# 3
# Crosslinking Strategies for the Study of Membrane Protein Complexes and Protein Interaction Interfaces

RASANJALA WEERASEKERA AND GEROLD SCHMITT-ULMS*

*Centre for Research in Neurodegenerative Diseases, University of Toronto, Ontario M5S 3H2, Canada*

## Introduction

Proteins are intimately involved in the ætiology of the vast majority of human diseases. It has long been recognized that individual proteins do not act in isolation but engage in complex and dynamic interactions with other proteins to fulfill their diverse cellular roles (Pandey and Mann, 2000; Aloy and Russell, 2002; Sobott and Robinson, 2002). In recognition of this, researchers in pursuit of novel therapeutic targets and diagnostic markers turn to strategies that provide insights into the molecular environments of established disease targets. Methodologies can be divided loosely into low resolution interactome mapping protocols and higher resolution interface mapping strategies. Low resolution interactome mapping strategies will either reveal the protein constituents of a protein assembly without establishing actual linkages, or provide data that establish direct interactions among proteins. Interface mapping, on the other hand, refers to experiments designed to map regions within a protein that contribute to internal contact sites, or that contribute to the binding to another protein and, therefore, can provide information about the topology of individual proteins or a protein complex.

There is no shortage of strategies for anyone intent on mapping protein interactions. Techniques range from genetic approaches such as the Two-Hybrid System

*To whom correspondence may be addressed (g.schmittulms@utoronto.ca)

Abbreviations: CID, collision-induced dissociation; EGF, epidermal growth factor; ESI, electrospray ionization; FGF, fibroblast growth factor; FRET, fluorescence resonance energy transfer; FT-ICR, Fourier transform ion cyclotron resonance; FTMS, Fourier transform mass spectrometry; MALDI, matrix-assisted laser desorption ionization; MIX, mixed isotope crosslinking; MS/MS, tandem mass spectrometry; PICUP, photo-induced crosslinking of unmodified proteins; PIR, protein interaction reporter; SDS–PAGE, sodium dodecylsulfate polyacrylamide gel electrophoresis; SEC, size-exclusion chromatography; SORI, sustained off-resonance irradiation; tcTPC, time-controlled transcardiac perfusion crosslinking; THS, Two-Hybrid System; TM, transmembrane.

*Biotechnology and Genetic Engineering Reviews* – Vol. 23, December 2006
0264–8725/07/23/41–62 $20.00 + $0.00 

(THS) (Fields and Song, 1989) and fluorescence resonance energy transfer (FRET) and their derivative strategies (Tsien, 1998; Aronheim, 2001; Hu, 2001; Legrain *et al.*, 2001; Stagljar and Fields, 2002; Deyev *et al.*, 2003; Lippincott-Schwartz and Patterson, 2003; Zal and Gascoigne, 2004; Wallrabe and Periasamy, 2005), to biochemical protocols that reveal the composition of a complex by immunoprecipitation or affinity-tag purification and mass spectrometry (Bauer and Kuster, 2003). However, close inspection of current technologies for the study of protein–protein interactions reveals at least two areas of investigation in which either progress has fallen behind or further advancements are imperative. These are the development of tools for the study of protein interactions involving membrane proteins, and the development of methodologies for the elucidation of protein complex topologies and the mapping of protein interaction interfaces. Upon cursory examination, these two areas of investigation may appear somewhat unrelated. However, common factors both in the reasons for these delays and in the strategies that may offer solutions, warrant their joint treatment in this review.

## INTRODUCTION TO INTERACTOME MAPPING STRATEGIES

Traditionally, protein interactions have been investigated by exploiting the unique physico-chemical properties of individual protein complexes. In such studies, a target protein would be subject to a complex biochemical purification scheme that often relied on cell fractionation and a series of conventional chromatography steps. If carried out under 'mild' conditions, this approach may lead to the co-purification of physiological interactors, and may also be applicable to membrane protein complexes. However, as this approach capitalizes on the unique properties of each complex, no purification scheme is the same, and the constant challenge remains the empirical adjustment of purification schemes to new targets. Another shortcoming of this strategy is its reliance on a few highly selective chromatography steps that afford sufficient enrichment of the target complex. An attractive remedy to this problem employs affinity purification steps that rely on the immobilization of either target-specific antibodies or high-affinity biological ligands. As antibodies can be generated against almost any protein target, this strategy virtually replaced the need for complex purification schemes. However, antibodies are costly and, if commercially unavailable, require significant efforts to generate. Therefore, for high-throughput interactome mapping, a strategy that relies on target proteins engineered with a tagged sequence might be more feasible. Several recent review articles provide detailed discussions of current methodologies for the low resolution mapping of protein interactions using the above-mentioned technique (Figeys *et al.*, 2001; Drewes and Bouwmeester, 2003; Warnock *et al.*, 2004; Piehler, 2005).

All of the above strategies suffer from the need to conduct purifications under conditions that preserve the physiological interactions of a given protein target. Typically, such mild purification conditions lead to the concomitant purification of a sizeable number of unspecific binders, predominantly proteins that are either highly abundant or known for their promiscuous binding to a large range of protein targets. The situation is worsened if a target protein complex is embedded in cellular membranes. In those applications, the need to solubilize the complex prior to affinity purification steps forces the use of detergents. It is, however, largely unpredict-

able whether a target protein complex will retain its integrity in the presence of a given detergent.

In this article, protein interaction mapping strategies will be presented only in the context of their strengths and weaknesses for the investigation of protein–protein interactions involving membrane proteins. Emphasis will be placed on manuscripts that report the use of chemical crosslinking for this purpose.

## INTRODUCTION TO INTERFACE MAPPING STRATEGIES

Apart from a few well-characterized protein assemblies, for which high resolution structural data are available (Sali *et al.*, 2003; Russell *et al.*, 2004), little is known about the topology of multi-constituent protein complexes. While progress with high resolution structural mapping technologies is continuously being made, it is likely that X-ray- or nuclear magnetic resonance (NMR)-based strategies will not provide routine access to interface data of multi-constituent membrane protein complexes or transient complexes for some time to come. THS, FRET, and its derivatives, are not only useful for the identification of protein interactors but also are generally considered powerful genetic tools for higher resolution interface mapping. However, a downside of any genetic approach for interface mapping is the hypothesis-driven nature of the experiment and the limited power it offers for the dissection of complex or non-linear interfaces. To minimize conceptual bias in such a study, the investigator is left with the choice of either generating a large number of expression constructs, or of narrowing down interfaces through repetitive cycles of expression construct cloning and testing. The same limitation also applies to biochemical surface plasmon resonance (SPR)-based interface mapping strategies that require the recombinant expression of a series of deletion constructs for individual protein complex constituents. An advantage of the SPR approach is that it not only reveals qualitative binding information, but also generates data that allow calculation of affinity constants (Nedelkov and Nelson, 2003; Buijs and Franklin, 2005).

Occasionally, valuable topology information can be gained by probing the solvent accessibility and hydrogen bonding characteristics of amino acids that are embedded in a protein complex. A well-suited methodology for this application is deuterium–hydrogen (D/H) exchange, followed by mass spectrometry (Eyles and Kaltashov, 2004; Busenlehner and Armstrong, 2005). Unfortunately, this approach provides limited information about actual interfaces and is poorly suited for linkage analysis of proteins within a complex.

Recently, the computational modelling of protein–protein interfaces and the use of sophisticated algorithms for protein structure prediction has received considerable attention. Some very good review articles are available on this topic (DeLano, 2002; Smith and Sternberg, 2002; Russell *et al.*, 2004; Aloy *et al.*, 2005).

