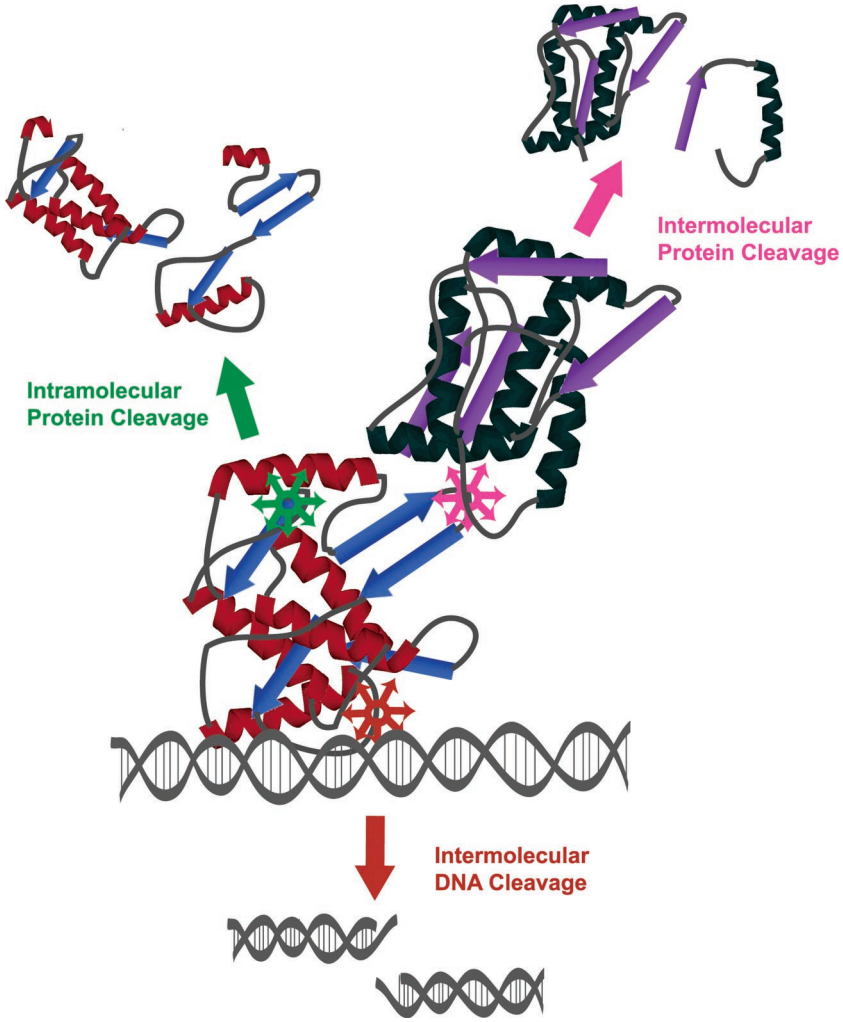

Figure 4. Site-directed cleavage agents such as FeBABE can be attached at site-specific cysteines or modified amines within a bait protein to detect intra- or intermolecular interactions. In the presence of ascorbate and peroxide, Fe^{3+} is reduced to Fe^{2+} and radicals cleave peptide bonds or phosphodiester bonds that are close in space. Detection of the resulting peptide or nucleic acid pattern can be analyzed by electrophoresis, Western blot, sequencing, or mass spectrometry and can provide information relating to the region of contact within the protein or the interacting complex.

specifically, and moving this unique cysteine around the protein enables interaction mapping.

Classically, label transfer used ^{125}I- (protein A-IgG) or ^{35}S- (gelatin-fibronectin) labeled reagents on lysine-reactive probes (*18*, *19*). Cleavage of the cross-linking group transfers the label to the interacting partner and was affected with DTT to cleave a disulfide or sodium dithionite to cleave an azo linkage. These early studies were enhanced when cysteine reactive probes were introduced. Chen and co-workers (*20*) incorporated ^{125}I-labeled *N*-(3-iodo-4-azidosalicyl)cysteine (IAC) into an engineered form of the catabolite gene activator protein (CAP). Following photo-cross-linking, label transfer, and DTT cleavage, the interaction of CAP, labeled in a region important for activation, was mapped to the carboxyl-terminal region of the α-subunit of RNA polymerase (RNAP) within the four-polypeptide RNAP. In the absence of the lac promoter DNA and/or CAP, no cross-linking or label transfer was observed, demonstrating the high specificity of the technique.

