Journal of Pharmaceutical and Scientific Innovation
www.jpsionline.com
Review Article

CHEMICAL CROSSLINKING OF PROTEINS: A REVIEW

Hemaprabha. E*
*Wayne State University School of Medicine, 421 E Canfield Street, Elliman Building, Detroit, Michigan-48201
Email: hemabioteck@gmail.com

Received on: 17/12/11 Revised on: 22/01/12 Accepted on: 25/01/12

ABSTRACT
Protein-protein interactions comprise the underlying molecular mechanism of a multitude of complex biological processes. Molecular chaperone complexes represent a very good example of multi-protein assemblies where the cooperative action of several components is required for folding/unfolding and transmembrane trafficking of proteins. Signal transduction pathways are illustrative of highly complex protein interaction networks displaying transient and semi-stable association of a series of proteins. Crosslinking is the process of chemically joining two or more molecules by a covalent bond. Crosslinking reagents contain two or more reactive ends that are capable of attaching to specific functional groups (primary amines, sulfhydryls, etc.) on proteins or other molecules.This article describes some of the different crosslinkers and their use in protein protein interactions.
KEYWORDS: Protein-protein interactions, Protein structure, Crosslinking, Crosslinker

INTRODUCTION

Crosslinking is the process of chemically joining two or more molecules by a covalent bond. Modification involves attaching or cleaving chemical groups to alter the solubility or other properties of the original molecule. The entire set of crosslinking and modification methods for use with proteins and other biomolecules in biological research is often called "bioconjugation". Covalent modification and crosslinking of proteins depends on the availability of particular chemicals that are capable of reacting with the specific kinds of functional groups that exist in proteins. In addition, protein function and structure are either the direct focus of study or they must be preserved if a modified protein is to be useful in a technique. Therefore, the composition and structure of proteins, and the potential effects of modification reagents on protein structure and function, must be considered. This review focuses on different crosslinking reagents used for protein crosslinking studies with a case study in neurobiology.

PROTEIN STRUCTURE:

Primary Structure:

The primary structure refers to amino acid sequence of the polypeptide chain. The primary structure is held together by covalent or peptide bonds, which are made during the process of protein biosynthesis or translation. The two ends of the polypeptide chain are referred to as the carboxyl terminus (C-terminus) and the amino terminus (N-terminus) based on the nature of the free group on each extremity. Counting of residues always starts at the N-terminal end (NH_2-group), which is the end where the amino group is not involved in a peptide bond. The primary structure of a protein is determined by the gene corresponding to the protein. Post-translational modifications such as disulfide formation, phosphorylations and glycosylations are usually also considered a part of the primary structure, and cannot be read from the gene[1].

Amino Acid Residues:

Each α-amino acid consists of a backbone part that is present in all the amino acid types, and a side chain that is unique to each type of residue. An exception from this rule is proline. Because the carbon atom is bound to four different groups it is chiral, however only one of the isomers occur in biological proteins. Glycine however, is not chiral since its side chain is a hydrogen atom.

Secondary Structure:

Secondary structure refers to highly regular local sub-structures. Two main types of secondary structure, the alpha helix and the beta strand, were suggested in 1951 by Linus Pauling and coworkers[2]. These secondary structures are defined by patterns of hydrogen bonds between the main-chain peptide groups. Both the alpha helix and the beta-sheet represent a way of saturating all the hydrogen bond donors and acceptors in the peptide backbone. Some parts of the protein are ordered but do not form any regular structures. They should not be confused with random coil, an unfolded polypeptide chain lacking any fixed three-dimensional structure. Several sequential secondary structures may form a "supersecondary unit"[3].

Tertiary Structure:

Tertiary structure refers to three-dimensional structure of a single protein molecule. The alpha-helices and beta-sheets are folded into a compact globule. The folding is driven by the non-specifichydrophobic interactions (the burial of hydrophobic residues from water), but the structure is stable only when the parts of a protein domain are locked into place by specific tertiary interactions, such as salt bridges, hydrogen bonds, and the tight packing of side chains and disulfide bonds. The disulfide bonds are extremely rare in cytosolic proteins, since the cytosol is generally a reducing environment.

Quaternary Structure:

Quaternary structure is a larger assembly of several protein molecules or polypeptide chains, usually called subunits in this context. The quaternary structure is stabilized by the same non-covalent interactions and disulfide bonds as the tertiary structure. Complexes of two or more polypeptides (i.e. multiple subunits) are called multimers. Specifically it would be called a dimer if it contains two subunits, a trimer if it contains three subunits, and a tetramer if it contains four subunits. The subunits are frequently related to one another by symmetry operations, such as a 2-fold axis in a dimer. Multimers made up of identical subunits are referred to with a prefix of "homo-" (e.g. a homotetramer) and those made up of different subunits are referred to with a prefix of "hetero-" (e.g. a heterotetramer, such as the two alpha and two beta chains of hemoglobin).