## WHAT DO MEMBRANE PROTEIN COMPLEXES AND PROTEIN INTERACTION INTERFACES HAVE IN COMMON?

With the advent of ever more powerful mass spectrometry tools, crosslinking approaches may be well suited to provide both a solution to shortcomings of current methodologies for the study of membrane protein interactions and a way to overcome

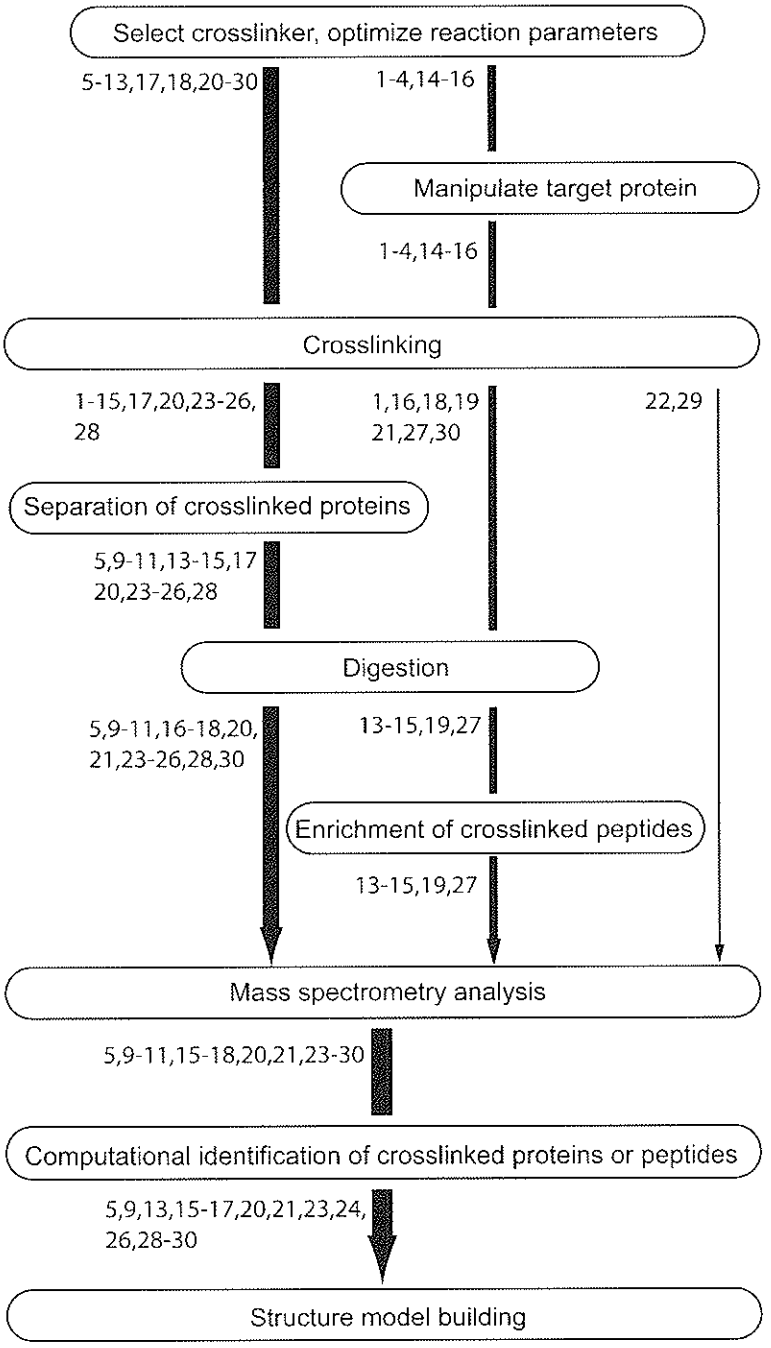


**Figure 3.1.** Flow-chart summarizing key elements of crosslinking strategies. The varying thickness of arrows that connect the steps indicates the popularity of a design feature in the combined literature, i.e. thicker arrows indicate more frequent use. Numbers listed alongside arrows point to publications listed in *Tables 3.1* and *3.2*.

the difficulty of providing low resolution topology data for protein complexes – a view taken by many that has spurred a renaissance of crosslinking methodologies for investigations into the interactome of membrane proteins and interface mapping applications in recent years. Relatively few chemical crosslinking studies have been reported so far with the aim to map protein–protein interactions involving membrane proteins. The key dilemma with membrane proteins is the need to either apply crosslinks *in vivo* prior to the solubilization of proteins or to find detergent conditions that do not disrupt the very protein interactions one intends to study. The latter requirement frequently results in a catch-22 situation as, without prior information on physiological protein interactors, it is impossible to test the suitability of a given detergent. A very similar situation is observed with protein interfaces where, without prior crosslinking, any manipulation aimed at studying the interfaces within a protein complex may cause dissociation or rearrangement of complex constituents.

Crosslinking approaches applied to the study of protein–protein interactions can be divided broadly into target-specific crosslinking and generic crosslinking strategies. Once crosslinked, complexes can be subject to high-stringency washing steps and, as a result, downstream procedures become insensitive to the use of detergents that may be needed for solubilization. The current literature is dominated by *in vitro* applications that report the characterization of interfaces of highly purified material with relatively large quantities of starting material and rarely more than two proteins as complex constituents. In addition, most current strategies aim to directly identify the crosslinked peptides, with or without an enrichment step, and therefore rely on non-trivial adjustments of algorithms for data mining. *Figure 3.1* outlines the main steps that are commonly followed in crosslinking strategies. The varying thickness of arrows that connect the steps indicates the popularity of a design feature in the combined literature, i.e. thicker arrows indicate more frequent use.

As could be expected, no 'one-beats-all' protocol is currently available. We will indicate the shortcomings of current protocols and will point towards directions for future improvements. Please note that reference numbers listed alongside arrows in *Figure 3.1* point to publications that are also listed in *Tables 3.1* and *3.2*. In addition, *Tables 3.1* and *3.2* list experimental details of the strategies reviewed here that should help the reader navigate within the body of literature on this topic. Clearly, the list of studies is not meant to be exhaustive. Instead, manuscripts have been included to outline concepts or reflect the uniqueness of an approach. For a more comprehensive review on selected strategies or more detailed discussions of commonly used chemical crosslinkers (Wong and Wong, 1992; Bennett *et al.*, 2000; Mueller *et al.*, 2001), general strategies of protein crosslinking experiments (Back *et al.*, 2003; Sinz, 2003; Melcher, 2004; Trakselis *et al.*, 2005), and different MS analysis methods (Jonscher and Yates, 1997; Burlingame *et al.*, 1998; Farmer and Caprioli, 1998; Marshall *et al.*, 1998; Chernushevich *et al.*, 2001; Mo and Karger, 2002; Aebersold and Mann, 2003), the reader is directed to selected literature in the text. We apologize to the many researchers who have contributed to this very exciting research field but whose work is not referenced in this review.

**Table 3.1.** Interactome mapping studies

| No | Protein(s) | [kDa] | Crosslinker | Quantity | Cleanup | Strategy for identification | Reference |
|---|---|---|---|---|---|---|---|
| 1 | RNAP α subunit + IAC(161)-CAP | 60 | Photoactivable crosslinking agent | ? | SDS–PAGE | 1D-SDS shift | Chen *et al.*, 1994 |
| 2 | His6-GST dimers | 55 | His6-Ni-mediated (Ni(Oac)2 + MMPP) | 0.1 nmol | SDS–PAGE | 1D-SDS shift | Fancy *et al.*, 1996 |
| 3 | His6-GST-IPCS dimer | ? | His6-Ni-mediated (Ni(Oac)2 + MMPP), DTSSP | 0.2 nmol | HPLC, SDS–PAGE | Formation of dityrosine, 1D-SDS shift | Fancy and Kodadek, 1998 |
| 4 | Grb2 + EGF receptor | 90 | *p*-benzoyl-L-phenyl-alanine + light | Endogenous | SDS–PAGE | 1D-SDS shift | Hino *et al.*, 2005 |
| 5 | Nup84p | >350 | SMCC, BS3 | 1 pmol | SDS–PAGE | 1D-SDS shift, MALDI MS, Western blot | Rappsilber *et al.*, 2000 |
| 6 | UvsY | ~111 | Ru(II)bpy3(2+) + APS + light | 0.4 nmol | SDS–PAGE | 1D-SDS shift, Coomassie stain | Fancy and Kodadek, 1999 |
| 7 | HA-PGRMC1 + Insig-1-myc | 110 | [1-14C] Photo-Met, Photo-Leu + UV | Transfected | SDS–PAGE | 1D-SDS shift, Western blot | Suchanek *et al.*, 2005 |
| 8 | Synaptophysin, VAMPII, SNAP-25 | 205 | Formaldehyde | Endogenous | SDS–PAGE | 1D-SDS shift, Western blot | Hannah *et al.*, 1998 |
| 9 | PrPc | 225 | Formaldehyde | Endogenous | SDS–PAGE, HPLC | 1D-SDS shift, Western blot, LC-MS/MS | Schmitt-Ulms *et al.*, 2001 |
| 10 | Myc-Ras | 250 | Formaldehyde | Endogenous | Immunoaffinity purification, SDS–PAGE | 1D-SDS shift, Western blot, LC-MS/MS | Vasilescu *et al.*, 2004 |
| 11 | PrPc | 225 | Formaldehyde (tcTPC) | Endogenous | Immunoaffinity purification | ESI MS/MS | Schmitt-Ulms *et al.*, 2004 |