As an alternative to radioactive visualization, an affinity label-transfer reagent, sulfo-SBED (Figure 3) was developed at Pierce (*21*). This reagent transfers biotin from the labeled protein to the cross-linked interaction partner. Biotin is then used as a visualization reagent: Western blots of cross-linking mixtures following label transfer are probed with horseradish peroxidase to reveal cross-linked proteins. Sulfo-SBED is lysine reactive and has been used to study many protein−protein interactions. To increase the utility of this reagent, it has been converted to cysteine-reactive reagent by attaching a thiopyridyl group (SCA1, Figure 3) (*22*). This modified reagent has been used to track protein−protein interactions in the bacteriophage T4 DNA replication complex (four polypeptides in a ca. 350 kDa complex with DNA) (*22*) and a 1.8 MDa RNA transcription complex with at least 14 unique polypeptides (*23*). The benefit of these trifunctional reagents is the rapid identification of interacting proteins which only require photo-cross-linking and Western blotting with a universal reagent, without resorting to having to raise antibodies to every potential protein in the complex and running multiple Western blots with these custom reagents.

Sample Preparation: Proteomics and ICAT. The investigations described above involve labeling or making a mutant of a protein of interest before cross-linking. Studying protein−protein interactions in vivo generally does not allow for such modifications before crosslinking. Using these highly complex mixtures of proteins requires significant effort in sample preparation to identify interacting proteins. This can be accomplished after cross-linking by electrophoresis or chromatography, either with

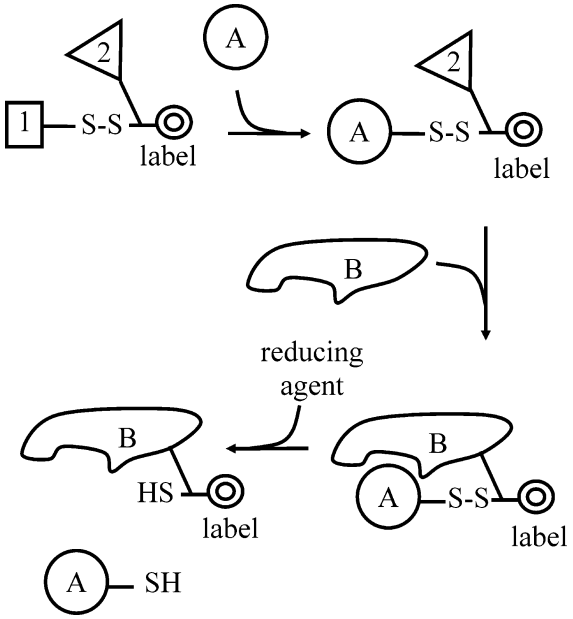

Figure 5. Label transfer cross-linking. Cross-linker consists of an amine or thiol reactive group (1), a photactivatable group (2), and a label (typically either a radioactive or biotin moiety). Protein A is first conjugated to the cross-linker at site 1. Protein B is then added and photo-cross-linked at site 2. Reduction of the disulfide bond with DTT results in the transfer of the label to protein B liberating free protein A. Separation of the protein mixture is performed by SDS−PAGE. Detection of the transferred label then occurs either by autoradiography or probing with streptavidin−horseradish peroxidase.

or without the benefit of affinity handles for enrichment and visualization.

The emergence of matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry within the past decade has greatly simplified the ability to identify proteins using an approach termed proteomics. Proteomics generally involves separating a mixture of proteins in a given system using techniques such as 2D gel electrophoresis followed by staining to detect protein bands. These bands are then excised, digested with a protease in-gel, and subjected to mass spectrometry. Each protein exhibits a specific digestion pattern with a given protease, which is manifested on MALDI as a specific mass fingerprint. This allows the identification of unknown proteins by comparing its digestion pattern to a database containing the proteolytic masses of known proteins. A variety of databases and algorithms have been constructed and developed to search for corresponding fragments within an organism's genome (*24*). The combination of these techniques with cross-linking has provided a powerful methodology to quickly determine interacting proteins in a complex.

This approach was exemplified by Bernhard et al. in the study of the transmembrane CD4 receptor (*25*). The authors used the homobifunctional lysine-lysine cross-linker DSP (Figure 3) to cross-link the proteins at the cell surface. Study of cell surface proteins and their associations with other proteins has traditionally been very difficult. In vitro analyses usually require removal of the protein from the membrane by proteolyitc cleavage, detergent-mediated release, or removal of the transmembrane domain by mutagenesis, all of which may alter its structure and function. However, the utilization of this cross-linker was beneficial because it allowed the protein to be studied in an in vivo setting, and the abundance of lysines on proteins in general allows for the maximal amount of cross-links between proteins in the complex. DSP also contains a disulfide linkage enabling cleavage of the linker if desired. The cross-linked proteins were released from the cell membrane and immunopreciptated using an antibody against the CD4 receptor to isolate only the cross-linked complex of interest. The cross-linker was then cleaved with a reducing agent to release the covalently linked complex, and the proteins were separated by SDS−PAGE. Individual bands were excised and digested, and the proteins were identified using MALDI MS. The authors were able to identify interacting proteins, including a trasferrin receptor, and a protein phosphatase, which provided valuable information about the function of the protein.