Domains, Motifs, And Folds In Protein Structure:

Proteins are frequently described as consisting from several structural units.

A structural domain is an element of the protein's overall structure that is self-stabilizing and often folds independently of the rest of the protein chain. Many domains are not unique to the protein products of one gene or one gene family but instead appear in a variety of proteins. Domains often are named and singled out because they figure prominently in the biological function of the protein they belong to; for example, the "calcium-binding domain of calmodulin". Because they are independently stable, domains can be "swapped" by genetic engineering between one protein and another to make chimeras.

The structural and sequence motifs refer to short segments of protein three-dimensional structure or amino acid sequence that were found in a large number of different proteins.

The super secondary structure refers to a specific combination of secondary structure elements, such as beta-alpha-beta units or helix-turn-helix motif. Some of them may be also referred to as structural motifs.

Protein fold refers to the general protein architecture, like helix bundle, beta-barrel, Rossman fold or different "folds" provided in the Structural Classification of Proteins database.

The complete structure of functioning proteins involves more than polypeptide chains in the four levels of structure. Various covalent modifications often occur, either during or after assembly of the polypeptide chain. Most proteins undergo co- and /or post-translational modifications. Examples include phosphorylation (of serine, threonine or tyrosine residues), glycosylation and ubiquination.

Knowledge of these native modifications is extremely important because they may alter physical and chemical properties, folding, conformation distribution, stability, activity, and consequently, function of the proteins. Because the structure of a protein dictates its biological activity, characterization of protein structure continues to be an important area of research. Proteins are relatively easy molecules to manipulate, and protein crosslinking and chemical modification methods are commonly used to determine the roles of individual amino acid side chains in the physical, chemical and biological properties of proteins. Just four protein chemical targets account for the vast majority of crosslinking and chemical modification techniques:

- Primary amines (–NH2): This group exists at the N-terminus of each polypeptide chain and in the side chain of lysine (Lys, K) residues.
- Carboxyls (–COOH): This group exists at the C-terminus of each polypeptide chain and in the side chains of aspartic acid (Asp, D) and glutamic acid (Glu, E).
- Sulfhydryls (–SH): This group exists in the side chain of cysteine (Cys, C). Often, as part of a protein's secondary or tertiary structure, cysteines are joined together between their side chains via disulfide bonds (–S–S–).
- Carbonyls (–CHO): These aldehyde groups can be created by oxidizing carbohydrate groups in glycoproteins. (Source: www.piercenet.com)

For each of these protein functional-group targets, there exist one to several types of reactive groups that are capable of targeting them and have been used as the basis for synthesizing crosslinking and modification reagents.

CROSSLINKING PROTEINS:

Chemical crosslinking offers a direct method of identifying both transient and stable interactions. This technique involves the formation of covalent bonds between two proteins by using bifunctional reagents containing reactive end groups that react with functional groups—such as primary amines and sulfhydryls—of amino acid residues. In the cell, a single protein may often engage in transient interaction with a variety of other partners in a given pathway. Where purified proteins are available, chemical crosslinking is the ideal strategy for an unambiguous demonstration of protein-protein interactions, *in vitro*. If two proteins physically interact with each other, they can be covalently cross-linked. The formation of crosslinks between two distinct proteins is a direct and convincing evidence of their close proximity. In addition to information on the identity of the interacting proteins, crosslinking experiments can reveal the regions of contact between them.

Thus, crosslinking is used for many purposes, including to:

- Stabilize protein tertiary and quaternary structure for analysis.
- Capture and identify unknown protein interactors or interaction domains.
- Conjugate an enzyme or tag to an antibody or other purified protein.
- Immobilize antibodies or other proteins for assays or affinity-purification.
- Attach peptides to larger "carrier" proteins to facilitate handling/storage.

Crosslinkers are selected on the basis of their chemical reactivities (i.e., specificity for particular function groups) and other chemical properties that facilitate their use in different specific applications:

1. Chemical specificity, including whether the reagent has the same or different reactive groups at either end (i.e., does it have a homobifunctional or heterobifunctional structure?)
2. Spacer arm length, including whether the arm is cleavable (i.e., can the linkage be reversed or broken when desired?)
3. Water-solubility and cell membrane permeability (i.e., can the reagent be expected to permeate into cells and/or crosslink hydrophobic proteins within membranes?)
4. Spontaneously reactive or photo-reactive groups (i.e., will the reagent react as soon as it is added to a sample or can its reaction be activated at a specific time?)