**Table 3.2.** Interface mapping studies

| No | Protein(s) | [kDa] | Crosslinker | Quantity | Cleanup | Mining | Inter | Intra | Reference |
|---|---|---|---|---|---|---|---|---|---|
| 12 | Na.K-ATPase α/β + γ | 80 | NHS-ASA, DST, EDC | ? (0.5 mg/ml) | SDS–PAGE | | 5 | | Fuzesi *et al.*, 2005 |
| 13 | Calmodulin + Melittin | 20 | EDAC | 15 nmol | RP HPLC | | 2 | | Scaloni *et al.*, 1998 |
| 14 | Rhodopsin + Transducin | 46 | SPDP, PEAS | 20–45 nmol | Avidin agarose | | 1 | | Itoh *et al.*, 2001 |
| 15 | Nebulin + Calmodulin | 31 | DB, EDAC | 24 nmol | SDS–PAGE, HPLC | NIH-XL | 6 | 10 | Sinz and Wang, 2001 |
| 16 | IL-6 dimer | 40 | BS3 | 10 nmol | | X-Link | 3 | 3 | Taverner *et al.*, 2002 |
| 17 | FGF-2 | 17 | BS3 | 100 µg | SEC | ASAP | | 15 | Young *et al.*, 2000 |
| 18 | CD28-IgG + CD80-Fab | 200 | DTSSP | 15 pmol | SDS–PAGE | GPMAW | 18 | | Bennett *et al.*, 2000 |
| 19 | Poly-ubiquitin | ? | None | 150 µg | Phenyl superose | | 1 | | Chen *et al.*, 1999 |
| 20 | Stathmin/OP18 + Tubulin | 90 | d0/d4-BSP, d0/d4-BSS | 3 µmol | SDS–PAGE | | 1 | | Mueller *et al.*, 2001 |
| 21 | Cytochrome *c* | ? | d0/d8-DSA, BS3 | 0.5 nmol | Dialysis | PAWS | | 5 | Pearson *et al.*, 2002 |
| 22 | Tyrlysbradykinin peptide | ? | BID | ? (6 mM) | ZipTip | | | 1 | Back *et al.*, 2001 |
| 23 | PHB1 + PHB2 | 66 | DTSSP, sBID | ? (0.5 mg/ml) | BNE | FindLink | 6 | 0 | Back *et al.*, 2002a |
| 24 | RNAse S | 14 | PIR | ? (10 µM) | SDS–PAGE | GPMAW | 1 | | Tang *et al.*, 2005 |
| 25 | Rad18 + Rad6-His | ? | sBID, BS3 | ? (0.5 mg/ml) | SDS–PAGE | FindLink | 8 | 11 | Back *et al.*, 2002b |
| 26 | BSA | 66 | DSS, DSG, BS3 | ? (10 mM) | SDS–PAGE | PAWS | | 19 | Huang *et al.*, 2004 |
| 27 | NC2 (α.β subunits) | 20 | 2 Modular crosslinkers | ? (49 µM) | Microconc. / Biotin AC | PeptideMap | 2 | 5 | Trester-Zedlitz *et al.*, 2003 |
| 28 | Calmodulin + M13 | 20 | Sulfo-NHS, DSA-d0/d8 | 10 nmol | SDS–PAGE | GPMAW, ASAP | 10 | 4 | Kalkhof *et al.*, 2005 |
| 29 | Ubiquitin | 8.5 | DSS | ? (1 mg/ml) | 1-step protein trap | MS2PRO | | 2 | Kruppa *et al.*, 2003 |
| 30 | Ffh + FtsY | 110 | DSS, BSSA | 0.35 nmol | SDS–PAGE | MS-BRIDGE | 9 | 4 | Chu *et al.*, 2004 |

## Interactome mapping by crosslinking

Protein crosslinking involves the introduction of a covalent bond between amino acid residues of two spatially proximal peptide strands. Commercial crosslinking reagents are classified as homo- or hetero-bifunctional crosslinkers, dependent on whether the crosslinker is equipped with identical or different reactive groups at its ends. Other important selection criteria for crosslinking reagents are the tether length, hydrophobicity, and whether the crosslinking chemistry can be reversed. For the study of membrane protein interactions, it is important to consider the membrane permeability of the crosslinking reagent. While there is an abundance of crosslinkers, the majority of these reagents will target one out of three functional groups present on proteins, i.e. amino groups, acidic side chains, or sulfhydryl groups. Overall, the majority of protein crosslinking work to date has been carried out with homo-bifunctional amino group specific crosslinkers. Chemistries employed frequently are rather unspecific. As a result, an amino group specific crosslinking reagent may, in addition to promoting crosslinking to lysines, also generate crosslink bonds with arginine, histidine, tyrosine, tryptophan, glutamine, or asparagine residues (Metz *et al.*, 2004).

### DIRECTED STRATEGIES

No large-scale crosslinking-based protein–protein interaction study has been reported in the literature to date. Two fundamentally different concepts dominate the crosslinking literature. The first, which we will refer to as the 'Directed strategy', exploits unique features of the target for either crosslinking, proteolysis, or detection of crosslinked peptides. The 'Generic strategy', in contrast, is based on crosslinking reagents that generically crosslink all proteins to their respective next neighbours, and does not rely on target-specific characteristics for downstream sample processing.

The directed crosslinking approach may offer greater control over the timing of the crosslinking reaction and may, in some implementations, provide access to a more aggressive chemistry for the capture of transient protein interactions. Various strategies for target-selective crosslinking have been proposed. Label transfer crosslinking, one of the better-known strategies, makes use of a hetero-bifunctional crosslinking reagent consisting of an electrophile that is connected through a cleavable linker to a radioactively labelled photo-activatable group. Following attachment of the reagent to a purified target protein through the electrophile, the derivatized protein is re-introduced into its natural protein environment. The protein is then crosslinked to its next neighbour(s) through a short light stimulus that triggers the attachment chemistry of the photo-activatable group. Upon opening of the cleavable linker, which commonly consists of a disulfide or an ester bond, the radioactive label is transferred to the next neighbour molecule, and assists in tracing protein identification steps (reviewed by Marriott and Ottl, 1998; Fancy, 2000). Should the primary structure of a protein target not provide adequate attachment sites, label transfer can be preceded by site-directed genetic manipulation to generate the required functional group (Chen *et al.*, 1994). A similar site-directed implementation of a directed crosslinking strategy requires that target proteins are

engineered with a Ni(II) peptide chelate, such as a 6-histidine tag (Fancy *et al.*, 1996). In the presence of an activating oxidant, a chemical crosslink is formed in a reaction that appears to involve a nearby electron donating tyrosine residue (Fancy and Kodadek, 1998).

Recently, two variations of a theme that take site-specific photo-crosslinking to the next level have been reported. Both strategies employ photo-activatable amino acid derivatives for crosslinking. For a cell to incorporate these altered amino acids into proteins during protein expression, the genetic code has to be expanded through modifications in tRNAs and tRNA synthetase, and the target codon has to be inserted into the protein of interest (Farrell *et al.*, 2005; Hino *et al.*, 2005). In one implementation of this strategy, an amber codon was inserted into the coding sequence of the adaptor protein Grb2 in the vicinity of its known epidermal growth factor (EGF)-receptor binding site. As a result of the expanded genetic code, and the presence of the amber codon, the addition of the photo-reactive amino acid *p*-benzoyl-L-phenylalanine to the medium caused its incorporation into Grb2 in the specified location. Sequential exposure of cells to an EGF stimulus and light resulted in the crosslinking of Grb2 to the EGF receptor (Hino *et al.*, 2005).

As only limited data have been reported, a full evaluation of this approach is not possible at this time. An obvious caveat of this strategy is its reliance on complex manipulations of the genetic code. However, it is conceivable that this approach, or a derivative thereof, may find application in the study of membrane protein interactions as these tools are made available to the wider research community.

## GENERIC STRATEGIES

The downside of any directed crosslinking strategy is the need to manipulate a target protein before the crosslinking reaction. Besides the extra effort involved in target preparation, such a step harbours well-established risks inherent in altering the expression level of a target, and may possibly generate an artificial biological molecule that no longer reacts with its physiological interactors, or is aberrantly processed within the cell.

One of the first publications to report on a generic strategy that combined chemical crosslinking with downstream mass spectrometry employed a series of homo-bifunctional crosslinkers to probe protein linkages within a six-member subcomplex of the yeast nuclear pore complex. Reaction products were subject to SDS–PAGE analysis and, following in-gel trypsinization of silver-stained crosslink bands, proteins were identified by peptide mass fingerprinting (Rappsilber *et al.*, 2000). While this study did not target membrane proteins, it provided a global strategy that can also be applied to membrane proteins. Fancy and Kodadek (1999) developed a metal-catalysed crosslinking strategy that enables photo-induced crosslinking of unmodified proteins (PICUP). In this strategy, a Ru(ii)-based complex initiates a rapid crosslink reaction among nearby proteins when subject to photo-activation in the presence of ammonium persulfate. While many alternative photo-crosslinking strategies suffer from poor yields, the authors reported surprisingly high crosslinking yields of up to 75% in a very fast reaction (<1 s) for PICUP. Unfortunately, so far, this approach has only worked *in vitro* and therefore does not yet offer a tool for the study of transient *in vivo* interactions.