A similar technique was employed by Meunier and coworkers (*26*) in the study of chaperone proteins within the ER of mammalian cells. This group also used DSP to cross-link the proteins and followed with isolation by affinity purification, separation by 2D gel electrophoresis, and identification with MS. The interaction of an immunoglobulin with a chaperone complex was probed. These studies recognized a novel organization of chaperone proteins in a preformed complex and gave insight into its interaction with unfolded proteins. Because of the ability of DSP to cross the cell surface and permeate cells, these experiments were conducted in intact cells, further delineating the importance of this technique in studying in vivo systems.

Proteomics and mass spectrometry also benefit from enrichment and visualization tools. The first tools designed for these purposes include isotope-coded affinity tags (ICAT) (Figure 3) (*27*), which were designed to modify proteins with noncleavable biotin groups containing isotopic labels. Biotin allows enrichment of the sample before analysis, and the use of multiple isotope labels allows different states of the biological system to be encoded with each label. Mass spectrometry enables rapid identification of labeled proteins from their unique masses. Extensions of this method include element-encoded affinity tags (ECAT) (*28*) and visible isotope-coded affinity tags (VICAT) (*29*). ECAT reagents combine the element tag and affinity handle into a single DOTA functional group (anti-DOTA antibodies can be used for sample enrichment), while VICAT incorporates a visual chromophore for tracking during chromatography.

To extend these techniques for the investigation of protein−protein interactions, Trester-Zedlitz and coworkers (*30*) combined label transfer and ICAT in a modular system. Label-transfer reagents were built such that spacer lengths, chemical reactivities, affinity handles, cleavage sites, and isotope tags were varied so that they could be optimized to the multiprotein complex to be studied. The authors point out that a general method of interaction identification is probably not achievable due to the diversity of protein structure and reactive groups for label attachment. Building a modular toolbox will extend these techniques to a much greater number of proteins.

MAPPING SPECIFIC PROTEIN−PROTEIN CONTACT POINTS

Once two proteins have been identified as interacting partners by cross-linking, the specific site of interaction can be determined through mass spectrometry. MS analysis can be simplified exponentially if the digested cross-linked peptide can be isolated and purified first. Incorporation of either fluorescent probes or affinity handles into traditional cross-linkers aid tremendously in the isolation of the suitable adducts. Once the peptide of interest is isolated, analysis by a variety of MS

techniques reduce the old saying of "finding a needle in a haystack" to just finding the hay. An important caveat to this approach is that significant amounts of cross-linked products are needed for MS analysis. Therefore, efficient cross-linking and/or enrichment of a cross-linked species are necessary for successful identification. Various techniques including cross-linking conditions, fluorescent cross-linkers, affinity labels, and the ever increasing sensitivity of mass spectrometers all help to overcome this limitation.

Bromobimanes are the most widely used fluorescent cross-linkers. They are thiol−thiol cross-linkers and have the advantage that they only fluoresce upon conjugation. Therefore, the reaction progress can be monitored directly, simplifying analysis. They are also extremely stable in solution and have high thiol selectivity (*31*). Dibromobimane has been used successfully to cross-link and identify intramolecular thiols in myosin (*32*), lactose permease (*33*), and ATPase (*34*). A cleavable, fluorescent, photoactive cross-linking agent, sulfosuccinimidyl-2-(7-azido-4-methylcoumarin-3-acetamido)-ethyl-1,3′-dithiopropionate (SAED, Figure 3) has also been used to identify specific protein cross-links (*35*, *36*). The cross-linker is initially attached to amino groups of the protein of interest, then photo-cross-linked via an azide to residues on an interacting protein. The complex is then reduced, cleaving the disulfide within the cross-linker and transferring the fluorescent group to the interacting protein in a label transfer reaction (Figure 4). Protease digestion will then liberate a peptide with the fluorescent group identifying the interacting site. Other fluorescent dyes can be incorporated into trifunctional cross-linkers as the third arm to provide a tracer for peptide isolation. After digestion with trypsin or other proteases, these fluorescent protein fragments can be separated chromatographically or by SDS−PAGE and specifically analyzed by MS.

Another method of identification of cross-linked products is the use of affinity handles for enrichment of cross-linked products. The most common affinity handle is biotin, which can be specifically isolated using an affinity column composed of monomeric avidin or streptavidin beads. The previously described biotin cross-linkers, typically used in a label transfer reaction for visualization (*21*, *22*), can be used in addition to other trifunctional reagents (*30*, *37*) as affinity purification tools. After cross-linking two proteins, the biotin moiety is transferred to the cross-linked interacting protein with addition of a reducing agent. After protease digestion, isolation of a cross-linked fragment occurs by passing the entire protease digestion over an avidin affinity column. After washing, the cross-linked peptide with the biotin moiety is eluted with free biotin and isolated for MS analysis.