Crosslinker Reactive Groups:

A number of chemical reactive groups have been characterized and used to target the main kinds of protein functional groups. Many different crosslinking reagents can be synthesized when different combinations of two are more of these reactive groups are incorporated into one molecule. When combined with different sizes and types of chemical "backbones" (called spacer arms because they comprise the space and distance between respective reactive ends), the number of possible crosslinking compounds is enormous. Table 1 provides the list of protein cross linker reactive group.

Table 1: The list of protein crosslinker reactive group

Functional Group Target	Reactive Group
Carboxyl (directly to amine)	Carbodiimide (e.g., EDC)
Amine	NHS ester Imidoester PFP ester Hydroxymethyl phosphine
Sulfhydryl	Maleimide Haloacetyl (Bromo- or Iodo-) Pyridyldisulfide Vinyl sulfone
Aldehyde (Carbonyls) i.e., oxidized carbohydrates	Hydrazide
Any Group (Nonselective)	Diazirine (Photo-reactive) Aryl Azide (Photo-reactive)
Hydroxyl (non-aqueous)	Isocyanate

(Source: www.piercenet.com)

Homobifunctional And Heterobifunctional Crosslinkers:

Crosslinkers (CL) are either homo- or hetero-bifunctional reagents with identical or non-identical reactive groups, respectively, permitting the establishment of inter- as well as intra-molecular crosslinkages. Inter-subunit crosslinks have been used for determination of the quaternary structure and arrangement of subunits within homo-oligomeric proteins and intra-subunit crosslinks for maintenance of stable tertiary structure. Ligand-induced conformational changes in proteins can be analyzed by a comparison of the rate/extent of crosslinking in the ligand-bound *versus* the unliganded states. In studies of hetero-oligomeric enzymes or multiprotein complexes—containing several different polypeptides-a crosslinking offers a reliable tool for unraveling spatial relationships of the components. A variety of crosslinkers are commercially available from major suppliers such as Pierce, Molecular probes, and Sigma. For an insightful treatment of the properties and underlying chemistry of crosslinkers, and methods of introducing reactive groups into proteins, the all-time classic by Means and Feeney (1971) can be consulted[4].

Homobifunctional Crosslinkers:

Homobifunctional crosslinkers have identical reactive groups at either end of a spacer arm, and generally they must be used in one-step reaction procedures to randomly "fix" or polymerize molecules containing like functional groups. For example, adding an amine-to-amine crosslinker to a cell lysate will result in random conjugation of protein subunits, interacting proteins and any other polypeptides whose lysine side chains happen to be near each other in the solution. Table 2 gives a list of homobifunctional crosslinkers.

Table 2: List of Homobifunctional crosslinkers

A list of crosslinking reagents that have the same type of reactive group at either end. Reagents are classified by what chemical groups they crosslink (left column) and their chemical composition (middle column). Products are listed in order of increasing length within each cell.

Crosslinking Target	Crosslinker Reactive Groups, Features	Example Products
Amine-to-Amine	NHS esters	DSG DSS BS3 TSAT (trifunctional)
	NHS esters, PEG spacer	BS(PEG)5 BS(PEG)9
	NHS esters, thiol-cleavable	DSP DTSSP
	NHS esters, misc-cleavable	DST BSOCOES EGS Sulfo-EGS
	Imidoesters	DMA DMP DMS
	Imidoesters, thiol-cleavable	DTBP
	Other	DFDNB
Sulfhydryl-to-Sulfhydryl	Maleimides	BMOE BMB BMH TMEA (trifunctional)
	Maleimides, PEG spacer	BM(PEG)2 BM(PEG)3
	Maleimides, cleavable	BMDB DTME
Nonselective	Aryl azides	BASED (thiol-cleavable)

(Source: www.piercenet.com)

Heterobifunctional Crosslinkers:

Heterobifunctional crosslinkers possess different reactive groups at either end. These reagents not only allow for single-step conjugation of molecules that have the respective target functional groups, but they also allow for sequential (two-step) conjugations that minimize undesirable polymerization or self-conjugation.Table 3 gives a list of homobifunctional crosslinkers.

Table 3: List of Heterobifunctional crosslinkers

A list of crosslinking reagents that have the different reactive groups at either end. Reagents are classified by what chemical groups they crosslink (left column) and their chemical composition (middle column). Products are listed in order of increasing length within each cell.