Very recently, the global *in vivo* incorporation of photo-activatable amino acid derivatives was reported. As opposed to the target-specific amino acid derivative strategy outlined above, which required expansion of the genetic code for the incorporation of modified amino acids, this global strategy employs amino acids that closely resemble natural amino acids and thereby evade the cellular identity control mechanisms (Suchanek *et al.*, 2005). As a result, photo-activatable amino acids are incorporated into proteins during the translation process in the same fashion as their natural amino acid counterparts. So far, this approach has worked successfully only for the insertion of photo-leucine and photo-methionine. For the study of membrane protein interactions, the utilization of photo-leucine as a leucine mimic is of particular interest as, due to its hydrophobic nature, leucine is highly represented in the membrane-spanning regions of transmembrane proteins. The authors report a 99% activation rate of the photo-activatable group following a 3 min UV exposure. A natural concern with a strategy that relies on global metabolic incorporation of amino acid derivatives is the cellular stresses that the presence of large amounts of chemically modified proteins may cause. Interestingly, no gross manifestations of toxicity were observed at levels of 0.7% photo-methionine incorporation in this study.

Mild formaldehyde crosslinking has been employed extensively for the study of nucleosomal protein interactions (Fragoso and Hager, 1997; Orlando *et al.*, 1997; Jackson, 1999; Wells and Franham, 2002) but, to date, has found limited use as a tool for the study of membrane protein interactions (Hannah *et al.*, 1998). Features that make formaldehyde crosslinking attractive are: (1) the water solubility of the reagent; (2) the absence of reagent-induced rearrangements of proteins; and (3) the short (2–3 Å) crosslink bonds that endure harsh, non-physiological treatments and are reversible (Jackson, 1999). A few groups have reported on the use of a strategy that employs mild formaldehyde crosslinking followed by SDS–PAGE analysis of crosslink products and tandem mass spectrometry to identify interactors of target proteins (Schmitt-Ulms *et al.*, 2001; Vasilescu *et al.*, 2004). A caveat in these studies has been the use of one-dimensional SDS–PAGE to resolve crosslinked proteins from their uncrosslinked counterparts. Frequently, crosslinked proteins do not resolve well on these gels, and with an upper molecular weight working limit of 500 kDa for SDS–PAGE analysis, many large protein complexes cannot be analysed in this manner.

A novel strategy covalently preserves protein interactions through time-controlled transcardiac perfusion crosslinking (tcTPC) prior to the disruption of tissue integrity. tcTPC combines transcardiac perfusion and mild formaldehyde crosslinking for the study of protein interactions in complex tissues. Crosslinked complexes are immunoaffinity purified, in-solution trypsinized, and finally subject to tandem mass spectrometry analysis to reveal *in vivo* interactors of selected target proteins. This method was validated with the γ-secretase enzyme complex as a target. Subsequently, a protein interaction map of the prion protein (PrP) was generated (Schmitt-Ulms *et al.*, 2004). The strategy enabled identification of more than 20 proteins residing in the vicinity of PrP, suggesting that PrP is embedded in specialized membrane regions with a subset of molecules that, like PrP, use a glycosylphosphatidyl inositol anchor for membrane attachment. Many of these proteins have been implicated in cell adhesion/neuritic outgrowth and harbour immunoglobulin C2 and fibronectin type III-like motifs (*Figure 3.2*). So far, tcTPC

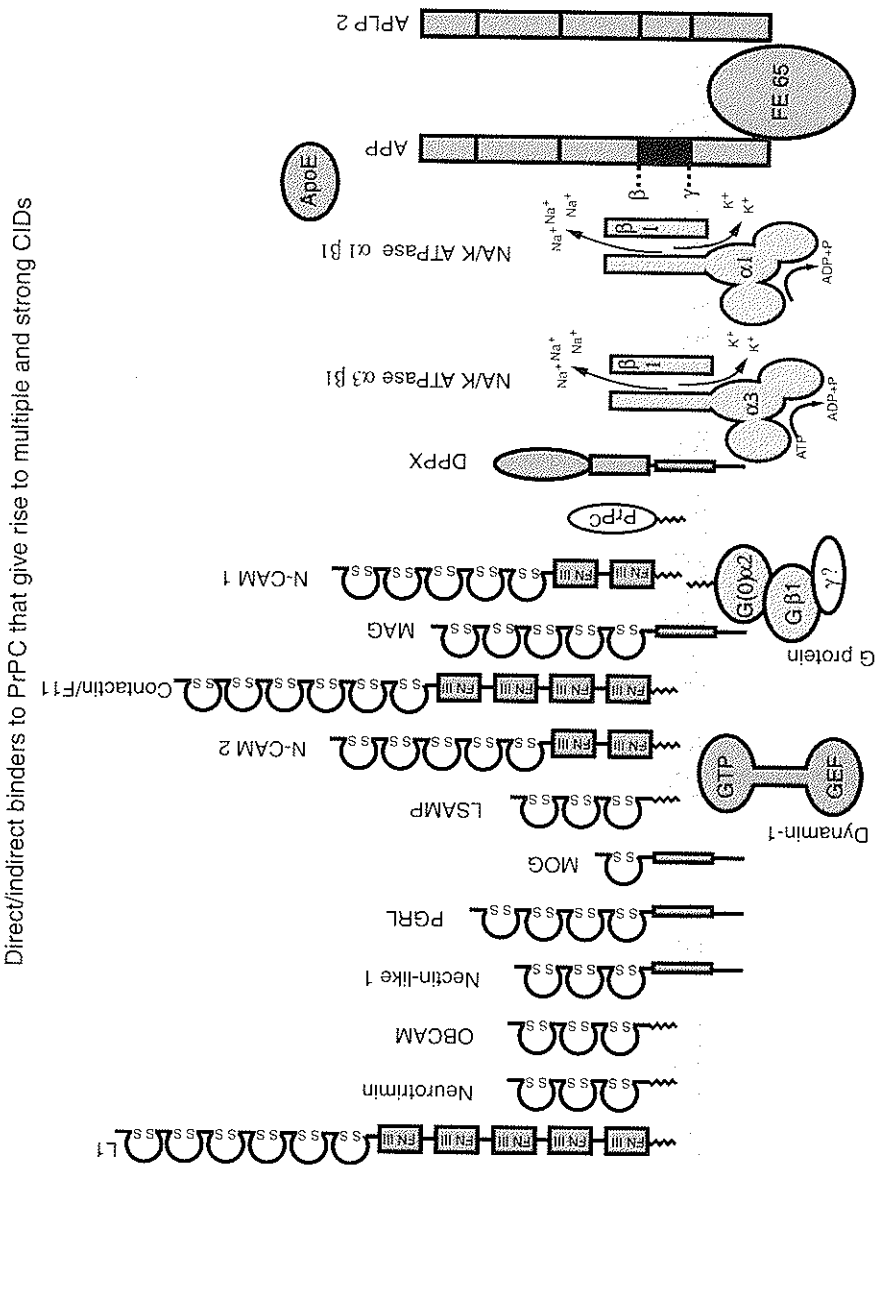


**Figure 3.2.** Cartoon representation of the interactome of the cellular prion protein (PrP) as determined by time-controlled transcardiac perfusion crosslinking. Spatial arrangement of proteins reflects their relative abundance in PrP interactome data set, i.e. proteins inserted closer to cellular PrP (PrPC) were represented with higher sequence coverage in tandem MS data set.

is the only crosslinking technique that crosslinks proteins in their physiologic milieu in a complex tissue. Thus, it can capture membrane protein interactions in-*trans* with proteins on neighbouring cells that frequently belong to a different cell type. A shortcoming of this strategy is that it does not provide information on whether any two proteins engage in direct interactions or are indirectly crosslinked. It also remains to be established whether this strategy can capture the more dynamic interactions such as those found in short-lived enzyme–substrate complexes.

## Interface mapping by crosslinking

The advent of soft ionization strategies for mass spectrometry and the increasing availability of genomic sequence depositories in the late 1980s and 1990s, respectively, set the stage for the rapid pace with which novel interface mapping strategies have become available within the past five to ten years. To date, however, no generic crosslinking-based strategy has been implemented that allows sensitive interface mapping of protein interactions *in vivo*. Major obstacles towards this goal are frequently low crosslinking efficiency, unsatisfactory ionization characteristics of the relatively large crosslinked peptides, and the high complexity of peptide mixtures derived from crosslinking experiments. Such a mixture may be populated by uncrosslinked peptides, incomplete digestion products, surface labelled peptides, crosslinked peptides, and various combinations of the above. For these reasons, identifying crosslinked peptides has been compared to the search for a needle in a haystack (Sinz, 2003), and frequently requires complex algorithms that rely on insights into the reactivity of a crosslinker to predict peptide derivative masses arising from all possible target/crosslinker combinations.