Upon isolation and purification of the cross-linked fragment, the mass of the individual peptide can be easily identified by MS through a single peak (*38*). Then the composition of the cross-linked peptide can be determined by comparison with the known sequence of the two interacting proteins. Advances in mass spectrometry have led to new methods of analysis including tandem mass spectrometry (*39*, *40*). This technique combines two rounds of MS analysis to identify the specific composition of a component. The first round separates the components based on mass, and the second round of MS breaks apart the individual peptides one amino acid at a time to yield a ladder of peaks identifying the sequence of a peptide. The mapping of specific protein−protein contact points by cross-linking along with individual structures of proteins and computer modeling will allow scientists to piece together protein complexes from a limited number of interaction points. As the sensitivity of MS detection and analysis increases, specific interactions in transient protein complexes should also become more clear.

USING CROSS-LINKING TO CREATE PROTEIN COMPLEX MODELS

Once interacting proteins in a complex are identified, with or without the determination of specific interaction points as described above, additional structural relationships and information about the architecture of the complex are necessary for a complete understanding of biological function. Sites of interaction or interfaces formed can be important in structure and function, and thus may be important in normal physiologic as well as pathologic processes, and generation of therapeutic agents. To reduce the bias and improve the resolution of the final model, a large number of cross-links should be formed across the interaction interface. Use of a homobifunctional lysine-lysine cross-linker allows the maximal amount of cross-linking due to the large number of lysines in proteins. Not only is there a greater chance that at least one of these lysines lie in or close to an interface, but the likelihood of multiple lysines at this site allows a map to be created of the interaction interface. In this way, unexpected or multiple modes of interaction may also be observed.

Young et al. (*41*) examined the intramolecular cross-links of fibroblast growth factor (FGF-2) which has a known three-dimensional structure. They were able to separate intramolecular cross-links from intermolecular cross-links by column chromatography or SDS−PAGE. Chromatographically separated fractions of the peptide digest of the intramolecular cross-links were then analyzed by MS. Cross-links were identified by analyzing the MS data with a custom software program, ASAP, and a model of the protein was generated. The data obtained were consistent with the crystal structure of FGF-2, with the interatomic distances experimentally determined by cross-linking close to the predicted distances. The structure of ubiquitin was also probed using a modified procedure in which the cross-linked protein was injected into an electrospray ionization (ESI)−Fourier transform (FT) mass spectrometer, and the cross-link positions were localized by multiple stages of fragmentation and mass spectrometry, instead of proteolysis followed by column chromatography and then MS (*42*). This approach, which yielded similar data, provided a simpler and quicker approach, which facilitates the application to high throughput experiments.

The structure of the yeast prohibin complex, a large multimeric assembly with chaperone-like activity was investigated without chromatographic separation techniques (*43*). The oligomer, which is composed of repeating heterodimers of a PHB1 and PHB2 subunits was cross-linked with DSP, digested with trypsin, and analyzed by either MALDI MS or ESI MS. The authors used a custom-made software tool, FindLink, to filter the data and identify cross-linked peptides. Since each protein contained multiple lysine residues close to the subunit interface, a total of nine heterodimeric cross-links were identified. This allowed a map of the subunit interface to be generated, and a model of the dimer was generated.

Similarly, a model of the signal recognition particle (FfH) and its receptor (FtsY) was built by chemical cross-linking with disuccinimydyl suberate (DSS, Figure 3) or bis(sulfosuccinimidyl) adipate (BSSA), proteolytic digestion, and tandem MS (*44*). In this study, they identified a total of nine separate intermolecular cross-linked

peptides. Using a novel model building approach which included geometric restraint optimization and macromolecular docking, they were able to identify a best fit for this complex based on the cross-linking data. After these findings, a partial crystal structure of the FfH−FtsY complex was determined and found to be in good agreement with the cross-linking model but lacked the M domain of FfH. Interestingly, four out of the nine cross-linked fragments provided information of the FfH M domain and provided new insights into its conformation relative to the rest of the complex which could not be determined in the crystal structure complex.

In a final example, the binding of calmodulin to mettilin, a polypeptide and principal component of honeybee venom, was also investigated without chromatographic separation techniques (*45*). The proteins were cross-linked with a number of different length amine−amine and amine−carboxylic acid cross-linkers, digested with trypsin, and analyzed by HPLC equipped with a high-resolution FT ion cyclotron resonance (ICR) mass spectrometer. The authors used distance constraints derived from the chemical cross-linking data in combination with computational methods three-dimensional structure models of the calmodulin−mettilin complex. From the numerous points of cross-linking obtained, an interface between the two proteins was mapped. The authors demonstrated for the first time that mettilin could bind to calmodulin in two different orientations (Figure 6).