Crosslinking Targets	Crosslinker Reactive Groups, Features	Example Products
Amine-to-Sulfhydryl	NHS ester / Maleimide	AMAS BMPS GMBS and Sulfo-GMBS MBS and Sulfo-MBS SMCC and Sulfo-SMCC EMCS and Sulfo-EMCS SMPB and Sulfo-SMPB SMPH LC-SMCC Sulfo-KMUS
	NHS ester / Maleimide, PEG spacer	SM(PEG)2 SM(PEG)4 SM(PEG)6 SM(PEG)8 SM(PEG)12 SM(PEG)24
	NHS ester / Pyridyldithiol, cleavable	SPDP LC-SPDP and Sulfo-LC-SPDP SMPT Sulfo-LC-SMPT
	NHS esters / Haloacetyl	SIA SBAP SIAB Sulfo-SIAB
Amine-to-Nonselective	NHS ester / Aryl Azide	NHS-ASA ANB-NOS Sulfo-SANPAH
	NHS ester / Aryl Azide, cleavable	Sulfo-SAND
	NHS ester / Diazirine	SDA and Sulfo-SDA LC-SDA and Sulfo-LC-SDA
	NHS ester / Diazirine, cleavable	SDAD and Sulfo-SDAD
Amine-to-Carboxyl	Carbodiimide	DCC EDC
Sulfhydryl-to-Carbohydrate	Maleimide / Hydrazide	BMPH EMCH MPBH KMUH
	Pyridyldithiol / Hydrazide	PDPH
Hydroxyl-to-Sulfhydryl	Isocyanate / Maleimide	PMPI
Amine-to-DNA	NHS ester / Psoralen	SPB

(Source: www.piercenet.com)

PROTEIN CROSSLINKING APPLICATIONS:

Crosslinking reagents are used in a variety of techniques to assist in determining partners and domains of protein interactions, three-dimensional structures of proteins, and molecular associations in cell membranes. They are also used to immobilize proteins on solid supports for affinity purification, to conjugate haptens to carrier proteins for immunization, and to prepare antibody-enzyme conjugates for detection procedures.

1. Protein Structural and Subunit Analysis
2. Protein Interaction Analysis
3.Immunotoxin Construction
4.Immunogen Preparation
5.Protein-Protein Conjugation
6.Solid-Phase Immobilization
7.DNA/RNA Crosslinking to Proteins
8.Other Applications

There are many additional applications for crosslinkers that are either antiquated methods, new technologies or for more specialized needs. Older methods for peptide synthesis involve use of carbodiimide crosslinkers such as DCC and EDC for the step-wise addition of individual amino acids to support bound peptides. Crosslinkers such as glutaraldehyde and dimethylpimelimidate have been used for tissue fixation.(Source: www.piercenet.com)

APPLICATIONS OF CROSSLINKING IN NEUROBIOLOGY- A CASE STUDY:

In the present study crosslinker DTSSP is used to crosslink the Human MPZ protein. Myelin protein-zero is the major structural protein of peripheral myelin[5,6]. Myelin protein zero (P0) is a glycoprotein associated with Charcot-Marie-Tooth disease and Dejerine-Sottas disease[7]. MPZ, a transmembrane protein of 219 amino acids, is a member of the immunoglobulin gene superfamily[8]. It has a single immunoglobulin-like extracellular domain of 124 amino acids, a single transmembrane domain of 25 amino acids, and a single cytoplasmic domain of 69 amino acids [9,10]. Smith, 1977 has described in detail the non-covalent crosslinking of lipid bilayers of Myelin basic protein[11].

3,3′-Dithiobis(sulfosuccinimidylpropionate) (DTSSP) contains an amine-reactive *N*-hydroxysulfosuccinimide (sulfo-NHS) ester at each end of an 8-carbon spacer arm. Sulfo-NHS esters react with primary amines at pH 7-9 to form stable amide bonds, along with release of the N-hydroxysulfosuccinimide leaving group.