For the purpose of clarity, we have again sorted manuscripts reviewed in this section into more directed and generic strategies that can be applied to many different target proteins. While some studies reviewed here were aimed at the identification of intramolecular crosslinks, we have not explicitly distinguished these, as their general concepts are also applicable to the analysis of intermolecular crosslinks.

### DIRECTED STRATEGIES

Using crosslinking to obtain insights into protein topologies is not a new idea. In the past, this approach was applied mainly as an auxiliary tool by structural biologists in order to provide initial glimpses into the structural organization of protein complexes. Thus, in the early days of protein interface mapping research, the aim was rarely to generate comprehensive topology data. Studies were based on a combination of crosslinking, limited proteolysis, and Edman sequencing or Western blotting analyses. With the advent of mass spectrometry, these first-generation strategies have been replaced largely by sophisticated methods that provide more comprehensive topology data. However, even in the recent literature, one can find occasional interface mapping reports that avoid the use of mass spectrometry technology. In one such study, Fuzesi and colleagues mapped interactions between $\alpha$, $\beta$, and $\gamma$ subunits of Na,K-ATPase using a combination of crosslinking, partial proteolysis, and Western blotting (Fuzesi *et al.*, 2005). A target-specific aspect of this study was

the use of Fe-catalysed oxidative cleavages that generated few and large fragments, which thereby aided the low resolution mapping of crosslinks. However, the requirement for a panel of antibodies that recognizes various epitopes within the target complex constituents reduces the potential use of this strategy to the small number of very well-characterized protein complexes for which more than a few antibodies – recognizing non-overlapping epitopes – may be available.

Many early studies that based topology mapping on mass spectrometry data retained target-specificity as a conceptual feature. An extreme example for a directed strategy is provided by the work of Scaloni and colleagues on the mellitin–calmodulin interaction (Scaloni *et al.*, 1998). In this work, the crosslinking and cleavage reagents chosen capitalized on the unusual absence of both acidic residues and methionines in melittin. In addition, the strategy for the detection of crosslinked fragments was only possible due to the presence of a single tryptophan within the complex, which could be traced by its absorbance at 280 nm. While this approach worked in this particular instance, it required prior knowledge of the primary structure of the complex constituents and relied heavily on unique biochemical features that will rarely be fulfilled by other complexes one may wish to study.

If no such unique biochemical features within target complex constituents are readily available, site-directed mutagenesis offers a means through which target molecules can artificially be rendered compatible to meet the needs of a topology mapping approach. A study that exemplifies this design concept investigated the interface of the rhodopsin–transducin complex (Itoh *et al.*, 2001). Following the introduction of cysteine residues into selected sites within the primary structure of rhodopsin, the complex interface could be mapped with the use of thiol-specific crosslinkers. Shortcomings of this approach are its hypothesis-driven nature and the need for a functional assay that enables the investigator to assess whether the introduction of a mutation interferes with the functional assembly of the target complex. A more sophisticated strategy does not require the expression of mutated complex constituents but, instead, renders a target complex susceptible to crosslinking through chemical derivatization of amino acid side chains. In an investigation of the topology of the complex formed by calmodulin and ND66, a C-terminal fragment of nebulin, lysine side chains were thiolated in the presence of 2-iminothiolane. This generated modified functional groups that could productively react with dibromobimane, a fluorescent thiol-specific crosslinking reagent (Sinz and Wang, 2001). The fluorescent characteristics of this reagent have facilitated tracing of labelled peptides during downstream HPLC enrichment steps, but did not provide a means to distinguish between *bona fide* crosslinked and merely end-labelled products.

Besides a need to distinguish crosslinked, derivatized, and uncrosslinked products, any crosslinking strategy will need to address the possibility that crosslinks can occur both intra- and inter-molecularly. Intermolecular crosslinking is immediately recognizable in instances where crosslinked peptides belong to non-identical molecules. However, there are various scenarios in which this distinction is not straightforward and needs to be resolved experimentally. The most obvious case is exemplified by a crosslinking study of a homo-dimeric or homo-oligomeric protein complex. This problem may also occur, though, when chemical crosslinking reactions are performed at high protein concentrations that promote artificial

crosslinking between identical non-interacting molecules. Finally, this issue may become important in studies of hetero-oligomeric protein complexes of unknown topology in which individual constituents are present at ratios that deviate from the simplest 1:1 stochiometry. Most frequently, the solution to this problem in the literature has been to follow crosslinking with a size fractionation step that will separate monomeric, intra-molecularly crosslinked proteins from the larger intermolecularly crosslinked protein complexes. The mixed isotope crosslinking (MIX) method developed by Taverner and colleagues may represent a more elegant solution to this problem (Taverner *et al.*, 2002). $^{15}$N-labelled and unlabelled ($^{14}$N) interleukin-6 (IL-6) monomers were generated and mixed 1:1 to produce $^{14}$N-, mixed $^{14}$N/$^{15}$N-, and $^{15}$N-labelled IL-6 homodimers prior to crosslinking. Intramolecular crosslinks form doublet peaks in this method, as crosslinked peptides contain either exclusively $^{14}$N or $^{15}$N. Intercrosslinked peptides, on the other hand, will give rise to additional peaks, as the crosslink may bridge peptides that originate from both $^{14}$N and $^{15}$N parent molecules.

## GENERIC STRATEGIES

In the conceptually simplest implementation of a generic interface mapping protocol, a purified protein is first crosslinked with a homo-bifunctional crosslinking reagent, then digested, and peptides are finally analysed by peptide mass fingerprinting in a single-stage mass spectrometry experiment. The presence of masses matching calculated masses of crosslinked peptides provides structural constraints, which can aid the generation of a simple topology model and may be utilized to restrict a given protein to possible fold families. For instance, in an investigation of the fibroblast growth factor (FGF)-2 structure, the amine-specific homo-bifunctional crosslinker $BS^3$ was utilized to form intramolecular crosslinks between lysine residues. Crosslinked monomers were separated from dimers by size-exclusion chromatography (SEC), and tryptic mass maps were recorded by peptide mass fingerprinting (Young *et al.*, 2000).

The simplicity of this strategy is compelling. However, this approach is limited to highly purified proteins of small size as, with an increase in the complexity of the mass spectrum, it becomes difficult to identify informative crosslinked peptides amongst the large number of uncrosslinked or derivatized peptides, and high confidence mass assignments to crosslinked peptides will no longer be possible. Various remedies to this problem have been proposed. A well-established approach relies on the use of cleavable crosslinkers. In one such study, a thiol-cleavable crosslinker was utilized to determine intermolecular crosslinks between monomers of the ParR dimer, as well as the glycoprotein fusion constructs, CD28–IgG and CD80–Fab. Differential peptide mass fingerprinting prior to and following reduction facilitated the assignment of putative crosslinks in this study (Bennett *et al.*, 2000). While providing an improvement, this protocol still could not distinguish between *bona fide* crosslinked and merely derivatized peptides in the mix. This approach is further compromised by its reliance on a highly efficient reduction step required for the disappearance of peaks during the MS analysis. A solution to this caveat may be provided by alternative protocols that base identification of crosslinks on the appearance, rather than the disappearance, of mass peaks, such as the strategy

presented in a study that explored the topology of the protein ubiquitin. Following crosslinking, ε-amino groups of lysines in ubiquitin were reductively methylated and thereby prevented from reacting with 2,4-dinitrofluorobenzene (DNFB). Ubiquitin was then partially hydrolysed with pronase and the newly formed α-amino termini were derivatized with a 1:1 mix of d0/d3-DNFB. As crosslinked peptides contain two N-termini following hydrolysis, they can incorporate two labels and give rise to a 1:2:1 isotope pattern that is clearly distinguishable from the 1:1 pattern of uncrosslinked peptides during MS analysis (Chen *et al.*, 1999). However, a shortcoming of this protocol is its reliance on efficient methylation and N-terminal labelling reactions. A more sophisticated strategy, therefore, may not rely on the labelling of all peptides, but instead on the attachment of the isotopic label to the crosslinker itself, thereby providing the opportunity for purification of isotopically labelled crosslinkers prior to the reaction with the target protein. How this can be put into practice was explored in a study that employed homo-bifunctional deuterium-labelled crosslinkers with reactivity towards primary amino groups. A 1:1 mixture of d0/d4 isotopomers of the crosslinking reagent was used to crosslink Op18 with tubulin (Mueller *et al.*, 2001). In the derived mass spectra, crosslinked peptides were identifiable as 4 amu spaced doublets of equal intensity.