The combination of cross-linking with mass spectrometric techniques used in proteomic identification has provided a powerful new tool to study the quaternary structure of multiprotein complexes. The development of tandem MS has allowed for more accurate identification and sequencing of peptides and the ability to detect femtomolar to nanomolar amounts of material, and increasingly sophisticated software tools may ultimately allow determination of cross-linking locations without the need for enrichment or separation techniques.

FUTURE DIRECTIONS

Genetically Encoded Cross-Linking Agents. One of the challenges of site-specific modification of proteins with cross-linking agents is the necessity for a single alkylatable amino acid, typically cysteine. This can be accomplished by mutagenesis, but large proteins with numerous cysteines pose a significant challenge for constructing genes with multiple mutations. The alternative is to genetically incorporate noncanonical amino acids into proteins to provide the cross-linking group. Incorporation of noncanonical amino acids has been an accessible technology for decades, such as the incorporation of selenocysteine by adding this amino acid to the growth medium during protein expression. The native cysteine-tRNA synthetase can aminoacylate $tRNA_{cys}$ with selenocysteine under conditions when intracellular cysteine concentrations are low. The diversity of noncanonical amino acids that can be mischarged is low and unfortunately does not include cross-linking groups.

In the past few years, tRNA synthetases have been engineered to be able to aminoacylate tRNAs with an impressive number of noncanonical amino acids (*46*), including photo-cross-linking amino acids. Benzophenone is a very high efficiency photo-cross-linker and can be incorporated into proteins with the amino acid *p*-benzoyl-L-phenylalanine (pBpa). A library of *Methanococcus jannaschii* tyrosyl-tRNA synthetases was selected for the ability to aminoacylate a mutant tyrosine amber suppressor tRNA with pBpa. By engineering an amber codon into glutathione S-transferase and expressing in the presence on the mutant tRNA synthetase/tRNA pair and pBpa, site-specific incorporation of a benzophenone group was accomplished (*47*).

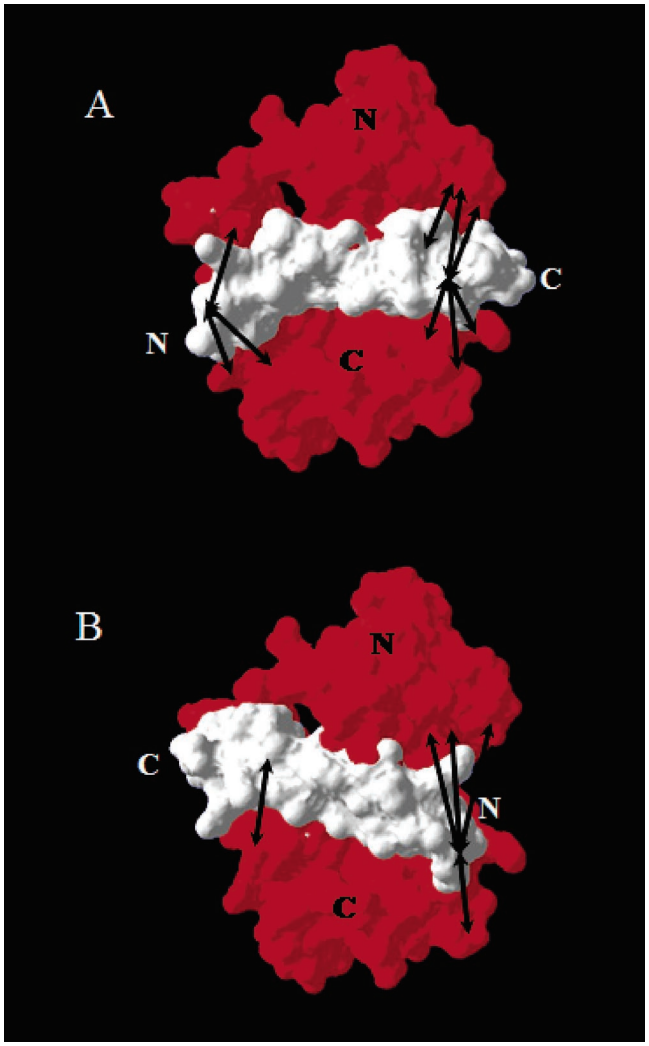

Figure 6. The complex between calmodulin (red) and mettilin (white) was analyzed using cross-linkers of different lengths. The arrows represent cross-links that were obtained and mapped. Calmodulin was found to bind mettilin in two different orientations, as shown in panels A and B. The amino- and carboxyl-terminal domains of calmodulin are labeled N and C, respectively. The amino and carboxyl termini of mettilin are similarly labeled.