DTSSP Highlights:

- Amine reactive sulfo-NHS ester reacts rapidly with any primary amine-containing molecule
- Disulfide bond in the spacer arm is readily cleaved by 10-50 mM DTT or TCEP at pH 8.5
- Spacer arm is also cleaved with 5% β-mercaptoethanol in SDS-PAGE sample loading buffer at 100°C for 5 minutes
- Cleavable crosslinker allows separation of crosslinked products
- DTSSP crosslinker is not membrane permeable, so cell-surface crosslinking can be done

Applications For DTSSP:

- Chemical crosslinking of cell surface proteins prior to cell lysis and immunoprecipitation
- “Fix” protein interactions to allow identification of weak or transient protein interactions
- Protein crosslinking to create bioconjugates via single-step reactions
- Immobilize proteins onto amine-coated surfaces

Cross-Linking of PO Protein:

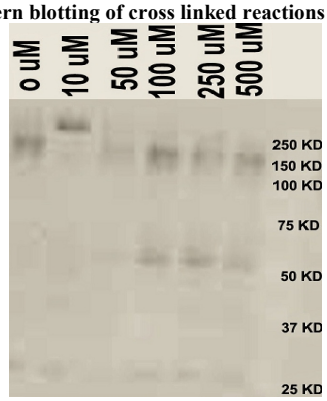
Cells were washed twice with reaction buffer to remove media. The reaction buffer contains PBS- Phosphate Buffered Saline (PBS; e.g., 0.1M phosphate, 0.15M NaCl; pH 7.2). The cells were incubated with 0.1 mM DTSSP [3,3′-dithiobis(sulfosuccinimidylpropionate); Pierce] for 30 min at room temperature. The reaction was quenched with 1 M Tristo a final concentration of 10-20mM for

15 min as stop solution, and the whole reaction was analyzed by SDS/PAGE[12].

P0, the major protein of myelin of theperipheral nervous system (PNS), functions to mediate membraneadhesion within the myelin sheath, also forms a head-totaildimer. Western Blot analysis of the crosslinking reactions showed

presence of dimer with increased concentration of the crosslinker DTSSP. P0 forms monomer at 25 kDa and dimer at 50 kDa (Fig 1). The samples will be purified by gel filtration and their masses will be determined through mass spectroscopy studies.

Figure 1: Western blotting of cross linked reactions of MPZ protein



ACKNOWLEDGMENT:

I would like to thank Dr. John Kamholz, School of Medicine,Wayne State University for his technical support and guidance with regard to the case study.

REFERENCES:

1. Govindarajan S, Recabarren R, Goldstein RA. Estimating the total number of protein folds. Proteins1999;35 (4): 408–414.
2. Pauling L, Corey RB, Branson HR.The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. ProcNatlAcadSciUSA1951; 37 (4): 205–211.
3. Chiang YS, Gelfand TI, Kister AE, Gelfand IM.New classification of supersecondary structures of sandwich-like proteins uncovers strict patterns of strand assemblage. Proteins 2007; 68 (4): 915–921.
4. Means GA., and Feeney RE. Chemical Modification of Proteins. Holden Day, Inc. San Francisco, Cambridge, London, Amsterdam.1971; p 220.
5. Gaboreanu AM., Hrstka R, Xu W, Shy M, Kamholz J, Lilien J *et al*. Myelin protein zero/P0 phosphorylation and function require an adaptor protein linking it to RACK1 and PKCα. JCB2007; 177: 707-716.
6. Kamholz, J, Menichella D, Jani, A, Garbern, J, Lewis, R.A, Krajewski*et al*.Charcot-Marie-Tooth disease type 1: Molecular pathogenesis to gene therapy. Brain2000; 123: 222-233.
7. Shy ME, JaniA, Krajewski K, Lewis RA, Li J, Shy RR*et al.*Phenotypic clustering in *MPZ* mutations. Brain, 2004; 127:371–384.
8. Kamholz J, Spielman R, Gogolin K, Modi W, O'brien S, and Lazzarini R. The Human Myelin-Basic-Protein Gene: Chromosomal Localization and RFLP Analysis. Am. J. Hum. Genet, 1987; 40:365-373.
9. Clements CS., Reid HH, Beddoe T, Tynan FE, Perugini MA, Johns TG*et al*. The crystal structure of myelin oligodendrocyte glycoprotein, a key autoantigen in multiple sclerosis.ProcNatlAcadSciUSA, 2003; 100: 11059–11064.
10. Uyemura K, Kitamura K and Miura M. Structure and molecular biology of P0 protein. In: Martenson RE, editor. Myelin: biology and chemistry. Boca Raton (FL): CRC Press. 1992; 481-508.
11. Smith R. Non-covalent cross-linking of lipid bilayers by myelin basic protein. A possible role in myelin formation. BiochimicaetBiophysicaActa (BBA) - Biomembranes, 1997; 470. 170-184.
12. Laemmli, U.K. Cleavage of structural proteins during assembly of the head of Bacteriophage T4. Nature (London) 1970; 227:680-685.