The above strategy works well in MALDI applications that predominantly give rise to singly charged ions, but makes for a difficult detection if crosslinked peptides are observed at higher charge states. In recognition of this caveat, Pearson and colleagues used an improved protocol based on d0 and d8 forms of crosslinkers to map the topology of RNAse A and cytochrome *c* (Pearson *et al.*, 2002). Downsides common to any deuteration-based strategies are: (1) a frequently observed subtle change in the chromatographic behaviour of analytes as a consequence of the hydrogen exchange; and (2) the splitting of parent ion signals that may jeopardize MS/MS-based peptide identification and thereby limit the sensitivity of this approach.

An alternative strategy, therefore, bases the detection of crosslinked peptides on signature peaks in MS/MS spectra. In one realization of this design concept, Back and colleagues devised the crosslinker N-benzyliminodiacetoyloxysuccinimid (BID) (Back *et al.*, 2001). Low-energy, collision-induced dissociation (CID) MS/MS was sufficient to release a marker benzyl cation from this crosslinker, a feature that could be exploited for the detection of crosslinker-containing ions through parent ion scanning. An additional benefit of BID-based crosslinking is that the modular design of this crosslinker enables studies that probe the structure of a protein complex with a series of crosslinkers differing only in their tether lengths. Whereas the first generation of BID crosslinkers was not water-soluble, the same group later presented work on the topology of the yeast prohibitin complex that employed improved water-soluble sulfo-BID reagents (Back *et al.*, 2002a). Often, MS/MS mass spectra contained in crosslinking data sets show a high degree of complexity. In an attempt to simplify the interpretation of tandem MS spectra, a novel crosslinker, named protein interaction reporter (PIR), was recently designed (Tang *et al.*, 2005). PIR acts in a similar manner to BID in that PIR fragmentation occurs with activation energy lower than that needed for peptide backbone fragmentation, and therefore gives rise to the release of signature ions that populate tandem MS spectra. In contrast to BID, the PIR crosslinker was equipped with additional weak linkages that break when low energy is deposited into the crosslinked molecule. Consequently,

the initial tandem MS experiment merely generates the signature ion and one ion for each peptide that engaged in a given crosslink. Consecutive third-stage MS analyses at higher energy then fragments these peptides to provide separate mass spectra that contain sequence information required for peptide identifications. The strategy enables distinction of modified, intra-crosslinked, and inter-crosslinked peptides, but will not be suitable for *in vivo* applications, and is reliant on MS instrumentation that can carry out MS/MS/MS experiments, such as ion traps or Fourier transform ion cyclotron resonance (FT-ICR) MS machines. A moiety for the enrichment of crosslinked peptides would be a desirable feature for a second-generation version of this crosslinker.

Instead of labelling peptides during the crosslinking reaction, the labelling step may be moved further downstream, and can occur concomitantly with the cleavage reaction. This strategy exploits the incorporation of water-derived oxygens into the C-terminal carboxy group of peptides during enzyme-mediated hydrolysis of proteins. As a result of the incorporation of two heavy oxygen isotopes, the mass of a peptide digested in $^{18}$O-water will differ by 4 Da from peptides digested in normal water. Thus, a crosslinked peptide that contains two C-termini will shift 8 Da, and can be distinguished from uncrosslinked peptides by this signature mass shift. An advantage seen with using $^{18}$O in contrast to deuterated compounds is that $^{18}$O-labelled analytes will not give rise to altered reverse phase retention times (Back *et al.*, 2002b; Huang *et al.*, 2004).

For large protein assemblies, the complexity of digest samples can represent a formidable challenge. Attempts have been made, therefore, to devise crosslinking strategies that simplify peptide mixtures prior to their introduction into the mass spectrometer. In one such approach, solid-phase chemistry was employed to synthesize modular crosslinkers which, in addition to an isotopic label, are equipped with a biotin affinity handle. The incorporation of the affinity handle greatly simplifies enrichment of crosslinked peptides, and thereby serves as an elegant solution for the replacement of the commonly applied SDS–PAGE step, which frequently is accompanied with smearing of crosslinking products over extended areas of the gel. Limitations of this strategy are the relatively large size of the modular crosslinker and an incompatibility with *in vivo* applications (Trester-Zedlitz *et al.*, 2003).

The development of crosslinking strategies is continuously paralleled by improvements in mass spectrometry instrumentation. In particular, FT-ICR mass spectrometry offers promise for interface mapping applications in the coming years due to the high mass resolution (low ppm to sub-ppm range) and mass accuracy that these instruments afford. In samples of low complexity, the mass defect inherent to a crosslinked peptide may be sufficient to resolve identification ambiguities with these instruments, and thereby replace the need for an external label. This, in turn, would lead to gains in sensitivity, as it would remove the need for larger analyte quantities that the acquisition of tandem MS data demands. So far, the literature reporting on a combination of crosslinking and FT-ICR for topology mapping is still small. Most studies available have not capitalized on the unique strengths of FT-ICR but are conceptually built upon established strategies (Dihazi and Sinz, 2003; Kalkhof *et al.*, 2005).

Another attractive feature of high-resolution FT-ICR technology is the ability to analyse relatively large molecules without prior fragmentation. This approach not

only eliminates the time-consuming clean-up steps and handling-associated losses of digested material, but also reduces the risk of artificially introducing modifications into peptides. In one of the first studies to explore this application, ubiquitin, a relatively small molecule of 8 kDa, was chemically crosslinked, infused into a FT-ICR mass spectrometer, and fragmented by sustained off-resonance irradiation collision-induced dissociation (SORI-CID). This work led to the identification of two internal crosslinks consistent with the known tertiary structure of ubiquitin (Kruppa *et al.*, 2003). Current drawbacks of this strategy include the formidable task of interpreting the very complex whole-protein fragmentation patterns it produces and the inefficient fragmentation of proteins of large size, which results in an overall dissatisfactory sensitivity.

Once crosslinking data for a protein complex become available, the spatial constraints they provide form the basis for the generation of a low resolution topology model. Model building strategies can, at this point, be borrowed from high resolution structure analysis techniques. A key objective during these steps is to derive a model that minimizes the internal energy of the assembled protein complex. Frequently, the high resolution structure of a complex constituent or a homologue thereof may be available and can provide important additional structural clues that provide a starting point for the assembly of the complex. Modern interface prediction algorithms can further facilitate the model building process. A recent manuscript describes the use of a combination of the above strategies to assemble a low resolution structural model for the bacterial signal recognition particle (Ffh) and its receptor (FtsY) from chemical crosslinking data (Chu *et al.*, 2004).

Despite its current limitations, chemical crosslinking offers many attractive features for the topology mapping of protein complexes. Among these, the sensitivity and speed, the ability to capture transient interactions, and the principal compatibility with *in vivo* applications are prominent. A rich repertoire of crosslinking strategies has been devised that forms the basis for explorations into the topology of virtually any protein complex, regardless of its size or protein composition.

## Future challenges

Naturally, the challenges in the field of study reviewed here largely overlap with the general challenges the proteomics research community faces in its quest to fully characterize the protein complement of cells and organisms. As such, key challenges are: (1) the spatial and temporal dynamics of protein assemblies; (2) the more than six orders of magnitude spanning dynamic range of expression levels observed; and (3) the complexity of protein regulation. Current technologies can only map relatively stable interactions, require a large quantity of material, and are largely blind towards the presence of post-translational modifications or splicing variants of individual proteins. Given the overwhelming task ahead, the research reviewed here can only be considered the first steps in a research field that is bound to expand. It is likely that research into protein–protein interactions in the coming years will be dominated by advancements in sample preparation protocols and crosslinking reagents combined with improvements in the sensitivity of mass spectrometry technology. On the sample preparation side, it will be necessary to improve generic strategies that will be applicable to a large number of targets and will overcome

sample loss issues. Crosslinking strategies need to be further refined to distinguish between inter- versus intra-crosslinks versus derivatizations, as well as to identify the subunit stoichiometry of protein complexes. On the mass spectrometry end, the very inefficient introduction of ions into contemporary mass spectrometers, and the relatively high chemical and electronic noise levels can be singled out as areas where improvements are to be expected.

The recent developments in top-down analysis strategies (Kelleher, 2004; Bogdanov and Smith, 2005) offer the exciting possibility that it may, some day, be possible to routinely subject large protein complexes to comprehensive characterizations of post-translational modifications/alternative splicing. Such an implementation will need to await much needed developments in FT-ICR or related mass spectrometry technologies. It will also require improvements to front-end interfaces for MS and fragmentation strategies that provide efficient internal fragmentation of large molecules. With current electron capture dissociation protocols, good sequence coverage of N- or C-terminal regions are routinely achieved, but the internal fragmentation of large proteins is hampered by strong hydrogen bonding forces of desolvated proteins in the gas phase.