Alternatively, a mutant *E. coli* phenylalanine-tRNA synthetase has been engineered to allow incorporation of *p*-azidophenylalanine. Incorporation of this photo-cross-linking amino acid was accomplished using an *E. coli* phenylalanine auxotroph. Dihydrofolate reductase was successfully expressed with *p*-azidophenylalanine in place of phenylalanine (*48*). The disadvantage of this system is that all of the phenylalanines in the protein will be replaced, which is no more desirable than chemical modification of multiple cysteines. It has been shown, however, that the *E. coli* phenylalanine codon degeneracy can be broken with a mutant yeast phenylalanine tRNA specific for only one of the two phenylalanine codons (*49*). Site-specific incorporation of *p*-azidophenylalanine could be accomplished after making a mutant gene encoding for only one copy of the target codon, similar to the amber suppressing method above.

Expressing site-specific photo-cross-linking amino acids in proteins opens up the possibility of probing protein−

protein interactions in vivo without having to purify a target protein, incorporate the cross-linker, and reconstitute the protein complex. Expression of proteins with noncanonical amino acids in yeast (*50*) may also enable the opportunity of studying eukaryotic protein complexes using these technologies.

Identification of Variable Protein Complex Interactions. Rearrangements of protein complex composition are known to occur in response to extremes that alter the cellular environment, during normal processes throughout the cell cycle of an organism, or in the presence of pathologic disease. Already, microarray technology has detected these phenomena at the mRNA level, but this technology does not demonstrate what happens at the functional protein level. Assembly and disassembly of protein complexes occur in response to cellular signals. Formation of DNA replication factories, modulation of signal transduction complexes, transcriptional activation and repression, and translational protein/RNA complexes, as well as other multiprotein enzymatic machines, are all assembled and disassembled in response to cellular signaling. In vivo cross-linking in combination with antibody arrays designed for a specific biological area of interest coupled with mass spectrometry will provide a way to monitor changes in protein−protein interactions in response to certain cellular signals.

The ability to detect complex networks of interacting proteins has been limited in the past to focusing on single pairs of interacting proteins or a limited study of single branches of pathways. With the advances in the sensitivity, accuracy, and speed of mass spectrometers, there has been a trend toward automation of gathering and analyzing data. Mass spectrometers equipped with robotic machinery can prepare, process, and analyze a vast number of samples, enabling the development of high throughput approaches to proteomic analysis. These advances, combined with current experimental techniques including cell cycle synchrony, in vivo cross-linking, and immunoprecipitation arrays will allow for study of large complexes and protein networks.

The benefit of this combination of techniques is that it will provide for complete protein complex analysis in response to a variety of cellular signals. Protein complex assembly and disassembly can be followed over time. Most interestingly, promiscuous proteins involved in a number of protein complexes can be specifically examined to determine their appropriate complex association under various conditions or times. The results of these studies will produce a complex traffic map of protein interactions during normal cellular function as well as in response to changing cellular conditions. This map can then be exploited in the development of therapeutics designed to disrupt or alter the assembly or disassembly of specific protein complexes in response to specific cellular signals.

ACKNOWLEDGMENT

The authors are grateful to Stephen J. Benkovic at The Pennsylvania State University for seeding our interest and knowledge of this topic. M.A.T. acknowledges the Royal Society for generous funding. We thank Nicole Okeley for critically reading this manuscript.

LITERATURE CITED

(1) Szegedi, S. S., and Garrow, T. A. (2004) Oligomerization is required for betaine-homocysteine S-methyltransferase function. *Arch. Biochem. Biophys. 426*, 32−42.

(2) Bunn, M. W., and Ordal, G. W. (2003) Transmembrane organization of the Bacillus subtilis chemoreceptor McpB deduced by cysteine disulfide crosslinking. *J. Mol. Biol. 331*, 941−9.

(3) Guan, L., Murphy, F. D., and Kaback, H. R. (2002) Surface-exposed positions in the transmembrane helices of the lactose permease of Escherichia coli determined by intermolecular thiol cross-linking. *Proc. Natl. Acad. Sci. U.S.A. 99*, 3475−80.

(4) Ermolova, N., Guan, L., and Kaback, H. R. (2003) Intermolecular thiol cross-linking via loops in the lactose permease of Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A. 100*, 10187−92.

(5) Giron-Monzon, L., Manelyte, L., Ahrends, R., Kirsch, D., Spengler, B., and Friedhoff, P. (2004) Mapping protein−protein interactions between MutL and MutH by cross-linking. *J. Biol. Chem. 279*, 49338−45.

(6) Brown, K. C., Yang, S. H., and Kodadek, T. (1995) Highly specific oxidative cross-linking of proteins mediated by a nickel-peptide complex. *Biochemistry 34*, 4733−9.