Once the next-generation tools become available, the sheer scale of the undertaking will require advancements in sample throughput. Here, proteomics research will have to apply knowledge gained from large-scale DNA sequencing and genome mapping projects. Amongst these, the need for concerted action, consistent data formats, and integrated and dynamic data storage are obvious areas for improvement. Eventually, we will hopefully see the emergence of novel imaging tools that narrow the gap between crosslinking-based, low resolution interactome mapping tools that are the focus of this article, and contemporary high resolution X-ray, electron microscopy, and NMR strategies.

Finally, research into protein–protein interfaces will be stimulated by the rapid advances in the field of computational research into the prediction of protein structures and protein–protein interfaces. However, even if technologies were available today that would provide full characterization of dynamic protein assemblies within an *in vivo* environment, the research community would still face the challenge of translating this knowledge into something useful for society, i.e. diagnostics to detect and therapeutics to manipulate and defeat diseases. Therefore, it will be important that advances in our understanding of protein assemblies are paralleled by equally productive translational and applied research.

## Acknowledgements

Work on this project was funded through grant support from the Canadian Institute for Health Research (MOP-74734) and the Canadian Foundation for Innovation (#10087). R.W. received fellowship support (#20059) from the University of Toronto.

## References

AEBERSOLD, R. AND MANN, M. (2003). Mass spectrometry-based proteomics. *Nature* **422**, 198–207.

ALOY, P. AND RUSSELL, R.B. (2002). The third dimension for protein interactions and complexes. *Trends in Biochemical Sciences* **12**, 633–638.

ALOY, P., PICHAUD, M. AND RUSSELL, R.B. (2005). Protein complexes: structure prediction challenges for the 21st century. *Current Opinion in Structural Biology* **15**, 15–22.

ARONHEIM, A. (2001). Protein recruitment systems for the analysis of protein–protein interactions. *Methods* **24**, 29–34.

BACK, J.W., HARTOG, A.F., DECKER, H.L., MUIJSERS, A.O., KONING, L. AND DE JONG, L. (2001). A new crosslinker for mass spectrometric analysis of the quarternary structure of protein complexes. *American Society for Mass Spectrometry. Journal* **12**, 222–227.

BACK, J.W., NOTENBOOM, V., DE KONING, L.J. *ET AL.* (2002a). Identification of crosslinked peptides for protein interaction studies using mass spectrometry and 18 O-labelling. *Analytical Chemistry* **74**, 4417–4422.

BACK, J.W., SANZ, M.A., DE JONG, L. *ET AL.* (2002b). A structure for the yeast prohibiting complex: structure prediction and evidence from chemical crosslinking and mass spectrometry. *Protein Science* **11**, 2471–2478.

BACK, J.W., DE JONG, L., MUIJSERS, A.O. AND DE KOSTER, C.G. (2003). Chemical crosslinking and mass spectrometry for protein structural modelling. *Journal of Molecular Biology* **331**, 303–313.

BAUER, A. AND KUSTER, B. (2003). Affinity purification mass spectrometry. *European Journal of Biochemistry* **270**, 570–578.

BENNETT, K.L., KUSSMANN, M., BJORK, P. *ET AL.* (2000). Chemical crosslinking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping – a novel approach to assess intermolecular protein contacts. *Protein Science* **9**, 1503–1518.

BOGDANOV, B. AND SMITH, R.D. (2005). Proteomics by FTICR mass spectrometry: top down and bottom up. *Mass Spectrometry Reviews* **24**, 168–200.

BUIJS, J. AND FRANKLIN, G.C. (2005). SPR-MS in functional proteomics. *Briefings in Functional Genomics and Proteomics* **4**, 39–47.

BURLINGAME, A.L., BOYD, R.K. AND GASKELL, S.J. (1998). Mass spectrometry. *Analytical Chemistry* **70**, 647R–716R.

BUSENLEHNER, L.S. AND ARMSTRONG, R.N. (2005). Insights into enzyme structure and dynamics elucidated by amide H/D exchange mass spectrometry. *Archives of Biochemistry and Biophysics* **433**, 34–46.

CHEN, X., CHEN, Y.H. AND ANDERSON, V.E. (1999). Protein crosslinks: universal isolation and characterization by isotopic derivatization and electrospray ionization mass spectrometry. *Analytical Biochemistry* **273**, 192–203.

CHEN, Y., EBRIGHT, Y.W. AND EBRIGHT, R.H. (1994). Identification of the target of a transcription activator protein by protein–protein photo-crosslinking. *Science* **265**, 90–92.

CHERNUSHEVICH, I.V., LOBODA, A.V. AND THOMSON, B.A. (2001). An introduction to quadrupole-time-of-flight mass spectrometry. *Journal of Mass Spectrometry* **36**, 849–865.

CHU, F., SHAN, S.O., MOUSTAKAS, D.T. *ET AL.* (2004). Unravelling the interface of signal recognition particle and its receptor by using chemical crosslinking and tandem mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 16454–16459.

DELANO, W. (2002). Unravelling hot spots in binding interfaces: progress and challenges. *Current Opinion in Structural Biology* **12**, 14–20.

DEYEV, S.M., WAIBEL, R., LEBEDENKO, E., SCHUBIGER, A.P. AND PLUECKTHUN, A. (2003). Design of multivalent complexes using the barnase–barstar module. *Nature Biotechnology* **21**, 1486–1492.

DIHAZI, G.H. AND SINZ, A. (2003). Mapping low-resolution three-dimensional protein structures using chemical crosslinking and Fourier transform ion-cyclotron resonance mass spectrometry. *Rapid Communications in Mass Spectrometry* **17**, 2005–2014.

DREWES, G. AND BOUWMEESTER, T. (2003). Global approaches to protein–protein interactions. *Current Opinion in Cell Biology* **15**, 199–205.

EYLES, S.J. AND KALTASHOV, I.A. (2004). Methods to study protein dynamics and folding by mass spectrometry. *Methods* **34**, 88–99.

FANCY, D.A. (2000). Elucidation of protein–protein interactions using chemical crosslinking or label transfer techniques. *Current Opinion in Cell Biology* **4**, 28–33.

FANCY, D.A. AND KODADEK, T. (1998). A critical role for tyrosine residues in $His_6$Ni-mediated

protein crosslinking. *Biochemical and Biophysical Research Communications* **247**, 420–426.

FANCY, D.A. AND KODADEK, T. (1999). Chemistry for the analysis of protein–protein interactions: rapid and efficient crosslinking triggered by long wavelength light. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6020–6024.

FANCY, D.A., MELCHER, K., JOHNSTON, S.A. AND KODADEK, T. (1996). New chemistry for the study of multiprotein complexes: the six-histidine tag as a receptor for a protein crosslinking reagent. *Chemistry and Biology* **3**, 551–559.

FARMER, T.B. AND CAPRIOLI, R.M. (1998). Determination of protein–protein interactions by matrix-assisted laser desorption/ionization mass spectrometry. *Journal of Mass Spectrometry* **33**, 697–704.

FARRELL, I.S., TORONEY, R., HAZEN, J.L., MEHL, R.A. AND CHIN, J.W. (2005). Photo-crosslinking interacting proteins with a genetically encoded benzophenone. *Nature Methods* **2**, 377–384.

FIELDS, S. AND SONG, O. (1989). A novel genetic system to detect protein–protein interactions. *Nature* **340**, 245–246.

FIGEYS, D., MCBROOM, L.D. AND MORAN, M.F. (2001). Mass spectrometry for the study of protein–protein interactions. *Methods* **24**, 230–239.

FRAGOSO, G. AND HAGER, G.L. (1997). Analysis of *in vivo* nucleosome positions by determination of nucleosome-linker boundaries in crosslinked chromatin. *Methods* **11**, 246–252.

FUZESI, M., GOTTSCHALK, K.-E., LINDZEN, M. *ET AL.* (2005). Covalent crosslinks between the γ-subunit (FXYD2) and α and β subunits of Na,K-ATPase. Modelling the α–γ interaction. *Journal of Biological Chemistry* **280**, 18291–18301.

HANNAH, M.J., WEISS, U. AND HUTTNER, W.B. (1998). Differential extraction of proteins from paraformaldehyde-fixed cells: lessons from Synaptophysin and other membrane proteins. *Methods* **16**, 170–181.

HINO, N., OKAZAKI, Y., KOBAYASHI, T., HAYASHI, A., SAKAMOTO, K. AND YOKOYAMA, S. (2005). Protein photo-crosslinking in mammalian cells by site-specific incorporation of a photoreactive amino acid. *Nature Methods* **2**, 201–206.

HU, J. (2001). Model systems: studying molecular recognition using bacterial n-hybrid systems. *Trends in Microbiology* **9**, 219–222.

HUANG, B.X., KIM, H.-Y. AND DASS, C. (2004). Probing three-dimensional structure of bovine serum albumin by chemical crosslinking and mass spectrometry. *Journal of the American Society for Mass Spectrometry* **15**, 1237–1247.

ITOH, Y., CAI, K. AND KHORANA, H.G. (2001). Mapping of contact sites in complex formation between light-activated rhodopsin and transducin by covalent crosslinking: use of a chemically preactivated reagent. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 4883–4887.

JACKSON, V. (1999). Formaldehyde crosslinking for studying nucleosomal dynamics. *Methods* **17**, 125–139.

JONSCHER, K.R. AND YATES, J.R., 3RD (1997). The quadrupole ion trap mass spectrometer – a small solution to a big challenge. *Analytical Biochemistry* **244**, 1–15.

KALKHOF, S., IHLING, C., MECHTLER, K. AND SINZ, A. (2005). Chemical crosslinking and high-performance Fourier transform ion cyclotron resonance mass spectrometry for protein interaction analysis: application to a calmodulin/target peptide complex. *Analytical Chemistry* **77**, 495–503.

KELLEHER, N.L. (2004). Top-down proteomics. *Analytical Chemistry* **76**, 197A–203A.

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 Communications in Mass Spectrometry* **17**, 155–162.

LEGRAIN, P., WOJCIK, J. AND GAUTHIER, J.M. (2001). Protein–protein interaction maps: a lead towards cellular functions. *Trends in Genetics* **17**, 346–352.

LIPPINCOTT-SCHWARTZ, J. AND PATTERSON, G.H. (2003). Development and use of fluorescent protein markers in living cells. *Science* **300**, 87–91.

MARRIOTT, G. AND OTTL, J. (1998). Synthesis and applications of hetero-bifunctional photocleavable crosslinking reagents. *Methods in Enzymology* **291**, 155–175.

MARSHALL, A.G., HENDRICKSON, C.L. AND JACKSON, G.S. (1998). Fourier transform ion cyclotron resonance mass spectrometry. *Mass Spectrometry Reviews* **17**, 1–35.

MELCHER, K. (2004). New chemical crosslinking methods for the identification of transient protein–protein interactions with multiprotein complexes. *Current Protein and Peptide Science* **4**, 287–296.

METZ, B., KERSTEN, G.F., HOOGERHOUT, P. *ET AL.* (2004). Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. *Journal of Biological Chemistry* **279**, 6235–6243.

MO, W. AND KARGER, B.L. (2002). Analytical aspects of mass spectrometry and proteomics. *Current Opinion in Chemical Biology* **6**, 666–675.

MUELLER, D.R., SCHINDLER, P., TOWBIN, H. *ET AL.* (2001). Isotope-tagged crosslinking reagents. A new tool in mass spectrometric protein interaction analysis. *Analytical Chemistry* **73**, 1927–1934.

NEDELKOV, D. AND NELSON, R.W. (2003). Surface plasmon resonance mass spectrometry: recent progress and outlooks. *Trends in Biotechnology* **21**, 301–305.

ORLANDO, V., STRUTT, H. AND PARO, R. (1997). Analysis of chromatin structure by *in vivo* formaldehyde crosslinking. *Methods* **11**, 205–214.

PANDEY, A. AND MANN, M. (2000). Proteomics to study genes and genomes. *Nature* **405**, 837–846.

PEARSON, K.M., PANNELL, L.K. AND FALES, H.M. (2002). Intramolecular crosslinking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. *Rapid Communications in Mass Spectrometry* **16**, 149–159.

PIEHLER, J. (2005). New methodologies for measuring protein interactions *in vivo* and *in vitro*. *Current Opinion in Structural Biology* **15**, 4–14.

RAPPSILBER, J., SINIOSSOGLOU, S., HURT, E.C. AND MANN, M. (2000). A generic strategy to analyse the spatial organization of multi-protein complexes by crosslinking and mass spectrometry. *Analytical Chemistry* **72**, 267–275.

RUSSELL, R.B., ALBER, F., ALOY, P. *ET AL.* (2004). A structural perspective on protein–protein interactions. *Current Opinion in Structural Biology* **14**, 313–324.

SALI, A., GLAESER, R., EARNEST, T. AND BAUMEISTER, W. (2003). From words to literature in structural proteomics. *Nature* **422**, 216–225.

SCALONI, A., MIRAGLIA, N., ORRU, S. *ET AL.* (1998). Topology of the calmodulin–melittin complex. *Journal of Molecular Biology* **277**, 945–958.

SCHMITT-ULMS, G., LEGNAME, G., BALDWIN, M.A. *ET AL.* (2001). Binding of neural cell adhesion molecules (N-CAMs) to the cellular prion protein. *Journal of Molecular Biology* **314**, 1209–1225.

SCHMITT-ULMS, G., HANSEN, K., LIU, J. *ET AL.* (2004). Time-controlled transcardiac perfusion cross-linking for the study of protein interactions in complex tissues. *Nature Biotechnology* **22**, 724–731.

SINZ, A. (2003). Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *Journal of Mass Spectrometry* **38**, 1225–1237.

SINZ, A. AND WANG, K. (2001). Mapping protein interfaces with a fluorogenic cross-linker and mass spectrometry: application to nebulin–calmodulin complexes. *Biochemistry* **40**, 7903–7913.

SMITH, G.R. AND STERNBERG, J.E. (2002). Prediction of protein–protein interactions by docking methods. *Current Opinion in Structural Biology* **12**, 28–35.

SOBOTT, F. AND ROBINSON, C.V. (2002). Protein complexes gain momentum. *Current Opinion in Structural Biology* **12**, 729–734.

STAGLJAR, I. AND FIELDS, S. (2002). Analysis of membrane protein interactions using yeast-based technologies. *Trends in Biochemical Sciences* **27**, 559–563.

SUCHANEK, M., RADZIKOWSK, A. AND THIELE, C. (2005). Photo-leucine and photo-methionine allow identification of protein–protein interactions in living cells. *Nature Methods* **2**, 261–267.

TANG, X., MUNSKE, G.R., SIEMS, W.F. AND BRUCE, J.E. (2005). Mass spectrometry identifiable crosslinking strategy for studying protein–protein interactions. *Analytical Chemistry* **77**, 311–318.

TAVERNER, T., HALL, N.E., O'HAIR, A.J. AND SIMPSON, R.J. (2002). Characterization of an antagonist interleukin-6 dimer by stable isotope labelling, cross-linking, and mass spectrometry. *Journal of Biological Chemistry* **277**, 46487–46492.

TRAKSELIS, M.A., ALLEY, S.C. AND ISHMAEL, F.T. (2005). Identification and mapping of protein–protein interactions by a combination of crosslinking, cleavage, and proteomics. *Bioconjugate Chemistry* **16**, 741–750.

TRESTER-ZEDLITZ, M., KAMADA, K., BURLEY, S.K., FENYO, D., CHAIT, B.T. AND MUIR, T.W. (2003). A modular crosslinking approach for exploring protein interactions. *Journal of the American Chemical Society* **125**, 2416–2425.

TSIEN, R.Y. (1998). The green fluorescent protein. *Annual Review of Biochemistry* **67**, 509–544.

VASILESCU, J., GUO, X.J. AND KAST, J. (2004). Identification of protein–protein interactions using *in vivo* crosslinking and mass spectrometry. *Proteomics* **12**, 3845–3854.

WALLRABE, H. AND PERIASAMY, A. (2005). Imaging protein molecules using FRET and FLIM microscopy. *Current Opinion in Biotechnology* **16**, 19–27.

WARNOCK, D., FAHY, E. AND TAYLOR, S.W. (2004). Identification of protein associations in organelles using mass spectrometry-based proteomics. *Mass Spectrometry Reviews* **23**, 259–280.

WELLS, J. AND FRANHAM, P.J. (2002). Characterizing transcription factor binding sites using formaldehyde crosslinking and immunoprecipitation. *Methods* **26**, 48–56.

WONG, S.S. AND WONG, L.J. (1992). Chemical crosslinking and the stabilization of proteins and enzymes. *Enzyme and Microbial Technology* **14**, 866–874.

YOUNG, M.M., TANG, N., HEMPEL, J.C. *ET AL.* (2000). High-throughput protein fold identification by using experimental constraints derived from intramolecular crosslinks and mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 5802–5806.

ZAL, T. AND GASCOIGNE, N.R. (2004). Using live FRET imaging to reveal early protein–protein interactions during T cell activation. *Current Opinion in Immunology* **16**, 674–683.