(7) Fancy, D. A., and Kodadek, T. (1999) Chemistry for the analysis of protein−protein interactions: rapid and efficient cross-linking triggered by long wavelength light. *Proc. Natl. Acad. Sci. U.S.A. 96*, 6020−4.

(8) Kim, K., Fancy, D. A., Carney, D., and Kodadek, T. (1999) Photoinduced Protein Cross-Linking Mediated by Palladium Porphyrins. *J. Am. Chem. Soc. 121*, 11896−11897.

(9) Amini, F., Denison, C., Lin, H. J., Kuo, L., and Kodadek, T. (2003) Using oxidative crosslinking and proximity labeling to quantitatively characterize protein−protein and protein-Peptide complexes. *Chem. Biol. 10*, 1115−27.

(10) Amini, F., Kodadek, T., and Brown, K. C. (2002) Protein affinity labeling mediated by genetically encoded peptide tags. *Angew. Chem., Int. Ed. 41*, 356−9.

(11) Fancy, D. A. (2000) Elucidation of protein−protein interactions using chemical cross-linking or label transfer techniques. *Curr. Opin. Chem. Biol. 4*, 28−33.

(12) Melcher, K. (2004) New chemical crosslinking methods for the identification of transient protein−protein interactions with multiprotein complexes. *Curr. Protein Pept. Sci. 5*, 287−96.

(13) Tullius, T. D., and Dombroski, B. A. (1986) Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to lambda repressor and Cro protein. *Proc. Natl. Acad. Sci. U.S.A. 83*, 5469−73.

(14) Rana, T. M., and Meares, C. F. (1990) Specific cleavage of a protein by an attached iron chelate. *J. Am. Chem. Soc. 112*, 2457−2458.

(15) Colland, F., Fujita, N., Ishihama, A., and Kolb, A. (2002) The interaction between sigmaS, the stationary phase sigma factor, and the core enzyme of *Escherichia coli* RNA polymerase. *Genes Cells 7*, 233−47.

(16) Chen, H. T., and Hahn, S. (2003) Binding of TFIIB to RNA polymerase II: Mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. *Mol. Cell 12*, 437−47.

(17) Schatz, P. J. (1993) Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in Escherichia coli. *Biotechnology (NY) 11*, 1138−43.

(18) Schwartz, M. A., Das, O. P., and Hynes, R. O. (1982) A new radioactive cross-linking reagent for studying the interactions of proteins. *J. Biol. Chem. 257*, 2343−9.

(19) Denny, J. B., and Blobel, G. (1984) 125I-labeled crosslinking reagent that is hydrophilic, photoactivatable, and cleavable through an azo linkage. *Proc. Natl. Acad. Sci. U.S.A. 81*, 5286−90.

(20) Chen, Y., Ebright, Y. W., and Ebright, R. H. (1994) Identification of the target of a transcription activator protein by protein−protein photocrosslinking. *Science 265*, 90−2.

(21) Hermanson, G. (1996) *Bioconjugate Techniques*, Academic Press, San Diego.

(22) Alley, S. C., Ishmael, F. T., Jones, A. D., and Benkovic, S. J. (2000) Mapping Protein−Protein Interactions in the Bacteriophage T4 DNA Polymerase Holoenzyme Using a Novel Trifunctional Photo-cross-linking and Affinity Reagent. *J. Am. Chem. Soc. 122*, 6126−6127.

(23) Brown, C. E., Howe, L., Sousa, K., Alley, S. C., Carrozza, M. J., Tan, S., and Workman, J. L. (2001) Recruitment of HAT

complexes by direct activator interactions with the ATM-related Tra1 subunit. *Science 292*, 2333−7.
(24) Sadygov, R. G., Cociorva, D., and Yates, J. R. (2004) Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book. *Nature Methods 1*, 195−202.
(25) Bernhard, O. K., Sheil, M. M., and Cunningham, A. L. (2004) Lateral membrane protein associations of CD4 in lymphoid cells detected by cross-linking and mass spectrometry. *Biochemistry 43*, 256−64.
(26) Meunier, L., Usherwood, Y. K., Chung, K. T., and Hendershot, L. M. (2002) A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins. *Mol. Biol. Cell 13*, 4456−69.
(27) Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol. 17*, 994−9.
(28) Whetstone, P. A., Butlin, N. G., Corneillie, T. M., and Meares, C. F. (2004) Element-coded affinity tags for peptides and proteins. *Bioconjugate Chem. 15*, 3−6.
(29) Lu, Y., Bottari, P., Turecek, F., Aebersold, R., and Gelb, M. H. (2004) Absolute quantification of specific proteins in complex mixtures using visible isotope-coded affinity tags. *Anal.Chem 76*, 4104−11.
(30) Trester-Zedlitz, M., Kamada, K., Burley, S. K., Fenyo, D., Chait, B. T., and Muir, T. W. (2003) A modular cross-linking approach for exploring protein interactions. *J. Am. Chem. Soc. 125*, 2416−25.
(31) Kim, J. S., and Raines, R. T. (1995) Dibromobimane as a fluorescent crosslinking reagent. *Anal. Biochem 225*, 174−6.
(32) Konno, K., Ue, K., Khoroshev, M., Martinez, H., Ray, B., and Morales, M. F. (2000) Consequences of placing an intramolecular crosslink in myosin S1. *Proc. Natl. Acad. Sci. U.S.A. 97*, 1461−6.
(33) Wu, J., Voss, J., Hubbell, W. L., and Kaback, H. R. (1996) Site-directed spin labeling and chemical crosslinking demonstrate that helix V is close to helices VII and VIII in the lactose permease of Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A. 93*, 10123−7.
(34) Bhattacharjee, H., and Rosen, B. P. (1996) Spatial proximity of Cys113, Cys172, and Cys422 in the metalloactivation domain of the ArsA ATPase. *J. Biol. Chem. 271*, 24465−70.
(35) Sharma, K. K., Kaur, H., and Kester, K. (1997) Functional elements in molecular chaperone alpha-crystallin: identification of binding sites in alpha B-Crystallin. *Biochem. Biophys. Res. Commun. 239*, 217−22.
(36) Wine, R. N., Dial, J. M., Tomer, K. B., and Borchers, C. H. (2002) Identification of components of protein complexes using a fluorescent photo-cross-linker and mass spectrometry. *Anal. Chem 74*, 1939−45.
(37) Alley, S. C., Trakselis, M. A., Mayer, M. U., Ishmael, F. T., Jones, A. D., and Benkovic, S. J. (2001) Building a replisome solution structure by elucidation of protein−protein interactions in the bacteriophage T4 DNA polymerase holoenzyme. *J. Biol. Chem. 276*, 39340−9.
(38) Hurst, G. B., Lankford, T. K., and Kennel, S. J. (2004) Mass spectrometric detection of affinity purified crosslinked peptides. *J. Am. Soc. Mass. Spectrom. 15*, 832−9.
(39) Heck, A. J., and Van Den Heuvel, R. H. (2004) Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev. 23*, 368−89.
(40) Gorman, J. J., Wallis, T. P., and Pitt, J. J. (2002) Protein disulfide bond determination by mass spectrometry. *Mass Spectrom. Rev. 21*, 183−216.
(41) Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W., and Dollinger, G. (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A. 97*, 5802−6.
(42) Kruppa, G. H., Schoeniger, J., and Young, M. M. (2003) A top down approach to protein structural studies using chemical cross-linking and Fourier transform mass spectrometry. *Rapid. Commun. Mass Spectrom. 17*, 155−62.
(43) Back, J. W., Sanz, M. A., De Jong, L., De Koning, L. J., Nijtmans, L. G., De Koster, C. G., Grivell, L. A., Van Der Spek, H., and Muijsers, A. O. (2002) A structure for the yeast prohibitin complex: Structure prediction and evidence from chemical crosslinking and mass spectrometry. *Protein Sci. 11*, 2471−8.
(44) Chu, F., Shan, S. O., Moustakas, D. T., Alber, F., Egea, P. F., Stroud, R. M., Walter, P., and Burlingame, A. L. (2004) Unraveling the interface of signal recognition particle and its receptor by using chemical cross-linking and tandem mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A. 101*, 16454−9.
(45) Schulz, D. M., Ihling, C., Clore, G. M., and Sinz, A. (2004) Mapping the topology and determination of a low-resolution three-dimensional structure of the calmodulin-melittin complex by chemical cross-linking and high-resolution FTI-CRMS: direct demonstration of multiple binding modes. *Biochemistry 43*, 4703−15.
(46) Link, A. J., Mock, M. L., and Tirrell, D. A. (2003) Non-canonical amino acids in protein engineering. *Curr. Opin. Biotechnol. 14*, 603−9.
(47) Chin, J. W., Martin, A. B., King, D. S., Wang, L., and Schultz, P. G. (2002) Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc. Natl. Acad. Sci. U.S.A. 99*, 11020−4.
(48) Kirshenbaum, K., Carrico, I. S., and Tirrell, D. A. (2002) Biosynthesis of proteins incorporating a versatile set of phenylalanine analogues. *Chembiochem 3*, 235−7.
(49) Kwon, I., Kirshenbaum, K., and Tirrell, D. A. (2003) Breaking the degeneracy of the genetic code. *J. Am. Chem. Soc. 125*, 7512−3.
(50) Chin, J. W., Cropp, T. A., Anderson, J. C., Mukherji, M., Zhang, Z., and Schultz, P. G. (2003) An expanded eukaryotic genetic code. *Science 301*, 964−7.

BC050043A